Evidence-based Practice Center Systematic Review Protocol

Project Title: Adverse Effects of First-line Pharmacologic Treatments of Major Depression in Older Adults

I. Background and Objectives for the Systematic Review

Depression is a common psychiatric disease in older adults. Approximately 15–20 percent of adults older than age 65 in the United States have experienced depression.[1]

Multiple systematic reviews have shown that antidepressant medications are better than placebo for treating depression in older patients.[2] However, effects are modest and side effects are common. Depression treatment in older patients may be complicated by their other comorbid conditions, age-related physiologic changes, and potential interactions with other medications. As a result, certain treatment options may be contraindicated, inadequately dosed, or poorly tolerated. In addition, clinicians must consider the balance of the risks and benefits of antidepressant medications, especially in comparison to other treatment options. While the effectiveness of interventions for treatment of depression in the elderly has been previously reviewed[2] the harms of commonly used treatments of depression have not been well quantified in older adults.

The American Psychiatric Association published guidelines for major depressive disorder (MDD) in 2010[3] and the American College of Physicians (ACP) published their guidelines in 2016.[4] Antidepressants are recommended as an initial treatment option. The guidelines cite similar efficacy within and between pharmacologic classes; thus the recommendation is to choose a medication based on adverse event profiles, patient preferences, dosing schedules, costs, and drug interactions. With all things considered, the guidelines suggest that selective serotonin reuptake inhibitors (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), bupropion or mirtazapine are optimal initial treatment choices for the majority of patients.[3] Although tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) are recognized as pharmacologic classes that may be used to treat depression, these classes are not considered first-line due to safety concerns and drug properties (e.g., interactions, dosing, dietary restrictions). Specific to treating depression in older patients, the APA (American Psychiatric Association) guidelines suggest treatment considerations follow those for younger patients.[3] However, several cautionary statements regarding side effect profiles for the primary pharmacologic treatments in older populations are made. Regimens should be adjusted for metabolic changes and potential drug interactions. SSRIs, SNRIs and other antidepressants are favored over TCAs and MAOIs due to orthostatic hypotension and cholinergic blockade. SSRIs are noted to increase the risk of syndrome of inappropriate antidiuretic hormone (SIADH) in older patients compared to other antidepressants.[3]

Impetus for the Review

The American Geriatric Society (AGS) regularly compiles a list of medications that should be used with caution in older individuals based on the Beers Criteria. Medications

on the list are best avoided by those with specific conditions, or used with caution, at lower doses, or with careful monitoring. In 2015, this list recommended that clinicians avoid prescribing SSRIs and TCAs in older adults with a history of falls or fractures.[5] However they noted that there may be situations when use of these medications may be appropriate and clinicians and patients must carefully weigh both benefits and potential harms.[6] The AGS suggests that SNRIs and bupropion are alternatives to TCAs and SSRIs.[7] However, the AGS also recommended using SSRIs and SNRIs with caution due to the potential to exacerbate or cause SIADH or hyponatremia.[5]

Given these concerns of potential adverse events in the older population with drugs commonly recommended to treat MDD, clinicians may be left selecting therapy based on comparative adverse effects. This review seeks to systematically review the comparative adverse effects of first-line pharmacologic antidepressants for treatment in MDD older adults.

II. The Key Questions

Draft key questions (KQs) and a contextual question (CQ) were posted for public comment in August 2017 prior to the topic refinement phase. Comments from the public, AHRQ, the AGS and Key Informant (KI) Panel were considered by the EPC during topic refinement and the following revisions were made. We removed the specific age threshold defining "older adults" from the question wording and instead specify age based criteria in the PICOTS given below. We replaced the term "harms" with "adverse effects" based on input from the KI panel suggesting that the term "harms" has a more negative connotation by the public. We added the term "pharmacologic" in front of "treatment" since this review is focused on pharmacologic treatments and will not include nonpharmacologic treatments or complimentary alternative medicines.

Key Question 1: In older adults with major depressive disorder, what are the adverse effects and comparative adverse effects of first-line pharmacologic treatments?

Key Question 2: In subgroups of older adults (e.g., age, sex, race, comorbidities) with major depressive disorder, what are the adverse effects and comparative adverse effects of first-line pharmacologic treatments?

Contextual Question: In older adults with major depressive disorder, including subgroups of interest, what is the effectiveness and comparative effectiveness of first-line pharmacologic treatments?

For both KQs and the CQ, the following PICOTS criteria apply:

Population(s):
The population of interest is "older adults", defined as 65 years of age and older with MDD. This age is consistent with the cutpoint used by the AGS in the Beers Criteria, the qualifying age for Medicare benefits, and input of the KI panel.

This review is focused on MDD. While diagnosis of MDD through DSM criteria or ICD codes would be most rigorous, we anticipate identification of "depression" in observational studies using a variety of validated tools and also patient self-report. Although these latter strategies are less rigorous, they will be considered for inclusion and described in evidence tables.

We will exclude studies that focus enrollment solely on one of the given patient populations 1) patients with MDD and comorbid seizures 2) patients with MDD and comorbid psychiatric conditions with the exception of anxiety; 3) patients with a specific subtype of MDD (e.g., catatonic, melancholic, psychotic, or atypical features) rather than MDD generally; or 4) patients with bipolar depression.

The subgroups of interest are those that were decided to possibly inform further stratification of older adults' risk for the adverse effects of interest. Subgroups include:
- Age group (65 to 74y, 75 to 84y, and ≥85y)
- Gender
- Race or ethnicity
- Risk of falls or history of fracture
- Dementia or cognitive impairment
- Nursing home setting
- ≥2 physical (i.e. non-psychiatric) comorbidities
- History of substance abuse
- Frailty
- Early versus late onset MDD
- Polypharmacy, defined as 5 or more concurrent prescription medications[8]
- Concurrent use of one other medication with CNS activity, defined as antipsychotics, benzodiazepines, nonbenzodiazepine hypnotics, and opioids[5]

Interventions:
We are interested in first-line pharmacologic antidepressant treatments of MDD, as single interventions (**Table 1**), categorized according to their mechanism of action. Interventions listed as an SSRI or SNRI will be evaluated on a class-basis. Interventions that are listed as "other" have a unique mechanism and will be evaluated individually, not as a class.

Table 1. First-line Pharmacologic Treatments for MDD in Older Adults

Class	Drugs
SSRI	Paroxetine, sertraline, citalopram, escitalopram, fluoxetine, fluvoxamine
SNRI	Venlafaxine, desvenlafaxine, duloxetine, milnacipran, levomilnacipran
Other	Bupropion, mirtazapine, trazodone, vilazodone, vortioxetine

We will exclude studies that evaluate non-pharmacologic interventions, complementary alternative medicines, pharmacologic therapies not listed in Table 1 or any combinations of therapies (pharmacologic or non-pharmacologic) for MDD treatment.

Comparators:
We are interested in direct comparisons of eligible interventions (Table 1) with a pharmacologic antidepressant for MDD (as listed in Table 1 or a TCA or MAOI) evaluated as a single intervention, placebo, or nonpharmacologic therapies. Nonpharmacologic therapies of interest include non-invasive psychotherapy-based interventions such as CBT, interpersonal psychotherapy, problem solving therapy, psychodynamic or supportive therapy, behavioral therapies, journaling as well as

exercise. We will include data for within class comparisons of SSRIs and SNRIs. We will exclude complementary and alternative medicine or combination therapies.

Outcomes:

We are interested in the following adverse effects for KQ1 and KQ2:

- Any adverse event, as in those who experienced an adverse events
- Withdrawal due to adverse events
- Serious adverse events, as defined per the study
- Hospitalization
- Emergency room visit
- Falls
- Fractures
- Syndrome of inappropriate anti-diuretic hormone (SIADH) or hyponatremia
- Bleeding (any reported bleeding or bruising)
- QTc prolongation
- Arrhythmias
- Changes in weight
- Changes in blood pressure and orthostatic hypotension
- Cognitive impairment/confusion
- Suicide/suicide attempt
- Suicidal thoughts
- Seizures
- Mortality

We are interested in the following efficacy outcomes for the CQ:

- Response: commonly defined as a 50% or more improvement on either the HAM-D or MADRS scales
- Remission: commonly defined as absence or near absence of depressive symptoms, or a HAM-D score of <8.
- Relapse: as defined per the individual study.
- Quality of life: as defined per the individual study
- Changes in functional status: as defined per the individual study, including activities of daily living (e.g., bathing, dressing, grooming, feeding oneself), instrumental activities of daily living (e.g., using the telephone, preparing meals, managing finances, taking medication, shopping, or physical measures (e.g., walking distance or stair and climbing).
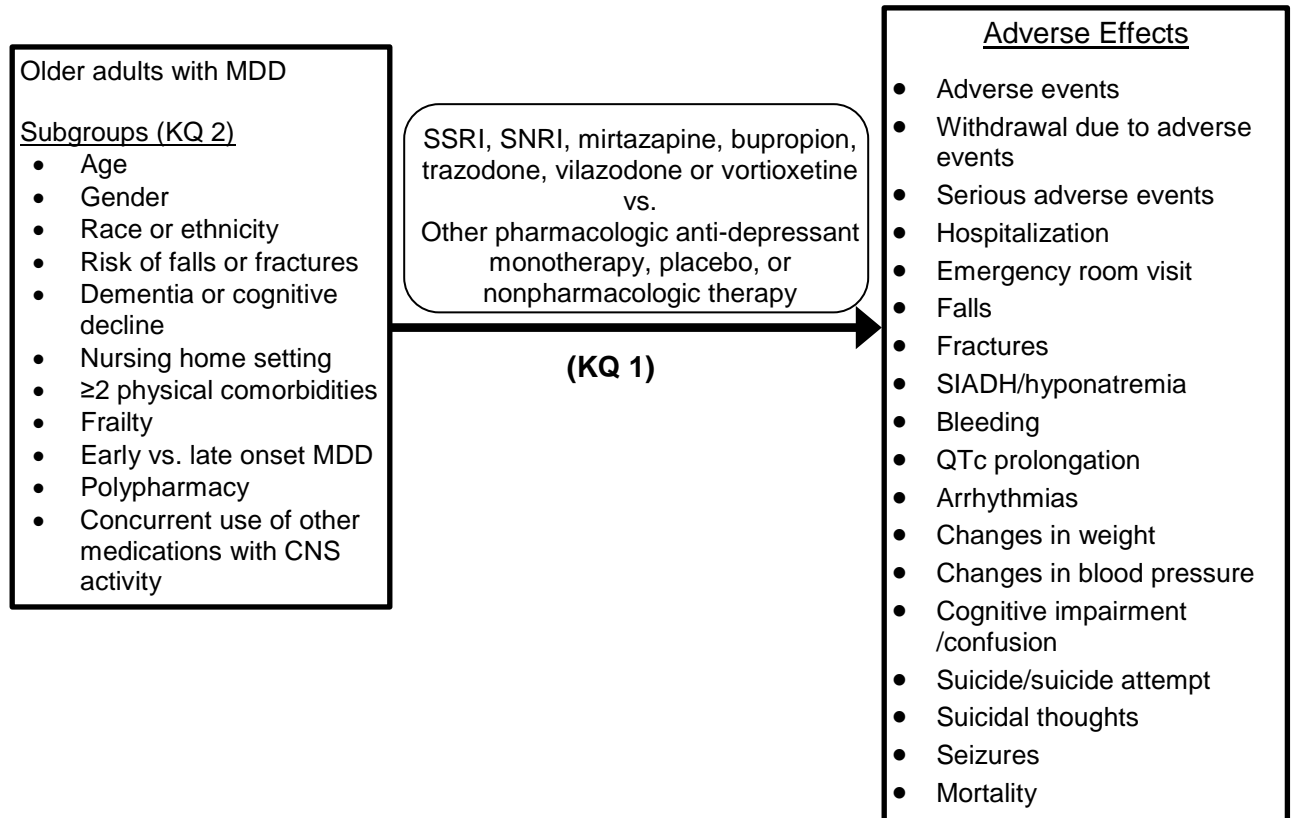
Timing:

We will have no limitations on study duration or length of follow-up, for KQs and CQ. We may consider study length for subgroup analysis if necessary.

Settings:

We are interested in non-acute care settings for KQs and CQ such as specialist or generalist outpatient setting, rehabilitation facility and nursing homes. Hospital or urgent care settings will be excluded.

III. Analytic Framework

Figure 1. Analytic Framework for Adverse Effects of First-Line Pharmacologic Treatments of Major Depression in Older Adults.

Older adults with MDD

Subgroups (KQ 2)
- Age
- Gender
- Race or ethnicity
- Risk of falls or fractures
- Dementia or cognitive decline
- Nursing home setting
- ≥2 physical comorbidities
- Frailty
- Early vs. late onset MDD
- Polypharmacy
- Concurrent use of other medications with CNS activity

SSRI, SNRI, mirtazapine, bupropion, trazodone, vilazodone or vortioxetine
vs.
Other pharmacologic anti-depressant monotherapy, placebo, or nonpharmacologic therapy

(KQ 1)

Adverse Effects
- Adverse events
- Withdrawal due to adverse events
- Serious adverse events
- Hospitalization
- Emergency room visit
- Falls
- Fractures
- SIADH/hyponatremia
- Bleeding
- QTc prolongation
- Arrhythmias
- Changes in weight
- Changes in blood pressure
- Cognitive impairment /confusion
- Suicide/suicide attempt
- Suicidal thoughts
- Seizures
- Mortality

Abbreviations: CNS=central nervous system; KQ=key question; MDD= major depressive disorder; SIADH=syndrome of inappropriate antidiuretic hormone; SNRI=selective serotonin norepinephrine inhibitor; SSRI=selective serotonin reuptake inhibitor

IV. Methods

The methods for this comparative effectiveness review follow the guidance provided in the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* for the Evidence-based Practice Center (EPC) program.[9]

Criteria for Inclusion/Exclusion of Studies in the Review- Inclusion and exclusion criteria for the KQs are listed in Table 2, consistent with the PICOTS above.

Table 2. Inclusion and exclusion criteria for KQs

Category	Inclusion Criteria	Exclusion Criteria
Population	Older adults age 65 years and older of all races and ethnicities with MDD. MDD will be determined as reported by the study, either with use of DSM, ICD codes, validated tools or patient self-report.	Patients younger than 65 years Studies that focus enrollment on patients with a subtype of MDD rather than general MDD; bipolar disorder; or patients with comorbid seizure disorder or comorbid psychiatric conditions with exception of anxiety
Intervention	SSRI, SNRI, bupropion, mirtazapine, trazodone, vilazodone or vortioxetine (Table 1) as a single intervention	Other pharmacologic therapies, non-pharmacologic therapies, complementary alternative medicines, or combinations of therapies
Comparator	A pharmacologic antidepressant for MDD (Table 1, or TCA or MAOI), as a single intervention, including within class comparisons of SSRIs and SNRIs. Placebo Nonpharmacologic interventions as specified in PICOTS	Other pharmacologic therapies, invasive nonpharmacologic interventions, complementary alternative medicines, combinations of therapies
Outcomes	As defined in the PICOTS criteria	Studies that do not include at least one of the outcomes listed in the PICOTS
Timing	All study durations and follow-up lengths will be included	None
Setting	Non-acute care setting (i.e. specialist or generalist outpatient setting, rehabilitation or nursing home)	Hospital or urgent care setting
Study Design	RCTs, nonrandomized controlled trials, prospective or retrospective controlled cohort studies, case-controlled studies	Case series, case reports, studies without an active comparator or non-active control group
Publication Language and Dates	No limits on publication date or language*	Abstracts without published study manuscripts. Non-English publications that do not have an English language abstract.

Abbreviations: DSM=Diagnostic and Statistical Manual of Mental Disorders; ICD=international classification of disease; MDD=major depressive disorder; PICOTS=population, intervention, comparator, outcomes, timing, setting; RCT=randomized controlled trial; SNRI=selective norepinephrine reuptake inhibitor; SSRI= selective serotonin reuptake inhibitor
*English language abstracts of non-English language articles will be reviewed at the abstract stage consistent with the process described by the Methods Guide.[8]

The following inclusion and exclusion criteria will be used for the CQ. We will consider peer-reviewed literature, systematic reviews, meta-analyses, and observational studies that are consistent with the defined PICOTS. We will also search www.guidelines.gov, www.tripdatabase.com, and the websites of the AGS, the American Association for Geriatric Psychiatry and the APA to access supplemental information such as key clinical practice recommendations.

Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions – Two search strategies will be implemented, one for the KQs and one for the CQ. The preliminary search strategies formatted for MEDLINE are shown in the Appendix and are comprised of medical subject heading (MeSH) terms and natural language terms reflective of the population and interventions. The search strategies will be adapted for the other databases as needed.

We will conduct an updated literature search (of the same databases searched initially) concurrent with the peer review process. We will investigate any literature that the peer reviewers or the public suggest and, if appropriate, will incorporate additional studies into the final review. The appropriateness of those studies will be determined using the methods described above.

To identify relevant published literature for KQ1 and KQ2, we will search the following databases: Ovid MEDLINE, Ovid MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, PsychInfo, and Cochrane Central Register of Controlled Trials via OVID. We will search cliniclatrials.gov and the World Health Organization International Controlled Trials Registry Platform (ICTRP) for ongoing studies as well as those completed with results, when available. Drug manufacturers will have the opportunity to submit supplementary evidence and data through the Effective Health Care website, coordinated by AHRQ. The reference list of key articles and systematic reviews or guidelines identified during the article screening process will be reviewed for additional eligible studies.

Articles retrieved through electronic database searching will be screened for inclusion in this review against the established PICOTS framework and inclusion/exclusion criteria. With citations retrieved through the search for KQs, the title and abstract of each article will be reviewed by two independent investigators and the article will be excluded if both reviewers agree that it meets one or more exclusion criteria. Articles identified for inclusion will advance to the full-text screening. Two independent reviewers will screen each article and agree upon the inclusion/exclusion decision. Disagreements will be resolved through consensus or adjudication in consultation with a third reviewer. Articles that meet inclusion/exclusion criteria will be eligible for data abstraction. When necessary, we may contact authors of candidate articles for clarification of reported study details in order to assess for inclusion/exclusion. For articles excluded at the full-text level, we will record the reason for exclusion and present a list of such studies in the review. Citations will be managed using Distiller.

Abstracts and meeting presentations will be considered for inclusion into the review if the abstract or presentation can be matched to an original publication that has been included into the review. The original full publication will always be used as the primary data source in the event discrepant data is reported in multiple publications. Post-hoc and subgroup analyses of included studies will be considered when they provide data on the outcomes of interest.

To identify relevant published literature for the CQ, we will search the following databases: Ovid MEDLINE, EMBASE, PsychInfo and the Cochrane Database for Systematic Reviews. We will also search www.guidelines.gov and the websites of the AGS, the American Association for Geriatric Psychiatry and the American Psychiatric Association to access supplemental information such as key clinical practice recommendations. Citations will be managed using Distiller.

Data Abstraction and Data Management – Data will be abstracted using Distiller by two trained researchers. The second reviewer will confirm the first reviewer's abstracted

data for completeness and accuracy. A third reviewer will audit a random sample of articles to ensure consistency of the process.

Articles referring to the same study will be abstracted on a single review form, assuming the populations are the same. Authors of individual studies may be contacted either for clarification or to request additional data, if necessary.

For all included studies, reviewers will extract data on study characteristics (e.g. study design, duration of follow-up), eligibility criteria, study population (e.g. age, gender, race/ethnicity, and depression severity), interventions (e.g. intervention drug(s), comparison, dose, frequency, and concomitant medications), outcome measures, and the results of each outcome, including measures of variability.

Assessment of Methodological Risk of Bias of Individual Studies - The assessment of risk of bias for included RCTs of pharmacologic interventions will be performed using the Cochrane Collaboration's Risk of Bias Tool with adaptation for harms outcomes.[10] For non-randomized studies, we will use the Newcastle Ottawa Scale.[11]

Two reviewers will independently assess the risk of bias of each included study, with disagreements resolved by either discussion or consultation with a third team member. The overall risk of bias for each study will be classified as low, moderate or high, according to the collective risk of bias per evaluated domain and the investigator's confidence in the study results given the identified limitations.[12]

Data Synthesis - For each KQ, we will create a set of detailed evidence tables containing all information extracted from included studies. Synthesis of data will be based on the pharmacologic class named as the intervention and comparator, not on the individual drug level, except in the case of bupropion, mirtazapine, trazodone, vilazodone and vortioxetine which are unique in their own mechanism and will be analyzed separately. We will perform random-effects meta-analysis using the Hartung-Knapp adjustment when sufficient data for a given outcome is available from at least two studies that are sufficiently homogenous with respect to key clinical (population characteristics, study duration, and intervention) and methodologic (based on risk of bias assessment) variables. Between-study variance will be estimated using the Paule-Mandel estimator.[13] We anticipate scenarios where outcomes will be rarely reported; thus, we will consider use of methods such as the Peto or Mantel-Haenzel odds ratios.[14] The choice between these methods will depend on factors such as the overall event rates and the balance of events between arms.[15,6] We also anticipate instances of either zero events in one study arm or in a study all together. In this instance, methods such as the arcsine difference[17] and continuity correction will be considered,[18] as well as inclusion of both-armed zero-event studies.[19] Statistical significance will be set at a two sided alpha of 0.05. All studies, including those that are not amenable to pooling, will be qualitatively summarized.

When quantitative pooling of studies is possible, we will evaluate for statistical heterogeneity using the Cochrane chi-square p-value and the I^2 statistic. A Cochrane p-value of <0.10 suggests the presence of statistical heterogeneity. The I^2 statistic assesses

the degree of inconsistency across studies and ranges from 0-100% with the higher percentage representing a higher likelihood of the existence of true heterogeneity as opposed to chance.[20] An I^2 value of greater than 50% will be considered substantial heterogeneity. We will attempt to determine potential reasons by conducting relevant subgroup analyses based on those subgroups listed in the analytic framework.

To assess for the presence of publication bias, visual inspection of funnel plots will be considered for each pooled analysis. Tests for funnel plot asymmetry, including Egger's weighted regression tests for continuous outcomes[21] and tests by Peters et al.[22] or Rucker et al.[23] (chosen depending on the amount of between-study heterogeneity) for dichotomous outcomes will be conducted when 10 or greater studies report the outcome. All analyses were performed using the 'meta' or 'metafor' packages in R (version 3.1.3; the R Project for Statistical Computing).

For the contextual question, the findings of the citations pertinent PICOTS will be discussed qualitatively. Formal analytic and reporting techniques will not be used.

Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes - We will grade the SOE based on the guidance established for the EPC program.[24] At the completion of the review, two reviewers will independently grade the SOE for critical outcomes which were selected with input from the Technical Expert Panelist (TEP). TEP were asked to rank outcomes from most to least important and the average rank of outcomes was discussed during a TEP call. The critical outcomes include suicide, mortality, hospitalization, serious adverse events, arrhythmias, QTc prolongation, falls, fractures, cognitive impairment and SIADH. If other outcomes are determined to be important for SOE grading during the review they too will be considered. Conflicts will be resolved either through consensus or third-party adjudication. The SOE approach incorporates five key domains: study limitations, directness, consistency, precision, and reporting bias of the evidence body. Additional domains (plausible confounding, dose-response, and magnitude of effect) will be considered when applicable. The SOE pertaining to each KQ will be classified into four categories:
1) High – We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable, i.e., another study would not change the conclusions.
2) Moderate – We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe the findings are likely to be stable, but some doubt remains.
3) Low – We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
4) Insufficient – We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of the effect for this outcome. No evidence is available

of the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Assessing Applicability – We will consider elements of the PICOTS framework when evaluating the applicability of evidence to answer our KQs as recommended by the EPC methods guide.[25] We will consider how important population characteristics (e.g. age, gender, race, ethnicity, and severity of depression), and intervention features (co-interventions) may cause heterogeneity of treatment effects and affect generalizability of the findings.

V. References

1. 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
2. Kok RM, Reynolds CF III, Management of Depression in Older Adults A Review. *JAMA*. 2017;317(20):2114-22.
3. Gelenberg AJ, Freeman MP, Markowitz JC et al. Practice guideline for the treatment of patients with major depressive disorder. 3rd ed. American Psychiatric Association 2010.
4. Qaseem A, Barry MJ, Kansagara D, et al. Nonpharmacologic versus pharmacologic treatment for adult patients with major depressive disorder: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2016;162:350-9.
5. American Geriatrics Society 2015 Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults by the American Geriatrics Society 2015 Beers Criteria Update Expert Panel (2015). *J Am Geriatr Soc.*2016;63:2227-46.
6. Steinmen MA, Beizer JL, DuBeau CE, Laird RD, Lundebjerg NE, Mulhausen P. How to Use the AGS 2015 Beers Criteria – A Guide for Patients, Clinicians, Health Systems, and Payors. *J Am Geriatr Soc* 2015;63(12):e1-e7.
7. Hanlon JT, Semla TP, Schmader KE. Alternative medications for medications in the use of high-risk medications in the elderly and potentially harmful drug-disease interactions in the elderly quality measures. *J Am Geriatric Soc.* 2015;63(12):e8-e18.
8. Levy HB. Polypharmacy reduction strategies. Tips on incorporating American Geriatrics Society Beers and screening tool of older people's prescriptions criteria. *Clin Geriatr Med*. 2017;33:177-87.
9. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
10. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. PMID: 22008217.
11. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed Nov 2, 2017.
12. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions; 2012. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ

Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

13. Paule RC, Mandel J. Consensus values and weighing factors. J Res Natl Bur Stand. 1982;87:377-85.

14. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. *JAMA*. 1995;273(18):1450-6.

15. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*. 2007;26(1):53-77.

16. Hernandez A V, Walker E, Ioannidis JPA, Kattan MW. Challenges in meta-analysis of randomized clinical trials for rare harmful cardiovascular events: the case of rosiglitazone. *Am Heart J*. 2008;156(1):23-30.

17. Rücker G, Schwarzer G, Carpenter J, Olkin I. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Stat Med*. 2009;28(5):721-38.

18. Li L, Wang X. Meta-analysis of rare binary events in treatment groups with unequal variability. *Stat Methods Med Res*. January 2017:962280217721246.

19. Cheng J, Pullenayegum E, Marshall JK, Iorio A, Thabane L. Impact of including or excluding both-armed zero-event studies on using standard meta-analysis methods for rare event outcome: a simulation study. *BMJ Open*. 2016;6(8):e010983.

20. Higgins JP, Thomas SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-60.

21. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-34

22. Peters JL, Sutton AJ, Jones DR, et al. Comparison of two methods to detect publication bias in meta-analysis. *JAMA*. 2006;295:676-80.

23. Rucker G, Schwarzer G, Carpenter JR. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stats Med*. 2008;27:746-63.

24. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update, 2013. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

25. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions; 2010. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

VI. Definition of Terms

Not applicable

VII. Summary of Protocol Amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol. Example table below:

Date	Section	Original Protocol	Revised Protocol	Rationale
This should be the effective date of the change in protocol	Specify where the change would be found in the protocol	Describe the language of the original protocol.	Describe the change in protocol.	Justify why the change will improve the report. If necessary, describe why the change does not introduce bias. Do not use justification as "because the AE/TOO/TEP/Peer reviewer told us to" but explain what the change hopes to accomplish.

VIII. Review of Key Questions

The Agency for Healthcare Research and Quality (AHRQ) posted the key questions on the AHRQ Effective Health Care Website for public comment. The Evidence-based Practice Center (EPC) refined and finalized the key questions after review of the public comments, and input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the key questions are specific and relevant.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The AHRQ Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are

selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and suggest approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The AHRQ TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than $1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder
This project was funded under Contract No. xxx-xxx from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The AHRQ Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

XIV. Registration
This protocol will be registered in the international prospective register of systematic reviews (PROSPERO).

Appendix

MEDLINE search for KQs
1. major depression.mp. or major Depression/
2. major depressive.mp.
3. 1 or 2
4. elderly.mp. or Aged/
5. "Aged, 80 and over"/ or late-life.mp.
6. later-life.mp.
7. older.mp.
8. geriatric.mp.
9. 4 or 5 or 6 or 7 or 8
10. (anti-depressant or antidepressant).mp
11. Antidepressant Agents/
12. paroxetine.mp. or Paroxetine/
13. sertraline.mp. or Sertraline/
14. citalopram.mp. or Citalopram/
15. escitalopram.mp.
16. fluoxetine.mp. or Fluoxetine/
17. fluvoxamine.mp. or Fluvoxamine/
18. selective serotonin reuptake inhibitor.mp. or Serotonin Uptake Inhibitors/
19. venlafaxine.mp. or Venlafaxine Hydrochloride/
20. desvenlafaxine.mp. or Desvenlafaxine Succinate/
21. duloxetine.mp. or Duloxetine Hydrochloride/
22. serotonin norepinephrine reuptake inhibitor.mp.
23. bupropion.mp. or Bupropion/
24. mirtazapine.mp.
25. trazodone.mp. or Trazodone/
26. vilazodone.mp. or Vilazodone Hydrochloride/
27. vortioxetine.mp.
28. milnacipran.mp.
29. levomilnacipran.mp.
30. Serotonin and Noradrenaline Reuptake Inhibitors/
31. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30
32. 3 and 9 and 31
33. Epidemiologic studies/
34. exp cohort studies/
35. exp case controlled studies/
36. Case control.tw.
37. (cohort adj (study or studies)).tw.
38. Cohort analy$.tw.
39. (Follow up adj (study or studies)).tw.
40. (observational adj (study or studies)).tw.
41. Longitudinal.tw.
42. Retrospective.tw.
43. Cross sectional.tw.

44. Cross-sectional studies/
45. or/33-44
46. Randomized Controlled Trials as Topic/
47. randomized controlled trial/
48. Random Allocation/
49. Double Blind Method/
50. Single Blind Method/
51. clinical trial/
52. clinical trial, phase i.pt.
53. clinical trial, phase ii.pt.
54. clinical trial, phase iii.pt.
55. clinical trial, phase iv.pt.
56. controlled clinical trial.pt.
57. randomized controlled trial.pt.
58. multicenter study.pt.
59. clinical trial.pt.
60. exp Clinical Trials as topic/
61. or/46-60
62. (clinical adj trial$).tw.
63. ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.
64. PLACEBOS/
65. placebo$.tw.
66. randomly allocated.tw.
67. (allocated adj2 random$).tw.
68. or/62-67
69. 61 or 68
70. case report.tw.
71. letter/
72. historical article/
73. or/70-72
74. 69 not 73
75. 45 or 74
76. 75 and 32