I. Background and Objectives for the Technology Assessment

Clinical and Epidemiological Issues. Depressive episodes can be seen in patients with either major depressive disorder (MDD) or bipolar disorder. In 2015, 6.6 percent of adults in the United States experienced a depressive episode in the past year. The bulk of these episodes are part of MDD, which more than 13 million U.S. residents experience each year. Of these individuals, one-half seek help for this condition; one in five of those seeking help receive adequate acute-phase treatment. Even for patients receiving adequate treatment, only 30 percent (i.e., 3% of patients with MDD) reach the treatment goal of full recovery, or remission.

The remaining 70 percent of MDD patients will either respond without remission (about 20%) or not respond at all (50%). Patients whose depressive disorder does not respond satisfactorily to adequate treatment clearly have harder-to-treat depression, which is generally (albeit not uniformly) referred to as treatment-resistant depression (TRD). Although often broadly defined this way, TRD is a complex phenomenon that is influenced by heterogeneity in depressive subtypes, psychiatric comorbidity, and coexisting medical illnesses. Such patients pose a common, challenging presentation to psychiatric and primary care clinicians. Although TRD is most commonly associated with MDD, treatment-resistant depressive episodes can also be seen in the depressed phase of bipolar disorder. Bipolar disorder affects 2.6 percent of the U.S. adult population each year. Much like MDD, bipolar depression can be treatment resistant. More than 30 percent of those suffering from bipolar disorder and receiving treatment do not experience sustained remission of depressive symptoms. Even among those who do achieve recovery for lengthy periods, depressive relapses are common; more than 20 percent will experience a depressive relapse within a year.
TRD has substantial effects on patients and major social impact, most of which has been described for MDD patients. Patients with TRD incur the highest direct and indirect medical costs among those with MDD. These costs increase with the severity of TRD. Treatment-resistant patients are twice as likely to be hospitalized, and their cost of hospitalization is more than six times the mean total cost for depressed patients who are not treatment resistant. TRD can nearly double both direct and indirect 2-year employer medical expenditures relative to expenditures for patients whose MDD responds to treatment ($35,500 for those with TRD and $18,600 for those with MDD). 

TRD is especially relevant for Medicare beneficiaries, for whom unsuccessfully treated depression has harmful sequelae. Mood disorders, which consist primarily of MDD and bipolar disorder, are the second leading cause of disability in Medicare patients under the age of 65. Furthermore, depression in the elderly is more associated with suicide than at any other age; although adults 65 and older make up 12 percent of the population, they constitute 16 percent of all suicide deaths. Indeed, the decrease in average life expectancy for those with depressive illness, including Medicare beneficiaries, is 7 to 11 years, similar to that in elderly smokers. Finally, depression is a major predictor of the onset of stroke, diabetes, and heart disease. Being depressed increases patients’ risk of developing coronary heart disease, and it raises the risk of dying from a heart attack nearly three-fold.

**Rationale for Review:** While broad agreement exists about the major impact of TRD, there is no universally accepted operational definition. Criteria for TRD have been variably defined in clinical research and practice, reflecting many difficulties and controversies about its definition. These definitional dilemmas limit the ability of systematic reviewers or other experts to synthesize information and generalize the findings of many TRD studies to the array of patient populations encountered in daily practice. A universal definition of TRD is needed to improve homogeneity among research samples—or at a minimum to permit adequate description of the heterogeneity among research subjects and patient populations (including those for which Medicare is the primary insurer). It is also required to guide the application of clinical research findings to clinical practice, including community populations of TRD patients.

Even further, these varying conceptualizations of TRD have made translation of research findings or systematic reviews into clinical practice guidelines challenging and inconsistent. Treatment guidelines reflect this variability: their definitions of TRD differ, agreement on what constitutes prior treatment adequacy is lacking, and recommended “next step” interventions can diverge.

Accordingly, we aim to review and inform the definition of TRD in clinical research as well as obtain information on the use of the definition of TRD in the context of coverage with evidence development (as defined by the Medicare Guidance Document) and treatment outcomes. The purpose of this report is not to determine outcomes associated with specific treatments of TRD but to comprehensively examine the study design issues affecting both outcomes and bias in the study of TRD.
II. The Key Questions

Narrative Review Questions: Based on a literature search for consensus statements, guidelines, materials from the U.S. Food and Drug Administration (FDA), the U.S. National Institutes of Health (NIH), and the U.S. Substance Abuse and Mental Health Services Administration (SAMHSA); systematic reviews; and on a review of UpToDate, an evidence-based, peer reviewed clinical information source, we will address the key questions (Key Questions [KQs] 1 through 5, with their subquestions) listed below. In addition, we will use information from the Medicare Evidence Development and Coverage Advisory Committee (MEDCAC) panel meeting on April 27, 2016,\textsuperscript{26} to augment our reporting on TRD definitions, study design issues, and the related topics.

The specific issues are:

1. What definitions of TRD are found in this literature? What consensus, if any, exists about the best definition(s) for this condition?

2. What methods do investigators use to diagnose this condition in clinical research? What consensus, if any, exists about the best measure(s) to use? Does the setting of the medical visit influence the choices that investigators make about the diagnostic tool they use?

3. What measures have been developed to determine the success and failure of treatment in clinical research studies of TRD?
   a. What consensus, if any, exists about the best measure(s) to investigate treatments for TRD? What are the main points of agreement about such measures?
   b. Are these measures physician-reported or patient-reported?
   c. What are the psychometric properties of these measures? Is the minimum significant clinical difference defined for these measures?
   d. Compare and contrast these measures in how they describe:
      i. Change in depression scores as measured by depression scales
      ii. Change in depressive symptomatology (e.g., sleep disorders, fatigue, weight change, cognition)
      iii. Change in measures of anhedonia
      iv. Change in measures of functional capacity (e.g., physical functioning, ability to care for self)
      v. Change in measures of quality of life
      vi. Change in measures of suicide ideation
      vii. Change in suicide attempts
      viii. Other

4. What types of research designs are used to study TRD?
   a. What consensus, if any, exists about the type of study design that best minimizes bias and the placebo effect in this field?
   b. If no consensus exists about study designs to accomplish these goals, what are the trends in study designs for assessing interventions for TRD? Do these
trends reflect long-lasting (e.g., traditional) designs or short-lived, evolving, or newly emerging designs?

c. What consensus, if any, exists about the appropriate length of a trial?

5. What are the risk factors for TRD?

**Systematic Review Questions**: From a systematic literature search for individual studies on TRD. We will address the KQs 6 through 11 with their subquestions as listed below.

6. What variables were considered for TRD patients in these studies? Specify at least the factors listed below.
   a. Patient Characteristics:
      i. Age
      ii. Type of depressive episode (unipolar, bipolar, psychotic, atypical, other)
      iii. Number of depression relapses and time to relapse
      iv. Psychiatric comorbidities
      v. Medical comorbidities (e.g., diabetes, cardiac disease, renal disease, dementia and other cognitive abnormalities)
      vi. Suicidal ideation
      vii. Suicide attempts
      viii. Duration of symptoms
      ix. Screening tools used to make the diagnosis
      x. Diagnostic tools to confirm the diagnosis
   b. Prior Treatments:
      i. The number, duration, dosage, or classes of antidepressants attempted for each trial of therapy
      ii. The number of failed trials of adequate therapy
      iii. The number of prior treatment trials that patients did not tolerate
      iv. The use of augmentation and combination pharmacological therapies for each attempted treatment trial
      v. The use of electroconvulsive therapy (ECT)
      vi. The use of psychotherapy
   c. Diagnostic characteristics
      i. The use of structured versus unstructured diagnostic assessments
      ii. Scores on standardized and validated depression rating instruments
      iii. Setting in which the diagnosis was made (i.e., primary care, generalized psychiatric setting, specialty psychiatric setting, other)

7. How do these inclusion criteria compare or contrast with the definition(s) of TRD noted in the Narrative Questions?

8. What were primary characteristics of included studies?
   a. What was the main design of each included study (e.g., randomized controlled trial [RCT] with blinding; interrupted time series; use of placebo, wait-list, or sham procedure)?
   b. Were run-in or wash-out periods (or both) used in included studies? If so, how long were they?
c. How long was each included study?

9. How were included studies designed to account for the risk factors for TRD (see (Narrative Question #5)? If the following characteristics are not noted above as risk factors, how did included studies account for at least the following: age, sex, race, socioeconomic status, duration of symptoms, disease severity, co-existing medical and psychiatric conditions, and placebo effect?

10. What are relationships between risk factors and various results of included studies?
   a. Using regression analysis or other statistical techniques, determine whether the risk factors for Narrative Review Question #5 and Systematic Review Question # 9 can be correlated with study results (i.e., the magnitude of treatment effects)?
   b. What is the influence of placebo response on the magnitude of treatment effects for different types of interventions?
   c. Does study duration moderate the influence of placebo response?

11. What variables or information did included studies report? Specifically:
   a. What measures are used to define end points in these TRD trials?
   b. In addition to the measures noted for Narrative Review Question #3, did these studies record:
      i. Adherence to treatment
      ii. Attrition from care
      iii. Changes in patient-selected factors of importance (i.e., outcome measures identified by patient as important)
      iv. Changes in employment or disability status
      v. Changes in use of medical resources (e.g., hospitalizations, emergency room or physician visits)
      vi. Time to relapse

**PICOTS.** For the above KQs, we will apply the following criteria for populations, interventions, comparators, outcomes, time frames, and settings:

**Population(s):** All adults (>18 years old) identified as having a depressive episode (including major depressive disorder [MDD] and bipolar disorder) who have not responded to treatment(s). The depressive episode must be part of a major depressive disorder or a bipolar disorder. Studies of people without a primary diagnosis of major depressive disorder or bipolar disorder, or without evidence of treatment nonresponse, will be excluded.
Interventions¹:
Any pharmacologic intervention tested as a treatment for TRD as a primary therapy or as an augmentation agent to an existing primary therapy.
- Antidepressants (e.g., selective serotonin reuptake inhibitors [SSRIs], serotonin-norepinephrine reuptake inhibitors [SNRIs], tricyclic antidepressants [TCA], monoamine oxidase inhibitors [MAOI], atypical agents)
- Atypical antipsychotics
- Anticonvulsants
- Mood stabilizers
- Psychostimulants
- Agents approved by the FDA for other indications but tested in TRD populations (e.g., ketamine, levothyroxine [T3], clonidine)

Any nonpharmacologic device or procedure tested as a treatment for TRD as a primary therapy or as augmentation to an existing primary therapy and identified as a TRD option by a consensus statement, guideline, the MEDCAC panel, or systematic review (e.g., ECT, repetitive transcranial magnetic stimulation [rTMS], vagus nerve stimulation [VNS], deep brain stimulation [DBS], cranial electrotherapy stimulation [CES])

Any nonpharmacologic intervention tested as a treatment for TRD as a primary therapy or as augmentation to an existing primary therapy and identified as a TRD option by a consensus statement, guideline, the MEDCAC panel, or systematic review.
- Complementary and alternative medication therapies (CAM)
- Psychotherapy
- Exercise

Comparators:
All comparative studies with a concurrent control group or a control group from an interrupted time-series study. These designs exclude pre/post studies that did not conduct interrupted time-series analyses.

Outcomes:
Mental health outcomes identified in previous depression comparative effectiveness review work as either critical or important for decisionmaking:
- Benefits that are reported as primary endpoints (or outcomes) for a trial. Such outcomes could include:
  o Reduction in suicidal ideation or suicide attempts
  o Quality of life
  o Response to treatment
  o Remission
  o Change in depressive severity

¹ A list of specific individual pharmacologic or nonpharmacologic interventions can be found in Attachment A at the end of the protocol. Eligible interventions include those that have both been tested as a treatment targeting TRD in adults and been identified by guidelines, consensus statements, the MEDCAC panel, or systematic reviews as alternatives for TRD treatment.
o Functional capacity (physical and cognitive functioning measured by validated scales)
o Speed of remission
o Speed of response
o Intervention durability (rates or counts of recurrence of a depressive episode for those who have remitted)
  • Adverse events from the intervention identified as either critical or important for decisionmaking.
    o Serious adverse events per FDA definition^27(rates or counts)
    o Overall adverse events (rates or counts)
    o Treatment discontinuations attributed to adverse events (rates or counts)

Timing:
  • Any study duration.

Settings:
  • All settings.

Our population of interest is adults 18 years of age or older with depression who have not responded to treatment(s). The depressive illness can be part of either MDD or a bipolar disorder, but one of these diagnoses must be a primary diagnosis; for example, schizophrenia with a secondary diagnosis of MDD, or dysthymia, would not be eligible for this report. If a study involves both eligible and ineligible patients and does not report data separately, that whole study will be excluded. Populations with no evidence of treatment nonresponse (e.g., a study in which the absence of treatment response is not part of the selection criteria) will not be eligible.

Eligible interventions include those that have both been tested as a treatment targeting TRD in adults and been identified by guidelines, consensus statements, the MEDCAC panel, or systematic reviews as alternatives for TRD treatment. These criteria ensure consideration of interventions with a minimum threshold amount of data addressing its effectiveness in TRD populations. Comparison groups include concurrent control groups (e.g., active, sham, or placebo) and a control group from an interrupted time series.

We will require outcomes to have been identified previously as the most meaningful to depression management decisionmaking. In our earlier comparative effectiveness work on depression,^28, 29 we asked our Technical Expert Panel and Key Informants to rank the relative importance of these outcomes following a process proposed by the GRADE Working Group.30 We used SurveyMonkey© for an anonymous ranking of the relative importance of outcomes. Participants used a 9-point Likert scale to rank outcomes into three categories: (1) critical for decisionmaking, (2) important but not critical for decisionmaking, and (3) of low importance for decisionmaking. They identified six outcomes as critical and five as important, and they supported the inclusion of an additional depressive outcome (change in depressive severity). For one of the adverse events outcomes, serious adverse events, we will use the FDA definition^27 and will consider physical, psychological, and cognitive events. We will require relevant studies for the current project to report on at least 1 of these 12 outcomes.
All study durations and all settings are eligible. Pre/post studies that do not use interrupted time series analyses will be excluded, because potential confounding from multiple sources renders questionable the ability of these study designs to support causal inferences. We will include English-language articles and exclude studies that are not published fully in English.

III. Methods

This Technology Assessment will be organized into sections addressing the Narrative Review KQs (1 through 5) and the Systematic Review KQs (6 through 11). Table 1 gives our selection (inclusion/exclusion) criteria and outlines methods to answer the KQs.

Criteria for Inclusion/Exclusion of Studies in the Review

Table 1. Inclusion/Exclusion Criteria

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Populations without a primary diagnosis of MDD or bipolar disorder will be excluded, as will those without evidence of treatment nonresponse.</td>
</tr>
<tr>
<td>All adult populations (&gt;18 years old) identified as having a primary diagnosis of depression (including MDD and bipolar disorder) who have had a depressive episode and have not responded to treatment(s).</td>
<td></td>
</tr>
<tr>
<td><strong>Interventions</strong></td>
<td>Interventions not targeting TRD</td>
</tr>
<tr>
<td>Any pharmacologic intervention tested as a treatment for TRD as a primary therapy or as an augmentation agent to an existing primary therapy.</td>
<td></td>
</tr>
<tr>
<td>• Antidepressants (e.g., SSRIs, SNRIs, TCAs, MAOIs, atypical agents)</td>
<td></td>
</tr>
<tr>
<td>• Atypical antipsychotics</td>
<td></td>
</tr>
<tr>
<td>• Anticonvulsants</td>
<td></td>
</tr>
<tr>
<td>• Mood stabilizers</td>
<td></td>
</tr>
<tr>
<td>• Psychostimulants</td>
<td></td>
</tr>
<tr>
<td>• Agents FDA-approved for other indications but tested in TRD populations (e.g., ketamine, levothyroxine, clonidine,)</td>
<td></td>
</tr>
<tr>
<td>Any nonpharmacologic device or procedure tested as a treatment for TRD as a primary therapy or as augmentation to an existing primary therapy.</td>
<td></td>
</tr>
<tr>
<td>• Devices (e.g., ECT, rTMS, VNS, DBS, CES)</td>
<td></td>
</tr>
<tr>
<td>Any nonpharmacologic intervention tested as a treatment for TRD as a primary therapy or as augmentation to an existing primary therapy.</td>
<td></td>
</tr>
<tr>
<td>• CAM</td>
<td></td>
</tr>
<tr>
<td>• Psychotherapy</td>
<td></td>
</tr>
<tr>
<td>• Exercise</td>
<td></td>
</tr>
<tr>
<td><strong>Comparators</strong></td>
<td>Pre/post studies where interrupted time-series analyses were not conducted</td>
</tr>
<tr>
<td>All comparative studies with concurrent control groups or control groups from an interrupted time series or pre/post studies with interrupted time series analyses.</td>
<td></td>
</tr>
</tbody>
</table>

EPC Protocol Version 17, 10/17/16
<table>
<thead>
<tr>
<th>Inclusion/Exclusion Criteria (continued)</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes</strong></td>
<td>Mental health outcomes critical or important for decisionmaking: Benefits that are reported as primary endpoints (or outcomes) for a study:</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>• Reduction in suicidal ideation or suicide attempts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Quality of life</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Response to treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Change in depressive severity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Functional capacity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Speed of remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Speed of response</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Intervention durability (i.e., rates or counts of recurrence of a depressive episode for those who have remitted)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adverse events from the intervention identified as either critical or important for decisionmaking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Serious adverse events per FDA definition (rates or counts)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Overall adverse events (rates or counts)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Drug interactions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Treatment discontinuations attributed to adverse events (rates or counts)</td>
<td></td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Any study duration; literature publication date from 1/1/95 to present</td>
<td>This date restriction provides literature relevant to contemporary definitions of TRD with diagnoses consistent with definitions in DSM-IV and later versions.</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>Study takes place in a highly developed country</td>
<td>None</td>
</tr>
<tr>
<td><strong>Study Designs</strong></td>
<td>For KQs 1–5: Consensus statements, guidelines, CMS/SAMHSA/FDA/NIH materials, UpToDate, information from the MEDCAC panel meeting on the definition of TRD on April 27, 2016, and systematic review articles.</td>
<td>For KQs 1–5: Evidence not meeting inclusion criteria. Note that individual trials will not be considered in this section.</td>
</tr>
<tr>
<td></td>
<td>For KQs 6–11: Randomized or prospective nonrandomized or observational studies (including concurrent controls and interrupted time series)</td>
<td>For KQs 6–11: Pre-post studies without interrupted time-series analyses. Any studies without a control group.</td>
</tr>
</tbody>
</table>
### Table 1. Inclusion/Exclusion Criteria (continued)

<table>
<thead>
<tr>
<th>Language</th>
<th>English only.</th>
</tr>
</thead>
<tbody>
<tr>
<td>We will exclude studies not published in English because their ability to provide meaningful information about the current understanding of TRD in a Medicare or Medicare-related population is limited.</td>
<td></td>
</tr>
</tbody>
</table>

CAM = complementary and alternative medication therapies; CES = cranial electrotherapy stimulation; CMS = Centers for Medicare & Medicaid Services; DBS = deep brain stimulation; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; ECT = electroconvulsive therapy; FDA = U.S. Food and Drug Administration; KQ = Key Question; MAOI = monoamine oxidase inhibitor; MDD = major depressive disorder; MEDCAC = Medicare Evidence Development and Coverage Advisory Committee; NIH = National Institutes of Health; rTMS = repetitive transcranial magnetic stimulation; SAMHSA = Substance Abuse and Mental Health Services Administration; SGA = second generation antidepressant; SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor; TCA = tricyclic antidepressant; TRD = treatment-resistant depression; VNS = vagus nerve stimulation.

* A list of specific individual pharmacologic or nonpharmacologic interventions can be found in Attachment A at the end of the protocol. Eligible interventions include those that have both been tested as a treatment targeting TRD in adults and been identified by guidelines, consensus statements, the MEDCAC panel, or systematic reviews as alternatives for TRD treatment.

* For serious adverse events, we will use the FDA definition and will consider physical, psychological and cognitive events.

* “Very High” on Human Development Index: Andorra, Argentina, Australia, Austria, Bahrain, Belgium, Brunei Darussalam, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong China (SAR), Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea (Republic of), Kuwait, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Qatar, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States.

### Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

**Assembling Articles.** An experienced EPC research librarian will develop the strategy for our comprehensive search of the literature. To ensure methodological quality, we will follow standard procedures for systematic literature searches specified in the Agency for Healthcare Research and Quality’s (AHRQ’s) Effective Health Care Program *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.* We will systematically search the published literature from January 1, 1995 to April 1, 2017 that is indexed in MEDLINE®, EMBASE, PsycINFO, and Cochrane Library and that addresses treatment of TRD in adults. The aim is to assemble literature relevant to contemporary definitions of TRD with diagnoses consistent with definitions in *Diagnostic and Statistical Manual of Mental Disorders*, Fourth edition and 5th edition. We also will review the reference lists of systematic reviews and protocols to identify any relevant citations that our electronic searches might have missed.

In addition, we will search for consensus statements, management guidelines, and relevant government materials from various Federal agencies. The last specifically include the following: FDA; NIH, including the National Institute of Mental Health (NIMH), Substance Abuse and Mental Health Services Administration (SAMHSA), and Centers for Medicare & Medicaid Services (CMS), including materials presented at the MEDCAC panel of April 27, 2016. We will also search other Websites such as [Clinicaltrials.gov](https://clinicaltrials.gov) and [Guideline.gov](https://guideline.gov) (AHRQ’s National Guidelines Clearinghouse) for relevant documents, and will search UpToDate, an evidence-based, peer reviewed clinical information source.
Trained members of the research team will dually review all titles and abstracts for eligibility based on the pre-established inclusion/exclusion criteria presented in Table 1. Studies marked for possible inclusion by either reviewer will undergo full-text review. Any study with inadequate information in the abstract also will proceed to full-text review. We will retrieve and review the full text of all articles included during the title/abstract review phase. Trained members of the research team will then dually review each full-text article for inclusion or exclusion on the basis of the eligibility criteria. We will document reasons for exclusion at this stage; we will also tag those selected for inclusion with the relevant KQ that the article addressed. Disagreements about inclusion will be resolved by discussion or consensus with review by the full research team as needed.

**Data Abstraction and Data Management**

Our data abstraction and management approaches are based on appropriate review methods. These include clear selection criteria based on PICOTS; dual, independent review of relevant titles/abstracts and full-text review of potentially relevant articles; and identification of articles meeting selection criteria. From included systematic reviews, consensus statements, guidelines, and other relevant materials, we will abstract the relevant information (e.g., definitions of TRD, study designs, methods, measures and risk factors) to answer Narrative Review KQs 1 through 5. We will hand-search the systematic reviews for eligible individual studies for the Systematic Review KQs 6 through 11 in addition to those identified from our general search. From all included individual studies, we will abstract relevant information to answer KQs 6 through 11 (see below). These steps will allow us to catalogue and describe the available controlled studies. We will track all literature screening results in the EndNote database. We will also record the reason that each excluded full-text publication did not satisfy the eligibility criteria.

We will abstract data from any studies that meet our inclusion criteria into a standardized template. For each study, we expect to capture the following: study characteristics (study design, sample size, interventions, comparators, duration, measures to define endpoints and accounting of risk factors, country, and setting); population characteristics (definition of TRD; coexisting psychiatric, substance abuse, and medical conditions; depression severity; prior TRD treatments; length of TRD; and age, sex, race, and ethnicity); and mental health outcomes (e.g., response, remission, depressive symptomatology). One member of the research team will collect the data, and another (senior) investigator will review the abstraction for accuracy and completeness.

**Assessment of Risk of Bias of Individual Studies**

Two investigators will independently assess the risk of bias of included individual studies. Disagreements will be resolved by discussion and consensus or by consulting an independent third party.

For RCTs, we will use the Cochrane Risk of Bias tool. Elements of risk of bias assessment for RCTs include, among others, randomization and allocation concealment, similarity of compared groups at baseline, masking of patients and study personnel, use of intent-to-treat analysis, and overall and differential loss to followup.
For nonrandomized trials and observational studies, we will employ the Newcastle-Ottawa Scale.\textsuperscript{37} Elements of this tool assess the comparability of baseline characteristics, the method of statistical adjustment for baseline confounding, and other factors.

We will use risk of bias in individual studies as a covariate in the regression model for KQ 10.

**Data Synthesis**

Our final technology assessment will follow the prescribed format for such reports for AHRQ and CMS. We will prepare a single report that documents “Narrative Review” findings for KQs 1 through 5 and then “Systematic Review” findings for KQs 6 through 11); we will use text and summary tables as appropriate for ease of presentation and readability. Detailed findings will appear in appendix tables.

For the Narrative KQs, our report will present summary text and a series of tables that will answer each KQ. For example, for KQ1, the summary table will document the variability of the definitions of TRD used which will let us identify where any consensus appears to lie. Similar to KQ1, separate summary tables for KQ2, KQ3 and KQ4, will present the various methods used to diagnose TRD and the measures and study designs used in TRD research. Again, these summary tables will allow us to identify any consensus for these issues. For the subquestions in KQs 2, 3 and 4, we will present the results in separate summary tables that address the specific characteristics called out in these questions; examples include diagnostic tools used in the different diagnostic settings, psychometric properties of measures used to determine efficacy or effectiveness, and study designs that have demonstrated effects on, for example, minimizing bias and placebo effects. For KQ5, we will report information on identified risk factors for TRD. For these KQs, we will have interpretative text summarizing the content of the tables. We will not do any quantitative analyses, but we will provide a qualitative synthesis of what these tables mean.

For the Systematic Review KQs, we will develop a similar series of tables addressing KQs 6 through 9 and KQ 11, again with summary text highlighting key table findings. For KQ 6, we will consider whether the variable is addressed as a criterion for inclusion, a criterion for exclusion, or is simply reported in the study. For KQ 10 (regression or other statistical analysis), we will first define patient- and study-level covariates that might be relevant in examining correlations. Because we will not have access to individual patient data, we will focus primarily on study-level characteristics (e.g., study design, study duration, risk of bias). To avoid issues of ecological fallacy, we will carefully consider which patient-level characteristics we can use. To ensure consistency, we will develop a data codebook and an analysis plan once we have selected the covariates. We will use Microsoft Excel and SAS software for data management, data cleaning, and graphical display of the data.

Regression or other statistical analyses will focus on interventions for which we have at least 10 studies. If necessary, we will combine interventions into categories (e.g., pharmacological interventions, behavioral interventions). We will classify comparator interventions as inactive (e.g., placebo, waiting list) or active. We will also select relevant outcome measures, focusing insofar as possible on patient-centered outcomes.
For computational reasons, we will focus on dichotomous outcomes (odds ratios). We will also recalculate the direction of effect, if necessary, so that an odds ratio >1 indicates a beneficial effect and an odds ratio <1 indicates a harmful effect.

To examine the influence of study characteristics, we will fit Bayesian hierarchical models with a binomial likelihood and allow for between-trial heterogeneity using the procedure described by Welton and colleagues (2009) for WinBUGS. This procedure will allow us to assess the influence of the study characteristic on the estimated intervention effect and variation in bias between studies. We will assume vague prior distributions for unknown parameters. For each outcome, we will first conduct univariate analyses for each characteristic of interest. We will then conduct multivariate analyses assuming no interaction of covariates and perform sensitivity analyses with assumed interactions for closely related variables.

**Assessing Applicability**

Applicability of findings may vary substantially by the PICOTS. For that reason, we will highlight how variability of PICOTS elements could influence applicability (i.e., generalizability or external validity). For example, a TRD definition may differ by population: a case in point is that the literature may differ according to what is relevant to patients 18 years of age or older who are not otherwise eligible for Medicare versus what is relevant to the Medicare population. Medicare applicability considerations might include eligibility because of end-stage renal disease or age; they might also involve clinical conditions such as cognitive impairment or long-standing coexisting chronic ailments. Also, a TRD definition relevant to specialty psychiatric settings may not be applicable (or feasible) in primary care settings. Furthermore, findings may differ depending on the definition of the primary outcome of interest (e.g., depression remission vs. improved function).

We note that CMS is interested in randomized and nonrandomized studies with control groups (including historical controls) and studies in settings that provide both diagnosis and treatment. We also note that recent clinical trial and surveillance literature considers how to identify diseases (including TRD) from claims data and electronic medical records, and we will look for evidence of the adequacy of computable phenotypes derived from such datasets.

**IV. References**


29. Gartlehner G, Gaynes BN, Amick HR, et al. Nonpharmacological Versus Pharmacological Treatments for Adult Patients With Major Depressive Disorder [Internet]. AHRQ Comparative
V. Definition of Terms

This section is not applicable for this Technology Assessment.

VI. 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:

EPC Protocol Version 17, 10/17/16
<table>
<thead>
<tr>
<th>Date</th>
<th>Section</th>
<th>Original Protocol</th>
<th>Revised Protocol</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 1,</td>
<td>Section II.</td>
<td>All comparative studies with a concurrent control group or a control group from an interrupted time-series study.</td>
<td>Added text: (which requires that data are collected at two or more time points before and after an intervention)</td>
<td>This clarification further defines what an interrupted time-series study is.</td>
</tr>
<tr>
<td>2017</td>
<td>PICOTS Comparators</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing</td>
<td></td>
<td>Any study duration</td>
<td>• Any study duration.</td>
<td>To reduce the size of the literature and make the yield more efficient, we limited the systematic reviews to be published 2005 to present. (Shojania KG, Sampson M, Ansare MT, Ji J, Doucetter S, and Moher D. How quickly do systematic reviews go out of date? A survival analysis. Annals of Internal Medicine. 2007: 147:224-233.)</td>
</tr>
<tr>
<td>Setting</td>
<td></td>
<td>All settings</td>
<td>All settings in very highly developed countries, according to the Human Development Index.</td>
<td>This focus will allow findings to be more applicable to the US.</td>
</tr>
<tr>
<td>Section III</td>
<td>Methods</td>
<td>Any study duration; literature publication date from 1/1/95 to present</td>
<td>Any study duration</td>
<td>To reduce the size of the literature and make the yield more efficient, we limited the systematic reviews to be published 2005 to present. (Shojania et al., 2007, see above) and 2005 to present for the KQ6-11 studies.</td>
</tr>
<tr>
<td></td>
<td>Table 1. Inclusion/Exclusion Criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Timing (inclusion)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Section</td>
<td>Original Protocol</td>
<td>Revised Protocol</td>
<td>Rationale</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Setting (exclusion)</td>
<td>None</td>
<td>Studies that take place in high, medium, or low human development countries.</td>
<td>This modification clarified those countries that will not be included.</td>
</tr>
<tr>
<td></td>
<td>Study Designs</td>
<td>For KQs 1–5: Consensus statements, guidelines, CMS/SAMHSA/FDA/NIH materials, UpToDate, information from the MEDCAC panel meeting on the definition of TRD on April 27, 2016, and systematic review articles</td>
<td>Added text: that (a) searched two or more literature databases, (b) included dual review of the literature and data abstraction, and (c) included quality or risk of bias assessments of included studies.</td>
<td>We provided a description of the quality assessment criteria of the systematic reviews for inclusion in this set of KQs.</td>
</tr>
<tr>
<td></td>
<td>Section III. Methods Searching for the Evidence</td>
<td>We also will review the reference lists of systematic reviews and protocols to identify any relevant citations that our electronic searches might have missed.</td>
<td>We also will review the reference lists of all systematic reviews that we include for KQs 1 through 5 and indexed protocols to identify any relevant citations that our electronic searches might have missed.</td>
<td>We clarified that the systematic reviews are only eligible for KQs 1 through 5 and that the protocols had to be indexed.</td>
</tr>
<tr>
<td></td>
<td>Searching for the Evidence (continued)</td>
<td>In addition, we will search for consensus statements, management guidelines, and relevant government materials from various Federal agencies . . . including materials presented at the MEDCAC panel of April 27, 2016. We will also search other Websites such as Clinicaltrials.gov, Guideline.gov (AHRQ’s National Guidelines Clearinghouse), and UpToDate, an evidence-based, peer-reviewed clinical information source.</td>
<td>In addition, we will search for consensus statements, management guidelines, and relevant government materials from various Federal agencies . . . including materials presented at the MEDCAC panel of April 27, 2016. Information relevant to KQs 1 through 5 will be abstracted and potentially relevant publications will be identified by reviewing the reference lists of these consensus statements, management guidelines, and government materials. We will also search other Websites such as Clinicaltrials.gov, Guideline.gov (AHRQ’s National Guidelines Clearinghouse), HSRProj (Health Services Research Projects in Progress database), and UpToDate, an evidence-based, peer-reviewed clinical information source, for potentially relevant publications.</td>
<td>We clarify the handsearching that will be done in KQs 1 to 5, and we add the HSRProj database to the list of websites to be searched.</td>
</tr>
<tr>
<td>Date</td>
<td>Section</td>
<td>Original Protocol</td>
<td>Revised Protocol</td>
<td>Rationale</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Searching for the Evidence (continued)</td>
<td>We will hand-search the systematic reviews for eligible individual studies for the Systematic Review KQs 6 through 11 in addition to those identified from our general search. From all included individual studies, we will abstract relevant information to answer KQs 6 through 11 (see below).</td>
<td>From all included individual studies, we will abstract relevant information to answer KQs 6 through 11 (see below). Because of the expected large size of the literature on pharmacologic treatments, we will randomly sample from this body of literature. We will use a sampling with replacement technique to achieve a representative sample of eligible studies. We will adopt a maximum error rate of 0.2 to calculate the necessary sample size.</td>
<td>We removed text about the KQs 1 to 5 handsearching, as it is now included above, and added text regarding the random sampling of the pharmacological treatments due to the large size of the literature search.</td>
</tr>
<tr>
<td></td>
<td>Assessment of Risk of Bias of Included Studies</td>
<td>Two investigators will independently assess the risk of bias of included individual studies included for KQ 9.</td>
<td>Two investigators will independently assess the risk of bias of individual studies included for KQ 10 only because we will use risk of bias as a covariate in the regression analyses.</td>
<td>We expanded the risk of bias assessment to KQ 10 as well.</td>
</tr>
<tr>
<td></td>
<td>Data Synthesis</td>
<td>Regression or other statistical analyses will focus on interventions for which we have at least 10 studies.</td>
<td>Regression or other statistical analyses will focus on interventions for which we have at least 10 studies using a similar control intervention. Our main focus is on interventions in general for TRD, and we are only secondarily concerned with the specific intervention type (e.g., rTMS vs. psychopharmacologic).</td>
<td>We provide further clarification of what will be included in our regression analyses.</td>
</tr>
<tr>
<td></td>
<td>Attachment A Pharmacological Interventions</td>
<td>Mianserin</td>
<td>Mianserin deleted from the list</td>
<td>We have deleted Mianserin because Mianserin is not FDA approved for use in the United States</td>
</tr>
</tbody>
</table>

AE = Associate Editor; TOO = Task Order Officer; TEP = Technical Expert Panel.

**VII. 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.

EPC Protocol Version 17, 10/17/16
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.

VIII. 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.

IX. Role of the Funder

This project was funded under Contract No. HHSA290201500011I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The 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.

X. Registration

This protocol will be registered in the international prospective register of systematic reviews (PROSPERO).
### Specific TRD Treatment Interventions by Category

#### TRD Treatment Interventions

<table>
<thead>
<tr>
<th>DEVICES</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Electroconvulsive therapy (ECT)</td>
</tr>
<tr>
<td>• Repetitive transcranial magnetic stimulation (rTMS), including theta burst stimulation</td>
</tr>
<tr>
<td>• Vagus nerve stimulation (VNS)</td>
</tr>
<tr>
<td>• Deep brain stimulation (DBS)</td>
</tr>
<tr>
<td>• Magnetic seizure therapy (MST)</td>
</tr>
<tr>
<td>• Transcranial direct stimulation (tDCS) (including theta burst stimulation)</td>
</tr>
<tr>
<td>• Low field magnetic stimulation (LFMS) (including transcranial pulsating electromagnetic field [tPEMF] stimulation)</td>
</tr>
<tr>
<td>• Cranial Electrotherapy Stimulation (CES)</td>
</tr>
</tbody>
</table>

#### PHARMACOLOGICAL INTERVENTIONS

| • Selective serotonin reuptake inhibitors (SSRIs): |
| • Serotonin-norepinephrine reuptake inhibitors (SNRIs): |
| • Noradrenergic and dopaminergic reuptake inhibitors: |
| • Monoamine oxidase inhibitors (MAOIs): |
| • 5-HT receptor antagonists (serotonin modulators): |
| • Atypical antipsychotics: |
| • N-methyl-D-aspartate (NMDA): |
| • Atypical antipsychotics: |
| • Anticonvulsants: |
| • Psychostimulants: |
| • Mood stabilizers: |
| • Other augmenters: |
| • Cognitive behavioral therapy: |

---

EPC Protocol Version 17, 10/17/16
- **assertiveness training**

- **Third wave cognitive behavioral therapies:**
  - acceptance and commitment therapy, behavioral activation, cognitive behavioral analysis system of psychotherapy, compassion focused, dialectical behavior therapy, functional analytic psychotherapy, metacognitive therapy, mindfulness-based cognitive therapy, mind training

- **Psychodynamic therapies:**
  - brief psychotherapy, countertransference, Freudian, group therapy, insight-oriented therapy, Jungian, Kleinian, object relations, person-centered therapy, psychoanalytic therapy, short-term psychotherapy, transference

- **Integrative therapies:**
  - cognitive analytical therapy, counseling, eclectic therapy, interpersonal therapy, psychodynamic interpersonal therapy, multimodal, transtheoretical

**COMPLEMENTARY AND ALTERNATIVE MEDICINE**

- Acupuncture
- Meditation (e.g., mindfulness-based stress reduction)
- Omega-3 fatty acids
- S-adenosyl-L-methionine (SAMe)
- St. John’s wort (Hypericum perforatum)
- Yoga
- Light therapy
- Sleep deprivation

**EXERCISE**

- Any formal exercise program