Washington State Health Care Authority

Health Technology Assessment

Nonpharmacologic Treatments for Treatment-Resistant Depression

Final Evidence Report

February 21, 2014

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126

http://www.hca.wa.gov/hta/Pages/

shtap@hca.wa.gov

Nonpharmacologic Treatments for Treatment-Resistant Depression

A Health Technology Assessment

Prepared for Washington State Health Care Authority

FINAL REPORT – January 31, 2014

Acknowledgement

This report was prepared by:

Hayes, Inc.
157 S. Broad Street Suite 200
Lansdale, PA 19446

P: 215.855.0615 F: 215.855.5218

TABLE OF CONTENTS

List of Tables

Key Abbreviations

AD, antidepressant (medication)

ATHF, Antidepressant Treatment History Form

DBS, deep brain stimulation

DLPFC, dorsolateral prefrontal cortex

ECT, electroconvulsive therapy

HAM-D, Hamilton Depression Rating Scale

MA, meta-analysis

MADRS, Montgomery-Åsberg Depression Rating Scale

MMD, major depressive disorder

OR, odds ratio

pt, patient

RCT, randomized controlled (or comparator) trial

RR, relative risk

rTMS, repetitive transcranial magnetic stimulation

SMD, standardized mean difference (also referred to as *effect size*)

tDCS, transcranial direct current stimulation

TRD, treatment-resistant depression

WMD, weighted mean difference

EVIDENCE SUMMARY

The **EVIDENCE SUMMARY** summarizes background information, the methods and search results for this report, findings with respect to the Key Questions, and payer policies and practice guidelines. Additional detail for all of these aspects of the report is provided, with citations, in the **TECHNICAL REPORT**. The **EVIDENCE SUMMARY** ends with an **Overall Summary and Discussion**, which is *not* repeated in the **TECHNICAL REPORT**.

Summary of Background and Technology Description

Epidemiology, Diagnosis, and Treatment of Depression

Major depressive disorder (MDD), or depression, affects approximately 7% of the adult population of the United States in any given year and 16.6% of adults over a lifetime. MDD is a leading cause of disability in the United States and many other developed countries. The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) categorizes MDD, along with bipolar disorder, as mood disorders. For a diagnosis of MDD, the fourth edition of the DSM (DSM-IV) requires clinically significant distress or impairment in social, occupational, or other important areas of life and continuation of symptoms for ≥ 2 weeks. These criteria did not change with the issuance in 2013 of the DSM-V. Treatment for depression typically consists of pharmacotherapy, psychotherapy, or a combination of these. If a particular antidepressant (AD) drug does not relieve depression symptoms or causes intolerable side effects, another class of AD may be prescribed. According to a review article, if there is a lack of adequate response after 4 to 8 weeks with an adequate dose of one of these first-line medications, the clinician adopts a switching or augmenting strategy. Switching to a different AD or a different combination of ADs is appropriate when there has been no response. Augmentation of the initial medication with an additional agent is appropriate when there has been a partial response. Different classes of AD medications may need to be tried. Treatment of a depressive episode is carried out over the *acute phase* until response is observed, through a *continuation phase* designed to prevent relapse, and eventually as part of a *maintenance phase* in which the goal is to prevent recurrence.

Treatment-Resistant Depression (TRD)

A multicenter study (Sequenced Treatment Alternatives to Relieve Depression trial [STAR*D]) of a particular algorithm for AD medication found that approximately one third of MDD patients achieved remission with an initial AD and cumulatively approximately half achieved remission after a second AD trial, provided the patients remained in treatment. These findings support a strategy of trying multiple AD medications until a response is achieved. However, analysis of the STAR*D trial results also found that with each new round of treatment that became necessary, the remission rates declined, falling

from 36.8% during the first treatment step to 13.0% during the fourth acute treatment step. Also, relapse was more common during the naturalistic follow-up phase in patients who had required multiple treatment steps. It is estimated that 4 million Americans suffer from severe depression that is refractory to multiple therapies.

There is no established definition for TRD. A common approach is to simply define TRD in terms of the number of previous AD failures. There is also no standard definition of AD *failure,* but a variety of TRD staging tools provide different ways of assessing the adequacy of prior treatment so that a clinician can determine treatment failure. These systems evaluate not only the number but also the adequacy of prior treatments. See **Appendix I** for a summary of findings from a 2007 systematic review of definitions for TRD, descriptions of staging systems, and a summary of a validation study for common staging systems. Both the 2007 systematic review and a 2011 evidence review conducted for the Agency for Healthcare Research and Quality (AHRQ) on nonpharmacologic treatments for TRD concluded that a growing consensus defines TRD as failure of ≥ 2 AD medications. Some practice guidelines set a threshold of 3 failures. None of these documents provide an evidence-based rationale for a definition of TRD. Failure of ADs from different classes is recommended by some experts and staging systems. Adequate trials are defined by different sources as those lasting from 4 to as long as 12 weeks, with some sources specifying maximum tolerable dose. According to a 2013 validation study, the most commonly used staging systems appear to be equally valid for differentiating between patients who have received adequate treatment and those who have not.

Treatment of TRD

The 2011 evidence review conducted for the AHRQ pooled data from 12 studies of different ADs for replacement of or in addition to failed pharmacotherapy in MDD patients. Mean within-group improvement was better with new pharmacologic therapies (switching or augmentation) than with no change in treatment (maintenance therapy), but overlapping confidence intervals suggest that there may not be a true difference. A small body of evidence analyzed in 2 systematic reviews, including the 2011 AHRQ evidence report, suggests that psychotherapy is effective in treating TRD, but the effectiveness of psychotherapy compared with a change in pharmacotherapy is not clear.

Continued efforts to find an effective medication or combination of medications for a patient increase the risk of drug-related adverse events and drug interactions. Thus, a number of nonpharmacologic neuromodulatory treatments for depression have been developed and tested clinically: electroconvulsive therapy (ECT), vagus nerve stimulation (VNS), deep brain stimulation (DBS), repetitive transcranial magnetic stimulation (rTMS), transcranial direct current stimulation (tDCS), magnetic seizure therapy (MST) (also called magnetic convulsion therapy), epidural cortical stimulation (ECS), and cranial electric (or electrotherapy) stimulation (CES).

This report addresses the effectiveness and safety of ECT, rTMS, tDCS, and DBS for TRD. The results of an update literature search conducted in August 2013 suggested that recent evidence is unlikely to alter the conclusions of the 2009 Washington HTA report on *Vagus Nerve Stimulation for Epilepsy and Depression*. No relevant RCTs with non-VNS control groups have been published since 2009, and a 2011 evidence review conducted for AHRQ concluded that the strength of evidence of VNS for TRD was low. Thus, VNS is not covered in the current report. MST, CES, and ECS have also been excluded due to the very small quantity and poor quality of available studies.

Measures of Treatment Outcome and Clinically Relevant Improvement

In general, trials that evaluate the effectiveness of treatments for unipolar or bipolar MDD measure outcomes in terms of *response, remission, recovery, relapse,* and *recurrence.* TRD can be thought of in general terms as referring to patients who do not remit or at least show a meaningful response after initial acute treatment.

Response refers to relative improvement on an index symptom scale and remission is typically defined as a reduction on the index scale to a particular level. Several such scales have been validated. **Appendix II** describes common scales. A common definition of meaningful response is a 50% reduction in score, relative to baseline, on a depression symptom scale; in the vast majority of studies reviewed for this report, the primary measurement scale was the Hamilton Depression Rating Scale (HAM-D or HDRS) or the Montgomery-Åsberg Depression Rating Scale (MADRS). A 50% reduction in HAM-D or MADRS score was the definition of response adopted in the 2011 evidence review of nonpharmacologic treatments for TRD prepared for AHRQ. Some older studies selected for the present report defined response as ≥ 60% on the HAM-D. Some studies defined a threshold of ≥ 25% improvement from baseline as an early indicator that a patient is responding to treatment. Remission is defined as reduction to a score below a certain point on one of the symptom scales, e.g., ≤ 8 on the HAM-$D_{17,}$ ≤ 10 on the HAM-D_{21}, or ≤ 8 on the MADRS. Practice guidelines offer varying definitions of both response and remission. However, although the scales have been validated, the literature reviewed for this report suggested that specific definitions of response and remission have not been empirically validated. The reviews and studies selected for the present report did not refer to definitions of response or remission specific to nonpharmacologic as opposed to pharmacologic treatment or to TRD as opposed to depression in general.

No standard definition of clinically relevant improvement was identified in the literature. By implication, definitions of response and remission might be assumed to denote clinically relevant improvement. However, as previously explained various definitions are in use and it appears that none have been empirically derived. Furthermore, definitions of response and remission do not answer the question of whether a smaller degree of improvement that does not meet the threshold for clinical response or remission might be considered clinically relevant. In a recent study of tDCS, a 3-point difference on the MADRS scale or an effect size of 0.5 was considered clinically relevant. Similarly, 1 of the RCTs included in the AHRQ review, which compared TMS both with sham stimulation and the AD escitalopram, identified an effect size of 0.40 as representing a minimal clinically important difference (MCID) for rTMS, based on the results of a previous placebo-controlled RCT of escitalopram (Bretlau et al., 2008). The authors of the AHRQ review suggested that pooled estimates of average improvement with *pharmacological* treatment of TRD (within-group estimates) provide an anchor against which to judge the magnitude of improvement in patients undergoing *nonpharmacologic* treatment for TRD. The following estimates are mean within-group changes on the MADRS (0 to 60 scale) for different pharmacologic approaches to TRD, where TRD involved failure of ≥ 2 AD trials:

- Switching strategies (replacement medication): 11.2-point improvement
- Augmentation strategies (add-on medication): 11.2-point improvement
- Maintenance strategies (no change in medication): 7.6-point improvement

The AHRQ author's point was that these estimates could be used to judge the within-in group effect of a nonpharmacologic treatment administered to patients with TRD, assuming patients would otherwise switch to a new medication, add a new medication, or continue with maintenance treatment using the same medication.

See **Appendix II** for a description of common scales used to measure quality of life (QOL), functional status, and disability for patients with depression. No definition of clinically relevant functional improvement was found in the literature reviewed for this report. Some studies have shown the HAM-D score to be inversely associated with function and QOL (lower HAM-D, higher function and QOL). Thus, functional improvement may progress in parallel fashion with symptom improvement. However, no studies mapping depression scores to a specific level of functional status were identified.

Technology Descriptions

Electroconvulsive Therapy (ECT)

ECT involves delivering electrical pulses to the brain via electrode pads positioned on the scalp above mood centers in the brain. These pulses cause an epileptic seizure, which results in global cerebral stimulation. An ECT session begins with a titration process in which stimulus intensity is slowly increased until it is strong enough to induce what is considered a clinically adequate seizure. This energy level is called the seizure threshold for that patient; some protocols involve stimulus at a small percentage above the seizure threshold. ECT has been considered the "control" treatment, i.e., the established therapy for TRD, in some studies of newer technologies. Its disadvantages include continued lack of acceptance, the need for anesthesia, the induction of seizures, and cognitive side effects.

Electrode placement is the most often studied parameter of treatment with ECT. Some studies have reported memory impairment following *bitemporal* ECT, also often referred to as *bilateral* ECT. Thus, unilateral ECT has been explored as a potential means of minimizing cognitive side effects. Bifrontal ECT, in which the electrodes are placed above the supraorbital ridge bilaterally, is another approach to reducing memory loss by avoiding exposure of the temporal lobes to the current.

Repetitive Transcranial Magnetic Stimulation (rTMS)

rTMS was developed as a physiologically similar but potentially more acceptable alternative to ECT. Another advantage of rTMS is that since it does not require anesthesia, it can be performed in an outpatient setting. rTMS is a noninvasive technique involving the generation of a magnetic field that penetrates the skull and induces low-level electric currents in underlying tissue, thereby altering local neuronal function without inducing seizure. This contrast with the global stimulation and induction of seizures associated with ECT.

The stimulation parameters for rTMS have evolved over time. Conventional rTMS involves either *high* frequency (up to 10 Hz) stimulation applied to the *left* dorsolateral prefrontal cortex (DLPFC) or *low* frequency stimulation (1 Hz or lower) applied to the *right* DLPFC. Recently, investigators have begun experimenting with bilateral sequential stimulation, with low-frequency (right side) stimulation applied first, followed by high-frequency (left side) stimulation. Calibration of rTMS intensity for an individual patient is based on the resting motor threshold (RMT), which is the minimum stimulus required to produce muscle twitches while the patient is at rest.

Transcranial Direct Current Stimulation (tDCS)

tDCS is a noninvasive neurostimulation method that delivers low-intensity electrical currents via 2 scalp electrodes to the cerebral cortex. Current protocols for tDCS, like those for rTMS, are designed to restore the balance in excitability between left and right DLPFC. tDCS may have advantages over rTMS in terms of cost, portability, and side effects.

Deep Brain Stimulation (DBS)

DBS has been investigated as a treatment for TRD because of its potential to rapidly modulate dysfunctional neural network activity and relieve symptoms, and because it can be switched on and off or readjusted when necessary. DBS requires the implantation of quadripolar electrodes that deliver electrical current directly into the brain. The DBS device consists of 4 components: the stimulating leads, a locking/anchoring device, extension wires, and a pulse generator. The stimulating leads are implanted through burr holes that are drilled into the skull. The extension wires are placed subcutaneously and the pulse generator is located subcutaneously in the chest/infraclavicular area. Over a period of weeks or months, the optimal stimulation, pulse duration, and amplitude are increased until the most significant therapeutic benefit, with the least side effects, has been achieved. Surgery to replace the pulse generator battery is necessary, typically every 12 months for constant stimulation treatments.

Regulatory Approval

Of the 4 technologies reviewed in this report, only ECT (Class III) and rTMS (Class II) have been approved or cleared for marketing by the FDA as treatments for depression.

Policy Context

Nonpharmacologic treatment for depression that does not respond to antidepressant (AD) medications was selected for review based on concerns about the safety, efficacy, and cost of the treatments. Depression is relatively common among adults and contributes to or is associated with higher rates of other disease processes, disability, and reduced quality of life. This review will help to identify safe and effective evidence-based care for TRD.

Summary of Review Objectives and Methods

Review Objectives

The scope of this report is defined as:

Population: Adults with major depressive disorder (MDD) or bipolar depression who have not responded to prior adequate pharmacologic treatments.

Interventions: Nonpharmacologic treatments for depression, including electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), transcranial direct current stimulation (tDCS), and deep brain stimulation (DBS).

Comparators: Sham treatment, treatment as usual, other nonpharmacologic treatment (including psychotherapy as a new treatment in response to treatment failure), pharmacologic treatment (a new medication to be tried in response to treatment failure), or combination therapy that does not include the nonpharmacologic therapy of interest.

Outcomes: Response, remission, depression severity, functional status, quality of life (QOL).

Key Questions

The following key questions will be addressed:

1. a. Are the following nonpharmacologic treatments effective for treatment-resistant depression (TRD)?

    - Electroconvulsive therapy (ECT)
    - Repetitive transcranial magnetic stimulation (rTMS)
    - Transcranial direct current stimulation (tDCS)
    - Deep brain stimulation (DBS)

   b. Does the effectiveness of these treatments vary according to treatment intensity, duration of treatment, use in an augmentation versus switch strategy, or any other variation in the manner in which TRD treatment was administered?

2. What adverse events, including withdrawal from treatment, are associated with nonpharmacologic treatments for TRD and what are the rates of withdrawal due to lack of benefit?

3. Does the effectiveness of nonpharmacologic treatments for TRD vary by subpopulation, defined by such factors as: age, race/ethnicity, gender, disease severity, disease duration, depression diagnosis (unipolar or bipolar depression), symptom type (e.g., psychotic, postpartum), comorbidities, or number and type of prior treatments (including other nonpharmacologic treatments)?

4. What are the cost implications and cost-effectiveness of nonpharmacologic therapies for TRD?

Methods

See the **Methods** section of the **TECHNICAL REPORT** and **Appendix III** (search strategies) for additional details.

Search Strategy and Selection Criteria

Systematic Reviews: Initially, evidence for this report was obtained by searching for systematic reviews and guidelines that had been published in the past *5 years* (as of July 2013). For additional evidence pertinent to Key Questions #1b (treatment parameters), #2 (safety), and #3 (differential effectiveness), the initial searches for systematic reviews and meta-analyses that were conducted in the Centre for Reviews and Dissemination database and PubMed were repeated for an earlier time frame (2003 to 2008) to identify reviews that might have included observational studies and addressed safety or differential effectiveness.

Primary Studies: Searches for primary studies published after the search time frames of the selected systematic reviews were conducted in the PubMed, Embase, and PsycINFO databases. An additional search without date limits was necessary to identify studies of patients with bipolar depression since 1 of the selected systematic reviews, an evidence review prepared for the Agency for Healthcare Research and Quality (AHRQ), excluded studies with > 20% of patients who had a diagnosis of bipolar disorder. The Excluded Studies list in the AHRQ evidence review was reviewed for studies excluded because of enrollment of > 20% of patients with bipolar depression and for comparator trials excluded because they had no sham control. See **Appendix III** for details.

Cost Studies: The National Health Service Economic Evaluation Database (NHS-EED) (2003 to 2013) and PubMed (August 2003 to August 2013) were searched for cost studies and economic evaluations published in the previous 10 years.

Final Search: An update search of all sources was conducted on November 12, 2013.

Inclusion/Exclusion Criteria: Different study designs were eligible for different technologies in accordance with the variable volume of literature across the 4 technologies of interest. Exclusion criteria included lack of information on what proportion of patients had at least 1 prior failure and other considerations. *However, systematic reviews that did not restrict study selection to TRD patients were considered if no other systematic review evidence or substantial trial data were available for a particular Key Question. Evidence from such sources was downgraded for uncertain applicability to the PICO statement.*

Quality Assessment

The process used by Hayes for assessing the quality of primary studies and bodies of evidence (see **Appendix IV**) is in alignment with the methods recommended by the GRADE Working Group. Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as AHRQ, use the phrase *strength of evidence*. The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool, along with a consideration of commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines.

Summary of Search Results

Evidence Pertaining to Key Questions

This report was based on evidence derived from the following:

- 15 systematic reviews (main source of data for approximately 70 studies not independently assessed)
- 23 randomized controlled or comparator trials (both referred to as RCTs) not included in or considered independently of the systematic reviews
- 1 post hoc analysis of RCTs
- 3 economic evaluations

The **SEARCH RESULTS** section in the **TECHNICAL REPORT** provides detail by Key Question (see **Table 8**).

Practice Guidelines

Six practice guidelines were identified.

Excluded Studies

Two studies and 1 systematic review of ECT and 6 systematic reviews of rTMS were initially considered but later excluded because of limited applicability to a TRD population, a comparator technology not named in the PICO (population-interventions-comparator-outcomes) statement for this report, and/or

the inclusion of unpublished data. The studies are identified in the corresponding section in the **TECHNICAL REPORT**.

Findings

Key Question #1a:
Are the following nonpharmacologic treatments effective for treatment-resistant depression (TRD)?

Study Characteristics

Only studies that explicitly stated or suggested that most patients had experienced ≥ 1 antidepressant (AD) medication failure were selected for evidence pertaining to Key Question #1a. The majority of the studies either only enrolled patients who had had ≥ 2 AD failures, reported that patients had experienced ≥ 2 AD failures, or reported that the mean number of prior AD failures was > 2. However, most studies did not indicate whether a priori criteria were used to determine the adequacy of prior AD trials, or whether adequacy was considered. When such criteria were described, studies often required ≥ 6 weeks for previous trials and sometimes added that maximum dose and/or 2 different classes had to have been tried. Most studies did not state explicitly whether failed AD trials had occurred in the current episode.

Some studies characterized the degree of medication resistance in terms of the Antidepressant Treatment History Form (ATHF), the Maudsley Staging Method (MSM), or Thase and Rush criteria. See **Appendix I** for descriptions of these systems. Most studies did not mention a systematic approach to assessing treatment resistance at baseline.

Nearly all studies that met inclusion criteria were assessments of acute therapy. A small number of studies also provided data from follow-up assessments at 2 weeks to 6 months after discontinuation of treatment with the technology of interest (although maintenance therapy with AD medication might have continued). One study evaluated maintenance treatment with ECT and another study of acute treatment with transcranial direct current stimulation (tDCS) evaluated a continued maintenance regimen of tDCS.

Where possible, the results selected for presentation in the **Literature Review** included data based on the Hamilton Depression Rating Scale (HAM-D) scale or the Montgomery-Åsberg Depression Rating Scale (MADRS) if the HAM-D scale was not used. All data pertaining to function and quality of life (QOL) are presented.

Electroconvulsive Therapy (ECT)

Two small double-blind, sham-controlled randomized trials (95 patients) suggested that ECT is an effective treatment for treatment-resistant depression (TRD) during the acute phase. A third unblinded RCT (n=39) showed that ECT in a switch strategy was more effective than a new AD medication. The difference in immediate posttreatment improvement between groups appeared to be substantial, although no definition of clinical significance was specified by the authors. The evidence of effectiveness was considered to be **low quality**. One- and 6-month follow-up data in 1 of the trials showed that while the ECT group had not declined, the sham group had improved and had scores comparable to those of the ECT group. Evidence of the durability of effect was **insufficient** because of the sparse data. No

controlled trials with data specific to ECT and bipolar depression were available. No trials presented data with respect to QOL or functional status (insufficient evidence). A fourth unblinded RCT (56 patients) evaluated ECT as an add-on treatment to pharmacotherapy during a maintenance phase in patients who had remitted in response to ECT treatment for medication-resistant depression, but a single small trial was considered **insufficient** to allow conclusions.

Magnitude of Effect

Neither of the sham-controlled studies evaluating the effectiveness of ECT for acute treatment reported response or remission rates. Therefore, a fair-quality systematic review of prospective, uncontrolled studies of ECT was consulted (Heijnen et al., 2010). Remission rates of 39% to 63% following ECT in 585 patients with medication resistance were reported. The review authors also reported an overall remission rate of 48%, calculated by a simple pooling of data across studies. This range might be compared with the within-group pooled remission rates that the 2011 AHRQ evidence review reported for continued pharmacologic treatment in patients with TRD: 22.3% for switching strategies, 27.2% for augmentation strategies, and 16.8% for maintenance strategies. (See Summary of Background and Technology Description, Treatment-Resistant Depression (TRD), *Treatment of TRD*; also, **Table 7** in the **TECHNICAL REPORT**.) This indirect evidence, based on studies that did not meet the inclusion criteria for the present report, suggests that the magnitude of benefit from ECT compares very favorably with and may exceed the magnitude of benefit from continued pharmacotherapy.

Older RCTs of uncertain applicability to a TRD population also suggest that ECT is effective. A fair-quality systematic review and meta-analysis published by the UK ECT Group in 2003 included several RCTs that were omitted from the 2011 AHRQ evidence review, and thus from evidence selected for Key Question #1a in the present report, because they were published prior to 1980. The results of the UK group's meta-analysis showed an effect size of 0.9, based on 6 sham-controlled studies, and an effect size of approximately 1.0 or 0.8, based on 13 studies comparing ECT with pharmacotherapy. These results suggest that ECT has a statistically large effect on depression. However, the review authors did not provide information about trial participants' medication history. The studies in these 2 meta-analyses were published between 1962 and 1989, except for 1 study published in 2000. Effect sizes for AD medication compared with placebo provide a very rough basis of comparison for the UK ECT Group results. Meta-analyses of placebo-controlled AD medication trials have yielded smaller effect sizes. One review reported 0.15 for trials submitted to the FDA but never published and 0.37 for published FDA trials (Turner et al., 2008). Another reported effect sizes of 0.21 to 0.47 by product all trials submitted to the FDA (Kirsch et al., 2008).

No studies that met inclusion criteria compared ECT with treatment as usual, psychotherapy, tDCS, DBS, or any active treatment other than rTMS (see following discussion).

Table 1. Summary of Findings, ECT, Key Question #1a

Key: AD, antidepressant medication; BD, bipolar disorder; BDI, Beck Depression Inventory; CI, confidence interval; ECT, electroconvulsive therapy; HAM-D, Hamilton Depression Rating Scale; HR, hazard ratio; NR, not reported; pharmacotx, pharmacotherapy; PICO, population-interventions-comparator-outcomes; pt(s), patient(s); QOL, quality of life; RCT, randomized controlled trial; TRD, treatment-resistant depression; tx, treatment

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
Depression outcomes, ECT vs sham or pharmacotx				
2 double-blind sham-controlled RCTs; 1 RCT comparing ECT with pharmacotx (fair) 134 pts	**Consistency**: ✓ **Applicability to PICO:** (✓). No explicit indication of TRD in 1 study. **Publication Bias:** Not tested	**Low** for study weaknesses, small quantity of data, and uncertain applicability in 1 study	Favored ECT	***Depression score change at end of tx:*** ~25 vs 18 points on HAM-D_{17} (0-54 scale) 15.6 vs 1.9 points on BDI (0-63 scale) 18.6 vs 9.6 on HAM-D_{21} (0-64 scale) (see text for discussion of uncontrolled study results)
Durability of effect: Insufficient (1 very small RCT)				
QOL/functional status: Insufficient evidence (no data)				
Maintenance tx with ECT				
1 unblinded RCT comparing ECT + pharmacotx w/ pharmacotx alone (fair) 56 pts	**Consistency:** ? **Applicability to PICO:** ✓ **Publication Bias:** ?	**Insufficient** (fair study quality and extremely small quantity of data)	Favored ECT	***Relapse rates:*** 32% vs 61% (*P*=0.036); HR, 2.32. (CI, 1.03-5.22)
ECT vs active tx other than pharmacotx: Insufficient evidence (no data)				
Clinical context Mean age 40-53 yrs; 52%-74% women Prior tx: Not fully known. In 1 trial, 70% had ≥1 AD failure and 15% of pts had prior ECT; in another, pts had 4-5 failures of prior adequate AD trials. Diagnosis: Moderate-severe depression, no BD Concomitant non-AD (psychotropic): Allowed in 1 study Tx strategy: 2 studies, switch strategy; 1 study, combination Tx parameters: 6-8 twice-weekly sessions bilateral ECT administered to inpts				

Repetitive Transcranial Magnetic Stimulation (rTMS)

The evidence is summarized in Table 2.

Effect on Symptoms

A meta-analysis representing 24 sham-controlled RCTs of rTMS (2011 AHRQ review) plus 3 additional RCTs provided favorable evidence of posttreatment efficacy (**moderate**-quality evidence). Although the 3 RCTs published after the AHRQ report did not consistently detect statistically significant differences between rTMS and sham stimulation, the overall body of evidence is consistent with regard to direction

of the results. A small quantity of data suggested that the durability of effect, i.e., the continued advantage of active rTMS over sham rTMS, may not last beyond 2 or 3 weeks after the end of treatment; rTMS may serve primarily to accelerate recovery (**low**-quality evidence).

Magnitude of Effect

The magnitude of placebo-controlled benefit suggested by the AHRQ meta-analysis (a 5.92-point between-group difference in HAM-D score reduction) can be assessed by comparing it with the scale size for the 2 instruments that were used—0 to a maximum score of 52 to 75 for HAM-D (depending on the version) and 0 to 60 for the MADRS. A relative improvement of 50% from baseline was the definition of clinical response adopted for the AHRQ review. Mean baseline scores for the studies included in the AHRQ report were in the range of 20 to 30. Comparing the weighted mean difference (WMD) of approximately 6 points with these baseline values suggests that the *mean between-group change difference* was < 50% of baseline scores. One study defined a clinically relevant response as a 3-point improvement on the MADRS (Brunoni et al., 2013b); according to this criterion, the pooled difference reported by the AHRQ review *could* be considered clinically relevant. These may not be valid methods of assessing the clinical relevance of trial results since they all entail comparing a *between-group* difference with *within-group* definitions. The AHRQ review did not provide a pooled estimate of either relative or absolute within-group reduction in symptom score for rTMS trial arms.

Another comparison that could be made is between within-group (rTMS arm only) data and estimates of within-group results in trials of new pharmacotherapy for TRD. The AHRQ meta-analysis included pooled estimates for pharmacotherapy arms (see **Table 7**). The pooled estimates of *response* reported by AHRQ were: 39.8% for switching strategies, 38.1% for augmentation strategies, and 27.3% for maintenance strategies. In 25 RCTs that were either included in the AHRQ review or reviewed independently for this report, response rates in rTMS arms ranged from 15% to 63.2%. The AHRQ estimates of *remission* for different pharmacotherapy strategies were 22.3%, 27.2%, and 16.8%. Another source of data on pharmacotherapy in TRD was the STAR*D study, which reported remission rates of 30.6% for patients treated at Step 2 (1 prior AD failure) and 14.3% for patients treated at Level 3 (2 prior AD failures) (Rush et al., 2006). By comparison, remission rates reported by 11 TMS trials ranged from 12% to 57%; this indirect comparison suggests at least comparable effects between continued pharmacotherapy and rTMS.

Number-needed-to-treat (NNT) calculations represent yet another assessment of clinical relevance. Pooled estimates derived by the AHRQ reviewers in an analysis of only those trials explicitly requiring either ≥ 1 or ≥ 2 prior AD failures yielded NNT values of 5 for response and 6 for remission were reported, suggesting that for 1 patient to experience treatment response (50% improvement from baseline), 5 patients would need to be treated with rTMS and for 1 patient to remit, 6 patients would need to be treated.

rTMS Versus Other Active Treatment

Five RCTs suggested that rTMS may be at least as effective as ECT under certain circumstances, but under other circumstances, ECT may be superior; this evidence is of **low** quality because of unexplained inconsistency in study results. Two RCTs suggested that rTMS combined with ECT is comparable to ECT alone; the 2 studies varied in the manner in which rTMS and ECT sessions were scheduled in the rTMS + ECT arms (**low**-quality evidence). *Other*

Five RCTs suggested that if rTMS has any effect on QOL or function, it is very small (**low**-quality evidence). No studies evaluated the use of rTMS as maintenance therapy after acute response (**insufficient** evidence).

Treatment Parameters

It should be noted that the bulk of evidence applies to a conventional strategy of high-frequency rTMS applied to the left DLPFC. See the details in Tables 10 to 13.

Table 2. Summary of Findings, rTMS, Key Question #1a

Key: AD, antidepressant (medication); AHRQ, Agency for Healthcare Research and Quality; BL, baseline; BD, bipolar disorder; CI, confidence interval; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; ECT, electroconvulsive therapy; HAM-D, Hamilton Depression Rating Scale; MA, meta-analysis; MDD, major depressive disorder; NNT, number-needed-to-treat; NR, not reported; PICO, population-interventions-comparator-outcomes-setting; posttx, posttreatment; pt(s), patient(s); RCT, randomized controlled trial; RR, relative risk; rTMS, repetitive transcranial magnetic stimulation; tx, treatment; WMD, weighted mean difference

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
Depression outcomes, rTMS vs sham				
1 MA (**good**; 24 **fair to good** RCTs) 3 additional RCTs **(fair)** 1372 pts total	**Consistency: (✓)** (Slight inconsistency between MA and recent RCTs) **Applicability to PICO:** ✓ **Publication bias:** None according to AHRQ analysis	**Moderate** for slight inconsistency	Favored rTMS	***WMD in change scores, depressive severity:*** –5.92 (CI, –8.15 to –3.70) (I^2=80%). WMD in individual studies consistently favored rTMS (24 RCTs) ***RR of response in trials requiring ≥1 or ≥2 AD failures:*** 2.68 (CI, 1.52-4.70; NNT=5) (16 RCTs) ***RR of remission in trials requiring ≥1 or ≥2 AD failures:*** 3.73 (CI, 1.23-11.30; NNT=6) (9 RCTs) (Pooled response/remission rates per group and risk differences NR.) ***Not represented in pooled estimates:*** Results favored rTMS but differences were not consistently significant
Durability of benefit, rTMS vs sham				
7 RCTs (fair to good)	***Consistency:*** Inconsistency in results beyond 2-3 wks **Applicability to PICO:** ✓ **Publication bias:** Not tested	**Low** for inconsistency and heterogeneity in measurement times	Possibly short-term only	Advantage over sham maintained 2-3 wks (3 RCTs) Inconsistent results at 3-6 mos (5 RCTs)

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
Depression outcomes, rTMS vs ECT				
5 RCTs (4 fair, 1 poor) 261 pts	***Consistency:*** Considerable inconsistency **Applicability to PICO:** ✓ **Publication bias:** Not tested	**Low** for study weaknesses, small volume of data, and inconsistency	Comparable or possible superiority of rTMS (2 RCTs; switch; ECT was unilateral) Favored ECT (3 RCTs; augmentation ; bilateral ECT for some pts)	***In the 3 RCTs favoring ECT (significant differences):*** Posttx HAM-D difference: CI, 3.40-14.05 (no point estimate) Difference in HAM-D change from BL: 36% points Risk difference, response: 37% points Risk difference, partial remission: 26% points Risk difference, remission: 42% points
Depression outcomes, rTMS+ECT vs ECT				
2 RCTs (fair) 44 pts	**Consistency:** ✓ **Applicability to PICO:** ✓ **Publication bias:** Not tested	**Low** for study weaknesses and sparse data	Comparable	-----
rTMS vs active tx other than ECT: Insufficient (no data)				
QOL/function				
5 RCTs (at least fair) 275 pts	***Consistency:*** Inconsistency, rTMS vs sham **Applicability to PICO:** ✓ **Publication bias:** Not tested	**Low** for small quantity of data and inconsistency	Conflicting findings, rTMS vs sham Comparable, rTMS vs ECT or rTMS+ECT vs ECT	Improvements, where observed, were very small, i.e., negligible to 2.2 points on 100-point scales.
rTMS as maintenance therapy: Insufficient evidence (no studies)				
Clinical context (rTMS vs sham trials) Mean age 40-58 yrs, 45%-68% women (not reported in AHRQ review). Diagnosis: Moderate-severe MDD (typically, according to DSM-IV); up to 20% pts w/ BD in 10 RCTs. Prior AD failures: ≥2 (21 RCTs); ≥1 (7 RCTS); unclear (6 RCTs). Explicit requirement of current episode (8 RCTs). Adequacy of prior AD trials: Where defined (infrequently), typically ≥6 wks, sometimes by ATFH score, and sometimes w/ specification of ≥2 classes and/or maximum dose. Other prior txs: Any ECT, 11.5%-64%%, where reported (infrequently). Psychiatric comorbidity: None or small percentage, where reported (NR in AHRQ review). Treatment strategy: Augmentation (20 RCTs), switch (9 RCTs), mixed (5 RCTs). Non-AD psychotropic medications allowed: Variable; mixed-strategy studies also varied as to whether new ADs were allowed. rTMS frequency and electrode placement: High left (23 RCTs), low right (3 RCTs), separate high left and low right arms (6 RCTs), bilateral sequential (1 RCT), high left and bilateral sequential arms (1 RCT). rTMS tx duration: 2-4 wks; ~10 sessions typically, 15 sessions in most recent RCTs.				

Transcranial Direct Current Stimulation (tDCS)

Pooled estimates of the effect of tDCS on depression symptoms suggested but did not prove an effect. In a good-quality systematic review published in 2013, the direction of results was found to be inconsistent across studies, and pooled estimates of response and remission were imprecise, with confidence intervals suggesting the possibility of both substantial reduction in the odds of response and remission as well as considerable increase in the odds of response and remission. Although a statistically significant effect was detected in the other review (published in 2012), the pooled percentage improvement from baseline for all tDCS groups (29%) was considerably less than the review authors' definition of response (50%). A recent RCT published after the 2 systematic reviews showed a posttreatment mean difference, favoring tDCS compared with sham, of 5.6 on the MADRS. This value exceeded the authors' definition of clinically relevant improvement (3 points) and may be reflective of the additional 2 sessions administered every other week following the initial 10 sessions over 2 weeks. The body of evidence regarding the efficacy of tDCS for TRD is of **low quality**. Evidence regarding durability of benefit was positive: follow-up results in 2 RCTs and 2 case series suggested that the benefit of tDCS persists or increases up to at least 1 month. However, this evidence is of **low quality** since only 1 of the RCTs had demonstrated a statistically significant posttreatment effect, which casts some doubt on the follow-up results, and case series are considered very poor quality because of the lack of control groups. Two RCTs suggested that maintenance treatment with continuing tDCS may be effective but the quality of this evidence is **insufficient**, given the number of patients involved and methodological weaknesses. No studies evaluated QOL or function.

No studies compared tDCS with treatment as usual, new pharmacotx, psychotherapy, rTMS, ECT, DBS, or any other active treatment.

Table 3. Summary of Findings, tDCS Versus Sham, Key Question #1a

Key: AD, antidepressant; BD, bipolar disorder; BL, baseline; CI, confidence interval; ECT, electroconvulsive therapy; MA, meta-analysis; MADRS, Montgomery-Åsberg Depression Rating Scale; NR, not reported; NS, not statistically significant; OR, odds ratio; PICO, population-interventions-comparator-outcomes; pt(s), patient(s); RCT, randomized controlled trial; sig, statistically significant; rTMS, repetitive transcranial magnetic stimulation; tDCS, transcranial direct current stimulation

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude and Clinical Relevance of Benefit
Depression outcomes, tDCS vs sham				
2 MAs (1 good [Berlim 2013c], 1 fair [Kalu 2012]), including 7 RCTs* 1 additional RCT	***Consistency:*** Inconsistency across studies and imprecise pooled estimates for response and remission ***Applicability to PICO:*** ✓ ***Publication bias:*** Not	**Low** for small quantity of data, inconsistency, and imprecision	Favored tDCS (but NS for pooled response/remission)	**ACUTE TREATMENT** ***Pooled tDCS-vs-sham effect size based on % change from BL:*** 0.74 (CI, 0.21 to 1.27; *P*=0.006); 6 RCTs (Kalu 2012) ***Pooled response (rTMS, sham, NNT† 10, pooled OR):*** 23.2%, 12.4%, OR, 1.97 (95% CI, 0.85-4.56; *P*=0.11); 6 RCTs (Berlim 2013c) ***Pooled Remission (tDCS, sham, NNT† 10, pooled OR):*** 12.2%, 5.4%, 2.13 (9.5% CI, 0.64 to 7.06; *P*=0.22); 6 RCTs (Berlim 2013c) ***Difference, mean MADRS (tDCS vs sham):*** –5.6 (CI, –1.30 to –10.01; *P*=0.01) (1 RCT)

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude and Clinical Relevance of Benefit
(good) 320 pts	tested			***Difference, mean MADRS (tDCS+sertraline vs sham+placebo):*** –11.5 (CI, –6.03 to –17.10; *P*<0.001) (1 RCT) ***Remission (OR, tDCS vs sham):*** 4.3 (CI, 1.2 to 15.6; *P*=0.02) (1 RCT) ***Remission (OR, tDCS+sertraline vs sham+placebo):*** 5.7 (CI, 1.6-20.3; *P*=0.007) (1 RCT)
Durability of benefit				
2 RCTs (not rated)*, 2 case series (very poor)	***Consistency:*** ✓ ***Applicability to PICO:*** ✓ ***Publication bias:*** Not tested	**Insufficient** (very small quantity of controlled data; posttx effect NS in 1RCT)	Sustained or additional benefit up to 1 mo	Positive results maintained at 1 mo (3 RCTs) Difference increased in favor of tDCS (1 RCT)
Maintenance tx w/ continuing tDCS				
2 RCT (1 good, 1 not rated) 30+ pts	***Consistency:*** ✓ ***Applicability to PICO:*** ✓ ***Publication bias:*** Not tested	**Insufficient** (very small quantity of data)	Sustained or additional benefit	Response persisted mean 11.7 wks (1 RCT); symptom scores increased (1 RCT)
Quality of life/functional outcomes: Insufficient evidence (no data)				
tDCS vs other active tx: Insufficient evidence (no data)				
Clinical context 46%-80% women; mean age 47-58 yrs. Diagnosis: Moderate-severe depression. 13% of pts had BD. Prior AD failures: 1.5-4.3 lifetime; only 2 RCTs required previous AD failure (≥2, different classes; current episode not specified); adequacy not defined. Other prior treatment: 33%-40% of pts had tried ECT (2 RCTs); response to ECT in current episode required (2 other RCTs). Psychiatric comorbidity: Substantial in 1 recent RCT; NR by review authors. Treatment strategy: Augmentation (4 studies), switch (1 study), switch and combination (1 study), unclear (2 studies). Treatment parameters: Most studies, 10 sessions; 1-2 MA. Concomitant non-AD psychotropic medications: Typically allowed.				

*No explicit assessment of study quality by the review authors. The RCTs selected for the Berlim review were all double-blind. Case series are generally considered very poor quality.

†Calculated on the basis of rate data supplied by the authors.

Deep Brain Stimulation (DBS)

Five very small prospective but uncontrolled studies reported response and remission rates suggesting that a substantial proportion of patients had improved at 6 months, and 2 studies showed an improvement in functional status at 1 or 2 years. However, the lack of control groups precludes a conclusion of causality. The evidence pertaining to relief of depression symptoms was positive but **insufficient** because of the small quantity of data and lack of controls. Follow-up data at 6 months after treatment and beyond were inconsistent and therefore evidence of durability of benefit is **insufficient**. The evidence pertaining to improvement in QOL/functional status was positive but **insufficient** because of the very small quantity of data.

No studies compared DBS with treatment as usual, new pharmacotherapy, psychotherapy, ECT, rTMS, tDCS, or any other active treatment.

Table 4. Summary of Findings, DBS, Key Question #1a

Key: AD, antidepressant; BL, baseline; ECT, electroconvulsive therapy; GAF, Global Assessment of Functioning; MDD, major depressive disorder; PICO, population-interventions-comparator-outcomes; pt(s), patient(s)

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
Depression outcomes				
5 prospective uncontrolled studies, including 1 w/ sham lead-in phase (86 pts) (**4 poor, 1 very poor**)	**Consistency:** ✓ **Applicability to PICO:** ✓ **Publication bias:** Not assessed	**Insufficient** for small quantity of data and poor/very poor studies	Improvement w/ respect to BL*	***Response rate:*** 40%-60% at 6 mos (4 studies); 29%-55% at 12 mos (3 studies). ***Remission rate*:** 18%-35% at 6 mos (1 study); 18%-36% at 12 mos (2 studies)
Durability of benefit				
5 studies (as above)	*Consistency:* Inconsistency Applicability to PICO: ✓ Publication bias: Not assessed	**Insufficient** for small quantity of data, poor/very poor studies, and inconsistency	Variable	Improvement vs decline after 6 mos was inconsistent across studies.
Quality of life/functional status				
2 prospective uncontrolled studies, including 1 w/ sham lead-in phase(34 pts) (2 poor)	**Consistency:** ✓ **Applicability to PICO:** ✓ **Publication bias:** Not assessed	**Insufficient** for very small quantity of data and poor/very poor studies	Improvement w/ respect to BL*	***Increase in GAF score:*** 18.4 points at 2 yrs (P=0.0009); 28.3 points at 1 yr (P<0.001) (1-100 scale)
DBS vs other active tx: Insufficient (no data)				
Clinical context Mean age 42-47 yrs; 36%-65% women. Diagnosis: Unipolar MDD (7 of 86 pts had bipolar depression); typically severe; current episode protracted (≥2 yrs or mean 5 to 9.3). Previous tx: ≥3 failed ADs (some studies explicitly required failures to have occurred in current episode). ECT tx and psychotx in most pts.				

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
Psychiatric comorbidities: NR Treatment strategy: Augmentation				

*But no difference in improvement in 1 study during 4 weeks of active stimulation compared with improvement during 4 weeks of sham stimulation.

Key Question #1b:

Does the effectiveness of these treatments vary according to treatment intensity, duration of treatment, use in an augmentation versus switch strategy, or any other variation in the manner in which TRD treatment was administered?

Findings for all 4 technologies are summarized in **Table 5**, following the discussion.

Electroconvulsive Therapy (ECT)

A good-quality systematic review of 8 double-blind randomized comparator trials reported effect sizes suggesting a small advantage to bifrontal stimulation compared with bilateral bitemporal stimulation., However, when bifrontal stimulation was compared with unilateral stimulation, a small difference favored unilateral stimulation. Most studies clearly involved TRD populations. This evidence was considered to be of **low quality** because pooled effect sizes were nonsignificant, the effect sizes for individual studies were inconsistent in direction, and studies lacked sham controls. Another fair-quality systematic review (UK ECT Group) found bilateral ECT to be superior in effectiveness to unilateral ECT (22 trials) and a higher dose to be more effective than low dose (6 trials), but the interaction between electrode placement and dose was unclear. This evidence was considered to be **moderate quality**, taking into account the lack of information regarding study quality and incomplete information on whether results are applicable to TRD populations. The UK ECT Group also estimated that a thrice-per-week regimen was more effective than a once-weekly regimen, but found that the thrice-per-week regimen did not offer an advantage over a twice-weekly regimen; this evidence was considered to be of **low quality**. Additionally, 2 small randomized comparator trials studies suggested that the superiority of bilateral ECT (1 study) and high dose ECT (1 study) may not apply when ultrabrief pulse ECT is used, but this evidence is **insufficient** to support conclusions.

Repetitive Transcranial Magnetic Stimulation (rTMS)

The only comparison of rTMS treatment parameters that was tested by more than 1 study entailed the relatively new approach of bilateral sequential application of low frequency (1 hertz [Hz]) to the right dorsolateral prefrontal cortex (DLPFC), followed by application of high frequency (10 Hz) to the left DLPFC. The comparator was standard unilateral high frequency rTMS applied to the left DLPFC. All 4 studies (total, n=373) were published in 2010 or later, were of at least fair quality, and with 1 exception had sham control arms. The study results were conflicting; with 3 studies showing no difference or a small potential difference favoring standard unilateral stimulation and 1 study suggesting that bilateral but not unilateral stimulation is effective. This evidence was considered to be **insufficient** because of the inconsistency in findings and the small quantity of data. Other treatment variations were investigated in single trials, but the evidence was **insufficient** to permit conclusions.

Transcranial Direct Current Stimulation (tDCS)

The 2 systematic reviews selected for Key Question #1a presented conflicting evidence regarding a differential effect according to number of sessions, strength of electrical current, or concurrent use of AD medication. One review detected no differences according to these 3 factors in metaregression. There was some evidence based on indirect comparisons in the other review that tDCS was more likely to be effective with fewer sessions, less intense current, and monotherapy (tDCS alone versus AD alone) compared with augmentation (tDCS added to AD). However, the paired pooled estimates in the latter review had overlapping confidence intervals. Given the inconsistency, indirect analyses, and lack of statistical significance in findings, the evidence is **insufficient** to allow a conclusion about the differential effectiveness of tDCS according to treatment parameters.

Deep Brain Stimulation (DBS)

Among the small number of clinical studies of DBS, no pattern of differential effectiveness according to treatment parameters was apparent. Given the poor quality of the studies and the very small quantity of data, the evidence is **insufficient** to allow conclusions.

Table 5. Summary of Main Findings, Key Question #1b

Key: DBS, deep brain stimulation; ECT, electroconvulsive therapy; MA, meta-analysis; PICO, patients-interventions-comparators-outcomes; RCT, randomized controlled/comparator trial; rTMS, repetitive transcranial stimulation; RUL, right unilateral; SES, standardized effect size; SR, systematic review; tDCS, transcranial direct current stimulation

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
ECT: Bifrontal stimulation				
1 SR/MA (good; 8 double-blind RCTs, but no sham controls)	**Consistency:** Inconsistent direction of findings across individual studies **Applicability to PICO:** ✓ **Publication bias:** Not assessed	**Low** for study, lack of sham controls, inconsistency, NS pooled effect sizes	Mixed	***Effect size:*** Bifrontal vs bitemporal: 0.102, favoring bifrontal but NS (5 RCTs) Bifrontal vs unilateral: –0.118 favoring RUL but NS (7 RCTs)
ECT: Bilateral vs unilateral				
1 SR/MA (good; 22 controlled trials)	**Consistency:** ✓ **Applicability to PICO:** incomplete information on treatment resistance **Publication bias:** Not assessed	**Moderate** for missing detail on study quality and some uncertainty about applicability	Favored bilateral	***Effect size:*** Fixed effects: –0.323 (CI, –0.446 to –0.1.99) Random effects: –0.322 (CI, –0.458 to –0.186)
ECT: High dose vs low dose				
1 SR/MA (good; 6 controlled trials)	**Consistency:** ✓ **Applicability to PICO:** Incomplete information on treatment	**Moderate** for missing detail on study quality and some uncertainty	Favored high dose	***Effect size:*** Fixed effects: 0.571 (CI, 0.352-

Quantity and Quality of Individual Studies	Other Quality Considerations	Quality Rating	Direction of Findings	Magnitude of Benefit
	resistance **Publication bias:** Not assessed	about applicability		0.790) (favors higher dose) Random effects: 0.575 (CI, 0.329-0.829)
ECT: Frequency of sessions				
1 SR/MA (fair) (6 RCTs, 210 pts; quality of individual studies NR)	**Consistency:** ✓ **Applicability to PICO:** ✓ **Publication bias:** Not assessed	**Low** for very small quantity of data and NS findings	Favored 3 times/wk over 1 time/wk. Favored 2 times/wk over 3 times/wk but small NS effect.	**Results are presented as SES's.** ***Once/wk vs thrice/wk (2 trials, 51 pts):*** Fixed effects: 0.841 (CI, 0.311-1.370) (favors thrice/wk) Random effects; 0.832 (–0.389 to 1.890) ***Twice/wk vs thrice/wk (SES) (4 trials, 159 pts):*** Fixed effects: –0.308 (CI, –0.629 to 0.014) (favors twice/wk) Random effects; –0.299 (–0.759 to 1.199)
rTMS: Bilateral vs unilateral high-frequency				
4 RCTs (373 patients; sham control in 3 studies; at least fair according to AHRQ or direct assessment)	**Consistency:** inconsistent **Applicability to PICO:** ✓ **Publication bias:** Not assessed	**Insufficient for** small quantity of data and inconsistency	Mixed	***Response rates (bilateral vs unilateral):*** • 20% vs 35% (NS) • 31% vs 48% (*P*=0.08) • No difference 38.5% vs 4.5% vs 10% (sham) (global *P*=0.006)
tDCS: Insufficient (conflicting results between 2 SRs/MAs); number of sessions, strength of current, concurrent use of ADs				
DBS: Insufficient (no differences detected in small number of poor-quality studies)				

Key Question #2:

What adverse events are associated with nonpharmacologic treatments for TRD and what are the rates of withdrawal due to lack of benefit?

Electroconvulsive Therapy (ECT)

Two systematic reviews (1 fair and 1 good quality) of pretest/posttest (before-and-after) data reported that ECT may result in cognitive decline over the course of treatment, although the results varied by study and by test. These effects are generally transient but according to 1 of the reviews, autobiographical memory loss may persist for several months in some individuals. The other review focused on studies of older patients (≥ 50 years, mean age across studies, 60 years); the authors were

unable to form any conclusions about the cognitive effects of ECT in older patients. A single sham-controlled RCT reported no difference in cognitive change between active and sham ECT as maintenance treatment. There is some evidence that cognitive effects can be minimized with the use of ultrabrief pulse unilateral ECT as opposed to brief pulse bilateral ECT, unilateral as opposed bilateral electrode placement, and other factors that diminish the intensity of ECT. One of the systematic reviews pointed out several issues relating to the validity of testing cognitive effects in trials of ECT for depression and the quality of the individual studies. Evidence pertaining to cognitive effects is of **low quality** due to the poor quality of studies, inconsistent findings, and problems inherent in measuring this outcome.

No large case series reporting reliable adverse event rates were identified. Among the 3 sham-controlled and 10 comparator trials of ECT selected for this report, the only serious events reported were a vascular incident involving the retina and a case of treatment-emergent mania. The evidence suggests that ECT is generally safe; this evidence is of **low quality** because of the lack of systematic analyses or large observational studies with comprehensive safety data. The only study reporting withdrawal due to lack of benefit showed a difference of 4.3% (ECT) versus 1.4% (sham), but this evidence from a single small (n=70) study was considered **insufficient**.

Repetitive Transcranial Magnetic Stimulation (rTMS)

Several of the sham-controlled RCTs and comparator trials provided safety data**.** Additional safety data were obtained from 3 systematic reviews that provided pooled estimates for rates of withdrawal and treatment-induced mania. A variety of noncognitive adverse events were reported by 7 RCTs, the most common being scalp discomfort or scalp pain, which occurred at rates of 2.1% to 35.8% in rTMS arms.

Compared with sham treatment, rTMS:

- Did not increase the risk of cognitive deterioration (6 fair- to good-quality RCTs).
- Generally did not increase the rate of withdrawal due the adverse events (16 fair- to good-quality RCTs).
- Was nonsignificantly associated with lower rates of overall withdrawals (2 fair-quality meta-analyses).
- Increase the incidence of local side effects, notably discomfort or pain in the scalp, compared with sham stimulation (7 fair- to good-quality RCTs).
- Was nonsignificantly associated with an increase (from 0.73% to 0.84%) in treatment-emergent mania, particularly in patients who had a diagnosis of bipolar disorder (1 fair-quality meta-analysis).
- Did not increase the risk of seizure. Across the 68 trials included in the AHRQ review or selected for this report, 1 case of posttreatment seizure was reported.

Compared with ECT:

- No overall difference in cognitive effects was demonstrated (3 RCTs).
- Data were insufficient to permit conclusions regarding withdrawals due to adverse events, overall withdrawals, or specific side effects due to rTMS versus ECT.

Four RCTs suggested that safety does not differ between high and low frequency rTMS or between bilateral sequential rTMS and unilateral rTMS when stimulation is delivered to conventional sites, but 1

study suggested that adverse events might increase with delivery of low frequency rTMS to the left DLPFC, which is a nonconventional protocol.

Overall, rTMS appears to be safe technology (**moderate-quality** evidence). There were no data specific to withdrawal due to lack of benefit, but indirect evidence suggesting no difference between rTMS and sham stimulation in terms of *overall* rates of withdrawal provides **low-quality** evidence that withdrawal due to lack of benefit may not differ substantially.

Transcranial Direct Current Stimulation (tDCS)

Minor reactions such as itching and skin redness are common with the use of tDCS (up to 39.3% for itching). Treatment-induced hypomania is a possibility in patients treated with tDCS. This event was reported at a collective rate of 7% in 3 studies that included a minority of patients with bipolar disorder and rates of 3% to 17% in the tDCS arms of a single RCT that recruited only patients with unipolar depression; rates of treatment-induced hypomania based on large populations were not available. The evidence pertaining to the safety of tDCS is of **moderate quality**, taking into account study quality and quantity (3 poor to good meta-analyses, 1 involving 209 studies, plus 1 good RCT), consistency of findings, and some uncertainty regarding applicability since the meta-analyses were not specific to depression. One of the meta-analyses reported dropout rates of 5.8% in active tDCS arms and 5.2% in sham arms, with a nonsignificant pooled OR, favoring sham stimulation (OR, 0.893; 95% CI, 0.259 to 3.079). The authors considered the dropout rates to be a measure of patient acceptability. This calculation provides indirect evidence suggesting that the rate of withdrawal due to lack of benefit may not differ substantially between tDCS and sham stimulation (**low-quality** evidence).

Deep Brain Stimulation (DBS)

Serious device-related events are possible with DBS, but the treatment-emergent rate of somatic events is uncertain. A narrative review article provided an estimate that hemorrhage can occur in up to 10% of patients (regardless of indication). A poor-quality systematic review of 546 English-language clinical studies and reports (≤ 10,339 patients; extent of overlapping populations unknown) tallied adverse events reported for patients undergoing DBS treatment for any indication. Of 6574 reported device-related events, 16% were due to infection, 15% involved explantation, 15% involved lead fracture, and 14% involved erosion. Less frequent device-related events included battery failure, intracranial hemorrhage, misplacement, and postoperative lead migration. Of 6573 reported somatic adverse events, a wide variety of events were reported, none of which accounted for at least 5% of events. Four non-suicide deaths and 11 cases (0.16% of all adverse events) of completed suicide were reported. The authors considered the incidence of completed suicide to be cause for concern. Two other systematic reviews of studies evaluating DBS for various psychiatric disorders reported either that no studies showed substantial cognitive decline or that cognitive side effects were generally transient. The clinical trials of DBS for TRD (5 uncontrolled studies, 86 patients) reported a variety of events, but the only common event was infection, which occurred in 5% to 20% of patients in 3 of the studies.

Evidence pertaining to safety is **insufficient** because of the lack of reliable per-patient event rates and the paucity and uncontrolled nature of data directly relevant to patients being treated for TRD. There were no data regarding withdrawal due to lack of benefit, and thus the evidence regarding this issue is **insufficient**.

Key Question #3:
Does the effectiveness of nonpharmacologic treatments for treatment-resistant depression vary by subpopulation defined by such factors as: age, race/ethnicity, gender, disease severity, disease duration, depression diagnosis (unipolar or bipolar depression), symptom type (e.g., psychotic, postpartum), comorbidities, or number and type of prior treatments (including other nonpharmacologic treatments)?

Electroconvulsive Therapy (ECT)

A fair-quality systematic review with meta-analysis (6 prospective or retrospective cohort studies, 1106 patients) reported a pooled OR of remission suggesting equivalent efficacy between patients with unipolar and bipolar major depressive disorder (MDD): OR, 1.08 (95% CI, 0.75 to 1.57). This evidence was considered to be **insufficient** because of inconsistency in direction of findings across the selected studies, high statistical heterogeneity in the pooled estimate, and unknown applicability to the population of interest since no information on treatment resistance was provided. Another fair-quality systematic review with meta-analysis (7 prospective cohort studies, 958 patients) reported a pooled OR of remission suggesting that unilateral low dose ECT is less effective in confirmed TRD than in major depression without a well-documented history of AD failure: OR, 0.52 (95% CI, 0.39 to 0.69). This evidence is of **low quality** because of poor study quality (lack of controls), inconsistency in the direction of the findings across the selected studies, and high statistical heterogeneity in the pooled estimate. Furthermore, the authors noted that the findings may not be generalizable to the current practice of using either bilateral stimulation or unilateral stimulation with a high dose. A post hoc analysis of 2 related randomized comparator trials (148 patients) found that the differential effectiveness of right unilateral ECT at a low dose (found to be relatively less effective in the 2 source trials) and right unilateral ECT at a high dose or bilateral ECT at high or low dose (found to be more effective treatment parameters) persisted in subgroups defined by psychosis, retardation, and agitation. This evidence was considered to be **low quality** because of the small quantity of data and lack of corroboration by analyses of other trials. Evidence regarding a differential effect according to age, race/ethnicity, gender, disease severity, disease duration, symptom type, or comorbidities is **insufficient** (no data).

Repetitive Transcranial Magnetic Stimulation (rTMS)

Three trials (total, n=321) suggested that the effectiveness of rTMS is *not associated* with duration of episode, and 2 trials (total, n=122) found *no association* with gender. One of the trials found *no association* of effect with unipolar versus bipolar depression. The AHRQ review calculated pooled estimates separately for trials in which study populations did or did not include patients with bipolar depression (up to 20% of study group), and estimates were very similar. One trial suggested *no difference* according to degree of medication resistance. The AHRQ review also reported pooled estimates that were somewhat smaller for trials of patients with ≥ 1 prior AD failure than for trials of patients with ≥ 2 prior AD failures, but CIs were largely overlapping, suggesting a nonsignificant difference. The evidence with regard to duration of episode, gender, unipolar versus bipolar depression, and medication resistance is of **low quality** because of the small quantity of data and/or indirect nature of some of the analysis.

A small number of trials presented conflicting evidence regarding an association with age and baseline depression severity. Other factors were investigated by single small trials and thus the evidence was

insufficient to support conclusions. Thus, the evidence regarding differential effectiveness according to age, race/ethnicity, disease severity, symptom type, comorbidities, or history of prior ECT is **insufficient**.

Trancranial Direct Current Stimulation (tDCS)

Metaregression conducted in the Kalu review suggested that the treatment effect does *not* vary according to baseline severity. However, given the small quantity of data represented by the 6 small studies and the indirect nature of metaregression, this evidence was considered to be of **low quality**. Evidence from case series suggested that among patients undergoing tDCS, response cannot be predicted on the basis of age, gender, or unipolar versus bipolar depression; this evidence was considered to be **insufficient** because of the very small quantity of data (no more than 2 studies reporting any 1 factor and sample sizes 23 to 32). Furthermore, case series do not allow an inference of a differential effect or lack of differential effect since they provide no comparison with untreated patients. For other factors (race/ethnicity, disease duration, symptom type, comorbidities, and number and type of prior treatments), the evidence was **insufficient** (no data or evaluated by single trials).

Deep Brain Stimulation (DBS)

The evidence was **insufficient** to allow conclusions. A single uncontrolled study evaluated certain response predictors and no controlled studies have been published.

Key Question #4:
What are the cost implications and cost-effectiveness of nonpharmacologic therapies for TRD?

Three economic evaluations of rTMS met criteria for review. An economic evaluation of rTMS versus sham treatment, as well as rTMS versus pharmacotherapy as usual, reported very favorable results (Simpson study). The studies serving as sources of effectiveness data for rTMS and the economic evaluation were sponsored by the manufacturer (Neuronetics). Estimates of the effectiveness of pharmacotherapy for patients with TRD came from the STAR*D (Sequenced Treatment Alternatives to Relieve Depression) trial. Reported results showed rTMS to be cost saving or to have a low cost-utility ratio compared with sham stimulation. Reported results also showed rTMS to be cost saving when compared with pharmacotherapy, regardless of whether healthcare costs only were considered or a societal perspective, considering work loss and caregiver time as well as healthcare costs, was assumed. The authors concluded that rTMS is a cost-effective treatment that may even result in cost savings, especially when used at earlier levels of treatment resistance. Numerous omissions from the study report published by Simpson and colleagues suggest reporting bias or methodological weaknesses and make interpretation of the findings very difficult. It was not clear that the evaluation supported the authors' conclusions.

A U.S. decision analysis study of rTMS, ECT, and combined rTMS-ECT assumed rTMS alone to be only slightly more effective than ECT alone for treatment of nonpsychotic patients with severe MDD (Kozel study). The study estimated a very high cost-utility ratio of $460,031 per quality-adjusted life-year (QALY) (base year unclear) for rTMS in acute and maintenance treatment versus ECT in acute and maintenance treatment. A strategy of initially treating patients with rTMS and then treating nonresponders with ECT dominated a strategy of ECT alone, i.e., the rTMS-then-ECT strategy was less expensive and more effective than ECT alone. The combination strategy might be considered a cost-effective alternative to rTMS alone ($31,783/QALY according to data supplied in the study report, costs collected 2004 or earlier). The study assumed a societal perspective and a 1-year time horizon. The

authors concluded that there may be a considerable cost advantage to a combination rTMS-ECT strategy compared with ECT alone. A key weakness of this study was the limited sensitivity analysis with respect to effectiveness estimates.

The UK trial of rTMS versus ECT, which was 1 of 3 studies that found ECT to be more effective than rTMS (see **Table 11)**, included a cost-effectiveness analysis as part of the study protocol. A separate publication (Knapp study) presented the trial-based cost-effectiveness results and was reviewed for this report. The trial showed rTMS to be considerably more effective than ECT. The findings suggested a very low probability (< 25%) that the cost utility of rTMS versus ECT would remain under £30,000/QALY ($55,282/QALY in 2013 U.S. dollars) when both cost and effectiveness assumptions were simultaneously varied. The authors concluded that it was unlikely decision makers would view rTMS as more cost-effective than ECT.

In summary, there is **insufficient** evidence to allow conclusions about the cost implications of any of the technologies of interest. A single poorly reported evaluation suggested that rTMS as an alternative to pharmacotherapy as usual could save costs. Two evaluations came to different conclusions about the economic advantages of rTMS over ECT, but these evaluations made different assumptions about the comparative effectiveness of rTMS and ECT and how rTMS might be fit into a comprehensive treatment strategy. As discussed in the findings from **Key Question #1b (Tables 2 and 11)**, the comparative effectiveness of rTMS and ECT has not been established. No economic evaluations of tDCS or DBS were identified.

Practice Guidelines

See **Practice Guidelines** in the **TECHNICAL REPORT** for additional detail.

Six relevant practice guidelines were identified. Two guidelines issued by the American Psychiatric Association (APA) on major depressive disorder (MDD) and bipolar disorder and guidelines from the Canadian Network for Mood and Anxiety Treatments (CANMAT) were considered to be of fair quality. Guidelines produced by the National Institute for Health and Care Excellence (NICE), the Institute for Clinical Systems Improvement (ICSI), and the Veterans Administration/Department of Defense (VA/DOD) were considered to be of good quality. All 5 organizations recommend electroconvulsive therapy (ECT) for medication-resistant depression. The NICE guidelines refer to repetitive transcranial magnetic stimulation (rTMS) as appropriate only in the context of research because of uncertain clinical efficacy. No other mention of rTMS and no mention of transcranial direct current stimulation (tDCS) were made in these guidelines. The CANMAT guidelines consider deep brain stimulation (DBS) to be investigational. The ICSI and VA/DOD guidelines consider failure of 3 previous antidepressant (AD) trials to establish the need for ECT but did not offer a basis for this threshold. Otherwise, the guidelines offered no definition of treatment-resistant depression (TRD).

Table 6. Summary of Practice Guidelines

Key: AD, antidepressant (medication); APA, American Psychiatric Association; BD, bipolar disorder; BDI, Beck Depression Inventory; CANMAT, Canadian Network for Mood and Anxiety Treatments; CBT, cognitive-behavioral therapy; ECT, electroconvulsive therapy; HAM-D, Hamilton Depression Rating Scale; ICSI, Institute for Clinical Systems Improvement; IPT, interpersonal therapy; MDD, major depressive episode; NICE, National Institute for Health and Care Excellence; NR, not reported; PHQ, Patient Health Questionnaire; pt(s), patient(s); rTMS, transcranial magnetic stimulation; TRD, treatment-resistant depression; tx, treatment; VA/DoD, Veterans Affairs and the Department of Defense; VNS, vagus nerve stimulation

Sponsor, Title	Relevant Recommendations	Quality*/ Comments
APA (2010) (MDD)	No definition of TRD, but guidelines imply that combination psychotherapy and AD medication should be tried before other txs are considered and cite "numerous" ineffective but adequately designed medication trials as a factor to take into account when considering ECT (e.g., pt considering suicide). ECT is considered the most effective acute phase tx for pts for whom medication and/or psychotherapy have been ineffective as acute phase txs and may be offered during the continuation phase; see text for other considerations. Light therapy is another option when medication and psychotherapy have failed. rTMS may be considered; less evidence than for ECT.	5 (no critical appraisal of evidence and unclear link between quality/quantity of evidence and recommendations)
APA (2002); APA (2005) (BD)	No definition of TRD. ECT may be considered for severe or tx-resistant bipolar depression.	5 (no critical appraisal of evidence and unclear link between quality/quantity of evidence and recommendations)
CANMAT (2009) (MDD)	No definition of TRD ECT is recommended for first-line tx for acute suicidal ideation, MDD with psychotic features, or TRD (Level 1 evidence) and for certain other indications (Level 3). Recommended as second-line treatment for patients who are otherwise treatment-resistant or who have medication intolerance. rTMS is recommended for second-line treatment (Level 1 for acute treatment and safety; Level 3 for relapse prevention). DBS is considered investigational.	5 (intended pt population for rTMS and whether it may be considered in the absence of failed ECT were unclear)
CANMAT (2013) (BD)	For depression in BD II (periods of hypomania and depression), ECT is recommended after failure of 3 prior AD trials, and for BD I (periods of mania and depression), as a third-line tx.	3 (no linking of recommendations w/ evidence)
ICSI (2012)	No definition of TRD. ADs and/or referral for psychotherapy for MDD. TRD defined as failure to achieve remission (HAM-D_{17} <7 or PHQ-9 <5) after 3 different classes of ADs. ECT, phototherapy, augmentation strategies, and hospitalization recommended for TRD. ECT may be recommended for special cases (see text).	6 (criteria for selecting evidence was not reported; body of evidence limitations NR; methods for formulating recommendations was NR)
NICE (2009)	No definition of TRD. A combination of AD medication and CBT is recommended for pts who	7

Sponsor, Title	Relevant Recommendations	Quality*/ Comments
	have not responded to drugs or psychotherapy. ECT is recommended for severe depression when other tx methods have failed. The routine use of ECT for moderate depression is not recommended, unless depression has not responded to multiple drug and psychological tx. rTMS should be reserved for research purposes only because of uncertainty about clinical efficacy.	
VA/DoD (2009)	Pts who do not respond to pharmacotherapy w/ a single agent may receive combination tx w/ pharmacotherapy and CBT or IPT. Pts who have not responded to 2 first-line ADs should either be switched to a new AD from a different class (venlafaxine is recommended, if not already tried) or receive augmentation w/ either medications or psychotherapy. Pts who have not responded to 3 different ADs should either receive augmentation with medications or psychotherapy or receive combination AD tx or ECT. Response/remission should be assessed at 8-12 wks after initiation of each new strategy. Significant response defined as 5-point reduction or score <10 on PHQ-9 or ≤25% reduction in score on an accepted standardized instrument. Remission defined as PHQ-9 ≤4, BDI ≤10, or HAM-D ≤7, maintained for ≥1 month.	6 (literature search only through December 2006; procedure for updating guideline NR)

*According to the Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool, along with a consideration of commercial funding and conflicts of interest among the guideline authors. Guidelines were scored on scale of 1 to 7 and judged to be good (6-7), fair (4-5), or poor (1-3).

Selected Payer Policies and Policy Guidance

See **Selected Payer Policies** in the **TECHNICAL REPORT** for additional detail and links to policy documents.

Electroconvulsive Therapy (ECT)

ECT is covered by Aetna for unipolar, bipolar, or mixed episode major depression under any of several very specific conditions, including lack of response to effective medication given for an adequate dose and duration (number of trials unspecified), favorable response to ECT in the past, or pregnancy. The Oregon Health Evidence Review Commission (HERC) recommends coverage of ECT for an episode of major depressive disorder (MDD) in patients who have failed ≥ 2 pharmacologic treatments. The New England Comparative Effectiveness Public Advisory Council (CEPAC) has concluded that the evidence is inadequate to support a conclusion that ECT is equivalent or superior to usual care for TRD.

No National Coverage Determination (NCD) by the Centers for Medicare & Medicaid Services (CMS) and no policies on the website for GroupHealth or Regence Group were identified.

Repetitive Transcranial Magnetic Stimulation (rTMS)

Aetna, GroupHealth, and Regence Group have noncoverage policies for rTMS. The Oregon HERC recommends coverage of rTMS for patients with an episode of MDD who have failed ≥ 2 pharmacologic

treatments. No NCD by CMS was identified. The New England CEPAC has concluded that rTMS is equivalent or better than both usual care and ECT as a treatment for TRD.

Transcranial Direct Current Stimulation (tDCS)

No policies or statements on tDCS could be identified on the Aetna, CMS, GroupHealth, Oregon HERC, or Regence Group website. This technology has not been addressed by The New England CEPAC.

Deep Brain Stimulation (DBS)

Aetna and Regence Group have noncoverage policies regarding DBS for depression. No policies or statements on DBS for depression could be identified on the CMS, GroupHealth, or Oregon HERC website. This technology has not been addressed by The New England CEPAC.

Overall Summary and Discussion

Evidence-Based Summary Statement

Electroconvulsive therapy (ECT) has been in use for decades, but its efficacy has not been extensively investigated in recent decades. Only 2 small sham-controlled trials of ECT published in 1980 and 1981 and a comparison of ECT with new pharmacotherapy published in 1997 were reviewed for this report. Trials published prior to 1980 were excluded from the 2011 AHRQ evidence report on *Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults* and hence were also excluded from the present report. Even the included RCTs of ECT may have limited generalizability to current practice because of differences in standard antidepressant (AD) medication regimens.

Some investigators have considered ECT the standard treatment for TRD against which to compare repetitive transcranial magnetic stimulation (rTMS) as an alternative more acceptable to patients. A large body of evidence has demonstrated the efficacy of rTMS compared with sham stimulation in patients with TRD. There is no evidence suggesting that rTMS is more effective than ECT, and the comparability of the 2 treatments has not been conclusively established. Although sham-controlled studies of rTMS showed no adverse effect on cognitive function, comparator trials did not prove rTMS to be safer than ECT in this regard. No serious safety issues were identified in trials of rTMS. Several variations are possible for the way ECT and rTMS are delivered, but optimal treatment parameters have not been identified. There is no evidence of at least moderate quality that establishes an association between the treatment effect of ECT or rTMS and the type of subgroup characteristics listed in Key Question #3.

These conclusions regarding ECT and rTMS are consistent with the conclusions of the 2011 Agency for Healthcare Research and Quality (AHRQ) evidence review, although the AHRQ conclusions were based only on studies in which patients had experienced ≥ 2 failed antidepressant (AD) trials. The authors of the AHRQ review noted that findings from the overall evidence base were consistent with findings specific to studies defining TRD as ≥ 2 AD failures but that the effect of combining patients meeting different definitions of TRD was unclear. The AHRQ review did not address treatment parameters and addressed only a limited number of patient characteristics (see **Appendix V**).

There is some positive evidence suggesting that transcranial direct current stimulation (tDCS) is efficacious as a treatment for TRD, but pooled estimates of response and remission were nonsignificant. Deep brain stimulation (DBS) has been studied only in very small case series with pretest/posttest

(before-and-after) analyses; although mean improvement was demonstrated, no conclusion regarding causality may be made. DBS can result in serious complications.

The literature provides very little direct evidence of improvement in quality of life (QOL) or functional status attributable to the technologies of interest. However, there is some evidence that lower symptom scores are associated with better function and QOL (this issue was not systematically investigated for this report). Thus, improvement in depression symptoms may also result in improvement in QOL and general function. Evidence of the durability of posttreatment benefits is sparse or missing for all of the technologies of interest.

One of the deficiencies of existing research evidence is the lack of a standard definition of TRD. The literature suggests a growing consensus that failure of 2 prior adequate trials of AD medication is an appropriate definition of TRD. The evidence reviewed for this report suggests that rTMS and ECT are effective even when TRD is defined as ≥ 1 failed AD trial. There was low-quality evidence suggesting that the effectiveness of ECT diminishes as medication resistance increases. The oft-cited STAR*D trial demonstrated that the effectiveness of new AD medication also diminishes with each round of failed treatment. No studies directly assessed the effectiveness of rTMS according to degree of medication resistance. Where reported in studies of tDCS, patients had generally failed ≥ 2 AD trials in their lifetime, but the evidence does not permit a conclusion regarding the effectiveness of tDCS by degree of medication. Studies of DBS required a failure of ≥ 3 prior AD trials.

Related to the lack of a standard definition of TRD is the need for greater standardization in the manner in which the adequacy of AD trials is defined and whether medication resistance should be demonstrated in the current episode. Regarding the duration of an adequate AD trial, ≥ 6 weeks was the most common threshold in studies that addressed this issue in inclusion criteria, American Psychiatric Association guidelines on the management of MDD advise that an AD trial should be continued 4 to 8 weeks, and other sources suggest 10 to 12 weeks for an adequate trial. No distinction was made in the studies reviewed for this report between patients who were and were not compliant in previous AD treatment. Very few studies clarified whether patients were enrolled on the basis of AD failure at any time in the past or on the basis of failure during the current episode.

No studies of any of the technologies of interest compared the technology with usual care or evaluated the technology as an add-on to usual care. Only 2 comparisons of nonpharmacologic treatment (ECT in 1 study, tDCS in the other) with a new AD medication as the control were identified. More studies providing a comparison with real-world alternatives are needed, ideally with a sham-controlled arm as well as a new pharmacotherapy or usual care arm.

Gaps in the Evidence

- Trials of sufficient size and design to determine the efficacy of tDCS and DBS.
- Trials comparing the technologies of interest with real-world alternatives.
- More randomized comparator trials addressing specific options for the manner in which treatments are delivered.
- A standard definition of TRD with criteria for judging the adequacy of previous AD pharmacotherapy, acknowledgment that AD failure can be due to intolerable side effects, and clarification of whether lifetime AD trials or only AD trials that took place in the current major depressive disorder (MDD) episode should be considered.

- More uniform reporting of all 3 forms of symptom outcomes: score change, response rate, and remission rate.
- Empirically derived definitions of clinically relevant improvement, response, and remission in patients with MDD.
- Randomized controlled trials (RCTs) and cohort studies powered to demonstrate differential effectiveness and safety according to patient characteristics.
- Additional cost-effectiveness analyses.

Limitations of this Report

- Study selection for trials of the efficacy of ECT and rTMS (Key Questions #1a and #1b) was restricted to randomized trials.
- The bulk of RCTs investigating ECT are not represented in the present report because trials published prior to 1980 were excluded from the key systematic review used as a source of evidence for the effectiveness of ECT.
- Results of extension studies (open-label trials in which nonresponders from both arms of an RCT may continue to receive active treatment or crossover to active treatment) were not reviewed. A small number of extension trials were excluded from the 2011 AHRQ report on nonpharmacologic treatments for TRD. Extension trial results for a recent RCT of rTMS (Fitzgerald et al., 2012) were also omitted.
- Descriptive studies, e.g., case series in which all patients receive the treatment of interest, were not considered as evidence for the differential effectiveness of ECT and rTMS according to patient characteristics (Key Question #3) even when response predictors were analyzed. Although such data do not establish differential effectiveness compared with sham or an alternative treatment, they may have implications for differential effectiveness.

TECHNICAL REPORT

Background and Technology Description

Epidemiology

According to a national survey conducted from 2001 to 2003, major depressive disorder (MDD), or depression, affects approximately 7% of the adult population of the United States in any given year and 16.6% of adults over a lifetime (Kessler et al., 2005a; Kessler et al., 2005b). The results from the 2007-2010 National Health and Nutrition Examination Survey (NHANES) revealed that nearly 8% of persons aged ≥ 12 years report current depression (defined as score ≥ 10 out of a possible 27) (CDC, 2012). MDD is a leading cause of disability in the United States and many other developed countries. If untreated, the frequency of depressive illness and the severity of symptoms typically increase over time, often leading to suicide (NAMI, 2013). The World Health Organization (WHO) estimates that the medical burden of MDD will be second only to that of ischemic heart disease by the year 2020 (Simpson et al., 2009). In the United States in 2000, the cost of depression was estimated to be $83 billion, with the cost of medical treatment at $26 billion and the social/personal economic costs, including reduced work productivity, chronic absences, and the value of lifetime earnings that were lost due to suicide, at $57 billion (NIMH, 2006).

Diagnosis of Depression and First-Line Treatment

Diagnosis

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) categorizes MDD, along with bipolar disorder, cyclothymic disorder, and dysthmic disorder as mood disorders (APA, 1994; AllPsych Online, 2013). For a diagnosis of MDD, the fourth edition of the DSM (DSM-IV) requires clinically significant distress or impairment in social, occupational, or other important areas of life and continuation of symptoms for ≥2 weeks. These criteria did not change with the issuance in 2013 of the DSM-V (APA, 2013).

Bipolar I Disorder is characterized by periods of mania, an intense high that may cause severe dysfunction, followed by periods of depression. Bipolar II Disorder is also characterized by periods of highs and depression, but the highs are hypomanic, meaning the highs are less severe and do not cause significant impairment (AllPsych Online, 2013). In the DSM-V, some changes have been made that allow a diagnosis of MDD with a specifier signifying that some features of hypomania are present but that the patient does not meet the full criteria for a diagnosis of Bipolar II Disorder (APA, 2013).

Treatment for depression typically consists of pharmacotherapy, psychotherapy, or a combination of these. Antidepressant (AD) medications help to relieve depression by normalizing levels of 1 or more neurotransmitters involved in regulating mood, particularly serotonin, norepinephrine, and dopamine (NIMH, 2011). AD medication may begin to relieve depression symptoms within a few weeks but may not achieve complete relief for 2 months and is usually continued for more than 6 months to prevent recurrence. If a particular AD drug does not relieve depression symptoms or causes intolerable side effects, another class of AD drug may be prescribed (NIMH, 2011).

Initial pharmacologic management usually consists of treatment with a selective serotonin reuptake inhibitor (SSRI), a serotonin and norepinephrine reuptake inhibitor (SNRI), bupropion, or mirtazapine. If there is a lack of adequate response after 4 to 8 weeks with an adequate dose of one of these first-line options, the clinician may adopt a switching or augmenting strategy. Switching to a different AD or a different combination of ADs is appropriate when there has been no response. Augmentation of the initial medication with an additional agent is appropriate when there has been a partial response. According to 1 review, after 2 failed AD trials, tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) rather than different first-line agents are considered (Mathys and Mitchell, 2011).

Treatment of a depressive episode is carried out over the acute phase until response is observed, through a continuation phase designed to prevent relapse, and eventually as part of a maintenance phase in which the goal is to prevent recurrence (Gaynes et al., 2011; Fountoulakis, 2012).

Treatment-Resistant Depression (TRD)

Clinical Course

A multicenter study (STAR*D) of a particular algorithm for AD medication treatment found that approximately one third of MDD patients achieved remission with an initial AD and cumulatively approximately half achieved remission after a second AD trial, provided the patients remained in treatment (Gaynes et al., 2008). These findings support a strategy of trying multiple AD medications until a response is achieved. However, analyses of the STAR*D trial results also found that with each new round of treatment that became necessary, the remission rates successively declined, falling from 36.8% during the first treatment step to 13.0% during the fourth acute treatment step. Also, relapse was more common during the naturalistic follow-up phase in patients who had required more treatment steps than in patients who were successful treated with fewer steps (Rush et al., 2006).

It is estimated that 4 million Americans suffer from depression that is refractory to multiple therapies (Ward and Irazoqui, 2010). Compared with those who respond to treatment, patients with treatment-resistant depression (TRD) have more outpatient visits, use more psychotropic medications, and are twice as likely to be hospitalized (Keitner and Mansfield, 2012). The medical costs associated with TRD in patients who are hospitalized are more than 6 times higher than the costs of patients who responded to treatment (Nemeroff, 2007). Multiple risk factors are associated with TRD, including genetic factors, comorbid medical conditions, alcohol or drug abuse, wrong medicine(s) or noncompliance, incorrect diagnosis, personality disorders, comorbid anxiety disorders, and poverty or low education level. Additionally, patients with severe or melancholic depression are more prone to developing TRD (Mathys and Mitchell, 2011; Al-Harbi, 2012). Treatment-resistant bipolar depression poses unique difficulties in the assessment of treatment response and treatment resistance (Fountoulakis, 2012).

Definition

There is no established definition for TRD (Berlim and Turecki, 2007; Gaynes et al., 2011). A common approach is to simply specify the number of previous AD failures. Several formal staging systems have been developed for systematically quantifying treatment resistance in terms of not only the number of prior failures but also whether previous treatment was adequate. These include the Antidepressant Treatment History Form (ATHF), the Maudsley Staging Method (MSM), the Massachusetts General Hospital Scale, and the Thase and Rush Scale. See **Appendix I** for a description of these systems. Authors of a recent Evidence Report prepared for the Agency for Healthcare Research and Quality (AHRQ) concluded that failure of 2 adequate trials of AD medication constitutes an emerging consensus definition (Gaynes et al., 2011). Authors of another systematic review came to a similar conclusion (Berlim and Turecki, 2007). The Berlim and Turecki review confirmed that trials have used widely differing terms for TRD, multiple assumptions about the number of AD failures that constitute TRD, and often very unclear criteria for determining that previous AD trials have been adequate. The findings from this review are described in **Appendix I**. In concert with the authors of the AHRQ report, Berlim and Turecki did conclude that a consensus is developing that TRD is defined by lack of improvement after 2 adequate trials of different classes of ADs. Berlim and Turecki also detected a consensus that adequate dose means the maximum tolerated dose and advised that an adequate trial of AD therapy may need to significantly exceed 4 weeks. They acknowledge that there is no standard definition of what constitutes *failure* in a previous AD trial and recommend that researchers adopt a common definition of remission, rather than current variable definitions of response (see **Measures of Treatment Outcome and Clinically Relevant Improvement**), to signify success. Despite the heterogeneity detected by Berlim and Turecki, a recent validation study (also summarized in **Appendix I**) concluded that the most commonly used staging systems are equally valid for documenting treatment failure in patients with depression (Hazari et al., 2013).

Regulatory and professional groups provide varying definitions that are not necessarily consistent with the conclusions of the systematic reviews that are described in the foregoing paragraph. The Food and Drug Administration (FDA) has approved repetitive transcranial magnetic stimulation (rTMS) for patients who have failed at least *1* AD medication administered in the current depressive episode at or above the minimal effective dose for at least the minimal effective duration (see following description of rTMS). The American Psychiatric Association does not endorse a definition of TRD but does describe an adequate trial of treatment as lasting 4 to 8 weeks (APA, 2010). Guidelines for diagnosis and management of major depression produced by the Institute for Clinical Systems Improvement (ICSI) define TRD as failure to achieve remission after *3* different classes of ADs (ICSI, 2012). Guidelines produced for the Department of Defense do not define TRD per se but offer electroconvulsive therapy (ECT) as an option for patients who have not responded to *3* different ADs after 8 to 12 weeks on each medication.

In bipolar depression, TRD may be defined in terms of lack of significant reduction in score on a depression symptom scale rather than in terms of the number of failed AD medications. Furthermore, the time frame required for an adequate trial of AD may need to be longer than with unipolar depression because of the greater natural fluctuation of the disease, which suggests that the clinician may need to observe a patient 2 to 4 weeks beyond the time frame usually considered adequate for an AD trial. Furthermore, since there is little class effect in the treatment of bipolar depression, the requirement of 2 medications from different classes is less relevant. The following definitions of refractoriness to depression treatment reflect criteria defined by the International Society for Bipolar Disorders (Fountoulakis, 2012):

Refractoriness: No significant reduction in score on the Montgomery-Åsberg Depression Rating Scale (MADRS) or Hamilton Depression Rating Scale (HAM-D). *Ideal trial duration 10 to 12 weeks.*

Refractoriness (maintenance phase): No change in the frequency of episodes, or MADRS/HAM-D scores > 6 > 7 between episodes. *Ideal trial duration 1 year.*

In summary, there is no standard definition of TRD. Failure of ≥ 2 adequate trials of ADs is a common definition, although some guidelines set a threshold of 3 failures. No source provided an evidence-based rationale for a definition of TRD. Failure of ADs from different classes is recommended by some experts. Adequate trials are defined by different sources as those lasting from 4 to as long as 12 weeks, with some sources specifying maximum tolerable dose. There is also no standard definition of *failure* or a stated consensus on whether failure must have occurred in the current MDD episode for there to be a diagnosis of TRD.

Continued Pharmacotherapy for TRD

A recent evidence review conducted for the AHRQ identified 12 randomized trials comparing different ADs for replacement of or in addition to failed pharmacotherapy in MDD patients (Gaynes et al., 2011). Five of the 12 studies included a maintenance arm as well, that is, a group who continued with the same pharmacologic treatment regimen. No attempt was made to synthesize the results in control arms (no AD medication) because of the lack of common control conditions. However, the results were averaged across studies in the active treatment arms. As the following pooled estimates show, mean within-group improvement was better with new pharmacologic therapies (switching or augmentation) than with no change in treatment (maintenance therapy), but overlapping confidence intervals suggest that there may not be a true difference.

Table 7. Within-Group Improvement, 12 RCTs Comparing Antidepressant Medications for Treatment of TRD (Gaynes et al., 2011)

Key: CI, confidence interval; MADRS, Montgomery-Åsberg Depression Rating Scale

Treatment Arm	Mean Change, MADRS Score	Mean Response (% Patients)	Mean Remission (% Patients)
Switching strategies	–11.2 (CI, –14.7 to –7.8)	39.8% (CI, 30.7-48.9)	22.3% (CI, 16.2-28.4)
Augmentation strategies	–11.2 (CI, –13.7 to –8.8)	38.1% (CI, 1.0-45.3)	27.2% (CI, 20.4-34.0)
Maintenance strategies	–7.6 (CI, –9.2 to –5.2)	27.3% (CI, 19.8-34.8)	16.8% (CI, 13.5-20.2)

These within-group findings provide an estimate of the degree of improvement that might be expected from new pharmacologic therapy (switch or augmentation) and from no change in treatment (maintenance therapy) as a response to TRD. As noted by Gaynes and colleagues, the estimates provide an anchor against which to judge the magnitude of improvement in patients undergoing nonpharmacologic treatment for TRD (Gaynes et al., 2011). In other words, these estimates could be used to judge the within-in group effect of a nonpharmacologic treatment administered to patients with TRD, assuming patients would otherwise switch to a new medication, add a new medication, or continue with the same medication.

There is some concern that prolonged use of psychotropic medications can lead to long-lasting changes in neurotransmitter and reception functions such that patients become susceptible to TRD (Keitner and Mansfield, 2012).

Psychotherapy for TRD

A small body of evidence analyzed in 2 systematic reviews suggests that psychotherapy is effective in treating TRD, but the effectiveness of psychotherapy compared with a change in pharmacotherapy is not clear. The evidence review conducted for AHRQ identified 4 studies comparing psychotherapy with no psychotherapy and 2 studies comparing psychotherapy with pharmacotherapy, e.g., a new AD, in patients with TRD (Gaynes et al., 2011). Cognitive-behavioral therapy (CBT) was the treatment of interest in all 6 studies. All patients had a diagnosis of MDD, generally of moderate severity. Outcomes consistently favored CBT, but statistical significance was unclear in 1 of the 2 studies comparing psychotherapy with pharmacotherapy. Another recent systematic review (Trivedi et al., 2011) of psychotherapy for patients with TRD identified 6 RCTs, 2 of which were different from any of the studies included in the AHRQ review. In 1 of these studies, augmentation with lithium resulted in better posttreatment scores than did augmentation with CBT. In the other study, no difference was observed between substitution of pharmacotherapy with CBT and continued pharmacotherapy.

Treatments for TRD Other than Pharmacotherapy and Psychotherapy

Continued efforts to find an effective medication or combination of medications for a patient not only have an increasingly smaller chance of being effective. Continuing to pursue pharmacotherapy also increases the risk of drug-related adverse events and drug interactions (Blumberger et al., 2012). Thus, several, neuromodulatory treatments have been developed and tested clinically: ECT, vagus nerve stimulation (VNS), deep brain stimulation (DBS), rTMS, transcranial direct current stimulation (tDCS), magnetic seizure therapy (MST) (also called magnetic convulsion therapy), epidural cortical stimulation (ECS), and cranial electric (or electrotherapy) stimulation (CES).

This report addresses the effectiveness and safety of ECT, rTMS, tDCS, and DBS. The results of an update literature search conducted in August 2013 suggested that recent evidence is unlikely to alter conclusions of the 2009 Washington HTA report on *Vagus Nerve Stimulation for Epilepsy and Depression*. No relevant RCTs with non-VNS control groups had been published since 2009, and a 2011 evidence review conducted for AHRQ (Gaynes et al., 2011) concluded that the strength of evidence of VNS for depression was of low quality. Thus, VNS is not covered in the current report. MST, CES, and ECS have also been excluded due to the very small quantity and poor quality of the available studies.

Measures of Treatment Outcome and Clinically Relevant Improvement

Symptom Relief

In general, trials that evaluate the effectiveness of treatments for unipolar or bipolar MDD measure outcomes in terms of *response*, *remission*, *recovery*, *relapse*, and *recurrence*. TRD can be thought of in general terms as referring to patients who do not remit or at least show a meaningful response after initial acute treatment. *Sustained* remission is a potential approach to evaluating the effectiveness of treatment during a maintenance phase, and sustained remission plus a return of function to near-normal levels might be thought of as *recovery* (Fountoulakis, 2012).

Response refers to relative improvement on an index symptom scale and remission is typically defined as a reduction on the index scale to a particular level. Several such scales have been validated. **Appendix II** describes common scales. A common definition of meaningful response is a 50% reduction in score, relative to baseline, on a depression symptom scale. In the vast majority of studies reviewed for this report, the primary measurement scale was the HAM-D (or HDRS) or MADRS. A 50% reduction in HAM-D or MADRS score was the definition of response adopted in the evidence review of nonpharmacologic

treatments for TRD prepared for the AHRQ (Gaynes et al., 2011). Some older studies defined response as ≥ 60% on the HAM-D (Sackeim et al., 1993; Sackeim et al., 2000). Some studies defined a threshold of ≥ 25% improvement from baseline as an early indicator that a patient is responding to treatment. The literature reviewed for the present report provided no indication that there is empirical evidence for these definitions of response.

Remission is defined as reduction to a score below a certain point on one of the symptom scales, e.g., ≤ 8 on the HAM-$D_{17,}$ ≤ 10 on the HAM-D_{21}, or ≤ 8 on the MADRS. However, the cutoff points for dichotomizing remission and nonremission and for defining categories of severity are consensus-based rather than empirically derived (Cusin et al., 2010; Zimmerman et al., 2012).

Practice guidelines offer varying definitions. APA guidelines for treatment of MDD do not offer a definition of response, remission, or clinical relevance other than to equate *effectiveness* with "moderate" relief of symptoms (*moderate* is not defined) (APA, 2010). ICSI guidelines define remission as HAM-D < 7 or Patient Health Questionnaire (PHQ)-9 < 5 (ICSI, 2012). Veterans Administration/Department of Defense (VA/DoD) guidelines similarly define remission as PHQ-9 ≤ 4, Beck Depression Inventory (BDI) ≤ 10, or HAM-D ≤ 7, maintained for ≥ 1 month. The VA/DoD guidelines define response as a 5-point reduction or score < 10 on the PHQ-9 or ≤ 25% reduction in score on an accepted standardized instrument (VA/DoD, 2009). (NOTE: The VA/DoD definition of response matches the definition of *partial response* in some studies.)

The studies assessed response and remission to the interventions of interest for this report similarly to the manner in which lack of response to AD medications usually determined rather than according to response/remission definitions specific to treatment of TRD. The following review of the evidence will discuss not only response and remission rates but also, where available, mean reduction in symptom score.

The American College of Neuropsychopharmacology Task Force offers a somewhat different definition of remission: complete absence of sad mood and anhedonia (inability to experience pleasure) *and* presence of ≤ 3 of the 9 diagnostic criteria for MDD (DSM-IV) (Mathys and Mitchell, 2011):

- Sad or depressed mood most days
- Decreased interest in activities that were pleasurable in the past
- Significant changes in weight or appetite (either increased or decreased)
- Sleeping too little or too much
- Psychomotor agitation or slowing
- Low energy
- Feelings of worthlessness or guilt
- Trouble concentrating or making decisions
- Thoughts of death or wanting to die

Clinical Relevance

No standard definition of clinically relevant improvement was identified in the literature. By implication, definitions of response and remission might be assumed to denote clinically relevant improvement. However, various definitions are in use and it appears that none have been empirically derived. Furthermore, definitions of response and remission do not answer the question of whether a smaller degree of improvement that does not meet the threshold for clinical response or remission might be considered clinically relevant. In a recent study of tDCS, a 3-point difference on the MADRS scale or an

effect size of 0.5 was considered clinically relevant (Brunoni et al., 2013b). The source of this assumption was a 2005 set of guidelines for treatment of depression, issued by the National Institute for Clinical Excellence (NICE) (currently referred to as the National Institute for Health and Care Excellence). Current NICE guidelines do not address clinical relevance (NICE, 2009). Similarly, 1 of the RCTs included in the AHRQ review, which compared TMS both with sham stimulation and the AD escitalopram, identified an effect size of 0.40 as representing a minimal clinically important difference (MCID) for rTMS, based on the results of a previous placebo-controlled RCT of escitalopram (Bretlau et al., 2008).

The authors of the AHRQ review suggested that estimates of average improvement with pharmacological treatment of TRD (within-group estimates) provide an anchor against which to judge the magnitude of improvement in patients undergoing *nonpharmacologic* treatment for TRD. As noted earlier, the review provided the following pooled estimates of mean changes on the MADRS (0 to 60 scale) for different pharmacologic approaches to TRD in trials where patients had experienced ≥ 2 AD failures (Gaynes et al., 2011):

- Switching strategies (replacement medication): 11.2-point improvement
- Augmentation strategies (add-on medication): 11.2-point improvement
- Maintenance strategies (no change in medication): 7.6-point improvement

Functional Improvement

See **Appendix II** for a description of common scales used to measure quality of life (QOL), functional status, and disability. No definition of clinically relevant functional improvement was found in the literature reviewed for this report. One of the selected studies found an 81% concordance between end-of-treatment HAM-D (symptom) score and Global Assessment of Functioning (GAF) score, comparing high (> median) with low (< median) values (Pridmore, 2000). Other studies have also shown that lower HAM-D scores are statistically associated with better QOL and functional status (Hung et al., 2009; Zimmerman et al., 2012). Thus, functional improvement may progress in parallel fashion with symptom improvement. However, no studies mapping depression scores to a specific level of functional status were identified.

Special Issues Associated with Bipolar Depression

The general concepts of response and remission are difficult to apply to bipolar depression. Return to high function and/or improvement on a symptom scale could signal response to AD treatment, a new manic episode, or transition to mixed or rapid cycling (combination of manic and depressive bipolar disorder). The following set of recommendations, reflecting criteria defined by the International Society for Bipolar Disorders, entails a graded approach to assessing treatment effectiveness for bipolar depression (Fountoulakis, 2012):

> Response (acute phase): < 25%, 25%-49%, 50%-74%, 75%-100% reduction in MADRS or HDRS scores. No significant increase in Young Mania Rating Scale (YMRS) or Mania Rating Scale (MRS) scores and YMRS and MRS scores stay below 5. *Ideal trial duration 10 to 12 weeks.*
> Response (maintenance phase): Significant change in the frequency of episodes. *Ideal trial duration 1 year.*
>
> Remission (acute phase): MADRS and HDRS scores stay below 6. No significant increase in YMRS or MRS scores and YMRS and MRS scores stay below 5. *No recommendation regarding duration.*

Remission (maintenance phase): Very rare new episodes, and MADRS/HDRS scores < 6 and YMRS/MRS scores < 7 between episodes. *Ideal trial duration 2 to 3 years (followed by a '?' in the article).*

Recovery (acute phase): MADRS and HDRS scores stay below 6. No significant increase in YMRS or MRS scores and YMRS and MRS scores stay below 5. *Ideal trial duration 8 weeks.*
Recovery (maintenance phase): No new mood episodes and MADRS/HDRS scores < 6 and YMRS/MRS scores < 7 between episodes. *Ideal trial duration 3 to 5 years (followed by a '?' in the article).*

None of the studies reviewed for this report used the definitions recommended by Fountoulakis. Only 1 study involving exclusively bipolar patients was identified.

Technology Descriptions

Electroconvulsive Therapy (ECT)

ECT involves delivering electrical pulses to the brain via electrode pads positioned on the scalp above mood centers in the brain (Gaynes et al., 2011; Mayo Clinic, 2012). These pulses cause an epileptic seizure, which results in global cerebral stimulation (Hansen et al., 2011). ECT procedures are performed under general anesthesia with a muscle relaxant, and trigger brief seizures, which alter brain chemistry as a way to alter mood. The anesthetic and muscle relaxant prevent severe bodily convulsions and awareness of the seizures (Mayo Clinic, 2012). An ECT session begins with a titration process in which stimulus intensity is slowly increased until it is strong enough to induce what is considered a clinically adequate seizure. This energy level is called the seizure threshold for that patient. Some protocols involve stimulus at a small percentage above the seizure threshold (Sackeim et al., 2000). ECT has been considered the "control" treatment, i.e., the established therapy, in some studies of newer technologies (Rosa et al., 2006). Its disadvantages include continued lack of acceptance, the need for anesthesia, the induction of seizures, and cognitive side effects (Eranti et al., 2007; Gaynes et al., 2011).

Electrode placement is the most often studied parameter of treatment with ECT. Some studies have reported memory impairment following *bitemporal* ECT, also often referred to as *bilateral* ECT. Thus, unilateral ECT has been explored as a potential means of minimizing cognitive side effects. Bifrontal ECT, in which the electrodes are placed above the supraorbital ridge bilaterally, is another approach to reducing memory loss by avoiding exposure of the temporal lobes to the current (Dunne and McLoughlin, 2012).

Because ECT was introduced prior to FDA device regulation, it was not subjected to formal review and approval as a device. Some ECT devices have been cleared for marketing, under the 510(k) process, for multiple mental disorders, including MDD and bipolar disorder. However, ECT is currently classified as a Class III device (21 CFR §882.5940). The FDA held a meeting in January 2011 to review the evidence for the effectiveness and safety of ECT for purposes of considering a possible reclassification as a formal 510(k) device. See FDA Executive Summary of Meeting to Discuss the Classification of Electroconvulsive Therapy Devices (ECT) (p. 9). A search of the FDA database indicates that no change in the FDA classification has been made to date; ECT devices are still considered Class III devices. See CFR – Code of Federal Regulations Title 21, Part 882: 21CFR882.5940. Class III applies to those devices for which "insufficient information exists to determine that general controls are sufficient to provide reasonable assurance of its safety and effectiveness." See CFR – Code of Federal Regulations Title 21, Part 860: 21CFR860.3.

Repetitive Transcranial Magnetic Stimulation (rTMS)

rTMS was developed as a physiologically similar but potentially more acceptable alternative to ECT. Another advantage of rTMS is that since it does not require anesthesia, it can be performed in an outpatient setting. (Pridmore, 2000; Keshtkar et al., 2011; Blumberger et al., 2012).

rTMS is a noninvasive technique that involves superficial but powerful magnetic stimulation of the brain. This is achieved by passing electrical energy through a handheld electromagnetic stimulation coil that is positioned on the scalp above the target cortical center. A pulsed electrical current passing through a coil generates a magnetic field. The magnetic field penetrates the skull and induces low-level electric currents in underlying tissue, thereby altering local neuronal function without inducing seizure, in contrast to the global stimulation and induction of seizures associated with ECT (McLoughlin et al., 2007; Hansen et al., 2011).

The stimulation parameters for rTMS have evolved over time. Conventional rTMS involves either *high*-frequency (up to 10 hertz [Hz]) stimulation applied to the *left* dorsolateral prefrontal cortex (DLPFC) or *low*-frequency stimulation (1 Hz or lower) applied to the *right* DLPFC. Recently, investigators have begun experimenting with bilateral sequential stimulation, with low-frequency (right side) stimulation applied first, followed by high-frequency (left side) stimulation. The rationale for bilateral stimulation with different frequencies is based on imaging studies showing an asymmetry of cortical excitability in patients with MDD: hyperfunction on the right side and depressed excitability on the left side. Low-frequency rTMS has been shown to induce transient inhibition, which may be therapeutic on the right side of the cortex, and high-frequency rTMS has been shown to increase excitability, which may be therapeutic on the left side. The bilateral approach may rectify a possible imbalance in prefrontal activities in depression (McLoughlin et al., 2007; Blumberger et al., 2012). Other aspects of the biological rationale behind TMS include the possibility of increased dopamine transmission, as suggested by preclinical studies (Aleman, 2013). Calibration of rTMS intensity for an individual patient is based on the resting motor threshold (RMT), which is the minimum stimulus required to produce twitches in a target muscle, e.g., abductor pollicis brevis/thenar muscle (part of the thumb musculature) contralateral to the side of the head that will receive stimulation (Kennedy et al., 2009).

Until recently, the only magnetic stimulator system cleared by the FDA for marketing as treatment for depression was the NeuroStar TMS Therapy System (Neuronetics Inc.). This system was approved in 2008 for treating adult patients with MDD only when the affected patient has failed to attain satisfactory improvement from ≥ 1 AD medication administered in the current depressive episode at or above the minimal effective dose for at least the minimal effective duration and only when TMS is prescribed by and performed under the supervision of a licensed psychiatrist. In January 2013, the Brainsway Deep TMS System (Brainsway Ltd.) was cleared for marketing as substantially equivalent to the NeuroStar system and for the same intended use. These devices are Class II devices. See 501k Premarket Notification Database, Product Code OBP.

Transcranial Direct Current Stimulation (tDCS)

tDCS is a noninvasive neurostimulation method that delivers low-intensity electrical currents via 2 scalp electrodes to the cerebral cortex. The resulting neural activity depends on the polarity of the current. Current protocols for tDCS, like those for rTMS, are designed to restore the balance between left and right DLPFC. Anodal tDCS stimulation is delivered to the left DLPFC, which is thought to be hypoactive during depression, to enhance cortical excitability. Cathodal stimulation through the return electrode is delivered to the right DLPFC, thought to be hyperactive during depression, to reduce cortical excitability.

tDCS has been studied as a treatment for mood and depressive diseases for several decades, but current treatment parameters have been used only since the late 1990s. tDCS may have advantages over rTMS in terms of cost, portability, and side effects (Kalu et al., 2012; Berlim et al., 2013c). The FDA has not approved any devices for tDCS. The studies represented by this report were conducted in Europe, Brazil, and Australia. Most did not identify a commercial product, but simply described the use of saline-soaked surface sponge electrodes and a constant current stimulator. Two studies referred to a stimulator marketed by a German company.

Deep Brain Stimulation (DBS)

DBS was initially investigated as a treatment for depression in 1954 and was first used to relieve Parkinsonian tremors in 1987. DBS shows promise as a treatment for TRD because of its potential to rapidly modulate dysfunctional neural network activity and relieve symptoms, and because it can be switched on and off or readjusted when necessary (Ward and Irazoqui, 2010).

DBS requires the implantation of quadripolar electrodes that deliver electrical current directly into the brain. The DBS device consists of 4 components: the stimulating leads, a locking/anchoring device, extension wires, and a pulse generator. The stimulating leads are implanted through burr holes that are drilled into the skull. This is generally performed under local anesthesia while the patient is awake. The extension wires and pulse generator are then implanted under general anesthesia. The extension wires are placed subcutaneously and the pulse generator is located subcutaneously in the chest/infraclavicular area. Over a period of weeks or months, the stimulation, pulse duration, and amplitude are increased until the most significant therapeutic benefit, with the least side effects, has been achieved. Surgery to replace the pulse generator battery is necessary, typically every 12 months for constant stimulation treatments (Volkmann et al., 2002; Ward and Irazoqui, 2010; Goodman and Alterman, 2012).

The FDA has not approved DBS as a treatment for depression. Surgically implantable devices for DBS are regulated by the FDA as Class III devices that are subject to the premarket approval (PMA) process. DBS systems approved through this process include implanted electrical stimulators approved for Parkinsonian tremor: PMA Database, Product Code MHY.

Washington State Utilization and Cost Data

Figure 1. All Agency Non-pharmacologic Treatments for TRD, 2009-2012. Non-pharma treatments for depression include electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), deep brain stimulation (DBS), transcranial direct current stimulation (tDCS), vagus nerve stimulation (VNS).

Agency/Year	2009	2010	2011	2012	4 Yr Overall Total**	Avg Annual % Change	
Public Employee Benefits/Uniform Medical Plan							
PEB/UMP Average Annual Members	210,501	213,487	212,596	212,684		0.3%	
PEB/UMP Members with Depression[1] (9.2% of members on average)	19,475	19,922	19,581	19,425		-0.4%	*
ECT Patients (all with depression[2] diagnoses)	**26**	**32**	**30**	**30**	**72**	**3.8%**	*
ECT Procedures (treatment days)	404	439	493	380	1716	-3.9%	*
Average Count per Patient	15.5	13.7	16.4	12.7	23.8	-4.9%	
Max Count per Patient	48	48	62	49	205	2.7%	
ECT Total Paid [3]	**$298,744**	**$288,606**	**$384,272**	**$312,751**	**$1,284,373**	-0.9%	*
Average Paid per Patient[3]	$11,490	$9,019	$12,809	$10,425	$17,839	-7.4%	
Average Paid per Patient, PEB/UMP Primary[4]	$16,756	$10,891	$16,508	$15,548	$23,067	-9.3%	
Maximum Paid (note outliers in 2010 and overalll)	$45,303	$31,562	$39,676	$29,458	$97,025		
95% Upper Limit (2 standard deviations above mean)	$45,247	$27,126	$40,103	$30,029	$61,312		
Average Paid per Procedure [3]	$739	$657	$779	$823	$748	4.4%	
Average Paid per Procedure, PEB/UMP Primary [4]	$1,160	$1,122	$1,138	$1,503	$1,214	10.1%	

Figure 1 continued

Agency/Year	2009	2010	2011	2012	4 Yr Overall Total**	Avg Annual % Change	
Medicaid							
Medicaid FFS*** Population	463,966	474,676	473,356	477,727		1.0%	
Medicaid claimants with Depression[1] (11.1% of members on average)	54,869	54,787	51,422	49,507		-4.3%	*
ECT Patients (depression and schizophrenia diagnoses)[5]	**43**	**55**	**45**	**28**	**134**	**-10.4%**	*
ECT Procedures (treatment days)	136	182	128	99	545	13.7%	*
Average Count per Patient	3.2	3.3	2.8	3.5	4.1	5.0%	
Max Count per Patient	8	9	10	10	22	7.9%	
ECT Total Paid [3]	**$26,017**	**$30,959**	**$14,574**	**$14,726**	**$86,275**	-13.4%	*
Average Paid per Patient[3]	$605	$563	$324	$526	$644	4.3%	
Average Paid per Patient, Non-Medicare[4]	$652	$686	$496	$925	$787	21.3%	
Maximum Paid (outliers)	$1,887	$3,095	$1,314	$2,506	$5,788		
95% Upper Limit (2 standard deviations above mean)	$1,605	$1,924	$1,017	$1,986	$2,474		
Average Paid per Procedure [3]	$191	$170	$114	$149	$158	-4.5%	
Average Paid per Procedure, Non-Medicare [4]	$215	$227	$231	$278	$230	9.1%	
Labor & Industry							
L&I Annual Claims	125,611	122,712	121,043	121,660		-1.1%	*
ECT Patients (all with Depression[2] Diagnoses)	**2**	**2**	**3**	**3**	**7**	**12.0%**	*
ECT Procedures (treatment days)	28	34	44	54	160	20.5%	*
Average Count per patient	14.0	17.0	14.7	18.0	22.9	10.1%	
Max Count per patient	16.0	30.0	20.0	26.0	60.0	28.1%	
ECT Total Paid* **	**$29,535**	**$36,331**	**$56,186**	**$62,181**	**$184,232**	22.0%	*
Average Paid per Patient[3]	$14,767	$18,165	$18,729	$20,727	$26,319	11.4%	
Average Paid per Procedure [3]	$1,055	$1,069	$1,277	$1,152	$1,151	3.7%	
Maximum Paid	$18,208	$28,255	$20,732	$28,004	$59,253		

Figure 1 footnotes

*Population adjusted average change

**Four year patient counts represent unique patients and are not necessarily the total of patient counts over 4 years.

***FFS: Medicaid Fee For Service

1 - all patients with any diagnosis code in the ranges below

Depression	290.2	Senile dementia with delusional or depressive	301.12	Chronic depressive personality disorder
	296.*	Episodic mood disorders	309.*	Adjustment disorders
	298.9	Depressive type psychosis	311.*	Depressive disorder

2 - Episodic and depression diagnosis codes

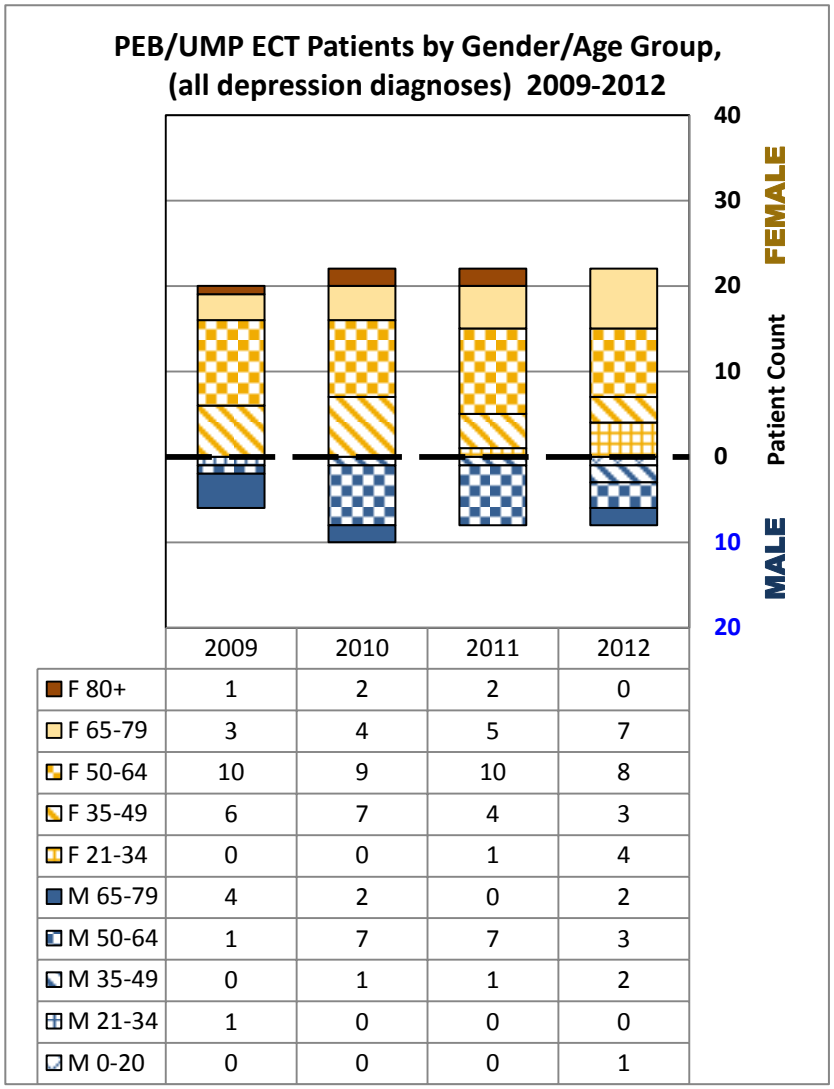
3 - Paid charges include only those reimbursed, and exclude payments by other carriers and patients. Payments for diagnosis related charges within 7 days of ECT delivery; anesthesia, facility charges, post ECT care, observation, pre and post ECT imaging are included.

4 - Only patients where PEB/UMP or Medicaid is the primary payer (e.g. non-Medicare), about 61% of PEB/UMP ECT claimants, about 65% of Medicaid ECT claimants.

5 – Medicaid ECT claims were mainly for episodic and depression diagnoses (about 85%), but 14 claims were for schizophrenia, and 3 were for psychosis.

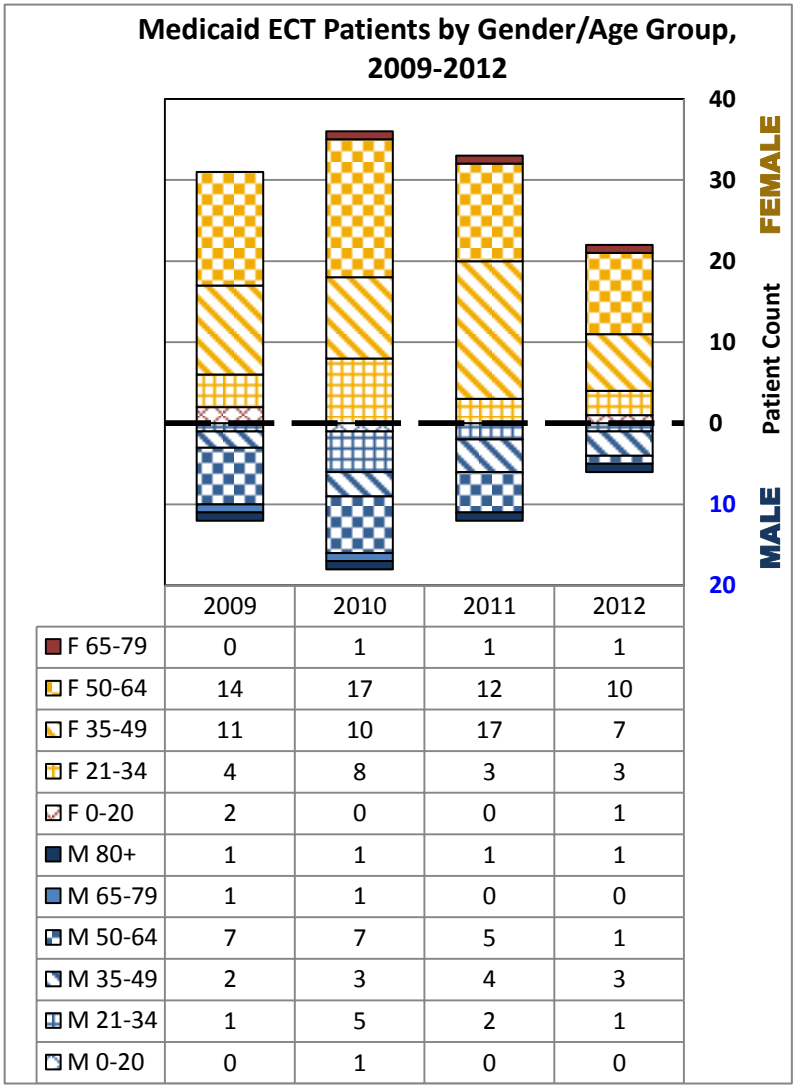
Note: only one claim was reported in PEB/UMP for rTMS (in 2012); not reimbursed. L&I and Medicaid report no rTMS claims. PEB, L&I and Medicaid had no claims for DBS, tDCS, or VNS, for depression. Average allowed costs for cranial neurostimulators for DBS and VNS, though for diagnoses unrelated to the current investigation, may be representative and are shown in Figure 3b.

Figure 2a. PEB/UMP ECT Patients by Gender and Age Group, 2009-2012



	2009	2010	2011	2012
F 80+	1	2	2	0
F 65-79	3	4	5	7
F 50-64	10	9	10	8
F 35-49	6	7	4	3
F 21-34	0	0	1	4
M 65-79	4	2	0	2
M 50-64	1	7	7	3
M 35-49	0	1	1	2
M 21-34	1	0	0	0
M 0-20	0	0	0	1

Figure 2b. Medicaid Patients by Gender and Age Group, 2009-2012



	2009	2010	2011	2012
F 65-79	0	1	1	1
F 50-64	14	17	12	10
F 35-49	11	10	17	7
F 21-34	4	8	3	3
F 0-20	2	0	0	1
M 80+	1	1	1	1
M 65-79	1	1	0	0
M 50-64	7	7	5	1
M 35-49	2	3	4	3
M 21-34	1	5	2	1
M 0-20	0	1	0	0

L&I Demographics for ECT Patients, 2009-2012	Gender	Age Group	Count	Gender	Age Group	Count
	Female	21-34	1	Male	35-49	2
		35-49	2		50-64	2

Figure 3a. All Agency Average Allowed Amounts for ECT Patients

ECT Average Allowed per Patient	PEB/UMP Medicare	PEB/UMP Primary	Medicaid Medicare	Medicaid Non-Medicare Primary
Patient Count	n=21	n=44	n=43	n=85
Treatment Count Average	38	18	5	3
Cost Breakdown 1				
Anesthesia	$4,063	$5,432	$251	$88
Treatment Delivery	$24,682	$15,067	$950	$335
Hospital Care	$4,160	$1,336	$21	$104
Imaging/Other	$2,574	$1,463	$175	$146
Cost Breakdown 2				
Facility	$26,668	$13,949	$842	$207
Provider	$8,810	$9,349	$554	$467
Total Allowed	**$35,479**	**$23,299**	**$1,396**	**$674**

Figure 3b. All Agency Average Allowed Amounts for Cranial Neurostimulators* 2009-2012.

NS* Average Allowed per Patient	PEB/UMP Medicare	PEB/UMP Primary	Medicaid Medicare	Medicaid Non-Medicare Primary
Patient Count	n=43	n=26	n=14	n=49
Cost Breakdown 1				
Anesthesia	$300	$1,217	$399	$49
Implantation	$20,887	$36,361	$6,739	$8,190
Device/Electrodes	$5,595	$8,550	$46	$493
Revision/Repair	$436	$678	$304	$115
Hospital Care	$17,290	$1,251	$0	$7
Imaging/Other	$1,389	$842	$9	$25
Cost Breakdown 2				
Facility	$43,967	$44,350	$6,963	$8,648
Provider	$1,931	$4,549	$534	$232
Total Allowed	**$45,898**	**$48,899**	**$7,497**	**$8,880**

***Cranial Neurostimulators for depression were not reported in PEB/UMP or Medicaid claims, but were reported for other diagnoses (epilepsy and tremor). The allowed amounts are presented to give an approximation of the cost of a similar procedure for the diagnoses of interest.

Note: L&I had too few claims to present average payment breakdowns.

Figure 4a. PEB/UMP Top Diagnosis Codes for ECT patients

Top Diagnoses	PEB/UMP ECT Patients n=72	Patients	% of All ECT Patients
ECT	RECUR DEPR PSYCH-SEVERE	38	52.8%
ECT	RECURR DEPR PSYCHOS-UNSP	31	43.1%
ECT	EPISODIC MOOD DISORD NOS	12	16.7%
ECT	BIPOL I SINGLE MANIC NOS	11	15.3%
ECT	DEPRESSIVE DISORDER NEC	11	15.3%
ECT	DEPRESS PSYCHOSIS-UNSPEC	10	13.9%
ECT	BIPOL I CURR DEP W/O PSY	8	11.1%
ECT	BIPOLAR DISORDER NEC	8	11.1%
ECT	BIPOL I CUR DEPRES NOS	7	9.7%
ECT	DEPRESS PSYCHOSIS-SEVERE	7	9.7%
ECT	REC DEPR PSYCH-PSYCHOTIC	7	9.7%

Figure 4b. Medicaid Top Diagnosis Codes for ECT patients

Top Diagnoses	Medicaid ECT Patients n=134	Patients	% of All ECT Patients
ECT	RECUR DEPR PSYCH-SEVERE	97	72.4%
ECT	SCHIZOAFFECTIVE DIS NOS	77	57.5%
ECT	BIPOL I SINGLE MANIC NOS	64	47.8%
ECT	BIPOL I CUR DEPRES NOS	53	39.6%
ECT	RECURR DEPR PSYCHOS-UNSP	44	32.8%
ECT	BIPOL I CURR DEP W/O PSY	40	29.9%
ECT	REC DEPR PSYCH-PSYCHOTIC	34	25.4%
ECT	EPISODIC MOOD DISORD NOS	25	18.7%
ECT	DEPRESSIVE DISORDER NEC	21	15.7%
ECT	BIPOLAR DISORDER NEC	13	9.7%
ECT	FOLLOW-UP EXAM NOS	13	9.7%
ECT	SIMPL SCHIZOPHREN-UNSPEC	13	9.7%

Related Medical Codes

Transcranial Magnetic Stimulation		
	90867	TRANSCRANIAL MAG STIMJ TX PLANNING
	90868	TRANSCRANIAL MAG STIMJ TX DLVR & MGMT
	90869	REPET TMS TX SUBSEQ MOTR THRESHLD W/DELIV & MNGT
	0310T	REPET TMS TX SUBSEQ MOTR THRESHLD W/DELIV & MNGT (new 1/1/2013)
Electroconvulsive Therapy		
	90870	ELEC-CONVULS THERAP; SNGL SEIZURE
	90871	DEL - ECT; MX SEIZURES PER DA
Neurostimulators (intracranial) 61850-61888 (DBS & VNS)		
Electrodes	61850	TWIST DRILL-IMPLNT ELECTROD; CORTIC
Electrodes	61855	DEL - TWIST DRILL-IMPLNT ELECTR; SUBCORTI
Electrodes	61860	CRANIECT IMPLNT ELECTROD; CORTICAL
Electrodes	61862	DEL - STEREOTACTIC IMPLANT NEUROSTIMULATOR ARRAY, SUBCORITCAL SITE, TWI
Electrodes	61863	STEREOTACT IMPLANT NEUROSTIM ELECTRO ARAY,SUBCORT SITE,W/O INTRAOP MER;1ST ARRAY
Electrodes	61864	STEREOTACT IMPLANT NEUROSTIM ELECTRO ARRAY,SUBCORT SITE,W/O INTRAOP MER;EA ADDL
Electrodes	61865	DEL - CRANIECT IMPLNT ELECTROD; SUBCORTIC
Electrodes	61867	STEREOTACT IMPLANT NEUROSTIM ELECTRO ARAY,SUBCORT SITE, \W/INTRAOP MER;1ST ARRAY
Electrodes	61868	STEREOTACT IMPLANT NEUROSTIM ELECTRO ARAY,SUBCORT SITE,W/INTRAOP MER;EA ADDL
Rev/rem lead	61880	REVIS/REMOV INTRACRAN ELECTRODES
Implant/Repl.	61885	INCS & PLCMT CRANIAL NEUROSTIM GEN
Implant/Repl.	61886	SUBQ PLACEMENT CRANIAL NEUROSTIMULATOR PULSE GENERATOR/RECEIVER;W
Rev/Rem	61888	REVIS/REMOV CRANIAL NEUROSTIM GEN
Percutaneous implantation of neurostimulator electrodes (DBS & VNS) Codes, any cranial nerve; new 2011		
Electrodes	64553	PERQ IMPLNT ELECTRODE; CRANIAL NERV
Electrodes	64568	INC IMPLTJ CRNL NRV NSTIM ELTRDS & PULSE GENER
Electrodes	64569	REVISION/REPLMT NSTIM CRNL ELTRDS
Electrodes	64570	REMOVAL CRNL NRV NSTIM ELTRDS & PULSE GENERATOR
Electrodes	64573	DEL - INCS IMPLNT ELECTRODE; CRANIAL NERV
Electrodes	64581	INCISION, IMPLANTATION NEUROSTIMULATOR ELECTRODES;
Electrodes	64585	REVIS/REMOV PERIPHERAL ELECTRODE
Neurostimulation programming and analysis		
Programming	95970	ANALY IMPLNT NEUROSTIM; WO REPROGRM
Programming	95971	ANALY IMPLNT NEUROSTIM; W/ PROGRM

Programming	95974	ANAL NEUROSTIM; CRAN NERV W/PROG-1
Programming	95975	ANALY NEUROSTIM; CRAN NRV W/PROG-RX
Programming	95978	ELECTRONIC ANLYS, IMPLANT NEUROSTIM SYS, COMPLEX PULSE GENERAT/TRANSMIT PROGR
Programming	95979	ELECTRONIC ANLYS, IMPLANT NEUROSTIM SYS, COMPLEX PULSE GENERAT/TRANSMIT PROGR
Imaging/plan.	70450	CT HEAD/BRAIN W/O CONTRAST
Imaging/plan.	70551	MRI BRAIN/BRAIN STEM W/O CONTRAST
Imaging/plan.	76376 76377	3D RENDERING W/INTERPRETATION OF CT, MRI, US OR OTHER

Review Objectives

Vagus nerve stimulation (VNS) will not be covered in this report. The Washington HTA Program previously reviewed VNS in 2009 (*Vagus Nerve Stimulation for Depression and Epilepsy*). An updated search for new literature conducted in August 2013 revealed that no randomized controlled trials (RCTs) comparing VNS for depression with non-VNS control groups have been published since the 2009 report and a 2011 evidence report produced for the Agency for Healthcare Research and Quality (AHRQ) (Gaynes et al., 2011) characterized the strength of evidence for VNS and treatment-resistant depression (TRD) as low.

The scope of this report is defined as:

Population: Adults with major depressive disorder (MDD) or bipolar depression who have not responded to prior adequate pharmacologic treatments.

Interventions: Nonpharmacologic treatments for depression, including electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), transcranial direct current stimulation (tDCS), and deep brain stimulation (DBS).

Comparators: Sham treatment, treatment as usual, other nonpharmacologic treatment (including psychotherapy as a new treatment in response to treatment failure), pharmacologic treatment (a new medication to be tried in response to treatment failure), or combination therapy that does not include the nonpharmacologic therapy of interest.

Outcomes: Response, remission, depression severity, functional status, quality of life (QOL).

Key Questions

The following key questions will be addressed:

1. a. Are the following nonpharmacologic treatments effective for TRD?
    - Electroconvulsive therapy (ECT)
    - Repetitive transcranial magnetic stimulation (rTMS)
    - Transcranial direct current stimulation (tDCS)
    - Deep brain stimulation (DBS)

   b. Does the effectiveness of these treatments vary according to treatment intensity, duration of treatment, use in an augmentation versus switch strategy, or any other variation in the manner in which TRD treatment was administered?

2. What adverse events, including withdrawal from treatment, are associated with nonpharmacologic treatments for TRD?
3. Does the effectiveness of nonpharmacologic treatments for TRD vary by subpopulation defined by such factors as: age, race/ethnicity, gender, disease severity, disease duration, depression diagnosis (unipolar or bipolar depression), symptom type (e.g., psychotic, postpartum), comorbidities, or number and type of prior treatments (including other nonpharmacologic treatments)?
4. What are the cost implications and cost-effectiveness of nonpharmacologic therapies for TRD?

Methods

Search Strategy and Selection Criteria

Systematic Reviews and Guidelines

Initially, evidence for this report was obtained by searching for systematic reviews and guidelines that had been published in the past *5 years* (as of July 2013). Searches were conducted in the following databases: Agency for Healthcare Research and Quality (AHRQ), Blue Cross Blue Shield TEC Assessments, Canadian Agency for Drugs and Technology in Health (CADTH), Centre for Reviews and Dissemination (York University), Cochrane Library, Hayes Knowledge Center, Institute for Clinical Systems Improvement (ICSI), National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (UK), U.S. Preventive Services Task Force (USPSTF), National Guidelines Clearinghouse (NGC), National Institute for Health and Care Excellence (NICE), VA/Department of Defense Clinical Practice Guidelines, and VA Technology Assessment Program (VA TAP).

The websites for the American Psychiatric Association, American Psychological Association, and American College of Neuropsychopharmacology were searched for guidelines. Additional systematic reviews were selected from a search of the PubMed database using filters for Practice Guidelines, Guidelines, Meta-analyses, and Systematic Reviews.

For additional evidence pertinent to Key Questions #1b (treatment parameters), #2 (safety), and #3 (differential effectiveness), the initial searches for systematic reviews and meta-analyses that were conducted in the Centre for Reviews and Dissemination database and in PubMed were repeated for an earlier time frame (2003 to 2008) to identify reviews that might have included observational studies and addressed safety or differential effectiveness. For safety evidence regarding deep brain stimulation (DBS), the searches were not restricted to DBS and depression, since safety in patients with Parkinson's disease might be applicable to patients being treated for depression.

Primary Clinical Studies

Searches were conducted in the PubMed, Embase, and PsycINFO databases. Initially, searches were conducted to identify primary studies published after the search time frames of the selected systematic reviews. Since 1 of the selected systematic reviews, an evidence review prepared for AHRQ (Gaynes et al., 2011), excluded studies with > 20% patients who had a diagnosis of bipolar disorder, a search without date limits but using publication type limits, was conducted for studies of electroconvulsive therapy (ECT) and repetitive transcranial magnetic stimulation (rTMS) in patients with bipolar

depression; in addition, the Excluded Studies list in the AHRQ evidence review was reviewed for studies excluded for this reason. The Excluded Studies list in the AHRQ report was also searched for comparator studies that had been excluded because of no sham control. See **Appendix III** for search details.

Cost Studies

The National Health Service Economic Evaluation Database (NHS-EED) (2003 to 2013) and PubMed (August 2003 to August 2013) were searched for studies published in the last 10 years. See **Appendix III** for search details.

Update Search

An update search of all sources was conducted on November 12, 2013.

Inclusion Criteria

- Consistency with PICOS (population-interventions-comparator-outcomes-setting) statement.
- Any of the following:
    - For ECT and rTMS: (a) randomized trial with a sham control, (b) randomized comparator trial (RCT) comparing rTMS with ECT or comparing either with another treatment, (c) a post hoc analysis of long-term follow-up or maintenance therapy following an RCT, or (d) observational study providing adverse event data for ≥ 100 patients or (e) an observational study assessing treatment effect (*not* response predictors) according to patient factors not addressed by a systematic review or by an RCT (see Key Question #3).
    - For transcranial direct current stimulation (tDCS) and DBS: any clinical study, including case series.
    - Systematic review.

Exclusion Criteria

- Systematic reviews, cost studies, and economic evaluations published before August 2003.
- No abstract.
- < 10 patients
- For RCTs of rTMS vs sham that were published after the search time frame of the AHRQ report: < 43 randomized patients.
- Studies that did not enroll patients on the basis of treatment-resistant depression (TRD) and did not provide information suggesting that a majority of patients had likely experienced ≥ 1 antidepressant (AD) failure. *However, systematic reviews that did not restrict study selection to TRD patients were considered if no other systematic review evidence or substantial trial data were available for a particular Key Question. Evidence from such sources was downgraded for uncertain applicability to the PICO statement.*
- Evaluation of different anesthesia products for ECT.
- Patients with bipolar disorder who were selected on the basis of a manic episode or mixed mania and depression.

The sample size restriction for RCTs of rTMS versus sham was considered justified since 24 trials representing this comparison were selected without sample size restriction for the AHRQ report and were included in pooled estimates. The cutoff value of 43 was based on a median sample size for the

studies published after the AHRQ report. Three RCTs of rTMS versus sham met the sample size threshold for studies published after the AHRQ report search, and 4 were excluded.

Quality Assessment

Appendix IV outlines the process used by Hayes for assessing the quality of individual primary studies and the quality of bodies of evidence. This process is in alignment with the methods recommended by the GRADE Working Group. Quality checklists for individual studies address study design, integrity of execution, completeness of reporting, and the appropriateness of the data analysis approach. Individual studies are labeled as *good*, *fair*, *poor*, or *very poor*. For individual studies included in systematic reviews, this report relies on the quality assessment by review authors. To aid in interpreting the assessment by review authors, a systematic review quality checklist, the Assessment of Multiple Systematic Reviews (AMSTAR) tool (Shea et al., 2007), was used.

Like the GRADE Working Group, Hayes uses the phrase *quality of evidence* to describe bodies of evidence in the same manner that other groups, such as AHRQ, use the phrase *strength of evidence*. The Hayes Evidence-Grading Guides assure that assessment of the quality of bodies of evidence takes into account the following considerations:

- Methodological quality of individual studies, with an emphasis on the risk of bias within studies.
- Applicability to the population(s), intervention(s), comparator(s), and outcome(s) of interest, i.e., applicability to the PICO statement.
- Consistency of the results across studies.
- Quantity of data (number of studies and sample sizes).
- Publication bias, if relevant information or analysis is available.

NOTE: Two terms related to applicability are *directness* and *generalizability. Directness* refers to how applicable the evidence is to the outcomes of interest (i.e., health outcomes versus surrogate or intermediate outcomes) or to the comparator of interest (indirect comparison of 2 treatments versus head-to-head trials). *Generalizability* usually refers to whether study results are applicable to real-world practice. If the setting is not specified in a PICO (population-interventions-comparator-outcomes) statement, the issue of generalizability to real-world settings is not typically treated as an evidence quality issue. Another term used by some organizations is *imprecision*, which refers to findings based on such a small quantity of data that the confidence interval surrounding a pooled estimate includes both clinically important benefits and clinically important harms or such a small quantity of data that any results other than large statistically significant effects should be considered unreliable.

Bodies of evidence for particular outcomes are labeled as being of *high*, *moderate*, *or low quality*, or they are deemed to be *insufficient* to permit conclusions. These labels can be interpreted in the following manner:

High: Suggests that we can have high confidence that the evidence found is reliable, reflecting the true effect, and is very unlikely to change with the publication of future studies.

Moderate: Suggests that we can have reasonable confidence that the results represent the true direction of effect but that the effect estimate might well change with the publication of new studies.

Low: We have very little confidence in the results obtained, which often occurs when the quality of the studies is poor, the results are mixed, and/or there are few available studies. Future studies are likely to change the estimates and possibly the direction of the results.

Insufficient: Suggests no confidence in any result found, which often occurs when there is a paucity of data or the data are such that we cannot make a statement on the findings.

Guidelines

The Rigor of Development domain of the Appraisal of Guidelines Research and Evaluation (AGREE) tool (AGREE Enterprise, 2012), along with a consideration of the items related to commercial funding and conflicts of interest among the guideline authors, was used to assess the quality of practice guidelines. The checklist and scoring guide have been included in **Appendix IV.**

Search Results

Evidence Selected to Answer Key Questions

This report was based on evidence derived from

- 15 systematic reviews (main source of data for approximately 70 studies not independently assessed)
- 22 randomized controlled or comparator trials (both referred to as RCTs) not included in or considered independently of the systematic reviews
- 1 post hoc analysis of RCTs
- 3 economic evaluations

Table 8 summarizes the quantity and type of evidence by Key Question and technology.

Table 8. Search Results

Key: AHRQ, Agency for Healthcare Research and Quality; DBS, deep brain stimulation; ECT, electroconvulsive therapy; EE, economic evaluation; MA, meta-analysis; RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; SR, systematic review; tDCS, transcranial direct current stimulation

ECT	rTMS	tDCS	DBS
KQ #1a - Effectiveness			
4 RCTs, (3 were reviewed in AHRQ report but independently assessed for the present report). 3 of 4 RCTs were sham-controlled.	1 SR/MA (AHRQ; pooled estimates based on 24 sham-controlled RCTs were used) 3 sham-controlled RCTs – published post-AHRQ review 1 ad hoc analysis of an RCT included in AHRQ review 5 RCTs, rTMS vs ECT (4 included in AHRQ review but independently assessed) 2 RCTs, rTMS+ECT vs ECT (included in AHRQ review but independently assessed)	2 SRs/MAs (7 RCTs and 4 case series) 1 RCT (2 publications)	1 SR (Hayes 2012; no pooled estimates; 5 uncontrolled studies, 9 publications)
KQ #1b – Effectiveness by treatment parameter			
1 SRs/MAs 7 randomized comparator trials	6 1a RCTs 4 randomized comparator trials	Data from the 1a SRs and RCT	---

ECT	rTMS	tDCS	DBS
KQ #2 – Safety			
Data from the 1a and 2b RCTs 2 SRs/MAs	Data from the 1a and 1b SR and RCTs 3 SRs/MAs	Data from the 1a SRs and RCT 1 SR w/ safety-only data	Data from 1a SR 3 SRs, safety-only data
KQ #3 – Differential effectiveness by patient characteristics			
2 SRs/MAs 1 post hoc analysis of 2 randomized comparator trials	Data from 1a SR and RCTs	Data from 1a SRs and RCT	Data from 1a study
KQ #4 – Cost			
---	2 EEs, rTMS vs ECT 1 EE, rTMS vs pharmacology	---	---

Excluded Studies

ECT

- A randomized comparator trial comparing magnetic seizure therapy (MST) with ECT (Kayser et al., 2011) since MST was not a technology of interest for this report.
- A study of the clinical and cost-effectiveness of ECT for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modeling studies conducted by the national health technology assessment (HTA) program in the UK (Greenhalgh et al., 2005). An economic model of ECT compared with pharmacotherapy for depression was constructed, but the authors stated that it "makes no assumptions about previous depressive episodes and previous treatment received" (p.50). In other words, neither the ECT studies nor the pharmacotherapy studies were selected or analyzed according to whether patients had TRD. A critique of the ECT-pharmacotherapy analysis cited several deficiencies, including (McDonald, 2006):
    - Use of data from disparate sources. A relative risk (RR) of response based on comparator trials of ECT and tricyclic antidepressants (TCAs), with publication dates primarily in the 1960s and 1970s, was applied to a TCA response rate derived from a single trial published in 1995.
    - Inclusion of medication trials in which patients were generally not treatment-resistant and were required to have only mild to moderate depression with ECT trials in which patients had severe depression, were psychotic, and/or were medication-resistant.
- A systematic review of ECT used as maintenance therapy in elderly patients (van Schaik et al., 2012). The review did not restrict selection to studies of patients with TRD. All 3 of the RCTs in this review were excluded from the AHRQ report; 2 RCTs only enrolled patients who were in remission.

rTMS

- Two systematic reviews with meta-analysis of RCTs using high-frequency rTMS (Berlim et al., 2013b) and low-frequency rTMS (Berlim et al., 2013a) were not used as evidence for effectiveness. Approximately half of the included studies in each review did not enroll patients on the basis of TRD, as evidenced by their rejection from the AHRQ review or by review for the

present report. The objective of the analysis of high-frequency rTMS (Berlim et al., 2013b) was to assess whether this technology can "hasten the therapeutic effects of standard antidepressants" (p. e123) rather than as a treatment for medication-refractory depression. However, both reviews provided data on acceptability, which are cited in the discussion of findings for Key Question #2.

- A systematic review and meta-analysis of effect modifiers in rTMS for depression (Herrmann and Ebmeier, 2006) was excluded for the same reasons the 2 reviews by Berlim and colleagues were thought to lack applicability to most of the Key Questions.
- A study that pooled data from 6 trials for purposes of analyzing predictors of response in the patients who received rTMS (Fregni et al., 2006c). Two trials had been excluded by the AHRQ review because they did not target patients with TRD, and 3 other trials were unpublished. An RCT of rTMS versus ECT (Grunhaus et al., 2000) that was included in the AHRQ report was excluded from detailed analysis in this report because, as noted by authors of the AHRQ review, there were serious risks to internal bias: (1) some patients in the rTMS group received 400 pulses and some received 1200 pulses and (2) ECT was administered as an augmentation strategy while rTMS was administered as a switch strategy. The AHRQ review characterized this as a "poor"-quality study.
- Two RCTs (Ray et al., 2011; Huang et al., 2012) that provided no information on past AD failures and may have enrolled patients without any past AD failures. The objective of the Huang study was to assess the ability of rTMS to accelerate the effect of an AD and achieve earlier improvement during a first episode of MDD.

rDCS

- A case series (n=84) that evaluated whether particular medications were associated with outcomes in a group of patients being treated for depression with tDCS (Brunoni et al., 2013a). The study was excluded because treatment lasted only 5 days, whereas treatment duration was 10 sessions in other studies.

Practice Guidelines

Six practice guidelines were identified. These were produced by the American Psychiatric Association (APA, 2002; APA, 2010; APA, 2011), the Canadian Network for Mood and Anxiety Treatments (CANMAT) (Kennedy et al., 2009), the National Institute for Health and Care Excellence (formerly the National Institute for Clinical Excellence) (NICE, 2009), the Institute for Clinical Systems Improvement (ICSI, 2012), and the Veterans Administration/Department of Defense (VA/DoD, 2009).

Literature Review

Key Question #1a:
Are the following nonpharmacologic treatments effective for treatment-resistant depression (TRD)?

Study Characteristics

Definitions of Treatment-Resistant Depression (TRD)

Only studies that explicitly stated or suggested that most patients had probably experienced ≥ 1 antidepressant (AD) medication failure were selected for evidence pertaining to Key Question #1a. The majority of studies either only enrolled patients who had had ≥ 2 AD failures, reported that patients had experienced ≥ 2 AD failures, or reported that the mean number of prior AD failures was larger than 2. However, most studies did not indicate whether any a priori criteria were used to determine the adequacy of prior AD trials or whether adequacy was considered. When such criteria were described, studies often required ≥ 6 weeks for previous trials and sometimes added that maximum dose and/or 2 different classes had to have been tried. Most studies did not state explicitly whether failed AD trials had occurred in the current episode.

Some studies characterized the degree of medication resistance in terms of the Antidepressant Treatment History Form (ATHF), the Maudsley Staging Method (MSM), or Thase and Rush criteria scale. See **Appendix I** for descriptions of these systems. Most studies did not mention a systematic approach to assessing treatment resistance at baseline.

Treatment Strategies

Studies used several strategies for introducing a new therapy in patients who had obtained inadequate relief from AD medication. Using the Agency for Healthcare Research and Quality (AHRQ) report as a guide but tailoring definitions to the studies actually included, the present report categorizes treatment strategies as follows:

Switch: All prior AD treatment was discontinued. Some of these studies allowed continuation of antipsychotics or other types of psychotropic medications.

Augmentation: Patients continued their prior AD pharmacotherapy; the new treatment was an add-on to current treatment. In some augmentation studies, a small proportion of patients were not taking any psychotropic medications before or during the trial.

Mixed: Patients were not required but were encouraged to discontinue their AD medications, so some were treated with a switch strategy and some, with an augmentation strategy.

Combination: Patients in both of 2 groups initiated a new psychotropic medication at the same time; 1 group also began a nonpharmacologic intervention.

Most studies used an augmentation strategy. Switch strategies were not uncommon. Only 2 studies compared a switch to nonpharmacologic treatment (ECT or tDCS) with a switch to a new AD; the other switch strategies involved a switch to active or sham nonpharmacologic treatment. A small number of studies used a mixed strategy. Only 1 study tested a combination strategy of nonpharmacologic treatment (tDCS) plus a new AD medication with the new AD medication alone.

Treatment Phase

Nearly all studies that met inclusion criteria were assessments of acute therapy. A small number of studies also provided data on follow-up assessments at 2 weeks to 6 months after discontinuation of treatment with the technology of interest (although maintenance therapy with AD medication might have continued). One study evaluated maintenance treatment with electroconvulsive therapy (ECT) and another study of acute treatment with transcranial direct current stimulation (tDCS) evaluated a continued maintenance regimen involving tDCS.

Outcomes Measurement

Where possible, the symptom data selected for presentation in the Literature Review were results according to the Hamilton Depression Rating Scale (HAM-D) scale, or the Montgomery-Åsberg Depression Rating Scale (MADRS) if the HAM-D scale was not used. All data pertaining to function and quality of life (QOL) are presented.

Electroconvulsive Therapy (ECT), Key Question #1a

The evidence review performed for the AHRQ (Gaynes et al., 2011) identified 2 double-blind, sham-controlled randomized controlled trials (RCTs) and 1 RCT of ECT versus new pharmacotherapy, which were reviewed independently for the present report. Key findings are presented in **Table 9**.

Efficacy**:** No ADs were allowed (switch strategy) during 1 sham-controlled study (Johnstone et al., 1980), whereas amitriptyline was administered (combination strategy) during the other sham-controlled study (West et al., 1981). Although neither study selected patients on the basis of AD failure, these studies were included since the authors of the AHRQ review considered the clinical characteristics of the study populations to suggest a high probability of ≥ 2 prior AD failures. In the Johnstone study, 15% of patients had previously undergone ECT, but the success of the previous ECT was not reported. In the Johnstone study, 70% of patients had received an AD, which was presumably ineffective, during the current episode. No other information on prior psychotherapy or other nonpharmacologic treatment was provided for the West study. The 2 studies showed that bilateral ECT accelerated improvement in depression scores, but in the Johnstone study, the sham group and ECT group had improved by a comparable degree at 6 months after the end of treatment, compared with study entry; no follow-up information was provided in the West study.

Since neither sham-controlled study reported response or remission rates, a fair-quality systematic review (Heijnen et al., 2010) was consulted for data on remission rates in patients treated with ECT. Studies were selected if they reported remission as ≤ 7 on the HAM-D_{17}, as ≤ 10 on the HAM-D_{24}, or ≤ 8 on the MADRS. Studies also had to assess medication resistance by using the Antidepressant Treatment History Form (ATHF) to evaluate the adequacy of previous AD trials (see **Appendix I** for more information on the ATHF). Across 7 prospective uncontrolled studies involving 545 patients with medication resistance, remission rates ranged from 39% to 63%. An overall remission rate of 48%, based on a simple pooling of data across studies, was reported. Heijnen et al. did not describe duration of treatment or other treatment parameters.

ECT, KQ #1a
Acute: 3 RCTs (Johnstone 1980, West 1981, Folkerts 1997) with supplemental data from an SR of uncontrolled studies (Heijnen 2010)
Maintenance: 1 unblinded RCT (Nordenskjöld 2013)

The comparison of ECT with a new AD medication suggested that unilateral ECT as part of a switch strategy was considerably more effective than a new AD medication in patients with clear TRD (Folkerts et al., 1997). Patients in the study by Folkerts et al. had experienced multiple AD failures.

Neither sham-controlled study mentioned the inclusion of any patients with bipolar depression, but 5 patients in the comparison with pharmacotherapy had bipolar disorder. A recent systematic review did not identify any controlled or comparative prospective studies addressing the efficacy of ECT in bipolar depression (Versiani et al., 2011).

Quality of Life (QOL)/Function: No studies evaluated QOL or functional outcomes.

Table 9. Evidence Overview, ECT vs Sham or Pharmacotherapy, Acute (see following discussion of 1 maintenance treatment trial)

Key: AD, antidepressant (medication); BD, bipolar disorder; BDI, Beck Depression Inventory; BL, baseline; ECT, electroconvulsive therapy; f/u, follow-up; grp, group; HAM-D, Hamilton Depression Rating Scale; ITT, intention-to-treat; MDD, major depressive disorder; NR, not reported; NS, not (statistically) significant; posttx, posttreatment; pt, patient; RCT, randomized controlled trial; RUL, right unilateral; SD, standard deviation; TCA, tricyclic antidepressant; tx, treatment

Authors/Study Details	Response/Remission Definitions	Main Posttreatment Results	Direction of Findings/Quality/Comments
Vs Sham			
Johnstone et al. (1980) Double-blind sham-controlled RCT 70 pts who required inpt tx 4 wks (8 sessions) bilateral Switch strategy (no ADs during trial) Benzodiazepines allowed	***Response:*** NR ***Remission:*** NR*	***Mean change HAM-D_{17} at end of tx (n=62):*** ~18 vs 25 (graph), P<0.01 favoring ECT ***Posttx f/u:*** No difference at 1 mo or 6 mos (in both grps, scores at 6 mos were similar to posttx scores in ECT grp)	4 wks of bilateral ECT in a switch strategy was effective. Fair 19% loss to 1- and 6-mo f/u.
West et al. (1981) Double-blind sham-controlled RCT 25 pts (inpts) 3 wks (6 sessions) bilateral Combination strategy	***Response:*** NR ***Remission:*** NR*	***Mean BDI (BL, end of tx) (n=22):*** ECT: 26.6±2.8, 10.8±2.6 (*P*<0.001) Sham: 24.1±3.5, 22.2±3.8 (NS) Change, ECT vs sham; *P*<0.002.	3 wks of bilateral ECT in an augmentation strategy was effective. Fair
Vs Pharmacotherapy			
Folkerts et al. (1997) RCT 39 pts 2 wks RUL ECT (6 sessions), 4 wks pharmacotx Switch strategy	***Response:*** ≥50% reduction in HAM-D score **Remission:** NR	**Mean HAM-D_{21} (BL, posttx) (mean±SD):** ECT: 31±1, 12.5±3.9 Pharmacotx: 32.6±5.4, 23.0±10.4 *P*=0.001, ECT vs pharmacotx at 2 wks. **% reduction in HAM-D_{21} posttx (ECT, pharmacotx):** 60%, 30% (P=0.001) **Response posttx (ECT, pharmacotx) (% pts):** 71.4%, 27.8% (P=0.006)	2 wks of unilateral ECT in a switch strategy was more effective than a new AD medication. Fair No pt or assessor blinding
Other clinical details: Mean age 40-53 yrs, 52%-74% women. Diagnosis: Moderate-severe MDD; no cases of BD mentioned in Johnstone 1980 and West 1981; 5 BD pts in Folkerts 1997. Other psychiatric diagnoses: NR. Prior AD failures: ≥1 in current episode in 70% of pts (Johnstone 1980). Unclear (West 1981). ≥2 different ADs, including ≥1 TCA, in trials lasting 8 wks, mean 4.3-4.9 (presumably lifetime) (Folkerts 1997). Other prior txs: ECT in previous episodes 21% (Johnston 1980); otherwise NR.			

*Remission rates of 39% to 63% in patients with medication resistance have been reported for prospective uncontrolled trials of ECT (Heijnen et al., 2010).

Maintenance Treatment with ECT: A multicenter RCT recruited 56 patients with unipolar MDD who had remitted in response to acute treatment with ECT in patients with medication-resistant depression

(Nordenskjöld et al., 2013). Remission was initially defined as score ≤ 10 on the MADRS, but because of low recruitment, the eligibility criteria were broadened to require ≤ 15 on the MADRS. In an unblinded protocol, patients were randomized to ECT plus pharmacotherapy or pharmacotherapy alone. Treatment resistance was defined as failure of ≥ 2 trials of AD medications in different classes. According to this definition, 57% of patients in the ECT arm and 46% of patients in the pharmacotherapy alone arm were treatment resistant prior to acute treatment with ECT. Patients were followed by telephone interview and were seen by the investigator for evaluation of relapse if MADRS score rose to above 20. There was a high rate of dropout and crossover. According to intention-to-treat (ITT) analysis, rates of relapse at 1 year favored ECT plus pharmacotherapy over pharmacotherapy alone: 32% versus 61% (*P*=0.036); Cox proportional hazard ratio 2.32 (95% confidence interval [CI], 1.03 to 5.22). Relapse rates at 2 and 6 months also favored the ECT group. In modified ITT analysis, which included all randomized patients who participated in ≥ 1 evaluation, 1-year relapse still favored the ECT group (39% versus 64%), but the difference was nonsignificant. This study was considered to be of fair quality because of the lack of blinding.

Repetitive Transcranial Magnetic Stimulation (rTMS), Key Question #1a

rTMS, KQ #1a
rTMS vs sham
1 systematic review: Gaynes 2011 (AHRQ) (good); 24 RCTs represented (5 good, 19 fair according to Gaynes et al.)
3 RCTs: Ullrich 2011 (good), Blumberger 2012 (fair), Fitzgerald 2012 (fair)
1 ad hoc analysis: Mantovani 2012 (relates to an RCT covered in AHRQ review)
rTMS vs ECT
5 RCTs: Grunhaus 2003, Rosa 2006, McLoughlin 2007, Hansen 2011, Keshtkar 2011 (all fair except Keshtkar [poor])
rTMS+ECT vs ECT
2 RCTs: Pridmore 2000, Chistyakov 2005 (both fair)

The evidence review performed for AHRQ (Gaynes et al., 2011) provided the bulk of evidence regarding rTMS versus sham stimulation. See **Appendix V** for an overview of the AHRQ report. The AHRQ review included pooled estimates for 24 RCTs that made a comparison with sham treatment. Three sham-controlled RCTs published after the AHRQ search time frame (Blumberger et al., 2012; Fitzgerald et al., 2012; Ullrich et al., 2012) were also selected. An additional 5 RCTs comparing rTMS with ECT (Grunhaus et al., 2003; Rosa et al., 2006; McLoughlin et al., 2007; Hansen et al., 2011; Keshtkar et al., 2011) and 2 RCTs comparing rTMS + ECT with ECT (Pridmore, 2000; Chistyakov et al., 2005) were selected. Most of the 7 trials making comparisons with ECT were included in the AHRQ review but were not subjected to meta-analysis; they were reviewed in detail for this report and an independent quality assessment was made. Study findings are presented in **Table 10**.

rTMS Versus Sham Stimulation, Key Question #1a

All data pertaining to rTMS versus sham stimulation are summarized in **Table 10** following this discussion. The AHRQ review reported a weighted mean difference (WMD) of 5.92, favoring rTMS for change in HAM-D score, based on data from 24 RCTs (1200 patients) comparing rTMS with sham stimulation. The results showed high statistical heterogeneity and thus, the authors did not include the forest plot for this overall analysis. However, the forest plots for stratified analyses showed that study-specific results consistently favored rTMS in terms of depression score changes. Both response (defined as ≥ 50% improvement from baseline) and remission (prespecified cutoff points for different scales) were *more likely* in the rTMS arms according to study-specific results. The AHRQ review did not provide a pooled estimate of response or remission for the entire set of studies. Numerous stratified analyses according to the number of required previous AD failures and whether ≤ 20% of patients had a diagnosis

of bipolar depression consistently resulted in pooled relative risks (RRs) of response and remission that favored rTMS. These stratified estimates will be presented in more detail in the discussion of Key Question #3. For the 16 studies that required failure of ≥ 1 or ≥ 2 prior AD trials, a strong association between rTMS and response (RR, 2.68; 95% CI, 1.52 to 4.70) and remission (RR, 3.73; 95% CI, 1.2 to 11.30) was demonstrated, and no heterogeneity was reported. The review authors reported a number-needed-to-treat (NNT) value of 5 for response and an NNT value of 6 for remission but did not provide risk differences.

The 3 RCTs published after the AHRQ review reported somewhat inconsistent findings, both with respect to whether standard unilateral rTMS is effective and whether bilateral sequential rTMS is superior to unilateral rTMS. However, the direction of findings in all of the studies favored rTMS over sham even though differences were small and not consistently significant. The issue of bilateral versus unilateral stimulation will be addressed in the discussion of Key Question #1b.

Table 10. Evidence Overview, rTMS Versus Sham, Symptom Outcomes in Acute Treatment (see following text for durability of benefits and QOL/functional outcomes)

Key: AD, antidepressant (medication); AHRQ, Agency for Healthcare Research and Quality; ATHF, Antidepressant Treatment History Form; BD, bipolar depression; BL, baseline; BPRS, Brief Psychiatric Rating Scale; CI, confidence interval; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; grp, groups; HAM-D, Hamilton Depression Rating Scale; hertz, Hz; MA, meta-analysis; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; N/A, not applicable; NR, not reported; pt(s), patient(s); RCT, randomized controlled trial; RR, relative risk; rTMS, repetitive transcranial magnetic stimulation; SD, standard deviation; SIGH-D, Structured Interview Guide for the Hamilton Depression Rating Scale; SR, systematic review; tx, treatment

Authors/Study Details	Response/Remission Definition	Main Post-treatment Results	Direction of Findings/Quality/ Comments
Gaynes et al. (2011) (AHRQ evidence review) SR: 24 RCTs, 1200 pts Augmentation (15 RCTs), switch (6 RCTs), mixed (3 RCTs) Left high frequency (18 studies), mix (5 studies), low (1 study) 1-6 wks, typically 10 sessions over 2 wks	***Response* definition:** ≥50% reduction in HAM-D or MADRS score ***Remission definition:*** ≤8 HAM-D_{17}, ≤10 HAM-D_{21}, or ≤8 on MADRS; or similar definitions	***WMD in change scores, depressive severity:*** –5.92 (CI, –8.15 to –3.70) (I^2=80%); WMD in individual studies consistently favored rTMS. ***RR of response in trials requiring ≥1 or ≥2 AD failures:*** 2.68 (CI, 1.52-4.70; NNT=5) (16 RCTs). All individual trial RRs favored rTMS. ***RR of remission in trials requiring ≥1 or ≥2 AD failures:*** 3.73 (CI, 1.23-11.30; NNT=6) (9 RCTs). All but 1 individual trial RR favored rTMS. Pooled RRs for response and remission not possible for entire set of 24 studies.	rTMS administered in a variety of clinical situations was effective. MA, good; studies, fair-good. Risk differences not reported. No evidence of publication bias.
Ullrich et al. (2011) Double-blind RCT 43 pts Augmentation Left high frequency (30 Hz) 3 wks (15 sessions)	***Response:*** >50% reduction in HAM-D_{21} **Remission:** HAM-D_{17} <8	No difference at end of tx in HAM-D_{21} score or in score reduction, after adjustment for concomitant use of lithium; 18.2% response rate in active rTMS vs 0 in sham (NS); no remission in either grp.	3 wks of unilateral high frequency rTMS for acute tx of pts w/ MDD (polarity unknown) and sham stimulation had comparable outcomes. Good
Blumberger et al. (2012) Double-blind RCT 63 pts Mix of augmentation and switch; no restrictions changes in AD use 2 active arms: sequential	***Response:*** >50% reduction in HAM-D_{17} ***Remission:*** HAM-D_{17} ≤10	***HAM-D_{17} score (BL, 3 wks, 6 wks):*** Bilateral: 25.1, 15.3, 14.4. Unilateral: 26.0, 19.6, 20.3. Sham: 25.2, 17.8, 18.9. Rate of change was not significantly different in pairwise comparisons between grps. ***Remission (bilateral, unilateral, sham) (% pts):*** 34.6%, 4.5%, 5%	3 wks of bilateral rTMS for acute tx of pts w/ unipolar MDD was associated w/ a tx effect, but unilateral rTMS was not. Fair

Authors/Study Details	Response/Remission Definition	Main Post-treatment Results	Direction of Findings/Quality/ Comments
bilateral and left high frequency arms 3 wks (15 sessions)		(global *P*=0.005) Bilateral vs sham, *P*=0.028, risk difference 29.6%, NNT 4. Bilateral vs unilateral, *P*=0.002, risk difference 30.1%, NNT 4. Unilateral vs sham, NS. ***Response (bilateral, unilateral, sham) (% pts):*** 38.5%, 4.5%, 10% (global *P*=0.006). Bilateral vs sham, *P*=0.003, risk difference 28.55, NNT 4. Bilateral vs unilateral, *P*=0.022, risk difference 34%, NNT 3. Unilateral vs sham, NS.	
Fitzgerald et al. (2012) Double-blind RCT 66 pts Augmentation; no change in AD use allowed Sequential bilateral and left high frequency arms 3 wks (15 sessions)	***Response***: 50% reduction in HAM-D ***Remission:*** NR	***HAM-D (baseline, 3 wks) (mean ± SD):*** Bilateral: 24.3±3.6, 22.2±6.0. Unilateral: 23.7±3.8, 19.6±4.2. Sham: 22.8±2.1, 22.6±5.0. Global *P*=0.05; bilateral vs sham, NS; bilateral vs unilateral, NS; unilateral vs sham, *P*=0.02. ***Response (bilateral, unilateral, sham) (% pts):*** 4.5%, NR, NR.	Compared w/ sham stimulation, both unilateral and bilateral rTMS resulted in better outcomes, but a statistically significant effect was detected only for unilateral rTMS. Fair
Other clinical details: Mean age 40-58 yrs, 54%-64% women (Blumberger 2012, Fitzgerald 2012); otherwise NR. Diagnosis: Moderate-severe MDD (typically, according to DSM-IV); ≤20% of pts had BD in 10 RCTs covered in AHRQ review. Prior AD failures: ≥2 (17 RCTs), ≥1 (7 RCTs), not specified (3 RCTs). Failures required to be in current episode: 8 RCTs. Adequacy of AD failure defined as ATHF score, either unspecified or ≥3 (2 RCTs in AHRQ review); 4 wks and augmentation w/ mood stabilizer in ≥1 trial (Rosa 2006); ≥6 wks (2 RCTs in AHRQ review; Blumberger 2012; Fitzgerald 2012); ≥8 wks (1 RCT in AHRQ review); ≥2 classes (Blumberger 2012). Other prior txs: Successful ECT 11.5%-15.0% (Blumberger 2012); otherwise NR. Psychiatric comorbidity: 0 in Ulrich 2011, 5%-12% active tx pts in Fitzgerald had anxiety; 0 w/ borderline or antisocial personality disorder in Blumberger 2012; NR in AHRQ review.			

Durability of Benefits, Depression Symptoms: A small number of sham-controlled studies (6 studies) included in the AHRQ review followed all or some patients for 2 to 24 weeks after the end of treatment. According to the AHRQ review, posttreatment differences persisted to 3 weeks after treatment but disappeared after 11 weeks of treatment in 1 study and persisted for 2 weeks to 3 months in another 2 studies. In other studies described by the AHRQ review, only responders were followed. In these studies, relapse rates were substantial (≥ 50%) for follow-up at 20 weeks to 6 months and did not differ statistically between rTMS and sham groups (2 studies), or relapse rates were relatively small and favored rTMS but the difference was not statistically tested (1 study). The AHRQ review did not offer any analysis of factors associated with durability of benefits; the variability of findings did not correspond to differences in the number of prior AD failures required for enrollment. In a follow-up study (Mantovani et al., 2012) published after the AHRQ review but related to 1 of the sham-controlled RCTs (George et al., 2010) included in the AHRQ review, 32 of 50 remitters started an AD medication for continuation treatment, underwent rTMS taper, and completed 3-month follow-up; of the 32 participants, 29 (90.6%) remained in remission. Another 5 remitters in the Mantovani study declined to participate in the continuation therapy protocol and all were in remission at 3 months.

QOL and Functional Assessments: According to data in the AHRQ review for 1 study of 30 patients, Global Assessment of Functioning (GAF) scores improved from baseline by 2.2 and 1.4 points by end of 2-week treatment in active treatment arms and improved by only 0.2 point in the sham arm (P=0.09 and P=0.03). In another study of 155 patients described in the AHRQ review, only negligible change in scores on the SF-36 Health Survey (QualityMetric Inc.) was observed at the end of 6 weeks of treatment. Improvement on the Quality of Life, Enjoyment and Satisfaction Questionnaire-Short Form in the second study favored rTMS but was very small: 2-point increase from 37.8 versus 1.3-point increase from 36.5 (P=0.035).

rTMS Versus ECT, Key Question #1a

Effectiveness: The results were inconsistent across 5 RCTs (total, n=261 patients) comparing rTMS with ECT alone. See **Table 11** following this discussion. Differences in treatment strategy may explain the inconsistent findings. In 2 RCTs comparing unilateral high-frequency rTMS with ECT in a switch strategy, posttreatment results were comparable between the 2 groups (Grunhaus et al., 2003) or slightly favored rTMS but the difference was nonsignificant (Rosa et al., 2006). In the other 3 RCTs, which represented an augmentation strategy, the results favored ECT (McLoughlin et al., 2007; Hansen et al., 2011; Keshtkar et al., 2011). There were also other differences in treatment parameters. In 2 (McLoughlin et al., 2007; Keshtkar et al., 2011) of the 3 trials favoring ECT, most or all patients in the ECT group received bilateral stimulation, i.e., a greater dose of ECT, whereas in studies suggesting comparable efficacy, ECT groups underwent unilateral stimulation. Additionally, in 2 trials favoring ECT (McLoughlin et al., 2007; Hansen et al., 2011), no minimum number of prior AD failures was required for enrollment, whereas the other trials required ≥ 1 or ≥ 2 failures. There is some evidence (low quality) that ECT is less effective in confirmed TRD than in MDD without a well-documented history of AD failure while the effectiveness of rTMS may not vary according to medication resistance (see findings for **Key Question #3**). The superiority of ECT in the McLoughlin and Hansen trials thus might be attributable to the inclusion of some patients with less well-established treatment resistance. The body of evidence is not large enough to allow strong conclusions about the reasons for inconsistency.

Durability of Benefit: In the 2 trials that reported posttreatment follow-up (McLoughlin et al., 2007; Hansen et al., 2011), differences between rTMS and ECT patients disappeared by the time of follow-up assessment (7 weeks and 6 months), due to continued improvement in the rTMS groups and some loss

of improvement or only negligible further improvement in the ECT groups. In the McLoughlin study, differences in secondary measures significantly favored ECT at 6-month follow-up.

QOL/Function: In the only 2 trials to assess functional status or QOL, no difference in end-of-treatment GAF score was observed (Grunhaus et al., 2003) or no difference in QOL derived from the SF-36 Health Survey at 6-month follow-up was observed (McLoughlin et al., 2007).

Table 11. Evidence Overview, rTMS Versus ECT in Acute Treatment

Key: AD, antidepressant (medication); BD, bipolar depression; BDI, Beck Depression Inventory; BL, baseline; BPRS, Brief Psychiatric Rating Scale; CI, confidence interval; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders; ECT, electroconvulsive therapy; GAF, Global Assessment of Functioning; HAM-D, Hamilton Depression Rating Scale; ITT, intention-to-treat; MDD, major depressive disorder; NR, not reported; NS, not statistically significant; pt(s), patient(s); posttx, posttreatment; RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; SD, standard deviation; SF-36, SF-36 Health Survey; tx, treatment; VAS, visual analog scale

Authors/Study Details	Response/Remission Definition	Main Results (Post-treatment Unless Otherwise Noted)	Direction of Findings/Quality/ Comments
Grunhaus et al. (2003) Single-blind (assessor) RCT, rTMS vs ECT, to replicate previous findings (Grunhaus et al., 2000) of no effect in nonpsychotic pts 40 pts Switch; unilateral, high-frequency rTMS, 4 wks (mean 10.25 sessions); unilateral ECT	***Response*:** ≥50% reduction in or final score ≤10 on HAM-D_{17} *and* final GAF ≥60 ***Remission:*** Score ≤8 on HAM-D_{17}	***HAM-D_{17} score (BL, end of tx) (mean ± SD):*** rTMS: 24.4±3.9, 13.3±9.2 ECT: 25.5±5.9, 13.2±6.6 ***Response at end of tx (rTMS, ECT) (% pts):*** 55%, 60% ***Remission at end of tx (rTMS, ECT) (% pts):*** 30%,30% ***GAF score (BL, end of tx) (mean ± SD):*** rTMS: 48.9±10.8, 58.3±17.1 ECT: 39.3±9.3, 60.6±13.5 Grp effect NS.	In a switch strategy, unilateral high-frequency rTMS for 4 wks was **comparable** w/ unilateral ECT. Fair
Rosa et al. (2006) Single-blind (assessor) RCT 42 pts Switch; unilateral, high-frequency rTMS, 4 wks (20 sessions); unilateral ECT	***Response***: 50% reduction in HAM-D_{17} ***Remission:*** HAM-D_{17} ≤8	***HAM-D_{17} (BL, 2 wks, 4 wks):*** rTMS: 30.1, 18 ECT: 32.1, 19 (4-wk HAM-D data estimated from graph) ***Response (rTMS, ECT) (% pts):*** Per protocol: 50%, 40% (NS) ITT: 45%, 30% (NS) ***Remission (rTMS, ECT) (% pts):*** Per protocol: 10%, 20% (NS) ITT: 9%, 15% (NS)	High-frequency unilateral rTMS may have been superior to unilateral ECT delivered in a switch strategy to a population w/ severe TRD. Poor to Fair Substantially greater dropout rate in ECT grp (but ITT analysis); underpowered to

Authors/Study Details	Response/Remission Definition	Main Results (Post-treatment Unless Otherwise Noted)	Direction of Findings/Quality/ Comments
			detect 10% or 15% difference*
McLoughlin et al. (2007); Eranti et al. (2007) UK study, 7 hospitals Single-blind (evaluators only) pragmatic RCT, rTMS vs ECT (equivalence hypothesis) 46 pts Generally, augmentation; unilateral, high-frequency rTMS, 3 wks (15 sessions); mix of bilateral and unilateral ECT	***Response***: 50% reduction in HAM-D_{17} ***Remission:*** HAM-D_{17} ≤7	***Between-grp difference in HAM-D:*** Posttx: Difference favored ECT (CI, 3.40-14.05; *P*=0.002) in pairwise testing w/ adjustment for multiple measurements). Point estimate NR. Effect size=1.44. 6 mos: No difference. (BDI mood [VAS], and BPRS were lower in the ECT grp.) ***Change in HAM-D score relative to BL (rTMS, ECT):*** 22%, 58% (absolute difference 36%; CI, 14%-58%) ***Remission at end of tx (rTMS, ECT) (% pts):*** 17%, 59% (*P*=0.005) ***Relapse at 6 mos (rTMS, ECT) (% posttx remitters):*** 50%, 42% (NS) ***QOL gains at 6 mos:*** No difference (societal utilities assigned to SF-36 scores).	Response and remission were more likely w/ unilateral/bilateral ECT than w/ high-frequency unilateral rTMS but the difference was not maintained at 6 mos. Fair (blinding seriously compromised; substantial loss to f/u at 6-mo f/u but ITT analysis). Some difficulty w/ recruitment because of unwillingness to be randomized to ECT.
Hansen et al. (2011) RCT 60 pts Primarily augmentation; unilateral, low-frequency rTMS, 3 wks (15 sessions); unilateral ECT	***Response***: 50% reduction in HAM-D_{17} ***Remission:*** HAM-D_{17} ≤8 ***Partial remission:*** HAM-D_{17} ≤12	***Response (rTMS, ECT) (% pts):*** Wk 3 (posttx): 20%, 57%, (difference 37%; CI, 14%-59%) (*P*=0.003) Wk 7 (4-wk f/u): 43%, 60% (difference 17%; CI, –8% to 42%) (NS) ***Partial remission (rTMS, ECT) (% pts):*** Wk 3: 27%, 53% (difference 26%; CI, 3%-51%) (*P*=0.035) Wk 7 (4-wk f/u): 40%, 57% (difference 17%; CI, –8% to 42%) (NS) ***HAM-D score (BL, 3 wks, 7 wks) (values estimated from graph):*** rTMS: 24, 16, 11 ECT: 24, 11, 10 *P*=0.001 for BL to 3 wks and BL to 7 wks in each grp.	Response and partial remission were more likely w/ unilateral ECT than with low frequency unilateral rTMS. Fair (lack of blinding and lack of power to detect differences in partial remission) ECT intensity was varied according to seizure response
Keshtkar et al. (2011) RCT (pseudo-randomized) 73 pts Augmentation;	***Response***: NR ***Remission:*** NR	***HAM-D_{21} (BL, end of tx) (mean ± SD:*** rTMS: 21.0±7.5, 15.1±5.6 ECT: 25.8±6.1, 8.4±6.1 *P*<0.001 for difference in change.	Improvement in symptoms was more likely w/ bilateral ECT than w/ unilateral rTMS of unknown

Authors/Study Details	Response/Remission Definition	Main Results (Post-treatment Unless Otherwise Noted)	Direction of Findings/Quality/ Comments
unilateral rTMS of unknown frequency, 2 wks (10 sessions); bilateral ECT			frequency. Poor (comparison of bilateral ECT w/ unilateral rTMS may not be valid; no true randomization).

Other clinical details: Mean age 48-62 yrs (34-36 yrs in Keshtkar 2011), 44%-80% women. Diagnosis: Moderate-severe MDD (according to DSM-IV); no pts w/ BD (4 RCTs), 13% BD (1 RCT). 16%-55% psychosis (2 RCTs); 0% psychosis (1 RCT); otherwise NR. Prior AD failures: Typically, >2 (≥2, 2 RCTs; mean 2.4-2.5, 1 RCT, median 2, 1 RCT, ≥1, 1 RCT). Only Rosa 2006 required ≥2 failures. Failures required to be in current episode (1 RCT). Adequacy of prior AD trials: ≥4 wks (3 RCTs); "full course" or "sufficient dose" (2 RCTs); otherwise, NR. Other prior txs: Any ECT, 55%-75% (1 RCT); otherwise NR. Psychiatric comorbidity: No other Axis I (2 RCTs); 65%-75% Axis II (1 RCT); otherwise NR.

*According to this online calculator: http://www.stat.ubc.ca/~rollin/stats/ssize/.

rTMS + ECT Versus ECT, Acute Treatment, Key Question #1a

As shown in **Table 12**, 2 RCTs (total, n=44) that combined rTMS and ECT according to different protocols both showed that posttreatment results were comparable between the combination treatments and ECT alone (Pridmore, 2000; Chistyakov et al., 2005). The Pridmore study also found no difference in overall psychological functioning, according to the GAF, at end of treatment. Neither study conducted follow-up assessments.

Table 12. Evidence Overview, rTMS + ECT Versus ECT in Acute Treatment

Key: AD, antidepressant (medication); BD, bipolar depression; BL, baseline; CI, confidence interval; ECT, electroconvulsive therapy; GAF, Global Assessment of Functioning; grp(s), group(s); HAM-D, Hamilton Depression Rating Scale; MDD, major depressive disorder; NR, not reported; pt(s), patient(s); RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; sx, symptom(s); tx, treatment

Authors/study design and pts/tx details	Response/Remission Definition	Main Results (Posttreatment Unless Otherwise Noted)	Direction of Findings/Quality/ Comments
Pridmore (2000) Single-blind (assessor) RCT comparing ECT alone w/ a regimen in which 4 out of 6 ECT sessions were replaced w/ double rTMS sessions 22 pts Primarily augmentation High-frequency left rTMS; unilateral ECT 2 ECT and 8 rTMS sessions	***Response:*** MADRS ≤12 *and* HAM-D_{17} ≤8. **Remission:** NR	***Response at end of tx (ECT-rTMS, ECT, risk difference) (% pts):*** 55%, 55% ***HAM-D_{17} score (BL, end of tx :*** ECT-rTMS: 28, 8 ECT: 30, 7 CI for ECT-rTMS minus ECT at 2 wks: –7 to 8 (NS) ***GAF score (BL, end of tx):*** ECT-rTMS: 41, 65 ECT: 41, 70 CI for ECT-rTMS minus ECT at 2 wks: –8 to 17 (NS)	Unilateral ECT partially replaced by unilateral high-frequency rTMS was associated w/ sx and functional outcomes comparable to those observed w/ ECT alone. Fair
Chistyakov et al. (2005) Single-blind(assessor) RCT in which ECT combined w/ rTMS was compared w/ ECT combined w/ sham stimulation 22 pts Primarily augmentation Low-frequency right rTMS; bilateral ECT 6 txs/wk over 3 wks; ECT on day 1 and 5, active or sham rTMS on days 2, 3, 4, and 6; total 6 ECT sessions and 12 active or sham rTMS sessions	***Response:*** ≥50% reduction in HAM-D_{17} score **Remission:** NR	***Response at end of tx:*** 86% overall study grp had a response; grp rates NR. ***HAM-D score (BL, 3rd ECT, end of tx):*** ECT+active rTMS: 43, 26, 17 ECT+sham rTMS: 43, 23, 15 Values approximated from graphs. NS difference between grps in improvement.	Bilateral ECT partially replaced by unilateral low-frequency rTMS was associated w/ sx outcomes **comparable** w/ those observed w/ ECT partially replaced by sham rTMS. Fair
Other clinical details: Mean age 57 yrs (where reported), 45%-68% women. Diagnosis: Moderate-severe MDD (according to DSM-IV); no pts w/ BD. Prior AD failures: ≥2 (Pridmore 2000); unclear (Chistyakov 2005). Adequacy of prior AD trials: "Maximum dose" (Pridmore 2000); otherwise, NR. Occurrence in current episode NR. Other prior txs: Any ECT, 45%-64%% (Pridmore 2000); otherwise NR. Psychiatric comorbidity: NR			

Transcranial Direct Current Stimulation (tDCS), Key Question #1a

Study Characteristics

tDCS, KQ #1a
2 systematic reviews: Kalu 2012 (fair), Berlim 2013c (good); 7 RCTs represented
1 RCT (2 publications): Brunoni 2013b (acute treatment, good), Valiengo 2013 (maintenance treatment, poor)

Two fair- to good-quality systematic reviews with meta-analyses (Kalu et al., 2012; Berlim et al., 2013c) and a more recently published RCT (Brunoni et al., 2013b) were selected. Seven RCTs, 1 of which was a crossover trial (Palm et al., 2012) and 2 of which were pilot studies reported as letters to the editor (Fregni et al., 2006a; Fregni et al., 2006b), were included in 1 or both systematic reviews. Thus, 8 RCTs are represented by the 2 systematic reviews and the recently published study. The Kalu review also considered case series eligible and included 4 case series. The 2 systematic reviews provided little clinical detail about the trial participants and limited information on treatment parameters. Individual RCTs, but not the case series included in the Kalu review, were retrieved and reviewed in order to create a clinical context for the findings. There was little patient information and limited information on clinical protocol in the 2 pilot studies.

Sample sizes in the 8 RCTs ranged from 10 to 120 patients, with a total of 320 patients represented by the 8 RCTs. Of those 320 patients, ≥ 43 (≥ 13 %) had bipolar depression. The proportion of women was 46% to 80%, and mean age ranged from 47 to 58 years. The duration of follow-up ranged from 0 to 30 days. Where reported, mean baseline scores on symptom severity scales were in the moderate (2 studies) to severe (4 studies) range. Patients typically had multiple AD failures (where reported, mean 1.5 to 4.3 lifetime), although AD failure was part of the inclusion criteria in only 2 RCTs (Blumberger et al., 2012; Palm et al., 2012). The Blumberger and Palm studies required failure of ≥ 2 AD trials from 2 different classes but included no specification that failure had to have occurred in the current episode. The most recently published RCT did not actually require patients to meet any definition of TRD, but 44% of patients had failed at least 2 AD trials (Brunoni et al., 2013b). *Adequacy* of previous AD trials was not defined. In 2 studies, 33% to 40% of patients had tried ECT at some time in the past; in 1 of these studies, 15% of patients had failed ECT in the current episode. In another 2 studies (Loo et al., 2010; Loo et al., 2012), patients were enrolled only if they had *responded* to ECT in *the current episode*; the rationale for this inclusion criterion was unclear. The odds ratio (OR) for response favored tDCS in 1 study requiring previous ECT and favored sham stimulation in the other. In most trials patients were allowed to continue psychotropic medications other than ADs. All RCTs compared active tDCS with sham stimulation. Treatment strategies included augmentation (4 studies), switch (1 study), switch or combination (1 study), and unclear (2 studies). Anticonvulsants and/or antipsychotics were sometimes disallowed during the study period. Electrical current of 1 milliampere (mA) or 2 mA was used.

Findings

Findings are summarized in **Table 13** following this discussion.

Acute Treatment with tDCS: The 2 meta-analyses of tDCS for MDD reached conflicting conclusions. The reviews did not include exactly the same studies and the authors used different outcome measures. The Kalu review calculated effect sizes based on percentage change in depression scores and concluded that tDCS was effective. The mean weighted percent change was 28.9% relative to baseline. This translated to an effect size of 0.75 (95% CI, 0.21 to 1.27); the authors described this as a medium to large effect. The effect size fell to 0.42 but was still significant (95% CI, 0.09 to 0.75) with no statistical heterogeneity

when the 2 pilot studies were omitted from the analysis. Study-specific effect sizes covered a wide range of values but were consistent in direction. Another group of investigators (Brunoni et al., 2013b) suggested that the lack of effect observed in 2 of the trials (Loo et al., 2010; Palm et al., 2012) was attributable to not only lack of statistical power but also their broader inclusion criteria (which allowed enrollment of Axis II disorders) and the lower dose of stimulation associated with the studies' particular treatment protocols.

The Berlim review calculated pooled ORs for response and remission and concluded that tDCS is *not* effective. The pooled OR for posttreatment response favored tDCS but was nonsignificant: OR, 1.97 (95% CI, 0.85 to 4.45). The pooled OR for remission also favored tDCS but was nonsignificant: OR, 2.13 (95% CI, 0.64 to 7.06). No heterogeneity or publication bias was detected for either pooled OR. Unweighted study-specific ORs for both response and remission were inconsistent in direction.

An RCT published after the latest systematic review randomized 120 patients to 4 groups: sham + placebo, sham + sertraline, tDCS + placebo, and tDCS + sertraline (Brunoni et al., 2013b). According to the definitions offered at the beginning of the **LITERATURE REVIEW**, these 4 groups allowed testing of both a switch strategy (tDCS + placebo versus sham + placebo) and a combination strategy (tDCS + sertraline vs sham + sertraline. As noted earlier, patients were not required to meet a definition of TRD, and the authors considered the study group as having "low treatment resistance." Although 44% of patients had experienced ≥ 2 AD failures, some patients (number unknown) had not used any AD during the current episode. Another distinctive feature of this study was the administration of 2 additional sessions at 2-week intervals following a typical 2 weeks of 5 sessions per week; in other words, patients had 12 sessions, the last of which was administered at 6 weeks from baseline. Mean MADRS scores improved in all groups, with the percentage change in mean score smallest in the sham + placebo group and greatest in the tDCS + sertraline group. Similarly, the largest between-group difference in mean MADRS score at 6 weeks was between tDCS + sertraline and sham + placebo. Response and remission rates varied across groups in a similar fashion. Response and remission rates at 6 weeks favored each of the 2 tDCS groups over both sham stimulation groups, whereas although response and remission rates in the sham + sertraline group were greater than those in the sham + placebo group, the differences were nonsignificant. Group differences in mean MADRS at the end of treatment exceeded the authors' definition of clinical relevance (3 points) for all comparisons except sham + sertraline versus sham + placebo.

Table 13. Evidence Overview, tDCS, Acute Treatment (see following text for discussion of durability of benefits, QOL, and maintenance treatment)

Key: AD, antidepressant (medication);BD, bipolar disorder; BL, baseline; CI, confidence interval; HAM-D, Hamilton Depression Rating Scale; ITT, intention-to-treat; MA, meta-analysis; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; NR, not reported; OR, odds ratio; pt(s), patient(s); RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; SR, systematic review; tDCS, transcranial direct current stimulation; tx, treatment

Authors/Study Details	Response/Remission Definition	Main Post-treatment Results	Direction of Findings/Quality/Comments
Kalu et al. (2012) SR/MA (6 RCTs, 194 pts; 4 case series, 80 pts) Acute tx Non-AD psychotropic medications typically allowed	***Response* definition:** ≥50% reduction in severity score ***Remission definition:*** Score ≤7 on HAM-D or ≤10 on MADRS	***Reduction in depression severity in rTMS arms (all studies) (weighted mean % relative to BL):*** 28.9%; range, 14.6%-60% ***Response in rTMS arms (% pts) (all studies):*** 19.8%; range by study, 0%-80% ***Remission in rTMS arms (% pts) (all studies):*** 8.5%; range by study, 0%-23.8% ***Pooled rTMS-vs-sham effect size based on % change from BL (RCTs only):*** 0.74 (CI, 0.21 to 1.27; *P*=0.006). Favored tDCS. Heterogeneity but no publication bias. Letters (Fregni 2006a, Fregni 2006b) excluded: 0.42 (CI, 0.09-0.75; *P*=0.013). No heterogeneity. Range by study: 0.014-2.111.	tDCS for acute tx of pts w/ unipolar or bipolar MDD was associated with a reduction symptom severity. Fair-quality SR/MA. No blinding in 3 RCTs w/ significant effect sizes; no other study quality information.
Berlim et al. (2013c) SR/MA (6 RCTs, 200 pts) Acute tx Non-AD psychotropic medications typically allowed	***Response:*** >50% improvement on HAM-D or MADRS at study endpoint ***Remission****:* Score ≤7 on HAM-D_{17}, ≤8 on HAM-D_{21}, or ≤6 on MADRS	***Response rate (tDCS rate, sham rate, pooled OR):*** 23.2%, 12.4%, OR 1.97 (95% CI, 0.85-4.56; *P*=0.11). No heterogeneity, low risk of publication bias. Unweighted OR range by study, 0.833-33.000. Favored tDCS in 3 RCTs ***Remission rate (tDCS rate, sham rate, pooled OR):*** 12.2%, 5.4%, OR 2.13 (9.5% CI, 0.64 to 7.06; *P*=0.22). No heterogeneity, low risk of publication bias. Unweighted OR range by study, 0.840-13.000. Favored tDCS in 2 studies. Fregni 2006b/2006c excluded from analysis; total	Overall results suggested but did not confirm a tx effect in terms of response and remission rates. Good-quality SR/MA. All RCTs were double-blind. NS heterogeneity and no suggestion of publication bias.

Authors/Study Details	Response/Remission Definition	Main Post-treatment Results	Direction of Findings/Quality/Comments
		n=190	
Brunoni et al. (2013b); Valiengo et al. (2013) Double-blind sham-controlled RCT (120 pts in acute tx; 42 pts in maintenance phase) Acute tx (10 sessions over 2 wks, then 2 sessions every other wk;12 sessions total) and open-label crossover, then maintenance phase (intermittent rTMS) Switch and combination strategies Non-AD psychotropic medications *not* allowed; exception, benzodiazepines	***Response:*** >50% decline in MADRS ***Remission:*** MADRS <10; ITT w/ correction for multiplicity	***Acute tx, MADRS at end of tx (difference in means):*** tDCS+sertraline vs sham+placebo: –11.5 (CI, –6.03 to –17.10; *P*<0.001) tDCS+sertraline vs sham+sertraline: –8.5 points (CI, –2.96 to –14.03; *P*=0.002) tDCS+sertraline vs tDCS+placebo: –5.9 (CI, –0.36 to –11.43; *P*=0.03) Sham+sertraline vs sham+placebo: NS tDCS+placebo vs sham+placebo: –5.6 (CI, –1.30 to –10.01; *P*=0.01) tDCS+placebo vs sham+sertraline: NS ***Acute tx, MADRS response (OR):*** tDCS vs placebo: 8.6 (CI, 2.5-29.1; *P*<0.001) Sertraline vs placebo: NS tDCS+sertraline vs placebo: OR 3.8 (CI, 1.1-12.7; *P*=0.03) ***Acute tx, MADRS remission (OR):*** tDCS vs placebo: 4.3 (CI, 1.2 to 15.6; *P*=0.02) Sertraline vs placebo: NS tDCS+sertraline vs placebo: 5.7 (CI, 1.6-20.3; *P*=0.007)	When used in a 6-wk switch strategy, tDCS was effective in reducing symptoms of MDD, and tDCS combined w/ sertraline was more effective than either tx alone. A difference between tDCS alone and sertraline alone was neither proven nor ruled out. In a grp of pts w/ unipolar MDD who responded to acute tx w/ tDCS, some of whom were also taking sertraline, 6 mos of maintenance tx w/ intermittent tDCS maintained remission in many pts. Good quality (acute); poor quality (maintenance). NS heterogeneity and no suggestion of publication bias
Other clinical details: Mean age 47-58 yrs. 46%-80% women. Diagnosis: MDD (13% pts had BD). Severity: moderate-severe. Mean # failed ADs: 1.5-4.3 lifetime; only 2 RCTs required previous AD failure (≥2 ADs, different classes; current episode not specified); adequacy not defined. Psychiatric comorbidity: dysthymia (26%), generalized anxiety disorder (50%), social phobia (12%), and panic disorder (14%) in Brunoni 2013b; NR by review authors.			

Durability of Depression Benefit: According to the Kalu review, 1 RCT and 2 case series reporting follow-up measurements at 1 month after treatment showed persistence of benefits, and in a fourth study (RCT), the response rate increased from end of treatment to 1 month later. However, this evidence does not permit a conclusion that benefits persist. Lack of control groups in the case series precludes a conclusion that benefits were caused by the tDCS, and the posttreatment effect of tDCS was statistically significant in only 1 of the 2 RCTs.

QOL/Function: No studies evaluated QOL or functional outcomes.

Maintenance Treatment with tDCS: A separate study (Valiengo et al., 2013) evaluated maintenance treatment with intermittent rTMS for 42 responders to tDCS originally enrolled in the RCT by Brunoni et al. (2013b). The study group represented 30 willing patients who were originally randomized to tDCS, who completed the trial, and who responded. The other 12 patients were originally randomized to sham treatment, completed the trial, did *not* respond to sham, agreed to cross over to active treatment at the of the trial, and responded to active treatment. Mean response duration was 11.7 weeks. The cumulative rate of survival to relapse was 60% at 12 weeks and 47% at 24 weeks. This was considered a poor quality study because of the lack of a control group, high loss to follow-up compared with the original trial participants, and other factors. In an RCT described in the Kalu review, patients who accepted 15 additional treatments after the end of the study experienced additional benefits.

Deep Brain Stimulation (DBS), Key Question #1a

Study Characteristics

DBS, KQ #1a
1 SR (Hayes 2012) covering 5 uncontrolled studies, 9 publications (Lozano 2008, Malone 2009, Bewernick 2010, Malone 2010, Grubert 2011, Kennedy 2011, Bewernick 2012, Holtzheimer 2012, Lozano 2012)

Five studies (9 publications) that evaluated the safety and efficacy of DBS for TRD were included in the Hayes report (Hayes, 2012). See **Appendix V** for an overview of the Hayes report. All 5 studies were prospective uncontrolled studies that compared depression scores after treatment with baseline depression scores. One trial included a sham lead-in phase, which was used as a comparison to active stimulation (Holtzheimer et al., 2012). All studies were in adults. Sample sizes ranged from 10 to 21 patients, with a total of 86 patients participating in the 5 studies. Patients were selected on the basis of treatment failure, typically with stringent criteria for assuring that prior treatments had been adequate, e.g., ≥ 3 different classes of AD medication for ≥ 6 weeks, ≥ 6 sessions of ECT, and ≥ 20 sessions of psychotherapy. Some studies also explicitly specified that these treatment failures had to have occurred in the current depressive episode. All studies employed an augmentation strategy, allowing patients to continue ADs and other psychotropic medications during the study period. **Table 14** presents study results.

Findings

The results suggested that DBS improves symptoms of depression in treatment-resistant patients. Where reported, response rates ranged from 41% to 60% at 6 months (4 studies) and 29% to 55% at 12 months (3 studies), and remission rates ranged from 20% to 36% at 12 months (2 studies). In 1 study (Holtzheimer et al., 2012), change in HAM-D scores following active stimulation were compared with change scores after a sham stimulation lead-in phase. Compared with the 4-week trial of sham stimulation, HAM-D scores improved to a slightly greater extent after 4 weeks of active stimulation;

however, the difference did not reach statistical significance (*P*=0.06). In this trial, a DBS discontinuation phase was initiated after 24 weeks of active stimulation. The first 3 patients to enter this phase of the trial suffered complete relapse within 2 weeks and did not improve immediately when stimulation was reinitiated. The AD effect of stimulation did slowly return. DBS was not discontinued in any other patients because of the first 3 patients' reactions. During both the lead-in phase and the discontinuation phase, patients were told they were being randomized to either active or sham stimulation, although all patients actually received sham stimulation.

Durability of Benefit: All studies reported follow-up data at 6 and 12 months. Some 2-year data were available, but dropout rates were very high at this point. Response and remission rates sometimes declined after 6 months and sometimes increased; thus, no conclusions can be drawn about the durability of benefit. No apparent pattern according to study and patient characteristics could be detected.

Quality of Life/Functional Status: Two studies (34 patients) measured overall psychological function with the GAF scale. They reported significant improvements of mean 18.4 points at last follow-up (mean 24 months) (Malone et al., 2009) and 28.3 points at 1-year follow-up (Holtzheimer et al., 2012). The GAF scale extends from 1 (maximum dysfunction) to 100 (no dysfunction).

Table 14. Evidence Overview, DBS, Acute Treatment

Key: AD, antidepressant (medication); ATHF, Antidepressant Treatment History Form; BD, bipolar disorder; BL, baseline; dx, diagnosis; f/u, follow-up; GAF, Global Assessment of Functioning; HAM-D, Hamilton Rating Scale for Depression; hx, history; MADRS, Montgomery-Åsberg Depression Rating Scale; MD, major depression; meds, medications; N/A, not applicable; NR, not reported; psychotx, psychotherapy; SC, subcallosal cingulate; stim, stimulation; STR/NAc, striatum/nucleus accumbens; tx, treatment/therapy

Authors/Study Details	Main Post-treatment Results	Direction of Findings/Quality/Comments
Lozano et al. (2008); Kennedy et al. (2011) (In: Hayes 2012) 20 pts Augmentation strategy Leads implanted in SC gyrus F/u, mean 42 mos	***Mean HAM-D (BL, 1 mo, 3 mos, 6 mos, 1 yr)*** 24.4, 15.4, 12.5, 11.8, 12.6 (all analyses; *P*<0.001) ***Response rates (1 wk, 6 mos, 1 yr) (% pts):*** 40%, 60%, 55% ***Remission rates (6 mos, 1 yr):*** 35%, 35%	6-mo improvement maintained at 1 yr. Poor. By 2 yrs >20% of pts had dropped out of study.
Malone et al. (2009); Malone (2010) (In: Hayes 2012) Augmentation strategy Leads implanted in STR/NAc, stim initiated 2-4 wks after surgery 17 pts F/u, mean 23.5 mos	***Mean HAM-D (BL, last f/u):*** 33.1, 14.4 (P=0.0007) ***Response rates (3 mos, 6 mos, last f/u) (% pts):*** HAM-D: 47%, 40%, 53% MADRS: 53%, 47%, 53% ***Remission rates (3 mos, 6 mos, last f/u) (%):*** HAM-D: 20%, 20%, 40% MADRS: 33%, 27%, 33% ***Mean GAF (BL, 3 mos):*** 43.4-58.4 (significance NR); 18.4-point improvement (*P*=0.0009) at last f/u	6-mo improvement w/ further gains at last f/u. Poor. 4 pts had medication changes during 1st 6 mos of study.
Bewernick et al. (2010); Grubert et al. (2011); Bewernick et al. (2012) (In:	***Mean HAM-D (BL, 1 mo, 1 yr, 2 yrs):*** 32.2, 23, 20, 20 (comparisons w/ BL, *P*<0.05 at each time point)	6-mo improvement maintained at 2 yrs.

Authors/Study Details	Main Post-treatment Results	Direction of Findings/Quality/Comments
Hayes 2012) 11 pts Augmentation strategy Leads implanted in STR/NAc F/u, ≤48 mos	Half of pts had 50% reduction in HAM-D at 1 yr. ***Mean MADRS (BL, 1 yr, 2 yrs):*** 32, 20, 19 (comparisons w/ BL, *P*<0.005 at each time point)	Poor.
Holtzheimer et al. (2012) (In: Hayes 2012) 17 pts F/u, 24 mos Augmentation strategy Leads implanted in SC, 4-wk sham stim phase, 24-wk active stim phase	***Mean HAM-D (BL, after 4-wk sham stim , after 4-wk active stim, 24 wks, 1 yr, 2 yrs):*** 23.9, 20.5 (*P*=0.02 vs BL), 17.9 , 13.1, 13.6, 7.1 (for change over time; global *P*<0.001) ***Change in HAM-D (BL to end of sham stim, end of sham to end of active tx):*** –3.3, –2.7 (P=0.06) ***Response rates (6 mos, 1 yr) (% pts):*** 41%, 36% ***Remission rates (6 mos, 1 yr) (%):*** 18%, 36% ***Mean GAF (BL, after 4-wk sham stim , after 4-wk active stim, 24 wks, 1 yr, 2 yrs):*** 33.9, 36.9, 43.9, 60.8, 62.2, 78.7 (change over time; global P<0.001)	Improvement, but sham lead-in phase casts some doubt on efficacy; unclear direction of change from 4-24 wks. Poor. By 2 yrs >20% of pts had dropped out of study.
Lozano et al. (2012) (In: Hayes 2012) Augmentation strategy Leads implanted in SC gyrus F/u, 12 mos	***% HAM-D score decrease from BL (2 mos, 6 mos, 12 mos):*** 40%, 43%, 41% ***Response rates (1 mo, 6 mos, 12 mos) (% pts):*** 57%, 48%, 29%	Improvement at 2 mos but declining rates after that. Very poor. No statistical analyses were reported.
Other clinical details: Mean age 42-47 yrs. 36%-65% women. Diagnosis: 0%-6% pts had BD (2 studies); 41% had BD in 1 study (Holtzheimer 2012). BL depression: severe. # failed ADs: mean 3.9-7.91; 1 study specified current episode. Other previous tx: Psychotx (100%, 3 studies; mean 316 hrs, Bewernick 2010); ECT (85%-94% in 2 studies; mean 21-31 previous sessions in 1 study). Psychiatric comorbidity: NR		

Key Question #1b:
Does the effectiveness of these treatments vary according to treatment intensity, duration of treatment, use in an augmentation versus switch strategy, or any other variation in the manner in which TRD treatment was administered?

Electroconvulsive Therapy (ECT), Key Question #1b

No comparison of treatment parameters was made in the selected sham-controlled RCTs. Two systematic reviews with meta-analysis (UK ECT Review Group, 2003; Dunne and McLoughlin, 2012) and 3 randomized comparator trials (Sackeim et al., 2008; Roepke et al., 2011; Mayur et al., 2013) were selected as evidence for Key Question #1b. For evidence pertaining to this question, an exception was made to the exclusion criterion of lack of information regarding the proportion of patients who had failed ≥ 1 previous AD trial. The rationale for this exception was the lack of studies with data pertinent to Key Question #1b specific to patients with TRD. The evidence was appropriately downgraded for possible lack of applicability to the population of interest (see **Table 5** in the **EVIDENCE SUMMARY**).

Where AD history was reported, patients had usually experienced ≥ 2 AD failures. See **Table 15** following this discussion for a presentation of study results.

Bifrontal Versus Bilateral (Bitemporal) or Unilateral Electrode Placement

A good-quality systematic review evaluated bifrontal versus bilateral (bitemporal) and unilateral electrode placement (Dunne and McLoughlin, 2012). Treatment resistance was *not* one of the inclusion criteria. However, the 8 included trials were retrieved and most were found to report that patients had failed ≥ 1 and typically multiple AD trials. The review was based on a search of several databases from their inception to March 2010. English-language studies comparing bifrontal ECT versus unilateral or bitemporal ECT in unipolar or bipolar adults with MDD (according to established criteria) were eligible for inclusion. Studies of patients with an Axis I comorbidity were excluded. *Bifrontal* placement was defined precisely as placement on a line at 2- to 3-centimeters (cm) above the supraorbital ridge, or 5-to 6-cm above the lateral canthus of the eye. Studies had to be double-blind with randomized treatment allocation. The authors did not comment on other aspects of study quality other than to report that the randomization method was described in only 1 study, and the method of allocation concealment in only 3 studies (4 studies were published in 2005 or earlier) and thus may not have been subject to journal requirements that allocation concealment be disclosed.

In terms of change in HAM-D score at the soonest assessment time following the last ECT session, the effect sizes calculated by Dunne and McLoughlin (2012) suggested a small advantage to bifrontal stimulation compared with bitemporal stimulation. However, when bifrontal stimulation was compared with unilateral stimulation, a small difference favored unilateral stimulation. Both estimates for change in HAM-D were nonsignificant. Although tests for heterogeneity were nonsignificant, the effect sizes for individual studies were inconsistent in direction. Changes in the Mini-Mental State Examination (MMSE) followed the same pattern, but the effect size for bifrontal versus bitemporal was large and statistically significant (0.889; P=0.037), which the authors said translated to a difference of 1 standard deviation on the MMSE. Analyses by the intensity of stimulation in terms of percentage of seizure threshold suggested that the advantage of unilateral over bifrontal stimulation for reduction of depression symptoms was present only at higher intensities but that the advantage of unilateral over bifrontal stimulation for effect on the MMSE was consistent at 250%, 500%, and 600% of seizure threshold. There was high statistical heterogeneity in the results for MMSE changes.

Dose, Bilateral Versus Unilateral Electrode Placement, and Pulse Width

A fair-quality systematic review and meta-analysis conducted in the UK evaluated a number of factors that might contribute to the effectiveness of ECT (UK ECT Review Group, 2003). Studies were *not* selected on the basis of whether they enrolled patients with TRD, and the review offered no analysis according to this factor. Studies were also not required to be randomized, and the authors did not provide an appraisal of study quality. Bilateral stimulation was found to be more effective than unilateral stimulation. A high dose of ECT was found to be more effective than low dose ECT. Individual trial results were inconsistent with respect to bilateral versus unilateral ECT, but consistently favored high dose over low dose. The authors explored the interaction between electrode placement and dose and found a suggestion that high dose leads to better outcomes in bilateral ECT but not in unilateral ECT; however, the interaction term was statistically nonsignificant.

The effect of dose and electrode placement may be different in techniques that use ultrabrief pulses as opposed to the more traditional brief pulses. In contrast to the UK review, a more recent trial found no association between dose and outcomes (Roepke et al., 2011). This inconsistency might be explained by

the use of ultrabrief pulses in the Roepke study, whereas all the studies included in the UK analysis of dose used brief pulse ECT. Similarly, another recent study contradicted the UK study by showing unilateral stimulation to be superior to bilateral stimulation, but only when ultrabrief pulses were used, not when brief pulses were used (Sackeim et al., 2008). The effect of ultrabrief versus brief pulse itself is not clear; 1 study found no difference according to pulse width in a study looking only at unilateral ECT (Mayur et al., 2013).

Frequency of Sessions

The UK review found that 3 times per week was more effective than once a week although statistical significance was demonstrated only in the fixed effects model. The review's analysis of studies comparing twice a week with 3 times per week yielded small nonsignificant differences favoring *twice* a week.

Table 15. Comparisons of Treatment Parameters, ECT

Key: AD, antidepressant (medication); BDI, Beck Depression Inventory; BiL, bilateral; BL, baseline; ECT, electroconvulsive therapy; grp(s), group(s); HAM-D, Hamilton Depression Rating Scale, Hz, hertz; ITT, intention-to-treat; MA, meta-analysis; mC, millicoulombs; MMSE, Mini-Mental State Examination; msec, millisecond; NR, not reported; RUL, right unilateral; SES, standardized effect size; sig, (statistically) significant; SR, systematic review; ST, seizure threshold; TCA, tricyclic antidepressant; tx, treatment

Author/Study Details	Main Results	Implications
UK ECT Review Group (2003) SR/MA Any controlled trial evaluating ECT for depression No other study selection criteria	**Results are presented as SES's.** ***Bilateral vs unilateral (22 trials, 1408 pts):*** Fixed effects: –0.323 (CI, –0.446 to –0.1.99) (favors bilateral) Random effects: –0.322 (CI, –0.458 to –0.186) ***Higher vs lower dose (6 trials, 337 pts):*** Fixed effects: 0.571 (CI, 0.352-0.790) (favors higher dose) Random effects: 0.575 (CI, 0.329-0.829) ***Once/wk vs thrice/wk (2 trials, 51 pts):*** Fixed effects: 0.841 (CI, 0.311-1.370) (favors thrice/wk) Random effects; 0.832 (–0.389 to 1.890) ***Twice/wk vs thrice/wk (SES) (4 trials, 159 pts):*** Fixed effects: –0.308 (CI, –0.629 to 0.014) (favors twice/wk) Random effects; –0.299 (–0.759 to 1.199)	Bilateral stimulation may be more effective than unilateral stimulation and higher dose may be more effective than lower dose. Optimal tx may require >1 session/wk, but it is not clear whether 3 as opposed to 2 sessions/wk are necessary.
Sackeim et al. (2008) Double-blind randomized comparator trial 90 pts Mean 5-6 AD trials during episode, 2-3 adequate AD trials; mean resistance score 3.1-3.6 (0-5 scale) Dose: 6.0 × ST (RUL), 2.5 × ST (BiL) 3×/wk; discontinuation at clinician discretion; mean 6.2-8.9 txs Response = ≥60% reduction in HAM-D. Remission = response and HAM-D≤10	***RUL ultrabrief (0.3 msec) vs BiL ultrabrief (1.5 msec) vs RUL brief vs BiL brief:*** Posttx response (% pts): 77%, 48%, 73%, 70%. 1-wk f/u response (% pts): 73%, 35%, 59%, 65% Posttx remission (% pts): 77%, 43%, 73%, 70% 1-wk f/u remission (% pts): 73%, 35%, 59%, 65% Response and remission were poorest in BiL ultrabrief grp and differences between BiL ultrabrief and each of the other 3 grps were sig at 1-wk f/u; no other sig between-grp differences Analysis of change in HAM-D: Effects of electrode placement, pulse width, and their interaction were sig	Ultrabrief RUL may be superior to ultrabrief BiL ECT, but RUL and BiL ECT do not differ when brief pulses are used, and pulse width does not influence the effectiveness of RUL.

Author/Study Details	Main Results	Implications
	Relapse among responders: Unrelated to tx parameters Cognitive effects: Least severe in RUL ultrabrief grp; no differences between other grps	
Roepke et al. (2011) Randomized comparator trial 40 pts ≥2 failed ADs 9 sessions, RUL ultrabrief	***Frequency:*** Remission (≤10 on HAM-D): 100 Hz (25%) vs 40 Hz (35%) Reduction in HAM-D: 100 Hz (5.8) vs 40 Hz (9.5) ((*P*=0.047 after adjustment for BL HAM-D; NS after adjustment for seizure threshold) (HAM-D version NR) Cognitive Tests: No sig differences ***Stimulation dose:*** No association w/ HAM-D	Neither stimulation frequency nor dose affects effectiveness of ultrabrief RUL ECT.
Dunne and McLoughlin (2012) SR and MA 8 double-blind RCTs (617 pts) Brief pulse in 7 RCTs, ultrabrief in 1 RCT 3×/wk (64% pts), 2×/wk (25% pts), NR (11% pts)	**Positive effect size for MMSE denotes less decline w/ bifrontal than w/ comparator.** ***Bifrontal vs bitemporal:*** Change in HAM-D: Effect size 0.102, favoring bifrontal but NS (5 RCTs) Change in MMSE: Effect size 0.889 favoring **bifrontal (*P*=0.037)** (4 RCTs) Other cognitive tests: Variable results, only 2-3 RCTs per test ***Bifrontal vs unilateral:*** Change in HAM-D: Effect size –0.118 favoring RUL but NS (7 RCTs) (separate analysis showed that at % seizure threshold of 100% or 250% [dose], results favored bifrontal placement; at 500% and 600%, results favored RUL) Change in MMSE: Effect size 0.101 favoring bifrontal but NS (# RCTs unclear). (results consistently favored RUL at 250%, 500%, and 600% seizure threshold)	Bifrontal stimulation has not been shown to have an advantage over bitemporal or unilateral stimulation.
Mayur et al. (2013) Double-blind randomized comparator trial 35 pts No information on past ADs RUL at 6 × ST, 3×/wk	***Ultrabrief (0.3 msec) vs brief (1 msec) pulse width:*** Reduction in MADRS score: No difference Median days to remission: 28 (CI, 17.9-38.0) vs 26 (CI, 18.6-33.4)	Pulse width does not influence the effectiveness of RUL ECT.

Repetitive Transcranial Magnetic Stimulation (rTMS), Key Question #1b

Four sham-controlled studies included in the AHRQ review reported results by subgroups defined by treatment parameters (Su et al., 2005; Garcia-Toro et al., 2006; Pallanti et al., 2010; Triggs et al., 2010). Two of the sham-controlled RCTs selected from studies published after the AHRQ review also compared different ways of applying rTMS (Blumberger et al., 2012; Fitzgerald et al., 2012). In addition, 5 randomized head-to-head comparator trials without sham controls were identified (Isenberg et al., 2005; Fitzgerald et al, 2008; Triggs et al., 2010; Galletly et al., 2012; Fitzgerald et al., 2013). The studies were generally of at least fair quality. See **Table 16** following this discussion for a presentation of study findings.

Seven (7) studies evaluated different strategies of low- and high-frequency rTMS applied to the right or left dorsolateral prefrontal cortex (DLPFC). The standard technique is unilateral high frequency rTMS applied to the left DLPFC. Four (4) of these studies (total, n=373), all published in 2010 or later and 3 of which had sham controls, compared bilateral sequential stimulation (low frequency [1 Hz] to the right, followed by high frequency [10 Hz] to the left) with standard unilateral high frequency rTMS to the left. The study results were conflicting, with 3 studies showing no difference or a small potential difference favoring standard unilateral stimulation and 1 study suggesting that bilateral but not unilateral stimulation is effective. Another 3 studies tested other combinations of frequency and electrode placement. Low frequency applied to the right and high frequency also applied to the right was found to be comparable in 1 study (Isenberg et al., 2005). A single study (Stern et al., 2007) compared low frequency with high frequency, both applied to the left DLPFC, and found low frequency rTMS to be ineffective, which is consistent with the rationale between left and right stimulation (see **BACKGROUND**). Another study (Triggs et al., 2010) compared left and right application of low frequency and observed comparable outcomes.

In a study that randomized 31 patients to high frequency TMS at an *intensity* of 100% motor threshold (MT), TMS at 90% MT, or sham stimulation, a difference across groups in reduction of depression symptoms was demonstrated when measured by the MADRS but not by the HAM-D scale (Padberg et al., 2002). However, intensity measured as continuous variable had no association with effect in 1 study (Su et al., 2005).

Three (3) small studies each evaluated a different aspect of treatment parameters, but conclusions are not permissible since the results were not corroborated by other randomized trials (Garcia-Toro et al., 2006; Fitzgerald et al., 2008; Galletly et al., 2012).

Table 16. Comparisons of Treatment Parameters, rTMS

Key: Hz, hertz; NS, not statistically significant; pt(s), patient(s); RCT, randomized controlled/comparator trial

Study Details	Low Frequency vs High Frequency, Both Right	Low Frequency, Left vs Right	Low Frequency vs High Frequency, Both Left	Sequential Bilateral vs Unilateral Stimulation	Other
Padberg et al. (2002) Randomized comparator trial 31 pts w/ sham control					Relative reduction in MADRS varied by intensity: 100% MT (–33.2%) vs 90% MT (–15.1%) vs sham (–4.1%) (global *P*<0.05). NS for HAM-D.
Isenberg et al. (2005) Randomized comparator trial 28 pts	1 Hz vs 20 Hz. No difference at last tx; 1-mo f/u scores favored low (*P*=0.059)				
Su et al. (2005) Double-blind sham-controlled RCT 30 pts					Intensity as continuous variable, no association
Garcia-Toro et al. (2006) Double-blind sham-controlled RCT 60 pts					Individualized vs standard electrode placement: no difference
Stern et al. (2007) Double-blind sham-controlled RCT 45 pts			Low (1 Hz) no better than sham while **high (10 Hz) was effective**		
Fitzgerald et al. (2008) Double-blind sham-controlled RCT 60 pts					Pretx priming of target region w/ low-intensity stimulation enhanced outcomes. For example, MADRS score at 4 wks was 23.4 (active prime) vs 28.9 (sham prime,

Study Details	Low Frequency vs High Frequency, Both Right	Low Frequency, Left vs Right	Low Frequency vs High Frequency, Both Left	Sequential Bilateral vs Unilateral Stimulation	Other
					P<0.001).
Pallanti et al. (2010) Double-blind sham-controlled RCT 60 pts				Response 20% vs 35% (vs 20%, sham) remission also favored unilateral but bilateral-unilateral differences NS, 10 Hz and 1 Hz	
Triggs et al. (2010) Double-blind sham-controlled RCT 48 pts		No difference, 5 Hz			
Blumberger et al. (2012) Double-blind sham-controlled RCT 68 pts				**Bilateral sequential** rTMS superior (38.5% response) to unilateral rTMS 10 Hz left (4.5%), and sham (10%) (global *P*=0.006)	
Fitzgerald et al. (2012) Double-blind sham-controlled RCT 66				Results favored unilateral 10 Hz left (response 48%) over bilateral (31%) but were inconclusive (*P*=0.08)	
Galletly et al. (2012) Randomized comparator trial 77 pts					No difference between 3 days/wk for 6 wks (18 sessions) and 5 days/wk for 4 wks (20 sessions)
Fitzgerald et al.				No difference	

Study Details	Low Frequency vs High Frequency, Both Right	Low Frequency, Left vs Right	Low Frequency vs High Frequency, Both Left	Sequential Bilateral vs Unilateral Stimulation	Other
(2013) Randomized comparator trial 179 pts					

Transcranial Direct Current Stimulation (tDCS), Key Question #1b

The 2 systematic reviews selected for Key Question #1a reported somewhat conflicting findings (Kalu et al., 2012; Berlim et al., 2013c). See **Appendix V** for an overview of these 2 systematic reviews. Metaregression in the Kalu review showed no effect according to the number of sessions, strength of electrical current, or concurrent AD use. In contrast, stratified pooled estimates in the Berlim review did suggest differences according to the same 3 factors: greater response with 5 versus ≥ 10 sessions (OR, 3.151 versus 1.623), greater response with the use of 1 mA versus 2 mA current (OR, 3.151 versus 1.623), and greater remission in trials using monotreatment (switch from AD to tDCS) than in trials using an augmentation strategy (tDCS added to AD). (The studies using 1 mA happened to be the same studies using 5 sessions.) However, confidence intervals overlapped between strata and all but 1 pooled OR were nonsignificant. The pooled OR was significant for response in trials using monotherapy (OR, 7.54; 95% CI, 1.63 to 34.8), whereas it was nonsignificant for trials using augmentation (OR, 1.88; 95% CI, 0.470 to 2.999). The authors of the Berlim review interpreted this observation as showing a greater effect with monotherapy, but this represents an indirect comparison and the difference between monotherapy and augmentation was likely nonsignificant since the confidence intervals overlapped.

An RCT published after the systematic reviews found that a combination of tDCS plus sertraline was more effective than tDCS alone. The difference in mean MADRS score at end of treatment was 8.5 points (CI, 2.96 to 14.03; *P*=0.002), and the OR of remission was 5.7 (CI, 1.6-20.3, *P*=0.007), favoring the combination treatment. The authors considered 3 points to indicate a significant clinically relevant improvement.

Deep Brain Stimulation, Key Question #1b

Among the small number of clinical studies of DBS, no pattern of differential effectiveness according to treatment parameters was apparent.

Key Question #2:
What adverse events are associated with nonpharmacologic treatments for TRD and what are the rates of withdrawal due to lack of benefit?

Electroconvulsive Therapy (ECT), Key Question #2

The UK study of rTMS versus ECT (McLoughlin et al., 2007; fair quality) cited observational studies showing that even when subjective mood improves during treatment, cognitive decline at posttreatment assessment has been observed, but other studies have documented significant

improvement at 6 months to 4 years after ECT. Still other studies have suggested that cognitive impairment attributable to ECT may endure over the long term.

A fair-quality systematic review of all prospective studies using right unilateral ECT published as of February 2011 identified 10 eligible studies and concluded from pretest/posttest (before-and-after) data that some loss of autobiographic memory may persist for 1 to 6 months after treatment, at least with brief as opposed to ultrabrief ECT (Verwijk et al., 2012). Other cognitive side effects were found to be transient.

Another systematic review (good quality) identified 27 studies of ECT and cognitive effects in which the minimum mean age of patients was 60 years and no single participant was younger than 50 years (Gardner and O'Connor, 2008). Pretest and posttest data were reviewed. Most studies detected little change, but tests of global cognitive function, such as the Mini-Mental State Examination (MMSE), showed mixed results. The authors surmised that the mixed results were attributable to undescribed differences in electrode placement, lack of discrimination between patients with and without dementia, small sample sizes, and tests that may be insensitive to subtle cognitive changes. Gardner and O'Connor also pointed out a number of deficiencies in the selected studies, as well as the questionable validity of memory tests in the elderly. They further noted, as have other authors, that since depression itself impairs cognitive performance, the lack of an effect on cognition could have resulted from a balancing of the adverse cognitive effects of ECT with improvement due to the antidepressant effect of ECT. They were unable to form conclusions about the adverse effects of ECT on cognitive function in the elderly.

A sham-controlled RCT included as evidence for Key Question #1a reported no difference in cognitive change between patients (mean age 52 years in the ECT arm) treated with unilateral ultrabrief ECT and patients receiving sham stimulation as maintenance therapy (Nordenskjöld et al., 2013).

The sham-controlled trial by Johnstone et al. (1980) reported that 4.3% of patients in the ECT (bilateral) arm withdrew because of failure to progress compared with 1.4% in the sham arm. The study also reported that 1.4% of patients in the ECT arm withdrew due to adverse events, compared with none in the sham arm. The adverse events included a vascular incident involving the retina and a case of treatment-emergent mania. No serious events were reported by any of the comparator trials reviewed for Key Question #1b.

Differential Safety According to Treatment Parameters: See **Table 15**. The UK study of ECT (UK ECT Review Group, 2003) cited findings from 2 comparator trials suggesting that while high dose unilateral and bilateral ECT are equally effective, adverse events are more common with bilateral ECT. In a later trial by (Sackeim et al., 2008), cognitive effects were less severe in a right unilateral ultrabrief pulse group than in a bilateral ultrabrief pulse group, but no differences were detected between unilateral and bilateral brief pulse groups or between unilateral ultrabrief pulse and unilateral brief pulse groups. Cognitive performance was equivalent whether patients were treated with 100 Hz or 40 Hz ECT in another recent study (Roepke et al., 2011). The systematic review by Dunne and McLoughlin (2012) reported variable results with respect to the comparative effect of bifrontal versus *bitemporal* stimulation on cognitive performance but a very small nonsignificant effect, favoring bifrontal versus *unilateral* stimulation, on Mini-Mental State Exam (MMSE) scores.

Repetitive Transcranial Magnetic Stimulation (rTMS), Key Question #2

Data reported by RCTs is summarized in **Table 17**. Additional safety data were obtained from 3 systematic reviews that provided pooled estimates for rates of withdrawal (Berlim et al., 2013a; Berlim et al., 2013b) and treatment-induced mania (Xia et al., 2008).

Cognitive Function: No risk of cognitive adverse events was demonstrated by 6 sham-controlled RCTs. In 3 comparisons of rTMS with ECT, changes in cognitive function were small and did not clearly favor 1 technology or the other. A comparison of rTMS + ECT with ECT alone reported a substantial incidence of memory problems in both groups (36% versus 55%), but the difference was nonsignificant and the reason for this high incidence is not clear.

Withdrawals Due to Adverse Events: In 6 RCTs, the rates were higher in rTMS arms (4.2% to 50%) than in sham arms (0% to 20%) in 6 RCTs. However, in another 10 sham-controlled RCTs, there were no withdrawals due to adverse events in either arm. Data from comparisons of rTMS with ECT were insufficient to permit conclusions.

Overall Withdrawals: Two fair-quality meta-analyses of RCTs evaluating high-frequency (Berlim et al., 2013b) or low-frequency (Berlim et al., 2013a) rTMS reported pooled dropout rates as a measure of tolerability. Calculations showed that dropout rates at final clinical assessment were *lower* in high-frequency rTMS (9.9%) than in sham arms (14.03%), but the OR of 0.7 was nonsignificant (5 RCTs) (Berlim et al., 2013b). Very similar results were reported in the meta-analysis of low-frequency rTMS versus sham: 5.3% versus 11.28%; OR, 0.53 (not significant) (8 RCTs). (Study selection for these meta-analyses was not restricted to TRD populations and thus these reviews were not used as evidence for Key Question #1a or #1b.) The AHRQ review also considered overall withdrawal rates a measure of tolerability. **Table 17** shows that in sham-controlled RCTs of rTMS for TRD, the overall withdrawal rate was typically lower in rTMS arms (0% to 20%) than in sham arms (0% to 30%). In comparisons of rTMS with ECT, the evidence was conflicting regarding the relative rates of withdrawal, and no data were available for comparisons of rTMS + ECT with ECT alone.

Seizures: Across the 68 trials included in the AHRQ review or selected for this report, 1 case of posttreatment seizure was reported.

Treatment-Emergent Mania: A fair-quality systematic review identified 10 trials (520 patients) that addressed the issue of treatment-emergent mania in the use of rTMS for depression (Xia et al., 2008). Collectively, 18.7% of patients had a diagnosis of bipolar disorder. The review was not restricted to studies of rTMS for TRD. Treatment-emergent mania occurred in 0.84% of patients undergoing active rTMS and 0.73% of patients undergoing sham stimulation; the difference was not significant. The authors noted that most patients who experienced treatment-emergent mania already had a diagnosis of bipolar disorder. The studies represented a typical range of treatment parameters, according to the review authors, and the only factor that appeared to have a possible relationship to mania was a twice-daily versus once-daily regimen.

Other Adverse Events: A variety of adverse events were reported by 7 RCTs, the most common being scalp discomfort or scalp pain, which occurred at rates of 2.1% to 35.8% in rTMS arms.

Adherence: Two studies reviewed in the AHRQ report provided evidence of full adherence, but no other studies reported adherence data.

Differential Safety According to Treatment Parameters: One comparator study reported no difference in local adverse event rates between low-frequency (right) and high-frequency (left) rTMS (Isenberg et al., 2005). Three sham-controlled studies observed no difference or a negligible difference in local adverse event rates between bilateral sequential rTMS and unilateral rTMS (Pallanti et al., 2010; Blumberger et al., 2012; Fitzgerald et al., 2012). The RCT by Pallanti et al. reported a higher incidence of cognitive complaints in the bilateral arm (15%) than in the unilateral arm (10%), but the statistical significance of this difference was not reported. The sham-controlled RCT by Stern et al. (2007) reported a 50% incidence of withdrawal due to adverse events in the arm receiving low-frequency rTMS to the left DLPFC (low frequency is typically delivered to the right DLPFC) and no withdrawals in the arm receiving high-frequency rTMS to the left DLPFC or in the arm receiving low frequency to the right DLPFC.

Table 17. Adverse Events Reported by RCTs of rTMS

NOTE: Where there is no reference to statistical significance, it was not reported.

Key: AHRQ, Agency for Healthcare Research and Quality; ECT, electroconvulsive therapy; GI, gastrointestinal; grp(s), group(s); NS, not statistically significant; pt(s), patient(s); RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation

Comparison	Cognitive Function	Withdrawals Due to Adverse Events	Overall Withdrawals	Seizures	Other Adverse Events
rTMS vs Sham	No difference (Blumberger 2012) No deterioration (Fitzgerald 2012) *10%-15% vs 30% (Pallanti 2010) In Gaynes 2011 (AHRQ report): NS differences in 3 RCTs Improvement in rTMS, no change in sham; *P*<0.05 (1 RCT)	2.1% vs 0% (Blumberger 2012) In Gaynes 2011: 0 in 10 RCTs 4.2%-9.1% (rTMS) vs 0%-3.4% (sham) in 3 RCTs	In Gaynes 2011: 0 in 6 RCTs 0%-20% (rTMS) vs 6.7%-30% (sham) in 9 RCTs	†0 (1 RCT in Gaynes 2011)	2.1% vs 0% (Blumberger 2012) (scalp discomfort, recurrent headache) †0%-5% vs 5%-10% (Pallanti 2010) (headache, pain/burning in scalp, dizziness, anxiety) 4 RCTs in Gaynes 2011, results by RCT: †33% vs 3% scalp pain (*P*<0.05); 0 seizures; no other adverse effects †Exacerbation of depression, 0.6% vs 1.9%; eye pain, 6.1% vs 1.9%; toothache, 7.3% vs 0.6%; application site discomfort, 10.9% vs 1.3%; application site pain, 35.8% vs 3.8%; facial pain, 6.7% vs 3.2%; muscle twitching, 20.6% vs 3.2%; skin pain, 8.5% vs 0.6%

Comparison	Cognitive Function	Withdrawals Due to Adverse Events	Overall Withdrawals	Seizures	Other Adverse Events
					†Headache, 60% vs 50%; difficulty starting urination, 2.0% vs 1.1% (*P*=0.03), but differences NS after correction for multiplicity. †NS differences or difference favored rTMS
rTMS vs ECT	Trend toward decline in ECT arm, slight improvement in rTMS; NS difference (Rosa 2006) Small changes in different directions and favoring different grps, depending on subscale, but all w/in and between-grp differences NS (McLoughlin 2007) Of 11 measures, 3 significantly favored rTMS (more impairment in ECT pts). Other measures showed no differences and all changes were small (Hansen 2011).	13% vs 0% (Hansen 2011)	9.1% vs 15.0% (Rosa 2006) 25% vs 0 (McLoughlin 2007) 33.3% vs 26.7% (Hansen 2011)	---	No major events (McLoughlin 2007) 17% vs 0 (Hansen 2011) (scalp discomfort)
rTMS+ECT vs ECT	†Memory problems, 36% vs 55% (NS) (Pridmore 2000)	---	---	---	†Headache, 55% vs 82%; muscle pains, 36% vs 55% (NS) (Pridmore 2000). (Events occurring less frequently were not reported by Pridmore.)

*Adverse event data from this study were missing from the safety chapter in the AHRQ review, except for withdrawals due to adverse events.

†A predefined list or tool was used to systematically assess adverse events.

Transcranial Direct Current Stimulation (tDCS), Key Question #2

A good-quality systematic review calculated adverse event rates based on 117 English-language studies that reported adverse events (Brunoni et al., 2011a). The 117 studies were identified from a set of 209 studies (3836 patients) of tDCS. Both controlled studies and case series were included. The studies were *not limited to depression as an indication* and the authors did not assess rates by diagnosis. Patients were on average 34 years of age and 50% were female. Across the 117 studies that entailed active tDCS, the following adverse events were reported: itching (39.3%); tingling (22.2%); headache (14.8%); burning sensation (8.7%); and discomfort (10.4%). Across the 82 studies that had sham groups, the following rates were reported in patients undergoing sham stimulation: itching (32.9%); headache (16.2%); tingling (18.3%); burning sensation (8.7%); and discomfort (10%). It appears that adverse events were more common in patients receiving active tDCS but the difference was small. Meta-analysis of itching was possible for 8 sham-controlled studies, 2 of which (Boggio et al., 2008; Loo et al., 2010) involved patients with depression and are discussed elsewhere in this report. The pooled OR for itching was 0.95 (95% CI, 0.28 to 3.24).

According to the Kalu review, headache, itchiness, and redness at the site of stimulation were experienced in both active treatment groups and sham groups (Kalu et al., 2012). The trial by Brunoni et al. (2013b) reported a higher incidence of skin redness in tDCS groups compared with sham groups: 25% versus 8% (P=0.03) at 2 weeks and 22% versus 8% (P=0.03) at 6 weeks. The frequency of itchiness was also substantially greater in the tDCS groups (37% versus 25% at 2 weeks; 34% versus 18% at 6 weeks). However, this and other differences in nonserious events were nonsignificant.

The Kalu review also reported 3 cases of tDCS-induced hypomania (low-intensity mania) among 42 (7%) patients receiving tDCS in 2 RCTs and 1 case series. The study populations represented by these 3 studies had a prevalence of bipolar disorder of 8.7%, but the review authors did not indicate whether hypomania occurred in patients with unipolar or bipolar depression at baseline. In the Brunoni trial, treatment-induced hypomania was observed in 17% of patients receiving tDCS + placebo, 3% of patients receiving sham sertraline, and 3% of tDCS + sertraline, but in none of the sham + placebo group; none of these patients had bipolar disorder at baseline.

The Berlim review reported dropout rates of 5.8% in active tDCS arms and 5.2% in sham arms, with a nonsignificant pooled OR, favoring sham stimulation, of 0.893 (95% CI, 0.259 to 3.079) (Berlim et al., 2013c). The review authors considered the dropout rates to be a measure of patient acceptability.

During maintenance treatment following the Brunoni trial, patients experienced minor adverse events similar to those reported in the original trial; 2 patients dropped out because they were feeling worse (Valiengo et al., 2013).

Deep Brain Stimulation (DBS), Key Question #2

There are several potential complications associated with DBS. As with any surgical procedure, the general anesthesia required for implantation of DBS components poses risks to the patient. According to a review article cited in the Hayes report, there is a risk of ≤ 10% for hemorrhage, which can range from being unnoticeable to the patient to causing neurological deficits or death (Goodman and Alterman, 2012). After surgery, there is a 0% to 15% risk of a device-related infection according to another sourced cited by the Hayes review (Lozano et al., 2008). The infection may be mild and treatable with antibiotics only, or could be more serious and require the removal of some or all of the hardware combined with intense antibiotic management. Other hardware-associated complications may include fracture or

breaking of hardware components, or erosion of the scalp tissue overlying the burr hole cap (Goodman and Alterman, 2012).

A poor-quality systematic review of 546 English-language clinical studies and reports (≤ 10,339 patients; extent of overlapping populations unknown) tallied adverse events reported for patients undergoing DBS treatment for any indication (Appleby et al., 2007). Patients were on average 54 years of age and 37% were female. The authors did not report how many, if any, of the studies or patients represented DBS for treatment-resistant depression. The authors did not attempt to identify unique patients, which would have allowed calculation of per-patient adverse event rates. They also did not provide information on follow-up intervals. Of 6574 reported device-related events, 16% were due to infection, 15% were due to explantation, 15% were due to lead fracture, and 14% were due to erosion. Less frequent device-related events included battery failure, intracranial hemorrhage, misplacement, and postoperative lead migration. Of 6573 reported somatic adverse events, a wide variety of events were reported, none of which accounted for ≥ 5% of events. Four non-suicide deaths and 11 incidents (0.16% of all adverse events) of completed suicide were reported. The authors considered the incidence of completed suicide to be cause for concern.

A recent fair-quality systematic review identified 26 studies of DBS in psychiatric patients (Bergfeld et al., 2013). Of the 130 patients, 28 were being treated for MDD. No study reported substantial decline following DBS and some studies reported an improvement in cognitive functioning. Another systematic review (poor quality) of DBS for psychiatric disorders reported that cognitive side effects were generally transient (Duits et al., 2013).

Adverse events reported in the primary studies of DBS for TRD included: infection, paresthesia, anxiety, mood changes, worsening depression, suicidal ideation, insomnia, dysphagia, headache, and psychotic symptoms. There were no consistent adverse events across the 5 studies included in the Hayes report (Hayes, 2012), with the exception of infection, which occurred in 5% to 20% of the patients in 3 of the studies. There was no evidence of cognitive decline in any study. In 1 trial, a DBS discontinuation phase was initiated after 24 weeks of active stimulation (Holtzheimer et al., 2012). The first 3 patients to enter this phase of the trial suffered complete relapse within 2 weeks, which did not improve immediately upon restimulation, although relief of symptoms did eventually occur. These patients experienced significant distress and increased suicidal ideation, and the DBS discontinuation phase was eliminated for subsequent patients based upon safety concerns. The authors speculated that the slow return of efficacy may have been partially due to patients' disappointment in not experiencing an immediate reduction in symptoms, as had been seen with discontinuation and resumption of DBS for other indications (e.g., Parkinson's disease).

Key Question #3:
Does the effectiveness of nonpharmacologic treatments for treatment-resistant depression vary by subpopulation defined by such factors as: age, race/ethnicity, gender, disease severity, disease duration, depression diagnosis (unipolar or bipolar depression), symptom type (e.g., psychotic, postpartum), comorbidities, or number and type of prior treatments (including other nonpharmacologic treatments)?

Electroconvulsive Therapy (ECT), Key Question #3

A fair-quality systematic review and meta-analysis calculated pooled odds ratios (ORs) for remission in prospective and retrospective cohort studies comparing the efficacy of ECT in patients with unipolar and bipolar MDD (Dierckx et al., 2012). Six studies were identified that reported remission as ≤ 7 on the 17-item HAM-D or as ≤ 10 on the 24-item HAM-D and met other inclusion criteria. The studies involved 1106 patients, 790 with unipolar depression and 316 with bipolar depression. The pooled OR of remission suggested equivalent efficacy: OR, 1.08 (95% CI, 0.75 to 1.57). There was high statistical heterogeneity in this estimate, and the study-specific ORs were inconsistent, ranging from 0.28 (significant, favors unipoloar) to 2.14 (nonsignificant, favors bipolar). The studies varied according to whether AD medications were allowed during the study, and the review authors did not attempt an analysis according to this factor. The authors also did not report whether patients were considered to have TRD. The review was considered since the 2011 evidence review for AHRQ (Gaynes et al., 2011) did not address differential effectiveness according to most of the factors of interest.

Another fair-quality systematic review and meta-analysis calculated pooled ORs for remission in prospective cohort studies comparing the efficacy of ECT in patients with and without documented medical resistance (Heijnen et al., 2010). Studies were selected if they reported remission as ≤ 7 on the HAM-D_{17}, as ≤ 10 on the HAM-D_{24}, or ≤ 8 on the MADRS. Studies also had to assess medication resistance by using the Antidepressant Treatment History Form (ATHF) to evaluate the adequacy of previous AD trials (see **Appendix I** for more information on the ATHF). Seven studies met these and other selection criteria. They involved 958 patients, 585 with previous AD failure and 373 without clear previous AD failure. The overall weighted OR of remission, comparing adequate versus inadequate prior AD treatment, was 0.52 (95% CI, 0.39 to 0.69); high heterogeneity was detected. Study-specific ORs ranged from 0.16 (significant) to 1.14 (nonsignificant). This suggests that ECT is *less effective* in confirmed TRD than in MDD without a well-documented history of AD failure. It should be noted that the ATHF guides the clinician to a rating based on the confidence with which medication resistance can be assumed, based on the available documentation, and does not rate patients according to a prospective assessment of treatment resistance. The authors point out that the results may not be generalizable to the current practice of using either bilateral stimulation or unilateral stimulation with a high dose. They also described considerable clinical heterogeneity among the studies in terms of diagnostic criteria, measurement instruments, and the proportion of patients with psychotic depression.

A post hoc analysis of 2 randomized comparator trials (148 patients) evaluated whether the difference between a relatively ineffective form of ECT, as identified in those trials, and more effective forms of ECT was present in different depression subtypes (Sobin et al., 1996). The 2 trials were conducted by the same group of researchers. The trials found that right unilateral ECT at a low dose was considerably less effective than right unilateral ECT at a high dose or bilateral ECT at high or low dose. A statistically significant difference, substantially favoring the more effective forms of treatment, was observed in subgroups defined by psychosis, retardation, and agitation, as well as in the combined overall study groups. According to the inclusion criteria for the 1993 source trial, all patients had failed ≥ 2 medications.

Repetitive Transcranial Magnetic Stimulation (rTMS), Key Question #3

Since the AHRQ review considered differential effectiveness only according to symptom type, age, and comorbidity, previous work by Hayes was used to identify studies that evaluated other factors. As **Table 18** shows, the evidence of a differential effect by age was conflicting, with 3 RCTs showing a positive association with younger age (Su et al., 2005), a positive association with older age (Jorge et al., 2008),

or no association (George et al., 2010). Two RCTs also presented conflicting data regarding whether an effect is associated with depression severity (Su et al., 2005; Jorge et al., 2008; Ullrich et al, 2013). Three trials (total, n=321) suggested that effectiveness is not associated with duration of episode (Su et al., 2005; Jorge et al., 2008; George et al., 2010), and 2 trials (total, n=122) found no association with gender (Su et al., 2005; Jorge et al., 2008).

One of the trials found no association of effect with unipolar versus bipolar depression (Su et al., 2005). The AHRQ review calculated pooled estimates separately for trials in which study populations did or did not include patients with bipolar depression (≤ 20% of study group), and estimates were very similar.

Stratified pooled estimates reported in the AHRQ review suggested little difference between trials requiring ≥ 2 prior AD failures and trials requiring ≥ 1 prior AD failure. Pooled estimates were somewhat smaller for trials of patients with ≥ 1 failure, but confidence intervals were largely overlapping, suggesting a nonsignificant difference. This comparison represents an indirect analysis. Only 1 trial (George et al., 2010) provided a direct comparison of patients within the same trial; the results suggested no difference according to medication resistance.

Other factors were investigated by single small trials and thus the evidence was insufficient to support conclusions. No trials restricted enrollment to particular subgroups such as the elderly or patients with psychiatric comorbidity.

Table 18. Comparison of Treatment Effect and Patient Factors in RCTs of rTMS

Key: AD, antidepressant (medication); AHRQ, Agency for Healthcare Research and Quality; BL, baseline; CI, confidence interval; MA, meta-analysis; NNT, number needed to treat; NR, not reported; pt(s), patient(s); RCT, randomized controlled trial; RR, relative risk; rTMS, repetitive transcranial magnetic stimulation; SR, systematic review; WMD, weighted mean difference

Study	Associated with Effect	Not Associated with Effect or Unlikely Association (Overlapping CIs)
Su et al. (2005) Double-blind sham-controlled RCT 30 pts (factors compared w/ response)	• Younger age (38.8 vs 50.3 yrs; *P*<0.05) • Younger age at disease onset (29.8 vs 43.6 yrs; *P*<0.05) • Menopausal status (pre- or peri- vs postmenopausal; *P*<0.05) • Lower BL HAM-D score (22.2 vs 28.9; *P*<0.05)	• Gender • Duration of current episode • # previous episodes • Unipolar vs bipolar depression
Jorge et al. (2008)* Double-blind sham-controlled RCT 92 pts (factors compared w/ response)	• Age ≥65 yrs • Greater gray matter volume	• Gender • Race • Marital status • Socioeconomic status • Severity or duration of current episode • Intensity of AD • Cognitive test score
George et al. (2010) Double-blind sham-		• Age • Duration of episode

Study	Associated with Effect	Not Associated with Effect or Unlikely Association (Overlapping CIs)
controlled RCT 199 pts (factors compared w/ response)		• Medication resistance (low vs high) • Treatment site
Gaynes et al. (2011) (AHRQ evidence review) SR w/ MA of RCTs		***Number of prior AD failures, WMD in score (negative value favors rTMS):*** All trials: –5.92 (CI, –8.15 to –3.70) (I^2=80%) (24 RCTs) ≥2: –5.74 (CI, –7.79 to –3.68) (11 RCTs) ≥1: NR Unspecified: Range –1 to –13.84 (3 RCTs) ***Number of prior AD failures, RR of response:*** ≥2: 3.34 (CI, 1.92-5.82; NNT=5) (11 RCTs) ≥1: 2.68 (CI, 1.52-4.70; NNT=5) (16 RCTs) ***Number of prior AD failures, RR of remission:*** ≥2: 6.12 (CI, 1.89-19.80; NNT=4) (5 RCTs) ≥1: 3.73 (CI, 1.23-11.30; NNT=6) (9 RCTs)
Ullrich et al. (2011) Double-blind sham-controlled RCT 43 pts		BL severity (no association w/ response)

*Denotes multivariate analysis, i.e., factors were found to be independent predictors after adjustment for other variables.

Special Populations: The AHRQ review identified 3 trials in older patients with vascular dementia. The results favored rTMS in all 3 trials, but 1 trial was underpowered to detect a significant effect.

Transcranial Direct Current Stimulation (tDCS), Key Question #3

Treatment Effect Modifiers

Meta-regression conducted by Kalu et al. (2012) showed that treatment effect did not vary by baseline severity. The authors noted that the metaregression considered patient factors separately and thus does not shed light on whether certain combinations of patient factors are associated with better outcomes. The authors of the Berlim review considered the number of trials too small to allow analysis of differential effect according to unipolar versus bipolar depression and did not comment on any other patient factors.

Participants in an open-label crossover and maintenance phase were more likely to relapse if they had had ≥ 2 failed AD medication failures prior to acute treatment with tDCS, compared with < 2 failures (Valiengo et al., 2013). In univariate analysis, there was no association between age, gender, baseline severity, benzodiazepine use, original treatment allocation, or sertraline use and relapse. The study was considered to be of poor quality.

Treatment Response Predictors

The trial by Brunoni et al. (2013b) found no association between response and either age or sex *within the entire study group*. Some case series have analyzed associations between patient characteristics and posttreatment scores. Data pertaining to factors not analyzed in either systematic review and addressed by ≥ 2 case series are summarized here. Sample sizes ranged from 23 to 32. The following results were reported:

- Age: No association (Ferrucci et al., 2009; Dell'Osso et al., 2012)
- Gender: No association (Ferrucci et al., 2009; Dell'Osso et al., 2012)
- Unipolar versus bipolar: No association (Brunoni et al., 2011b; Dell'Osso et al., 2012)

NOTE: The study by Dell'Osso et al. appears in the Kalu review with a date of 2011, but is currently indexed in PubMed as a 2012 study.

Although the cases series controlled for some but not all potential confounders and reported blinded evaluation in some cases, the evidence does not directly address the question of a differential effect.

Deep Brain Stimulation (DBS), Key Question #3

The available studies of DBS for TRD were very small and uncontrolled. There were no analyses of the outcomes of DBS for subpopulations, except in the study by Holtzheimer et al. (2012). The study showed small differences favoring greater improvement in unipolar patients (n=10) compared with bipolar patients (n=17) according to HAM-D and BDI scores, but mixed results according to GAF score; statistical significance was not reported. However, this analysis speaks to response predictors, not differential effects.

Key Question #4: What are the cost implications and cost-effectiveness of nonpharmacologic therapies for TRD?

Three economic evaluations of rTMS met criteria for review. **Table 19** presents findings to support the following discussion.

rTMS Versus Sham or Pharmacotherapy as Usual

Simpson Study, U.S.

An economic evaluation of rTMS versus sham treatment, as well as rTMS versus pharmacotherapy as usual, reported very favorable findings (Simpson et al., 2009). Estimates of the effectiveness of rTMS in patients with TRD came from 3 trials sponsored by the manufacturer (Neuronetics): a multicenter double-blind sham-controlled trial (n=301) (O'Reardon et al., 2007), an open-label trial (n=158) (Avery et al., 2008) in which nonresponders from both the active and sham arms of the RCT participated, and an open-label study (unpublished data referenced by Janicak et al., 2008) of AD maintenance treatment with the possibility of TMS as an add-on rescue treatment for patients who responded to TMS (≥ 25% improvement in HAM-D score) in either of the first 2 studies. The economic evaluation was also sponsored by Neuronetics. Estimates of the effectiveness of pharmacotherapy for patients with TRD came from the STAR*D trial (Rush et al., 2006).

Reported results showed rTMS to be cost-saving or to have a low cost-utility ratio compared with sham stimulation. However, the authors noted that an incremental cost-effectiveness ratio (ICER) derived from a sham-controlled trial would have little practical meaning and thus emphasized results from the models that combined data from the open-label study for nonresponders with data from the STAR*D trial for patients who had to move to Level 2 or Level 3 because of medication resistance. Reported results showed rTMS to be cost saving when compared with pharmacotherapy, regardless of whether healthcare costs only were considered or a societal perspective, considering work loss and caregiver time as well as healthcare costs, was assumed. A subgroup analysis was conducted, considering only those patients with the lowest level of treatment resistance at baseline, i.e., patients whose prior AD failure was simply an inadequate trial of medication. In this subgroup, cost savings were even greater than for the overall population of patients. Conclusions remained qualitatively the same when the cost per rTMS and costs associated with suicide were varied. The authors concluded that rTMS is a cost-effective treatment that may even result in cost savings, especially when used at earlier levels of treatment resistance.

Several omissions from the study report published by Simpson and colleagues suggest reporting bias or methodological weaknesses and make interpretation of the findings very difficult. It was not clear which quantitative measures of benefit were included in the models and what value was assumed for those measures. Though not mentioned in the Simpson report, the primary outcome measure in the original RCT (O'Reardon et al., 2007), which was difference in MADRS score change at 4 weeks adjusted for baseline score and medication resistance, did not show a statistically significant effect. The difference became significant when 6 patients with mild baseline depression were excluded, but the magnitude of improvement in that subset was not reported. The unadjusted difference in improvement in the overall group was very small (from 32.8 to 27, 5.8 points, in the rTMS arm; from 33.9 to 29.8, 5 points, in the sham arm). (The mean number of fully adequate AD treatments prior to study enrollment was exactly the same in the 2 arms.) It was not clear whether rTMS was assumed to be more effective or as effective as the comparator treatments; Simpson and colleagues reported effect sizes suggesting that rTMS had an effect when compared with sham stimulation, but effect size data did not appear in the original trial report. The economic evaluation did not include a sensitivity analysis of effectiveness estimates, and some data seemed to be missing from the results of sensitivity analyses of the 2 cost assumptions selected for testing. Additionally, Simpson and colleagues did not identify the outcomes data that were used in what they identified as the most meaningful comparison, open-label rTMS study participants versus treatment-resistant patients in the STAR*D trial.

Other deficiencies were noted in the review of this economic evaluation. The durability of effect data used in this evaluation had not been published. Utilization data were collected at entry into the Neuronetics RCT and at the end of maintenance treatment, but it was unclear how utilization was determined for patients who did not respond well enough to enter the maintenance phase, or for patients in the usual pharmacotherapy (STAR*D) group. Lastly, there were small numerical discrepancies between the results as presented in the text of the economic evaluation report and in the results table. It is not clear that this evaluation supports the authors' conclusions.

rTMS Versus ECT

Kozel Study (U.S.)

A U.S. decision analysis study estimated a very high cost-utility ratio of $460,031 per quality-adjusted life-year (QALY) (base year unclear), when rTMS for both acute and maintenance treatment was assumed to be only slightly more effective (64% response rate) than ECT (60% response rate) for acute

and maintenance treatment in nonpsychotic patients with severe MDD (Kozel et al., 2004). The study assumed a societal perspective and a 1-year time horizon. The study that served as the source of response rates in the base case (Grunhaus et al., 2000) was not included in the present report because of poor quality. However, the findings from this study were less favorable to rTMS than the findings in some of the other studies and thus provide a conservative estimate for the an analysis of the cost-effectiveness of rTMS versus ECT. A strategy of initially treating patients with rTMS and then treating nonresponders with ECT dominated a strategy of ECT alone, i.e., the rTMS-then-ECT strategy was less expensive and more effective. Although the authors did not report a cost-utility ratio for the rTMS-then-ECT strategy versus rTMS alone, data supplied in the study report yield a ratio of $31,783/QALY for this comparison. Such a ratio would likely be considered acceptable, even taking into account that cost data were collected in 2004 or earlier. The authors concluded that there may be a considerable cost advantage to a combination rTMS-ECT strategy compared with ECT alone. A key weakness of this study was failure to conduct multi-way sensitivity analyses to determine whether varying the estimates of rTMS and ECT effects in opposite directions at the same time would lead to the same conclusion. This type of sensitivity analysis would have been especially useful since the comparative effectiveness of rTMS and ECT has not been established (**see Tables 2 and 11**).

Knapp Study (UK)

The UK trial of rTMS versus ECT (Eranti et al., 2007; McLoughlin et al., 2007) included a cost-effectiveness analysis as part of the study protocol. A separate publication (Knapp et al., 2008) presented the cost-effectiveness results and was reviewed for this report. The trial compared 15 sessions of unilateral, high-frequency rTMS with biweekly, primarily bilateral ECT (≤ 10 sessions, as determined by clinician). This trial was rated as fair quality. Posttreatment remission was 59% in the ECT group and 17% (*P*=0.05) in the rTMS group according to intention-to-treat (ITT) analysis (see **Table 11**). Costs were collected during treatment and during the 6-month follow-up period. Functional status was measured at baseline and at 6-month follow-up with the Mental Health Component of the SF-36 Health Survey. Scores were then translated to utility values (0 = death to 1 = perfect health) according to societal values.

As a form of sensitivity analysis and to measure uncertainty around their estimate of costs compared with benefit, the authors constructed cost-effectiveness acceptability curves (CAECs). A CAEC plots the probability (y-axis) that the cost per unit of benefit will equal or be less than a particular cost-effectiveness threshold (x-axis). The CAECs suggested that the ICER would probably not exceed approximately £500 per unit difference in HAM-D score from the perspective of the national health and social services plan and would probably not exceed approximately £700 per unit difference in HAM-D from a societal perspective. The societal perspective differed from the national health and social services plan by the addition of the cost of informal (unpaid) care given to the patient after return home. Assuming a price year of 2004, these estimates translate to $921 and $1290 in U.S. 2013 dollars per unit difference in HAM-D score. The authors did not identify a desirable cost-per-effectiveness-unit ratio. Another CAEC suggested that the probability of cost utility remaining under £30,000/QALY ($55,282/QALY in 2013 U.S. dollars) was very low—13% to 22%, depending on the cost/QALY value, which ranged from £0 to £30,000. The authors concluded that it was unlikely that decision makers would view rTMS as more cost-effective than ECT.

(NOTE: The currency and date conversions represent approximate translations of results to current U.S. values, based on November 9, 2013 use of the CCEMG-EPPI-Centre web-based cost converter with the International Monetary Fund [IMF] dataset for Purchasing Power Parity [PPP] values, and assuming pre-

Euro currency. The calculator is available at: http://eppi.ioe.ac.uk/costconversion. See Shemilt et al. [2010].)

Institute for Clinical and Economic Review (ICER) (U.S. study)

In its work with the New England Comparative Effectiveness Public Advisory Council (CEPAC), ICER conducted a cost-effectiveness modeling study, assuming that rTMS and ECT have equivalent efficacy (Emond et al., 2013; New England CEPAC, 2013a). This assumption was based on results from the only comparator trial in the AHRQ report in which patients were required to have failed ≥ 2 AD trials prior to study enrollment (Rosa et al., 2006). In the Rosa study, results favored rTMS but were not statistically significant. As an estimate of the clinical outcome of usual care, ICER used the inverse of AHRQ's pooled relative risks (RRs) in its meta-analyses of rTMS-versus-sham trials. Patients were assumed to have experienced ≥ 2 adequate AD trials, and the hypothetical cohort included adults aged 18 to 64 years (Medicare population excluded). The model predicted a cost-utility ratio of $216,468/QALY from a payer perspective and $321,880/QALY from a societal perspective. Further analysis showed that the cost of each rTMS session would have to decrease by 50% to achieve a cost-utility ratio of $100,000/QALY. This cost-effectiveness study was not reviewed in detail in the present report or considered in forming conclusions regarding Key Question #4 since full details have not been published in a peer-reviewed journal.

Table 19. Evidence Overview, Economic Evaluations of rTMS

Key: AD, antidepressant (medication); AHRQ, Agency for Healthcare Research and Quality; ATHF, Antidepressant Treatment History Form; CEAC, cost-effectiveness acceptability curve; ECT, electroconvulsive therapy; ED, emergency department; f/u, follow-up; grp(s), group(s); HAM-D, Hamilton Depression Rating Scale; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; NR, not reported; pharmacotx, pharmacotherapy; pt(s), patient(s); QALY, quality-adjusted life-year; rTMS, repetitive transcranial magnetic stimulation; SF-36; SF-36 Health Survey; STAR*D trial, Sequenced Treatment Alternatives to Relieve Depression trial; TRD, treatment-resistant depression; tx, treatment (or therapy)

Authors/Location/ Study Design/Patients	Treatment Groups	Effectiveness and Utility Estimates	Source of Costs/Costs Included/Base Year and Adjustments	Findings/Sensitivity Analysis	Comments
Kozel et al. (2004) U.S. Decision analytic model. Societal perspective. 1-yr time horizon. 39,486 nonpsychotic pts w/ severe MDD	Bilateral, brief-pulse ECT: 8 sessions Unilateral rTMS: 20 sessions (total 8000 pulses) in first 8 pts, 20 sessions (24,000 pulses) in next 12 pts rTMS followed by ECT in nonresponders	***Response rates, acute phase:*** rTMS, 64%. ECT, 60%. Source: RCT (Grunhaus 2000) ***Continued response in maintenance phase:*** rTMS, 50%. ECT, 93%. Sources: Unpublished abstract (rTMS), case-control study (ECT). ***Utility (scale 0-1.0):*** Depressed, 0.25; recovered, 0.91. Source: Published literature.	Census and insurance data. Direct tx-related costs, productivity loss (including companion for ECT txs), travel costs. Base yr NR.	***ECT vs rTMS:*** ICER, $460,031/QALY ***rTMS-then-ECT vs ECT:*** rTMS-then-ECT dominated ECT (less expensive by $61,425,778 and more effective by 1538 QALYs) ***rTMS-then-ECT vs rTMS:*** ICER, ~$31,783/QALY (calculated w/ data supplied in article) ***1-way sensitivity analyses:*** Conclusions held when rTMS acute response decreased to 40%, ECT acute response rose to 80%, rTMS maintenance response fell to 25%, rTMS costs increased, or ECT costs decreased.	***Strengths:*** Sensitivity analysis spans rTMS response rates and exceeds ECT rates in other comparator trials selected for this report (see **Table 11**). In the source study for response rates, most pts were medication resistant. ***Weaknesses:*** 1-way sensitivity analysis.
Knapp et al. (2008) UK Trial-based study. National health/social services perspective *and* societal	Unilateral rTMS (15 sessions; 15,000 pulses) Unilateral (18%) /bilateral (82%) ECT (≥10 sessions)	***Remission (rTMS, ECT):*** 17%, 59%. ***Relapse at 6 mos (rTMS, ECT):*** 50%, 42%. ***Utility gain at 6 mos (rTMS, ECT) (scale***	Standard checklist for service utilization data. Unit costs from annual, national source. Direct tx-related costs, including annualized costs associated w/	***Mean difference in QALY gain (ECT minus rTMS):*** 0.0003 (NS) ***Mean total cost at 6 mos (ECT minus rTMS):*** –£4329 ***Sensitivity analyses:*** Different assumptions for informal	***Strengths:*** Empirical. Benefit and costs from same source. Inclusion of capital costs. In the trial, pts had failed

Authors/Location/ Study Design/Patients	Treatment Groups	Effectiveness and Utility Estimates	Source of Costs/Costs Included/Base Year and Adjustments	Findings/Sensitivity Analysis	Comments
perspective. 6 mos time horizon. 46 pts w/ unipolar or bipolar MDD and referred for ECT		***0-1.0):*** 0.0300, 0.0297 All estimates were from a fair-quality single (assessor)-blind RCT (Eranti et al., 2007; McLoughlin et al., 2007). Utility values calculated according to societal weights attached to trial-based SF-36 scores. See Table 11.	facility and machinery, and maintenance costs (national health/social services perspective). Tx-related costs plus informal (unpaid) care costs after return to home (societal perspective). Costs collected 2002-2004; no adjustments reported.	care costs yielded similar results. CEACs suggested cost/HAM-D improvement would remain ≤£700, depending on perspective and cost/QALY gain had very low probability (<25%) of remaining at £30,000/QALY.	mean >2 ADs. ***Weaknesses:*** High loss to f/u at 6 mos (but ITT analysis); no inclusion of work loss; no transportation costs for ECT
Simpson et al. (2009) U.S. Markov modeling study. Health system perspective (related medical costs) and societal perspective (productivity costs [work loss and caregiver time] added). 1-yr time horizon. 301 pts w/ unipolar MDD and ≥1 AD failure in current/most recent episode (Neuronetics studies*) + 1669 pts in STAR*D trial	Unilateral rTMS for up to 6 wks (20-30 sessions; 30,000-45,000 pulses), followed by 3-wk taper Sham rTMS Pharmacotx as usual in pts who had 1 or 2 prior AD failures	***Depression outcomes:*** Quantitative assumptions unclear. Sources: Neuronetics studies* and STAR*D trial. *Utility:* No depression (0.83), mild (0.73), moderate (0.63), and severe (0.30); in-hospital failure (0.09). Source: Other studies in the literature.	Neuronetics study: Work loss time, healthcare utilization, and caregiver support according to self-report questionnaire (at entry to initial RCT and at end of maintenance tx study). Red Book (2006) and Medicaid (2004) data: Inpatient, ED, outpt, and drug costs. Cost/session rTMS, $300 in base case.	***Acute rTMS vs sham at rTMS cost/session $300 (n=301):*** ICER <$10,000/QALY (societal perspective) and <$40,000 (health system perspective) in overall population; cost-saving in low resistance subgrp ($5092, societal; $29,556, health system). ***Open label trial in nonresponders (n=158) vs STAR*D, cost/session $300:*** rTMS was cost-saving from both perspectives and in the overall study grp as well as in the low resistance subgrp (savings $746-$10,516 per pt per yr, depending on analysis).	***Strengths:*** Multisite, larger sample than other studies, 3-phase trial, modest loss to f/u. ***Weaknesses:*** Numerous reporting omissions. ***Generalizability:*** # sessions and total # pulses considerably greater than in other studies. Funded by

Authors/Location/ Study Design/Patients	Treatment Groups	Effectiveness and Utility Estimates	Source of Costs/Costs Included/Base Year and Adjustments	Findings/Sensitivity Analysis	Comments
undergoing Level 3 or Level 4 tx.			Base yr 2006 (2004 costs inflated to 2004 according to medical care consumer price index).	**Break-even point (time at which costs for an rTMS pt are offset by costs saved from add'l txs averted) (rTMS cost $300/session):** Overall: 37 wks. Low resistance subgrp: 29 wks. ***Sensitivity analyses:*** Varying cost/rTMS session from $250 to $400 and cost of suicide from $40,000 to $60,000 had little effect.	Neuronetics.

*Sham-controlled RCT (n=301) (O'Reardon et al., 2007). Nonresponders at week 4 from either arm (total, n=158) could then start a new course of rTMS, same protocol in open-label study arms (Avery et al., 2008). Pts with severe depression (MADRS >27) end of RCT and open-label trial received a new pharmacotherapy. Remitters from the first 2 studies underwent maintenance pharmacotherapy, with the possibility of additional rTMS if needed (maintenance pharmacotherapy for all remitters (Janicak et al., 2008); these data were not used in the cost-effectiveness study.

Practice Guidelines

American Psychiatric Association (APA)

Major Depressive Disorder

The third edition of the APA guideline on the treatment of patients with major depressive disorder (MDD) was rated as fair quality (APA, 2010). The document advises that initial treatment for an acute episode of MDD may include pharmacotherapy, depression-focused psychotherapy, or a combination of medications and psychotherapy. The choice of initial treatment modality is based on clinical features and other factors, including patient preference and prior treatment experiences. The guidelines present psychotherapy alone or an antidepressant (AD) medication alone as alternatives for initial treatment of patients with mild-to-moderate MDD. A combination of psychotherapy and AD medication is recommended as an option for patients with moderate-to-severe MDD, and ADs are definitely recommended for patients with severe MDD. According to the guidelines, patients who have an incomplete response to adequate trials of either psychotherapy alone or AD medication may benefit from combined treatment with medication and a depression-focused psychotherapy.

The authors stated that initial treatment should not be considered unsuccessful until no improvement is observed after 4 to 8 weeks of treatment. At this point, the treatment plan may be adjusted after reassessing the diagnosis and reviewing comorbid conditions and psychosocial factors. If psychotherapy alone is the initial treatment and is found to be ineffective, the clinician is advised to consider increasing the intensity of treatment or prescribing an AD. Adjustments to dosing are recommended if 4 to 8 weeks of AD medication is not effective. Effectiveness is equated with bringing at least "moderate" relief of symptoms. "Moderate" is not defined, although the use of depression scales is recommended, and no standard definition of response is offered. If 4 to 8 weeks of optimally designed treatment with AD medication is ineffective, augmentation and switching strategies are recommended. The guidelines do not define the number of AD medication trials that must be tried before therapies such as electroconvulsive therapy (ECT) are tried but imply that a combination of AD medication and psychotherapy should first be tried.

ECT is considered by the APA to be the most effective therapy for patients who have not responded to psychotherapy and/or adequately prescribed AD medications. Other factors that the APA says may increase the need for ECT are functional impairment, numerous medical trials, psychotic or catatonic features, urgent need for response (as when patients are suicidal or refusing food), and patient preference for ECT or a previous positive response to ECT. Light therapy is also presented as an option. The APA guidelines also state that repetitive transcranial magnetic stimulation (rTMS) may be considered, although there is less evidence to support this modality compared with ECT. Vagus nerve stimulation (VNS) is considered an option for patients who have not responded to at ≥ 4 trials of treatment, including ECT.

According to the APA recommendations regarding continuation treatment, patients who responded to ECT as an acute-phase treatment may receive continuation pharmacotherapy or continuation ECT, particularly when medication or psychotherapy has been ineffective in maintaining remission.

Most of the recommendations given in the APA guidelines were considered to be based on Level I (substantial clinical confidence) evidence; some recommendations were considered to be supported by Level II evidence. However, recommendations were not directly paired with study references.

Bipolar Disorder

Guidelines published in 2002 include a Level I recommendation (recommended with substantial clinical confidence) that ECT be considered for patients with severe depression or with treatment-resistant depression (TRD) (APA, 2002). However, no definition of treatment resistance is offered. The guideline further states that maintenance ECT may be considered if there is a response to acute treatment with ECT and describes ECT as a reasonable alternative to antipsychotic medication for patients who have psychotic features during a depressive episode, but these recommendations are not rated. The APA website includes the following disclaimer:

> *"This guideline is more than 5 years old and has not yet been updated to ensure that it reflects current knowledge and practice. In accordance with national standards, including those of the Agency for Healthcare Research and Quality's National Guideline Clearinghouse, this guideline can no longer be assumed to be current." (http://psychiatryonline.org/guidelines.aspx)*

A *Guideline Watch* was published in 2005 to provide a review of studies published since the 2002 guidelines were prepared (Hirschfeld, 2005). This document does not refer to any studies of ECT, rTMS, transcranial direct current stimulation (tDCS), or deep brain stimulation (DBS). The 2002 guidelines, combined with the *Guideline Watch*, were rated as being of fair quality.

Canadian Network for Mood and Anxiety Treatments (CANMAT)

Section IV of the CANMAT guidelines on management of MDD in adults relates to neurostimulation therapies, including rTMS, ECT, vagus nerve stimulation (VNS), and DBS (Kennedy et al., 2009). Recommendations are characterized by the strength of supporting evidence, ranging from Level 4 (expert opinion/consensus) to Level 1 (≥ 2 RCTs with adequate sample sizes and/or meta-analysis with narrow confidence intervals). The guidelines make these recommendations:

> ECT: Recommended for first-line treatment for acute suicidal ideation, MDD with psychotic features, or TRD (Level 1 evidence for acute treatment, relapse prevention, and safety). Also recommended as first-line treatment for catatonia or when there has been a prior favorable response, repeated medication intolerance, or rapidly deteriorating physical status (Level 3). Recommended as second-line treatment for patients who are otherwise treatment-resistant or who have medication intolerance.
>
> rTMS: Second-line treatment (Level 1 for acute treatment and safety; Level 3 for relapse prevention).
>
> DBS: Investigational

The CANMAT guidelines offer no definition of TRD or medication resistance. The intended patient population is also not clearly defined for rTMS. The guidelines acknowledge conflicting evidence regarding the comparative effectiveness of rTMS and ECT and are not explicit about whether rTMS should be considered a second-line treatment after medication or only after ECT. Because of this ambiguity, the guidelines were considered to be of fair quality.

An update of "CANMAT Guidelines for the Management of patients with Bipolar Disorder" was published in 2013 (Yatham et al., 2013). The algorithm for management of bipolar I depression lists ECT

as a consideration after 3 trials of pharmacotherapy, and the list of recommendations for treatment of bipolar II depression lists adjunctive ECT as a possible third-line treatment. However, the 2013 version of the guidelines do not reference or discuss evidence related to ECT in bipolar patients. Because of the lack of any explicit discussion of evidence regarding ECT, these guidelines were considered to be of poor quality for the purpose of this report. This publication provides no guidance on TMS, tDCS, or DBS.

Institute for Clinical Systems Improvement (ICSI)

The ICSI recently published good-quality guidelines for the diagnosis, assessment, and management of adults with major depression (Trangle et al., 2013). AD medications and/or referral for psychotherapy are recommended as treatment for major depression. If patients are not responding to treatment, clinicians are advised to assess the dose, duration, type, and compliance. The diagnosis may need to be reconsidered should the impact of comorbidities. TRD is defined as lack of remission after 3 different classes of AD medications at adequate duration (undefined) and dosage. The guideline defines remission as the absence of depressive symptoms, or the presence of minimal depressive symptoms such as a Hamilton Depression Ration Scale (HAM-D) score < 7 or a Patient Health Questionnaire (PHQ-9) score < 5 and complete response as ≥ 50% reduction in symptoms (as measured on a standardized rating scale).

Recommended interventions for TRD are: medication augmentation strategies, hospitalization, ECT, and phototherapy. ECT may be indicated in cases of geriatric depression, intolerance of AD medication, unsuccessful trials of AD medications, catatonia, previous successful treatment with ECT, need for rapid response, depression with psychoses, and predominant melancholic symptoms.

National Institute for Health and Care Excellence (NICE)

NICE has published 1 guideline relevant to the scope of this report. This guideline makes recommendations regarding the diagnosis, treatment, and management of depression as a primary diagnosis in adults (NICE, 2009). The quality of this guideline was rated as good. NICE recommends combination therapy with AD medication and cognitive-behavioral therapy (CBT) for patients who have not responded to AD medication or any form of psychotherapy. ECT is recommended for patients with moderate-to-severe depression that has not resolved after trials of multiple (number unspecified) AD medications and psychological treatments. NICE recommends against the routine use of ECT for moderate depression. NICE suggests that the use of rTMS be reserved for investigational use only. Although there are no major safety concerns associated with rTMS, there is uncertainty about the procedure's clinical efficacy.

Veterans Affairs and Department of Defense (VA/DoD*)*

The VA/DoD published good-quality guidelines for the best care of adults with MDD (VA/DoD, 2009). The recommended initial treatment for moderate depression is AD medication and/or psychotherapy. The recommended initial treatment for severe depression is combination pharmacotherapy and psychotherapy or multiple drug therapy. Patients who have not responded to pharmacotherapy with a single agent after 8 to 12 weeks may receive combination treatment with pharmacotherapy and either CBT or interpersonal therapy (IPT). For patients who have not responded to 2 first-line ADs, the VA/DoD recommends a trial of a new AD from a different class (venlafaxine is recommended, if not already tried) or augmentation with either medication or psychotherapy. Patients who have not responded to 3 different classes of ADs should either receive augmentation with other psychotropic medications or psychotherapy or receive combination AD treatment or ECT. The guidelines also support the use of ECT

as a first-line treatment for pregnant women, patients with psychotic depression, catatonic patients, and patients who have severe self-neglect issues. rTMS was not considered because of the lack of FDA approval at the time the guideline was last updated.

Selected Payer Policies and Policy Guidance

Aetna

Aetna considers electroconvulsive therapy (ECT) medically necessary for members diagnosed with unipolar, bipolar, or mixed episode major depression. More than 20 sessions of ECT in a treatment series is considered to be rarely medically necessary. Members are eligible for ECT if they are ≥ 12 years of age and meet 1 of the following criteria:

- Member is unresponsive to effective medication, given for an adequate dose and duration, that are indicated for major depression; or
- Member is unable to tolerate effective medications or has a medical condition for which medication is contraindicated; or
- Member has had a favorable response to ECT in the past; or
- Member is unable to safely wait until medication is effective; or
- Member is experiencing severe mania or depression during pregnancy; or
- Member prefers ECT as a treatment option in consultation with the psychiatrist.

See Electroconvulsive Therapy: Aetna Clinical Policy Bulletin No. 0445.

Aetna considers rTMS experimental and investigational because its value and effectiveness have not been established. TMS is not covered for any indication, including depression.

See Transcranial Magnetic Stimulation and Cranial Electrical Stimulation: Aetna Clinical Policy Bulletin No. 0469.

Aetna considers DBS experimental and investigational for depression because there is insufficient evidence to support its effectiveness for this indication.

See Deep Brain Stimulation: Aetna Clinical Policy Bulletin No. 0208.

No coverage policy for transcranial direct current stimulation (tDCS) was identified on the Aetna website on September 17, 2013 (searched *direct current stimulation*) (Aetna Clinical Policy Bulletins).

Centers for Medicare & Medicaid Services (CMS)

No CMS National Coverage Determinations (NCDs) were identified for ECT, rTMS, tDCS, or DBS for the treatment of depression on September 17, 2013 (search National Coverage Documents, National Coverage Determinations, by keywords *electroconvulsive therapy*, *transcranial magnetic stimulation*, *direct current stimulation*, or *deep brain stimulation* and in entire document at: CMS Advanced Search Database). In the absence of an NCD, coverage decisions are left to the discretion of local Medicare carriers. CMS has a noncoverage policy for multiple-seizure ECT (M-ECT), in which patients undergo repeated stimulations designed to induce multiple adequate seizures in the same session rather than a single seizure, as with conventional ECT.

See Decision Memo for Multiple-Seizure ELECTROCONVULSIVE THERAPY: CMS Decision Memo (CAG-00134N).

GroupHealth

GroupHealth does not cover rTMS for TRD because there is insufficient evidence in the published medical literature to show that rTMS is as safe as standard therapies and/or provides better long-term outcomes than current standard therapies.

See Repetitive Transcranial Magnetic Stimulation (rTMS) for Patients with Treatment-Resistant Major Depression: Group Health Clinical Review Criteria.

No policies for ECT, tDCS, or DBS for the treatment of depression were identified on the GroupHealth website on September 17, 2013 (searched by keywords *electroconvulsive therapy*, *direct current stimulation*, *or deep brain stimulation* (Group Health Clinical Guidelines).

New England Comparative Effectiveness Public Advisory Council (CEPAC)

The New England CEPAC provides guidance on the use of comparative effectiveness research in the region. CEPAC partners with the Institute for Clinical and Economic Review (ICER) to promote adaptation and dissemination of federally-produced comparative effectiveness information. CEPAC's mission is to produce actionable information to aid regional policymakers in the medical policy decision-making process (New England CEPAC, 2013b). ICER prepared a supplementary report to the 2011 AHRQ systematic review of *Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults* (Gaynes et al., 2011). The supplementary report reviewed evidence published after the AHRQ report and included no new RCTs. As part of its report, ICER conducted a budget impact analysis for the region to assess the economic effect of coverage by Medicaid and the 3 largest private payers. ICER also conducted a cost-effectiveness analysis. Both the AHRQ report and the ICER report were discussed at a December 11, 2011 meeting, after which the Council agreed by vote on the following conclusions (Emond et al., 2013; New England CEPAC, 2013a):

- The evidence was adequate to demonstrate that rTMS was as good as or better than usual care for the treatment of patients with TRD.
- The evidence was adequate to demonstrate that rTMS has outcomes that are equivalent or superior to those of ECT.
- The use of rTMS represented "reasonable value" when compared with usual care and a "low value" when compared with ECT.
- The evidence was *inadequate* to demonstrate that ECT was equivalent or superior to usual care.

  Usual care was defined as general supportive psychotherapy with or without continued use of AD medication.

Following the CEPAC meeting, the Medicare Administrative Contractor for most of New England reversed its non-coverage decision and became the first player in the nation to cover rTMS, according to a summary of the meeting (Emond et al., 2013).

Oregon Health Evidence Review Commission (HERC)

The Oregon HERC has concluded that rTMS and ECT should be covered for patients with an episode of MDD who have failed at least 2 pharmacologic treatments. This conclusion is based on a review of a

2011 Evidence Report prepared for the Agency for Healthcare Research and Quality (AHRQ) and a 2009 systematic review by the Department of Veterans Affairs. Coverage guidance decisions by HERC are intended to guide public and private purchasers in Oregon in making informed decisions about healthcare services.

> *See* Non-pharmacologic Interventions for Treatment Resistant Depression: Health Evidence Review Commission Coverage Guidance.

No recommendations concerning tDCS or DBS were identified (searched by keyword *transcranial direct current stimulation* and keyword *deep brain stimulation*) (Health Evidence Review Commission).

Regence BCBS

Regence considers rTMS investigational as a treatment for all indications, including depression. Although evidence regarding rTMS compared with sham demonstrates significant improvement in depression, the clinical significance of these findings to change health outcomes has not been demonstrated.

> *See* Transcranial Magnetic Stimulation as a Treatment of Depression and Other disorders: Regence Group Medical Policy No. 148.

Regence considers DBS investigational for depression and bipolar disorder. There is insufficient evidence to determine the safety and efficacy of DBS for conditions other than Parkinson's disease, essential tremor, or primary dystonias.

> *See* Deep Brain Stimulation: Regence Group Medical Policy No. 84.

No policy on ECT or tDCS stimulation was identified on the Regence website on September 17, 2013 (searched by keyword *electroconvulsive therapy* and keyword *direct current stimulation*) (Regence Group Medical Policies).

References

Aas IH. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. *Ann Gen Psychiatry*. 2010;9:20.

AGREE Enterprise. Appraisal of Guidelines for Research and Evaluation (AGREE) II. 2009. Available at: http://www.agreetrust.org/. Accessed July 15, 2013.

Aleman A. Use of repetitive transcranial magnetic stimulation for treatment in psychiatry. *Clin Psychopharmacol Neurosci*. 2013;11(2):53-59.

Al-Harbi KS. Treatment-resistant depression: therapeutic trends, challenges, and future directions. *Patient Prefer Adherence*. 2012;6:369-388.

AllPsych Online. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV)*. 2013. Available at: http://allpsych.com/disorders/dsm.html. Accessed November 11, 2013.

American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*. Washington, DC: American Psychiatric Association;1994.

American Psychiatric Association (APA). DSM-5 Development. Highlights of Changes from DSM-IV-TR to DSM-5. 2013. Available at: http://www.dsm5.org/Pages/Default.aspx. Accessed August 18, 2013.

American Psychiatric Association (APA). Practice Guideline for the Treatment of Patients with Bipolar Disorder, Second Edition. April 2002. Psychiatryonline [website]. Available at: http://psychiatryonline.org/content.aspx?bookid=28&sectionid=1669577. Accessed October 31, 2013.

American Psychiatric Association (APA). Guideline Watch: Practice Guideline for the Treatment of Patients with Bipolar Disorder, Second Edition. 2005. Available at: http://psychiatryonline.org/content.aspx?bookid=28&sectionid=1682557. Accessed October 31, 2013.

American Psychiatric Association (APA). Online Assessment Measures. 2012. Psychiatry.org [website]. Available at: http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures. Accessed October 30, 2013.

American Psychiatric Association (APA). Practice Guideline for the Treatment of Patients with Major Depressive Disorder, Third Edition. 2010. Psychiatryonline [website]. Available at: http://psychiatryonline.org/content.aspx?bookid=28&sectionid=1667485. Accessed November 11, 2013.

American Thoracic Society (ATS). Hamilton Rating Scale for Depression (HRSD). 2013. Available at: http://www.thoracic.org/assemblies/srn/questionaires/hrsd.php. Accessed October 31, 2013.

Appleby BS, Duggan PS, Regenberg A, Rabins PV. Psychiatric and neuropsychiatric adverse events associated with deep brain stimulation: a meta-analysis of ten years' experience. *Mov Disord*. 2007;22(12):1722-1728.

Avery DH, Isenberg KE, Sampson SM, et al. Transcranial magnetic stimulation in the acute treatment of major depressive disorder: clinical response in an open-label extension trial. *J Clin Psychiatry*. 2008;69(3):441-451.

Bergfeld I, Mantione M, Hoogendoorn ML, Denys D. Cognitive functioning in psychiatric disorders following deep brain stimulation. [References]. *Brain Stimulat*. 2013;6(4):532-537.

Berlim MT, Turecki G. What is the meaning of treatment resistant/refractory major depression (TRD)? A systematic review of current randomized trials. *Eur Neuropsychopharmacol*. 2007;17(11):696-707.

Berlim M, Van den Eynde F, Daskalakis Z. Clinically meaningful efficacy and acceptability of low-frequency repetitive transcranial magnetic stimulation (rTMS) for treating primary major depression: a meta-analysis of randomized, double-blind and sham-controlled trials. *Neuropsychopharmacology*. 2013a;38(4):543-551.

Berlim M, Van den Eynde F, Daskalakis Z. High-frequency repetitive transcranial magnetic stimulation accelerates and enhances the clinical response to antidepressants in major depression: a meta-analysis of randomized, double-blind, and sham-controlled trials. *J Clin Psychiatry*. 2013b;74(2).

Berlim MT, Van den Eynde F, Daskalakis ZJ. Clinical utility of transcranial direct current stimulation (tDCS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. *J Psychiatr Res*. 2013c;47(1):1-7.

Bewernick BH, Hurlemann R, Matusch A, et al. Nucleus accumbens deep brain stimulation decreases ratings of depression and anxiety in treatment-resistant depression. *Biol Psychiatry.* 2010;67(2):110-116.

Bewernick BH, Kayser S, Sturm V, Schlaepfer TE. Long-term effects of nucleus accumbens deep brain stimulation in treatment-resistant depression: evidence for sustained efficacy. *Neuropsychopharmacology.* 2012;37(9):1975-1985.

Blumberger DM, Mulsant BH, Fitzgerald PB, et al. A randomized double-blind sham-controlled comparison of unilateral and bilateral repetitive transcranial magnetic stimulation for treatment-resistant major depression. *World J Biol Psychiatry*. 2012;13(6):423-435.

Boggio PS, Bermpohl F, Vergara AO, et al. Go-no-go task performance improvement after anodal transcranial DC stimulation of the left dorsolateral prefrontal cortex in major depression. *J Affect Disord*. 2007;101(1-3):91-98.

Bretlau LG, Lunde M, Lindberg L, Undén M, Dissing S, Bech P. Repetitive transcranial magnetic stimulation (rTMS) in combination with escitalopram in patients with treatment-resistant major depression: a double-blind, randomised, sham-controlled trial. *Pharmacopsychiatry*. 2008;41(2):41-47.

Brunoni AR, Amadera J, Berbel B, Volz MS, Rizzerio BG, Fregni F. A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. *Int J Neuropsychopharmacol*. 2011a;14(8):1133-1145.

Brunoni AR, Ferrucci R, Bortolomasi M, et al. Interactions between transcranial direct current stimulation (tDCS) and pharmacological interventions in the Major Depressive Episode: findings from a naturalistic study. *Eur Psychiatry*. 2013a;28(6):356-361.

Brunoni AR, Ferrucci R, Bortolomasi M, et al. Transcranial direct current stimulation (tDCS) in unipolar vs. bipolar depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2011b;35(1):96-101.

Brunoni AR, Valiengo L, Baccaro A, et al. The sertraline vs. electrical current therapy for treating depression clinical study: results from a factorial, randomized, controlled trial. *JAMA Psychiatry Chic Ill*. 2013b;70(4):383-391.

Centers for Disease Control and Prevention (CDC). QuickStats: Prevalence of Current Depression* Among Persons Aged ≥12 Years, by Age Group and Sex—United States, National Health and Nutrition Examination Survey, 2007-2010. *MMWR Morb Mortal Wkly Rep*. 2012;60(51);1747. Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6051a7.htm?s_cid=mm6051a7_w. Accessed November 18, 2013.

Chistyakov AV, Kaplan B, Rubichek O, et al. Effect of electroconvulsive therapy on cortical excitability in patients with major depression: a transcranial magnetic stimulation study. *Clin Neurophysiol*. 2005;116(2):386-392.

Cusin C, Yang H, Yeung A, Fava M. Chapter 2: Rating Scales for Depression. In: Baer L, Blais MS, eds. *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*. 2010. Available at: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=17&ved=0CFsQFjAGOAo&url=http%3A%2F%2Fwww.springer.com%2Fcda%2Fcontent%2Fdocument%2Fcda_downloaddocument%2F9781588299666-c1.pdf%3FSGWID%3D0-0-45-802906-p173751812&ei=QX1yUsXdOpaw4AOY8YHYAg&ug. Accessed October 31, 2013.

Dell'Osso B, Zanoni S, Ferrucci R, et al. Transcranial direct current stimulation for the outpatient treatment of poor-responder depressed patients. *Eur Psychiatry*. 2012;27(7):513-517.

Dierckx B, Heijnen WT, van den Broek WW, Birkenhäger TK. Efficacy of electroconvulsive therapy in bipolar versus unipolar major depression: a meta-analysis. *Bipolar Disord*. 2012;14(2):146-150.

Dunne RA, McLoughlin DM. Systematic review and meta-analysis of bifrontal electroconvulsive therapy versus bilateral and unilateral electroconvulsive therapy in depression. *World J Biol Psychiatry*. 2012;13(4):248-258.

Duits A, Temel Y, Ackermans L, Visser-Vandewalle V. The cognitive safety of deep brain stimulation in refractory psychiatric disorders. *Behav Neurol*. 2013;26(3):195-197.

Emond SK, Ollendorf DA, Colby JA, Reed SJ, Pearson SD. Evaluating the evidence on comparative effectiveness and value of management options for treatment-resistant depression. *Postgrad Med*. 2013;125(6):7-16.

Eranti S, Mogg A, Pluck G, et al. A randomized, controlled trial with 6-month follow-up of repetitive transcranial magnetic stimulation and electroconvulsive therapy for severe depression. *Am J Psychiatry*. 2007;164(1):73-81.

Fekadu A, Wooderson SC, Markopoulou K, Cleare AJ. The Maudsley Staging Method for treatment-resistant depression: prediction of longer-term outcome and persistence of symptoms. *J Clin Psychiatry*. 2009;70(7):952-957.

Ferrucci R, Bortolomasi M, Brunoni A, et al. Comparative benefits of transcranial direct current stimulation (TDCS) treatment in patients with mild/moderate vs. severe depression. *Clin Neuropsychiatry*. 2009:246-251.

Fitzgerald PB, Hoy KE, Herring SE, et al. A double blind randomized trial of unilateral left and bilateral prefrontal cortex transcranial magnetic stimulation in treatment resistant major depression. *J Affect Disord*. 2012;139(2):193-198.

Fitzgerald PB, Hoy KE, Singh A, et al. Equivalent beneficial effects of unilateral and bilateral prefrontal cortex transcranial magnetic stimulation in a large randomized trial in treatment-resistant major depression. *Int J Neuropsychopharmacol*. 2013;16(9):1975-1984.

Fountoulakis KN. Refractoriness in bipolar disorder: definitions and evidence-based treatment. *CNS Neurosci Ther*. 2012;18(3):227-237.

Folkerts HW, Michael N, Tölle R, Schonauer K, Mücke S, Schulze-Mönking H. Electroconvulsive therapy vs. paroxetine in treatment-resistant depression -- a randomized study. *Acta Psychiatr Scand*. 1997;96(5):334-342.

Fregni F, Boggio PS, Nitsche MA, Marcolin MA, Rigonatti SP, Pascual-Leone A. Treatment of major depression with transcranial direct current stimulation. *Bipolar Disord*. 2006a;8(2):203-204.

Fregni F, Boggio PS, Nitsche MA, Rigonatti SP, Pascual-Leone A. Cognitive effects of repeated sessions of transcranial direct current stimulation in patients with depression. *Depress Anxiety*. 2006b;23(8):482-484.

Fregni F, Marcolin MA, Myczkowski M, et al. Predictors of antidepressant response in clinical trials of transcranial magnetic stimulation. *Int J Neuropsychopharmacol*. 2006c;9(6):641-654.

Galletly C, Gill S, Clarke P, Burton C, Fitzgerald PB. A randomized trial comparing repetitive transcranial magnetic stimulation given 3 days/week and 5 days/week for the treatment of major depression: is efficacy related to the duration of treatment or the number of treatments? *Psychol Med*. 2012;42(5):981-988.

Garcia-Toro M, Salva J, Daumal J, et al. High (20-Hz) and low (1-Hz) frequency transcranial magnetic stimulation as adjuvant treatment in medication-resistant depression. *Psychiatry Res*. 2006;146(1):53-57.

Gardner BK, O'Connor DW. A review of the cognitive effects of electroconvulsive therapy in older adults. *J ECT*. 2008;24(1):68-80.

Gaynes BN, Lux L, Lloyd S et al. Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults. Comparative Effectiveness Review No. 33. (Prepared by RTI International-University of North Carolina (RTI-UNC) Evidence-based Practice Center under Contract No. 290-02-0016I.) AHRQ Publication No. 11-EHC056-EF. Rockville, MD: Agency for Healthcare Research and Quality. September 2011. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/76/787/Treatment-Resistant-Depression_executivesummary.pdf. Accessed November 8, 2013.

Gaynes BN, Rush AJ, Trivedi MH, Wisniewski SR, Spencer D, Fava M. The STAR*D study: treating depression in the real world. *Cleve Clin J Med*. 2008;75(1):57-66.

George MS, Lisanby SH, Avery D, et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch Gen Psychiatry*. 2010;67(5):507-516.

Goodman WK, Alterman RL. Deep brain stimulation for intractable psychiatric disorders. *Annu Rev Med*. 2012;63:511-524.

Greenhalgh J, Knight C, Hind D, Beverley C, Walters S. Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies. *Health Technol Assess Winch Engl*. 2005;9(9):1-156, iii-iv.

Grubert C, Hurlemann R, Bewernick BH, et al. Neuropsychological safety of nucleus accumbens deep brain stimulation for major depression: effects of 12-month stimulation. *World J Biol Psychiatry.* 2011;12(7):516-527.

Grunhaus L, Dannon PN, Schreiber S, et al. Repetitive transcranial magnetic stimulation is as effective as electroconvulsive therapy in the treatment of nondelusional major depressive disorder: an open study. *Biol Psychiatry*. 2000;47(4):314-324.

Grunhaus L, Schreiber S, Dolberg OT, Polak D, Dannon PN. A randomized controlled comparison of electroconvulsive therapy and repetitive transcranial magnetic stimulation in severe and resistant nonpsychotic major depression. *Biol Psychiatry*. 2003;53(4):324-331.

Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62.

Hansen PE, Ravnkilde B, Videbech P, et al. Low-frequency repetitive transcranial magnetic stimulation inferior to electroconvulsive therapy in treating depression. *J ECT*. 2011;27(1):26-32.

Hayes, Inc. Hayes Medical Technology Directory. *Deep Brain Stimulation for Treatment-Resistant Depression*. Lansdale, PA: Hayes, Inc.; November 5, 2012.

Hazari H., Christmas D., Matthews K. The clinical utility of different quantitative methods for measuring treatment resistance in major depression. *J Affect Disord*. 2013:231-236.

Heijnen WT, Birkenhäger TK, Wierdsma AI, van den Broek WW. Antidepressant Pharmacotherapy failure and response to subsequent electroconvulsive therapy: a meta-analysis. *J Clin Psychopharmacol*. 2010;30(5):616-619.

Herrmann LL, Ebmeier KP. Factors modifying the efficacy of transcranial magnetic stimulation in the treatment of depression: a review. *J Clin Psychiatry*. 2006;67(12):1870-1876.

Hirschfeld RM. *Guideline Watch: Practice Guideline for the Treatment of Patients with Bipolar Disorder*. 2nd Ed. Arlington, VA: American Psychiatric Association; 2005. Available at: http://psychiatryonline.org/content.aspx?bookid=28&sectionid=1682557. Accessed October 31, 2013.

Holtzheimer PE 3rd, Russo J, Claypoole KH, Roy-Byrne P, Avery DH. Shorter duration of depressive episode may predict response to repetitive transcranial magnetic stimulation. *Depress Anxiety*. 2004;19(1):24-30.

Huang M, Luo B, Hu J, et al. Repetitive transcranial magnetic stimulation in combination with citalopram in young patients with first-episode major depressive disorder: a double-blind, randomized, sham-controlled trial. *Aust N Z J Psychiatry*. 2012;46(3):257-264.

Hung CI, Wang SJ, Liu CY. Validation of the Depression and Somatic Symptoms Scale by comparison with the Short Form 36 scale among psychiatric outpatients with major depressive disorder. *Depress Anxiety*. 2009;26(6):583-591.

Institute for Clinical Systems Improvement (ICSI). May 2012. Major Depression in Adults in Primary Care Guideline. July 11, 2013. Available at: https://www.icsi.org/_asset/lfhfzf/DeprES0512.pdf. Accessed July 11, 2013.

Isenberg K, Downs D, Pierce K, et al. Low frequency rTMS stimulation of the right frontal cortex is as effective as high frequency rTMS stimulation of the left frontal cortex for antidepressant-free, treatment-resistant depressed patients. *Ann Clin Psychiatry*. 2005;17(3):153-159.

Janicak PG, O'Reardon JP, Sampson SM, et al. Transcranial magnetic stimulation in the treatment of major depressive disorder: a comprehensive summary of safety experience from acute exposure, extended exposure, and during reintroduction treatment. *J Clin Psychiatry*. 2008;69(2):222-232.

Johnstone EC, Deakin JF, Lawler P, et al. The Northwick Park electroconvulsive therapy trial. *Lancet*. 1980;2(8208-8209):1317-1320.

Jones SH, Thornicroft G, Coffey M, Dunn G. A brief mental health outcome scale-reliability and validity of the Global Assessment of Functioning (GAF). *Br J Psychiatry*. 1995;166(5):654-659.

Jorge RE, Moser DJ, Acion L, Robinson RG. Treatment of vascular depression using repetitive transcranial magnetic stimulation. *Arch Gen Psychiatry*. 2008;65(3):268-276.

Kalu UG, Sexton CE, Loo CK, Ebmeier KP. Transcranial direct current stimulation in the treatment of major depression: a meta-analysis. *Psychol Med*. 2012;42(9):1791-1800.

Kayser S, Bewernick BH, Grubert C, Hadrysiewicz BL, Axmacher N, Schlaepfer TE. Antidepressant effects, of magnetic seizure therapy and electroconvulsive therapy, in treatment-resistant depression. *J Psychiatr Res*. 2011;45(5):569-576.

Keitner GI, Mansfield AK. Management of treatment-resistant depression. *Psychiatr Clin North Am*. 2012;35(1):249-265.

Kennedy SH, Milev R, Giacobbe P, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. IV. Neurostimulation therapies. *J Affect Disord*. 2009;117(Suppl 1):S44-53.

Keshtkar M, Ghanizadeh A, Firoozabadi A. Repetitive transcranial magnetic stimulation versus electroconvulsive therapy for the treatment of major depressive disorder, a randomized controlled clinical trial. *J ECT*. 2011;27(4):310-314.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005a;62(6):593-602.

Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005b;62(6):617-627.

Knapp M, Romeo R, Mogg A, et al. Cost-effectiveness of transcranial magnetic stimulation vs. electroconvulsive therapy for severe depression: a multi-centre randomised controlled trial. *J Affect Disord*. 2008;109(3):273-285.

Kozel FA, George MS, Simpson KN. Decision analysis of the cost-effectiveness of repetitive transcranial magnetic stimulation versus electroconvulsive therapy for treatment of nonpsychotic severe depression. *CNS Spectr*. 2004;9(6):476-482.

Loo CK, Alonzo A, Martin D, Mitchell PB, Galvez V, Sachdev P. Transcranial direct current stimulation for depression: 3-week, randomised, sham-controlled trial. *Br J Psychiatry J Ment Sci*. 2012;200(1):52-59.

Loo CK, Sachdev P, Martin D, et al. A double-blind, sham-controlled trial of transcranial direct current stimulation for the treatment of depression. *Int J Neuropsychopharmacol*. 2010;13(1):61-69.

Lozano AM, Mayberg HS, Giacobbe P, Hamani C, Craddock RC, Kennedy SH. Subcallosal cingulate gyrus deep brain stimulation for treatment-resistant depression. *Biol Psychiatry*. 2008;64(6):461-467.

Lozano AM, Giacobbe P, Hamani C, et al. A multicenter pilot study of subcallosal cingulate area deep brain stimulation for treatment-resistant depression. *J Neurosurg*. 2012;116(2):315-322.

Malone DA Jr. Use of deep brain stimulation in treatment-resistant depression. *Cleve Clin J Med*. 2010;(77 Suppl 3):S77-S80.

Malone DA Jr, Dougherty DD, Rezai AR, et al. Deep brain stimulation of the ventral capsule/ventral striatum for treatmentresistant depression. *Biol Psychiatry.* 2009;65(4):267-275.

Mantovani A, Pavlicova M, Avery D, et al. Long-term efficacy of repeated daily prefrontal transcranial magnetic stimulation (TMS) in treatment-resistant depression. *Depress Anxiety*. 2012;29(10):883-890.

Mathys M, Mitchell BG. Targeting treatment-resistant depression. *J Pharm Pract*. 2011;24(6):520-533.

Mayo Clinic. Electroconvulsive therapy (ECT). May 2012. Available at: http://www.mayoclinic.com/health/electroconvulsive-therapy/MY00129. Accessed November 11, 2013.

Mayur P, Byth K, Harris A. Acute antidepressant effects of right unilateral ultra-brief ECT: a double-blind randomised controlled trial. *J Affect Disord*. 2013:1-3.

McDonald WM. Is ECT cost-effective? A critique of the National Institute of Health and Clinical Excellence's report on the economic analysis of ECT. *J ECT*. 2006;22(1):25-29.

McLoughlin DM, Mogg A, Eranti S, et al. The clinical effectiveness and cost of repetitive transcranial magnetic stimulation versus electroconvulsive therapy in severe depression: a multicentre pragmatic randomised controlled trial and economic analysis. *Health Technol Assess.* 2007;11(24):1-54.

National Alliance on Mental Illness (NAMI). What is depression? 2013. Available at: http://www.nami.org/template.cfm?section=Depression. Accessed November 11, 2013.

National Institute for Health and Care Excellence (NICE). *Depression in adults. The treatment and management of depression in adults*. London, UK: National Institute for Health and Care Excellence; 2009. NICE Clinical Guideline No. 90. Available at: http://www.nice.org.uk/nicemedia/live/12329/45888/45888.pdf. Accessed November 11, 2013.

National Institute of Mental Health (NIMH). Questions and Answers about the NIMH Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study. January 2006. Available at: http://www.nimh.nih.gov/health/trials/practical/stard/backgroundstudy.shtml. Accessed November 11, 2013.

National Institute of Mental Health (NIMH). What Is Depression? Revised 2011. NIH Publication No. 11-3561. Available at: http://www.nimh.nih.gov/health/publications/depression/index.shtml. Accessed November 11, 2013.

Nemeroff CB. Prevalence and management of treatment-resistant depression. *J Clin Psychiatry*. 2007;68(Suppl 8):17-25.

New England Comparative Effectiveness Public Advisory Council (CEPAC). New England Comparative Effectiveness Public Advisory Council. 2013a. Available at: http://cepac.icer-review.org/. Accessed November 22, 2013.

New England Comparative Effectiveness Public Advisory Council (CEPAC). The New England Comparative Effectiveness Public Advisory Council Public Meeting. December 9, 2011. 2013b. Available at: http://cepac.icer-review.org/wp-content/uploads/2011/04/Final-Report-TRD_FINAL2.pdf. Accessed November 22, 2013.

Nordenskjöld A, von Knorring L, Ljung T, Carlborg A, Brus O, Engström I. Continuation electroconvulsive therapy with pharmacotherapy versus pharmacotherapy alone for prevention of relapse of depression: a randomized controlled trial. *J ECT*. 2013;29(2):86-92.

O'Reardon JP, Solvason HB, Janicak PG, et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry*. 2007;62(11):1208-1216.

Padberg F, Zwanzger P, Thoma H, et al. Repetitive transcranial magnetic stimulation (rTMS) in pharmacotherapy-refractory major depression: comparative study of fast, slow and sham rTMS. *Psychiatry Res*. 1999;88(3):163-171.

Pallanti S, Bernardi S, Di Rollo A, Antonini S, Quercioli L. Unilateral low frequency versus sequential bilateral repetitive transcranial magnetic stimulation: is simpler better for treatment of resistant depression? *Neuroscience*. 2010;167(2):323-328.

Palm U, Schiller C, Fintescu Z, et al. Transcranial direct current stimulation in treatment resistant depression: a randomized double-blind, placebo-controlled study. *Brain Stimulat*. 2012;5(3):242-251.

Pridmore S. Substitution of rapid transcranial magnetic stimulation treatments for electroconvulsive therapy treatments in a course of electroconvulsive therapy. *Depress Anxiety*. 2000;12(3):118-123.

Psychiatric Times. Patient Health questionnaire, including PHQ, PHQ-9, PHQ-Brief, and PHQ-SADS. April 11, 2013. Available at: http://www.psychiatrictimes.com/clinical-scales-patient-health-questionnaire/patient-health-questionnaire. Accessed November 1, 2013.

Ray S, Nizamie SH, Akhtar S, Praharaj SK, Mishra BR, Zia-ul-Haq M. Efficacy of adjunctive high frequency repetitive transcranial magnetic stimulation of left prefrontal cortex in depression: a randomized sham controlled study. *J Affect Disord*. 2011;128(1-2):153-159.

RAND. 36-Item Short Form Survey from the RAND Medical Outcomes Study. 2013. Available at: http://www.rand.org/health/surveys_tools/mos/mos_core_36item.html. Accessed November 19, 2013.

Roepke S, Luborzewski A, Schindler F, et al. Stimulus pulse-frequency-dependent efficacy and cognitive adverse effects of ultrabrief-pulse electroconvulsive therapy in patients with major depression. *J ECT*. 2011;27(2):109-113.

Rosa MA, Gattaz WF, Pascual-Leone A, et al. Comparison of repetitive transcranial magnetic stimulation and electroconvulsive therapy in unipolar non-psychotic refractory depression: a randomized, single-blind study. *Int J Neuropsychopharmacol*. 2006;9(6):667-676.

Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905-1917.

Sackeim HA, Prudic J, Devanand DP, et al. A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Arch Gen Psychiatry*. 2000;57(5):425-434.

Sackeim HA, Prudic J, Devanand DP, et al. Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *N Engl J Med*. 1993;328(12):839-846.

Sackeim HA, Prudic J, Nobler MS, et al. Effects of pulse width and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *Brain Stimulat*. 2008;1(2):71-83.

SF-36.org. The SF Community - SF-36® Health Survey Update. 2013. Available at: http://www.sf-36.org/tools/SF36.shtml. Accessed November 20, 2013.

Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS One*. 2007;2(12):e1350.

Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. *Evid Policy J Res Debate Pract*. 2010;6(1):51-59.

Simpson KN, Welch MJ, Kozel FA, Demitrack MA, Nahas Z. Cost-effectiveness of transcranial magnetic stimulation in the treatment of major depression: a health economics analysis. *Adv Ther*. 2009;26(3):346-368.

Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). *Arthritis Care Res*. 2011;63(S11):S454-S466.

Sobin C, Prudic J, Devanand DP, Nobler MS, Sackeim HA. Who responds to electroconvulsive therapy? A comparison of effective and ineffective forms of treatment. *Br J Psychiatry*. 1996;169(3):322-328.

Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999;282(18):1737-1744.

Stern WM, Tormos JM, Press DZ, Pearlman C, Pascual-Leone A. Antidepressant effects of high and low frequency repetitive transcranial magnetic stimulation to the dorsolateral prefrontal cortex: a double-blind, randomized, placebo-controlled trial. *J Neuropsychiatry Clin Neurosci*. 2007;19(2):179-186.

Substance Abuse and Mental Health Services Administration (SAMHSA). Patient Health Questionnaire (PHQ-9). SAMHSA-HRSA Center for Integrated Health Solutions. October 4, 2005. Available at: http://www.integration.samhsa.gov/images/res/PHQ%20-%20Questions.pdf. Accessed November 1, 2013.

Su TP, Huang CC, Wei IH. Add-on rTMS for medication-resistant depression: a randomized, double-blind, sham-controlled trial in Chinese patients. *J Clin Psychiatry*. 2005;66(7):930-937.

Thase ME, Friedman ES, Biggs MM, et al. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. *Am J Psychiatry*. 2007;164(5):739-752.

Trangle M, Dieperink B, Gabert T, et al. Institute for Clinical Systems Improvement. Adult Depression in Primary Care. Updated September 2013. Available at: https://www.icsi.org/_asset/fnhdm3/Depr-Interactive0512b.pdf. Accessed November 11, 2013.

Triggs WJ, Ricciuti N, Ward HE, et al. Right and left dorsolateral pre-frontal rTMS treatment of refractory depression: a randomized, sham-controlled trial. *Psychiatry Res*. 2010;178(3):467-474.

Trivedi RB, Nieuwsma JA, Williams JW Jr. Examination of the utility of psychotherapy for patients with treatment resistant depression: a systematic review. *J Gen Intern Med*. 2011;26(6):643-650.

UK ECT Review Group. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. *Lancet*. 2003;361(9360):799-808.

Ullrich H, Kranaster L, Sigges E, Andrich J, Sartorius A. Ultra-high-frequency left prefrontal transcranial magnetic stimulation as augmentation in severely ill patients with depression: a naturalistic sham-controlled, double-blind, randomized trial. *Neuropsychobiology*. 2012;66(3):141-148.

Valiengo L, Benseñor IM, Goulart AC, et al. The sertraline versus electrical current therapy for treating depression clinical study (select-TDCS): results of the crossover and follow-up phases. *Depress Anxiety*. 2013;30(7):646-653.

van Schaik AM, Comijs HC, Sonnenberg CM, Beekman AT, Sienaert P, Stek ML. Efficacy and safety of continuation and maintenance electroconvulsive therapy in depressed elderly patients: a systematic review. *Am J Geriatr Psychiatry*. 2012;20(1):5-17.

Versiani M, Cheniaux E, Landeira-Fernandez J. Efficacy and safety of electroconvulsive therapy in the treatment of bipolar disorder: a systematic review. *J ECT*. 2011;27(2):153-164.

Verwijk E, Comijs HC, Kok RM, Spaans HP, Stek ML, Scherder EJA. Neurocognitive effects after brief pulse and ultrabrief pulse unilateral electroconvulsive therapy for major depression: a review. *J Affect Disord*. 2012;140(3):233-243.

Veterans Administration/Department of Defense (VA/DoD). Management of Major Depressive Disorder (MDD). VA/DoD Clinical Practice Guideline. May 2009. Available at: http://www.healthquality.va.gov/mdd/MDDFULL053013.pdf. Accessed July 21, 2013.

Volkmann J, Herzog J, Kopper F, Deuschl G. Introduction to the programming of deep brain stimulators. *Mov Disord*. 2002;17 Suppl 3:S181-187.

Ward MP, Irazoqui PP. Evolving refractory major depressive disorder diagnostic and treatment paradigms: toward closed-loop therapeutics. *Front Neuroeng*. 2010;3:7.

Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.

West ED. Electric convulsion therapy in depression: a double-blind controlled trial. *Br Med J Clin Res Ed*. 1981;282(6261):355-357.

Xia G, Gajwani P, Muzina DJ, et al. Treatment-emergent mania in unipolar and bipolar depression: focus on repetitive transcranial magnetic stimulation. *Int J Neuropsychopharmacol*. 2008;11(1):119-130.

Yatham LN, Kennedy SH, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipolar Disord*. 2013;15(1):1-44.

Zimmerman M, Martinez J, Attiullah N, et al. Further evidence that the cutoff to define remission on the 17-item Hamilton Depression Rating Scale should be lowered. *Depress Anxiety*. 2012;29(2):159-165.

APPENDICES

APPENDIX I. Definitions of Treatment-Resistant Depression (TRD)

Results of a Systematic Review of Definitions For Treatment-Resistant Depression (TRD) (Berlim and Turecki, 2007)

This review was designed to evaluate how the concept of TRD was applied in trials that enrolled patients with a primary diagnosis of unipolar major depressive disorder (MDD) that was resistant to treatment at the time of enrollment. The authors selected 47 randomized trials published in English in peer-reviewed journals after a search of the PubMed, PsycINFO, Cochrane Library, and Embase databases. Key findings include:

- Wide variety of terms for TRD: treatment-resistant (n=17), medication-resistant (n=10), resistant (n=5), refractory (n=4), treatment-refractory (n=3), therapy-resistant (n=2), drug-resistant (n=2), antidepressant-refractory (n=1), pharmacotherapy-resistant (n=1), antidepressant-resistant (n=1), and pharmacotherapy-refractory (n=1).
- 6 different definitions of TRD (% studies).
    - 1 previous antidepressant medication (AD) trial (10.6%).
    - ≥ 1 previous AD trial (17%).
    - 2 previous AD trials (12.8%).
    - 2 previous trials with different ADs (6.4%).
    - ≥ 2 previous AD trials (17%).
    - ≥ 2 previous AD trials with different ADs (17%).
- Typically (38 trials), no indication of whether the failed AD trials occurred during the current episode.
- Typically (37 trials), no information on how the adequacy of previous AD trials was evaluated.
    - 4 studies referred to systematic approaches: Antidepressant Treatment History Form (ATHF) (n=3), and the Harvard Antidepressant Treatment History (n=1).
- Typically (22 trials), no specification of the maximum doses at which previous AD trials were deemed to be failures, and where doses were mentioned, they were inconsistent.
- Different definitions of adequate duration of an AD trial:
    - Where described: ≥ 4 weeks (n=18), ≥ 6 weeks (n=8), ≥ 8 weeks (n=3).
    - 10 studies provided no definition.
    - No clarification of whether titration periods were considered in trial duration estimates.
- Inconsistent descriptions about what constituted failure and success in previous AD trials.
    - Inconsistent language (e.g., lack of response, failure, lack of clinical improvement).
    - ≤ 50% score reduction on the Hamilton Depression Rating Scale (HAM-D) (n=2). (This was by far the most common definition of lack of response in the studies selected for the current report.)
    - < 20% reduction in HAM-D score (n=1).
    - < 30% reduction in Montgomery-Åsberg Depression Rating Scale (MADRS) (n=1).
    - Maintenance of full DSM-IV criteria for an episode of MDD (n=1).

  - Scores of ≥ 3 and ≥ 4 on the Clinical Global Impression – Improvement Subscale as indicators of *minimal improvement* and *refractoriness* (n=2).
- No consideration of compliance or adherence.
- Typically (25 trials), use of the HAM-D for assessing enrollment eligibility. Next most common, MADRS (5 trials).

In addition to the need for researchers to standardize terminology and criteria, Berlim and Turecki (2007) concluded that their findings had the following implications:

- Diagnosis of the underlying MDD should entail structured interview instruments, rather than relying only on *Diagnostic and Statistical Manual of Mental Disorders* (DSM) checklists so that psychiatric comorbidity can be identified.
- A consensus is developing that TRD is defined by lack of improvement after 2 adequate trials of different classes of ADs.
- Ascertainment of resistance should be based on prospective follow-up to avoid patient recall bias. If retrospective data are used, instruments allowing systematic evaluation (e.g., the ATHF) should be used.
- A consensus is developing that adequate dose means the maximum tolerated dose.
- An adequate trial of AD should probably exceed 4 weeks, may require as long as 10 weeks, and may require 12 weeks in the elderly.
- Failure to achieve remission according to a standard definition, rather than varying definitions of response, should become the "gold standard" for defining failure of previous AD trials.
- Compliance should be considered a key component of TRD.

Tools for Staging TRD

Antidepressant Treatment History Form (ATHF)

A small number of typically older studies selected for this report used a scale called the Antidepressant Treatment History Form (ATHF) to rank medication resistance according to the adequacy of the most potent previous trial. According to the developers of the ATHF, adequate dose is equated with the minimal dosage at which RCTs have shown the agent to be effective in major depression, or two thirds of the maximal dose recommended in the *Physicians' Desk Reference (PDR)*. The ATHF guides the clinician to record detailed information, including dose, duration, and patient compliance. The scale extends from 1 to 5; levels 1, 3, and 5 are defined as follows (Sackeim, 2001; Sackeim et al., 2008):

> 1 = definitely inadequate due to insufficient dose and/or duration (clinician has no confidence in adequacy).
>
> 3 = trial meets threshold criteria for adequacy of dose and duration using established AD medication (clinician has moderate confidence; information largely from 1 reliable source).
>
> 5 = definitely adequate trial of sufficient dose and duration with established AD medication and an established augmentation strategy (clinician has high confidence based on excellent documentation of all relevant information).

The accompanying Instruction Guide includes duration and dosage guides for specific drugs, by class (Sackeim, 2001).

Maudsley Staging Method (MSM)

The MSM score ranges from 3 to 15, and categories of mild, moderate, and severe have been defined. Three domains are assessed: duration, symptom severity at baseline, and treatment failures. The intensity of treatment failure is scored according to the number of AD medications, whether augmentation was used, and whether electroconvulsive therapy (ECT) was used. An original validation study found that hospitalized patients with a higher MSM score were more likely to be symptomatic when discharged, rather than in remission (Fekadu et al., 2009).

Additional validation research has shown that the MSM is useful for predicting longer-term outcomes and persistence of symptoms (Fekadu et al., 2009). In a longitudinal study after discharge of 62 of the patients in the original validation sample for the MSM, MSM independently predicted duration of the episode as well as functional impairment, while AD count and the Thase and Rush system were not independent predictors of either depression persistence or functional impairment.

Massachusetts General Hospital Method

This method relies primarily on the number of prior AD failures but also takes intensity of treatment, augmentation strategies, and prior failed ECT treatment into account. It has been criticized for the lack of clear evidence supporting the weight attached to ECT failure (Fekadu et al., 2009). Multivariate analysis testing of the Massachusetts General Hospital method and the Thase and Rusch has shown the Massachusetts General Hospital method to be more predictive of nonremission (Berlim and Turecki, 2007).

Thase and Rush Scale

Some studies selected for this report described patients as having Stage II TRD on the Thase and Rush scale, which is a 5-part categorical scale. Stage II is equated with failure to achieve remission or inability to tolerate 2 trials of AD medication from separate classes used in an adequate dose for a sufficient period of time (Thase et al., 2007; Blumberger et al., 2012; Hazari et al., 2013).

An Analysis of the Clinical Utility for Different Staging Systems (Hazari Et Al., 2013)

This validation study determined scores according to a variety of treatment-resistance staging tools for patients who were undergoing treatment for depression in a variety of settings: primary care, a depression clinic, electroconvulsive therapy (ECT) in a secondary center, a tertiary care center, vagus nerve stimulation (VNS) treatment, and neurosurgical treatment. The tools used were the Thase and Rush criteria, the MSM, the Massachusetts General Hospital methods, and the ATHF system. All 4 tools clearly differentiated between patients receiving treatment in the different settings, with scores lowest for those patients being treated in primary care centers. The authors interpreted these findings as a validation of the tools' ability to differentiate patients according to the adequacy of their prior treatment and thus to provide a measure of treatment resistance. They found that some tools were easier to use than others but did not recommend any one tool over another in terms of accuracy. The ATHF tool was described as being intended primarily for use in research settings.

APPENDIX II. Outcome Measurement Tools

SYMPTOM SCALES

Brief Psychiatric Rating Scale (BPRS): An 18-item validated scale used to assess depressive and psychotic symptoms (Ray et al., 2011).

Beck Depression Inventory (BDI): This scale is considered the gold standard of self-rating scales for depression. It measures current symptom severities experienced by the patient. The BDI has 21 items that are scored as 0 to 3 by the patient for a maximum possible score of 63. BDI has undergone a few revisions and the most recent version is BDI-II, in which 4 items have been replaced to achieve better alignment with the DSM-IV criteria of MDD (Cusin et al., 2010). Domains on the scale include cognitive, affective, somatic, and vegetative symptoms of depression (ATS, 2013). Internal consistency of BDI is 0.86 and convergent validity is 0.72 with clinical ratings and 0.73 with HAM-D (Cusin et al., 2010). Suggested interpretations of BDI-II include (Smarr and Keefer, 2011):

- 0-13: no depression
- 14-19: mild depression
- 20-28: moderate depression
- > 29: severe depression

Link to BDI: http://www.thecommunityhouse.org/wp-content/uploads/2012/01/Beck-Depression-Inventory-and-Scoring-Key1.pdf

Hamilton Rating Scale for Depression (HAM-D): This 17-item scale is widely used in clinical studies and in clinical practice. Designed as a 21-item tool, 4 of the questions were subsequently dropped because they were uncommon or did not reflect depression severity. The test is administered as a structured interview by 1 or 2 clinician(s) who is/are skilled at extracting the desired information from the patient. Domains of the scale include mood, suicide, work, loss of interest, agitation, gastrointestinal symptoms, somatic symptoms, hypochondriasis, insight, and weight loss. Each item is measured on a 3- or 5-point scale and the score for each item is summed to obtain the final patient score, with a possible total score of 52 (Hamilton, 1960). The score ranges for other versions of the HAM-D are as follows: 0 to 64 (HAM-D_{21}), 0 to 75 (HAM-D_{24}), and 0 to 52 (HAM-D_{25}) (Gaynes et al., 2011).

The validity of the 17-item HAM-D has been established through comparisons with global measures of depression severity and with other depression scales, e.g., the Montgomery-Åsberg Depression Rating Scale (MADRS) (Cusin et al., 2010). The convergent validity and discriminant validity are adequate (ATS, 2013). The main limitations of the 17-item scale include: failure to include all symptom domains of MDD, presence of items measuring different constructs, and uneven weight attributed to different symptom domains (Cusin et al., 2010). The HAM-D score has been found to be significantly correlated with each of the 8 SF-36 subscales (Hung et al., 2009). (See description of the SF-36 Health Survey in the next section.)

Cutoff points for diagnostic conclusions have not been empirically derived but, rather, represent a consensus among clinicians. For the 17-point HAM-D scale, the maximum possible score is 54 and typically scores are interpreted as (Cusin et al., 2010):

- 0-6: no depression
- 7-17: mild depression

- 18-24: moderate depression
- > 24: severe depression

Link to an online presentation of the HAM-D_{21}: http://healthnet.umassmed.edu/mhealth/HAMD.pdf

Montgomery-Åsberg Depression Rating Scale (MADRS): This scale was designed to be sensitive to the effects of ADs, particularly tricyclic ADs. It is focused on the psychological aspects of depression. MADRS is administered by a clinician and is commonly used in clinical studies and clinical practice. This scale has never been modified and does not target reverse neurovegetative symptoms. MADRS has a very high internal consistency and has a high correlation with HAM-D. The scale is composed of 10 questions and the score for each item is summed to obtain the final patient score. The maximum possible score is 60 and scores of 10 or below indicate no depression (or remission) and scores greater than 30 (or sometimes 35) indicate severe depression (Cusin et al., 2010).

Link to the MADRS: www.sfaetc.ucsf.edu/docs/MADRS.pdf

Patient Health Questionnaire (PHQ): The PHQ is a self-administered scale that is designed for use in the primary care setting. PHQ includes anxiety, mood, alcohol, eating, and somatoform modules and has 16 questions. Shorter and less specific versions are available, including a 9-item scale (PHQ-9) and one that is focused on a common triad of symptoms (somatic, anxiety, and depressive) called PHQ-SADS (Psychiatric Times, 2013). Each question asks the patient how much he or she has been bothered by particular symptoms. The 3 available responses are: "not bothered," "bothered a little," or "bothered a lot." The final score is derived from diagnostic algorithms. Interpretation of PHQ-9 scores is as follows (SAMHSA, 2005):

- 1-4: minimal depression
- 5-9: mild depression
- 10-14: moderate depression
- 15-19: moderately severe depression
- 20-27: severe depression

PHQ has validity that is comparable to the larger, clinician-administered screening instrument Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer et al., 1999).

Link to the PHQ: http://www.pdhealth.mil/guidelines/downloads/appendix2.pdf

Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D): A validated guide for improving the reliability of the Hamilton Depression Rating Scale (Ray et al., 2011).

QUALITY OF LIFE AND FUNCTION/DISABILITY SCALES

Global Assessment of Functioning (GAF) Scale: In the chapter on Multiaxial Assessment, the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), advises use of the GAF scale for the clinician's assessment of a patient's overall functioning (Axis V). The GAF allows a rating of overall psychological functioning on a 0 to 100 scale. Example definitions of scale segments include (APA, 1994):

0: Inadequate information

1-10: Persistent danger of severely hurting self or others (e.g., recurrent violence) OR persistent inability to maintain minimal personal hygiene OR serious suicidal act with clear expectation of death.

11-20: Some danger of hurting self or others (e.g., suicide attempts without clear expectation of death, frequently violent; manic excitement) OR occasionally fails to maintain minimal personal hygiene (e.g., smears feces) OR gross impairment in communication (e.g., largely incoherent or mute).

71-80: If symptoms are present, they are transient and expectable reactions to psychosocial stressors (e.g., difficulty concentrating after family argument); no more than slight impairment in social, occupational, or school functioning (e.g., temporarily falling behind in schoolwork).

91-100: Superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought out by others because of his or her many positive qualities. No symptoms.

A GAF score ≤ 50 signifies severe symptoms and/or psychosocial dysfunction, scores of 51 to 60 are considered moderate scores, scores of 61 to 70 are considered mild scores, and scores ≥ 71 signify absent or only transient symptoms and/or minimal dysfunction (Holtzheimer et al., 2012). Although this scale is widely used, at least one expert reports that it is not well validated (Aas, 2010) but at least 1 validation study was identified ((Jones et al., 1995).

SF-36 Health Survey: The SF-36 was developed by the RAND Corporation as part of the Medical Outcomes Study (MOS), a multi-year, multi-site study designed to explain variations in patient outcomes (RAND, 2013). The SF-36 is a validated set of 36 generic, self-reported items organized into 8 domains, originally described as: 1) physical functioning; 2) role limitations because of physical health problems; 3) bodily pain; 4) social functioning; 5) general mental health (psychological distress and psychological well-being); 6) role limitations because of emotional problems; 7) vitality (energy/fatigue); and 8) general health perceptions. Most of the items were adapted from longstanding instruments (Ware and Sherbourne, 1992). The 8 scales can be combined into two summary measures—Physical Health (sometimes referred to as the Physical Component Score) and Mental Health (Mental Component Score). Version 2.0 of the instrument was released in 1996 (SF-36.org). Each scale is scored from 0 to 100 scale according to a norm-based algorithm, with 100 representing perfect health. The scale scores are averaged for an overall score (RAND, 2013; SF-36.org, 2013).

World Health Organization Disability Assessment Schedule (WHODAS): This generic disability scale is one of several "emerging measures" that have been featured in the new DSM-V and about which the American Psychiatric Association (APA) is soliciting clinician and researcher feedback. The WHODAS is intended for use in the initial patient interview and for monitoring treatment progress but is not intended to serve as the sole basis for a diagnosis. The adult self-administered version has 36 items. Patients rank each item according to a 5-point scale (1 = none to 5 = extreme) on the basis of how much difficulty the item has presented within the previous 30 days. The items address the following 6 domains: understanding and communicating, getting around, self-care, getting along with people, life activities (i.e., household, work, and/or school activities), and participation in society. The producers of the DSM have conducted some field trials for clinician use in the U.S. and Canada (APA, 2012).

APPENDIX III. Search Strategy

Exact search strategies, as used in PubMed, are presented here. Searches in Embase and PsycINFO were designed to match PubMed searches as closely as possible.

INITIAL SEARCH, SYSTEMATIC REVIEWS AND PRACTICE GUIDELINES (conducted July 8 to August 1, 2013)

Initially, evidence for this report was obtained by searching for systematic reviews and guidelines that had been published in the past *5 years*. Searches were conducted in the following databases: Agency for Healthcare Research and Quality (AHRQ), Blue Cross Blue Shield TEC Assessments, Canadian Agency for Drugs and Technology in Health (CADTH), Centre for Reviews and Dissemination (York University), Cochrane Library, Hayes Knowledge Center, Institute for Clinical Systems Improvement (ICSI), National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (UK), U.S. Preventive Services Task Force (USPSTF), National Guidelines Clearinghouse (NGC), National Institute for Health and Care Excellence (NICE), VA/Department of Defense Clinical Practice Guidelines, and VA Technology Assessment Program (VA TAP).

The websites for the American Psychiatric Association, American Psychological Association, and American College of Neuropsychopharmacology were searched for guidelines.

Additional systematic reviews were selected from a search of the PubMed database using filters for Practice Guidelines, Guidelines, Meta-analyses, and Systematic Reviews.

PRIMARY CLINICAL STUDIES PUBLISHED AFTER THE SYSTEMATIC REVIEWS

Databases Searched

PubMed

Embase

PsycINFO

PubMed Search Strategies for Primary Studies Published After the Search Time Frames of Selected Systematic Reviews (conducted August 1, 2013)

Electroconvulsive Therapy

1. Electroconvulsive therapy
2. ((treatment resistan*) OR refractory)) AND depression
3. ((randomized controlled trial[Publication Type] OR (randomized[Title/Abstract] AND controlled[Title/Abstract] AND trial[Title/Abstract]))
4. 1 and 2 and 3
5. Filters: Publication date from 2010/11/01 to 2013/12/31

Transcranial Magnetic Stimulation

1. repetitive transcranial magnetic stimulation
2. depression AND ((refractory OR (treatment resistan*))
3. ((randomized controlled trial[Publication Type] OR (randomized[Title/Abstract] AND controlled[Title/Abstract] AND trial[Title/Abstract]))
4. 1 and 2 and 3
5. Filters: Publication date from 2010/11/01 to 2013/12/31

Direct Current Stimulation

1. direct current stimulation
2. depression AND ((treatment resistan*) OR refractory))
3. 1 and 2
4. Filters: Publication date from 2011/05/01 to 2013/12/31

Deep Brain Stimulation

1. deep brain stimulation
2. depression AND ((treatment resistan*) OR refractory))
3. 1 and 2
4. Filters: Publication date from 2012/10/01 to 2013/12/31

Search for Studies in Patients with Bipolar Disorder (conducted August 4, 2013)

The purpose of this search was to identify studies that had been omitted from the AHRQ evidence review (Gaynes et al., 2011).

1. "Bipolar Disorder"[Mesh]
2. “bipolar disorder”
3. 1 and 2
4. (treatment resistan* or refractory)
5. 3 and 4

(No date filter. The previous search phrases for the treatment modalities of interest were used.

Studies Excluded from the AHRQ Evidence Review

The Excluded Studies list in the AHRQ evidence review (Gaynes et al., 2011) was reviewed for studies excluded because > 20% of the participants had bipolar depression and for studies that entailed a comparison of treatment parameters but were excluded because there was no sham or alternative treatment group.

Additional Searches for Primary Studies not Included in Selected Systematic Reviews (conducted October 12 to 13, 2013)

The previously described searches were found to have missed some eligible studies because of the search phrase *treatment resistan* or refractory*. The original searches were repeated, new searches without this search phrase were run, and the Boolean operator *NOT* was used to identify missed citations.

Searches for Additional Evidence Pertinent to Key Questions #1b (Treatment Parameters), #2 (Safety), and #3 (Differential Effectiveness) (conducted October 12 to 13 and October 27, 2013) – rTMS and ECT only

The initial searches for systematic reviews and meta-analyses that were conducted in the Centre for Reviews and Dissemination database and PubMed were repeated for an earlier time frame (2003 to 2008) on October 12 and 13 to identify reviews that might have included observational studies and addressed safety or differential effectiveness/safety.

Additionally, to augment the RCTs selected for rTMS and ECT, the following search string, borrowed from the 2011 AHRQ evidence review (Gaynes et al., 2011), was used to identify observational studies published since November 2010 that might have safety data for ≥ 100 patients, follow-up/maintenance treatment data, or differential treatment effect data (October 27):

"Longitudinal Studies"[Mesh] OR "Comparative Study "[Publication Type]) OR "Cohort Studies"[Mesh] OR "observational studies"[tw]

(This search could not be repeated in Embase or PsycINFO.)

Lastly, safety data from registries were sought using the following search phrase (PubMed) (October 27):

"Registries/statistics and numerical data"[Mesh]

A search using "registr*" as a keyword was also conducted in Embase and PsycINFO.

Searches for Cost Studies or Economic Evaluations (conducted August 2, 2013)

The National Health Service Economic Evaluation Database (NHS-EED) was searched with the terms *treatment resistant depression*, *bipolar disorder*, *electroconvulsive therapy*, *transcranial magnetic stimulation*, *transcranial direct current stimulation*, and *deep brain stimulation*, restricted to Title. Publication year was set 2003 to 2013.

In addition, PubMed was searched using this strategy:

1. "bipolar disorder" OR "Bipolar Disorder"[Mesh] OR depression
2. treatment resistan* or refractory
3. (((((((((economic analysis) OR (economic evaluation)))) OR (((((cost AND (analysis OR benefit OR effective* OR consequence OR minimization)))) OR (("Costs and Cost Analysis"[MeSH] OR "Cost-Benefit Analysis"[MeSH])))))))
4. 1 and 2 and 3
5. Filters: published in the last 10 years

Miscellaneous

On September 3, 2013, the following search was conducted in the MEDLINE, Embase, and PsycINFO databases (OVID platform):

1. magnetic seizure therapy or magnetic convulsion therapy
2. depression and (treatment resistan* or refractory)
3. 1 and 2

   Limits: Human, Humans, year 2002-2013

MST/MCT is a variation of ECT.

Update Searches: November 12, 2013

ORIGINAL SEARCH FOR EVIDENCE ON VAGUS NERVE STIMULATION (Conducted August 1, 2013)

1. vagus nerve stimulation
2. depression and (treatment resistan* or refractory)
3. 1 and 2

   Limits: Human, Humans, 2009-2013

APPENDIX IV. Overview of Evidence Quality Assessment Methods

Clinical Research

Tools used include internally developed Quality Checklists for evaluating the quality (internal validity) of different types of studies, a checklist for judging the adequacy of systematic reviews used instead of de novo analysis, and Hayes Evidence-Grading Guides for evaluating bodies of evidence for different types of technologies. Hayes methodology is in alignment with the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) system, which was developed by the GRADE Working Group, an international collaborative body.

Step 1	Individual study appraisal a. Initial rating according to study design *Good:* Randomized Controlled Trials *Fair:* Nonrandomized Trial (controlled, parallel group, quasi-randomized) *Poor:* Observational Analytic Studies (prospective or retrospective trials involving historical controls, pretest posttest control trial [patients legitimately serve as their own controls], case-control, registry/chart/database analysis involving a comparison group) *Very Poor:* Descriptive Uncontrolled Studies (case reports, case series, cross-sectional surveys [individual-level data], correlation studies [group-level data]) b. Consider the methodological rigor of study execution according to items in a proprietary Quality Checklist c. Repeat for each study
Step 2	Evaluation of each body of evidence by outcome, key question, or application a. Initial quality designation according to *best* study design in a body of evidence b. Downgrade/upgrade *Downgrade factors:* Study weaknesses (Quality Checklists), small quantity of evidence, lack of applicability, inconsistency of results, publication bias *Possible upgrade factors:* Strong association, dose-response effect, bias favoring no effect c. Assign final rating: High-Moderate-Low-Insufficient d. Repeat for each outcome/question/application
Step 3	Evaluation of overall evidence a. Rank outcomes by clinical importance b. Consider overall quality of evidence for each *critical* outcome c. Assign overall rating based on lowest-quality body: High-Moderate-Low-Insufficient
Step 4	Evidence-Based Conclusion Overall quality of evidence + Balance of benefits and harms

Practice Guidelines (checklist taken from AGREE Tool and approach to scoring used in this report)

Rank each item on a scale of 1-7.

Decide on overall quality (1 = lowest to 7 = highest), giving strongest weight to items 7-14 (Rigor of Development Domain) and items 22-23 (Editorial Independence). For qualitative labels:

Very poor = 1; *Poor* = 2-3; *Fair* = 4-5; *Good* = 6-7

1. The overall objective(s) of the guideline is (are) specifically described.
2. The health question(s) covered by the guideline is (are) specifically described.
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.
4. The guideline development group includes individuals from all relevant professional groups.
5. The views and preferences of the target population (patients, public, etc.) have been sought.
6. The target users of the guideline are clearly defined.
7. Systematic methods were used to search for evidence.
8. The criteria for selecting the evidence are clearly described.
9. The strengths and limitations of the body of evidence are clearly described.
10. The methods for formulating the recommendations are clearly described.
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.
12. There is an explicit link between the recommendations and the supporting evidence.
13. The guideline has been externally reviewed by experts prior to its publication.
14. A procedure for updating the guideline is provided.
15. The recommendations are specific and unambiguous.
16. The different options for management of the condition or health issue are clearly presented.
17. Key recommendations are easily identifiable.
18. The guideline describes facilitators and barriers to its application.
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.
20. The potential resource implications of applying the recommendations have been considered.
21. The guideline presents monitoring and/or auditing criteria.
22. The views of the funding body have not influenced the content of the guideline.
23. Competing interests of guideline development group members have been recorded and addressed.

AGREE Enterprise. Appraisal of Guidelines for Research and Evaluation (AGREE) II. 2009. Available at: http://www.agreetrust.org/. Accessed May 29, 2013.

APPENDIX V. Overview of Systematic Reviews Selected for Key Question #1

The systematic reviews described in the following table were used in lieu of literature searches (i.e., to identify studies) and as a source of study details and/or pooled estimates.

Key: AD, antidepressant, AHRQ, Agency for Health Care Research and Quality; BD, bipolar depression/disorder; BL, baseline; CI, confidence interval; DBS, deep brain stimulation; ECT, electroconvulsive therapy; f/u, follow-up; grp(s), group(s); HAM-D, Hamilton Depression Rating Scale; MA, meta-analysis; MADRS, Montgomery-Åsberg Depression Rating Scale; MDD, major depressive disorder; NR, not reported; NS, not (statistically significant); OR, odds ratio; psychotx, psychotherapy; pt(s), patient(s); QOL, quality of life; RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; sig, (statistically) significant; SR, systematic review; tDCS, transcranial direct current stimulation; TRD, treatment-resistant depression; tx, treatment (or therapy)

Authors and Methods	Methods; Studies and Patients	Main Findings; Authors' Conclusions	Quality of the Evidence; Comments
Gaynes et al. (2011) (AHRQ) Comparative effectiveness review of psychotx, ECT, rTMS, and VNS. MEDLINE, Embase, Cochrane Library, PsycINFO, International Pharmaceutical Abstracts. 1980 to November 18, 2010. Conference proceedings and abstracts also sought.	***Eligible studies:*** Good/fair-quality RCTs of TRD in populations of 100% MDD or MDD and ≤20% BD, published 1980 or later. Observational controlled or comparative studies also eligible for issues related to safety, differential efficacy, and health-related outcomes. ***Definition of TRD considered most relevant:*** ≥2 prior failures of adequate trials of AD (categorized as *Tier 1*). Studies defining TRD as ≥1 failure (*Tier 2*) or providing no definition but likely including pts meeting the preferred criteria (*Tier 3*) were also selected and analyzed separately. ***Outcomes of interest:*** Depressive severity, response, remission, harms (including	FINDINGS AND QUALITY ASSESSMENT, MOST RELEVANT BODIES OF EVIDENCE (TIER 1) The following was abstracted from Tables 96-106 in the AHRQ report. "Eligible" in this context means studies using the Tier 1 definition of TRD (≥2 tx failures). Other studies were not included in quality assessments for bodies of evidence. ***Acute phase tx:*** Psychotx vs control: No eligible studies. Psychotx vs pharmacotx: No eligible studies. ECT vs sham: No eligible studies. ECT vs pharmacotx: 1 RCT (fair). ECT associated w/ greater improvement in depressive severity and response rate (low quality for both outcomes). ECT vs rTMS: 1 fair-quality RCT. No sig differences (low quality). rTMS vs sham: 7 RCTs (3 good, 4 fair). rTMS associated w/ greater change in depressive severity and higher response rate (high quality, both outcomes) and greater remission rate (moderate quality). ***Durability of remission:*** 3 RCTs (fair) of rTMS vs sham. No sig differences (insufficient evidence due to study limitations). No other eligible studies. ***Comparative effectiveness according to symptom types:*** No eligible studies. ***Safety:*** *Cognitive functioning:*	***Quality of included studies:*** All studies included in analysis were rated as good or fair. Most studies did not assess both remission and response rates. ***Quality of SR:*** Good ***Comments:*** 22% of included studies supported by industry. Publication bias assessed but findings NR; authors commented that tests would have had low sensitivity, given small # studies.

Authors and Methods	Methods; Studies and Patients	Main Findings; Authors' Conclusions	Quality of the Evidence; Comments
	adherence and withdrawals), health-related outcomes (e.g., functional status, QOL). ***Response definition**:* ≥50% improvement from BL on HAM-D or MADRS. ***Remission definition:*** HAM-D_{17}, ≤8, HAM-D_{21} ≤10, MADRS ≤8, or anything comparable. ***Evidence grading:*** Risk of bias in studies, overall strength of evidence, and applicability (generalizability), according to published AHRQ methods. ***Included studies:*** 64 studies (79 publications); all but 2 studies were RCTs. ***Studies excluded because of poor quality:*** 8 otherwise eligible studies ***Pt characteristics:*** Most pts were severely depressed. No other aggregated data.	ECT vs rTMS: 1 RCT (fair) and 1 cohort study (fair). Conflicting results (insufficient evidence). ECT vs ECT+rTMS: 1 RCT (fair). No differences (insufficient evidence). rTMS vs sham: 3 RCTs (1 good, 2 fair). Conflicting results (insufficient evidence). *Specific adverse events:* 1 RCT each for ECT vs ECT+rTMS (no differences), rTMS vs sham (more scalp pain w/ rTMS) (low quality for all comparisons). *Withdrawals due to adverse event:* ECT vs rTMS: 1 cohort study (fair). No difference (low quality). rTMS vs sham: 7 RCTs (1 good, 6 fair). Mixed results (insufficient evidence). *Overall withdrawals:* ECT vs rTMS: 1 RCT (fair) and 1 cohort study (fair). More withdrawals w/ ECT (low quality). rTMS vs sham: 8 RCTs (fair). Mixed results (insufficient evidence). ***Differential efficacy/safety for subpopulations:*** 2 RCTs (fair) of rTMS vs sham suggested better outcomes in young adults (ages 18-37 yrs) or in older adults w/ post-stroke depression (low quality). ***Health-related outcomes:*** No differences in ECT vs ECT+rTMS. Results favored rTMS compared w/ sham (1 RCT; low quality). PHARMACOTX VS CONTROL (FOR INDIRECT COMPARISONS) Controls were not comparable; estimates of effect could not be pooled. *W/in-grp averages for pharmacotx arms only:* Switching: Mean change in MADRS, –11.2 (CI, –14.7 to –7.8); response, 39.8% (CI, 30.7-48.9); remission, 22.3% (CI, 16.2-28.4). Augmentation: Mean change in MADRS, –11.2 (CI, –13.7 to –8.8); response, 38.1% (CI, 13.0-45.3), remission 27.2% (CI, 20.4-34.0). Maintenance: Mean change in MADRS, –7.6 (CI, –9.2 to –5.2); response,	

Authors and Methods	Methods; Studies and Patients	Main Findings; Authors' Conclusions	Quality of the Evidence; Comments
		27.3% (CI, 19.8-34.8); remission, 16.8% (CI, 13.5-20.2). META-ANALYSES, NONPHARMACOLOGIC TXS, ALL TIERS "The evidence base combining data for Tiers 1-3 on the whole produced findings that were consistent with Tier 1 TRD data and also appear applicable to TRD populations. . . Use of multiple definitions makes synthesis of the available information difficult, as the effect of combining patients with one treatment failure with those of two or more (or four or more) remains unclear" (p. 159). ***Authors' conclusion:*** Comparative effectiveness research in this area is in its early stages, w/ the greatest volume of evidence for ECT and rTMS. Head-to-head trials using a consistent definition and specifying the # adequate tx failures in the current episode are needed.	
Hayes 2012 SR of studies on DBS for TRD PubMed and MEDLINE. January 2000 to October 2012. Manual search of bibliographies of retrieved articles. **NOTE:** For the present report, the Embase and PsycINFO databases were searched to assure that no studies were missed; no missed studies were identified.	***Eligible studies:*** Prospective studies w/ ≥10 pts, English language, and abstracts. ***Definition of TRD:*** Each trial had different definition, most required failure of ≥4 tx methods (nonpharmacological or pharmacological). ***Outcomes of interest:*** Depression severity, response, remission, harms. ***Included studies:*** 5 prospective, uncontrolled, open-label (before-and-after) studies (3 of subcallosal singulate stimulation and 2 of striatum/nucleus accumbens stimulation) (86 pts). ***Pt characteristics:*** MDD; current depressive episode for >12	***Efficacy/Effectiveness:*** Response rate ranged from 40%-60% at 6 mos and 29%-55% at 12 mos (3 studies) Remission rate ranged from 18%-35% at 6 mos (1 study) and 18%-36% at 12 mos (3 studies). Significant improvement in the depression scores vs BL. ***Safety:*** Several potential complications are possible but no evidence of cognitive decline. ***Authors' conclusion:*** There is some preliminary evidence that DBS reduces depressive symptoms in pts w/ TRD. (No overall statement about safety.)	***Quality of included studies:*** All studies included in analysis were rated as poor or very poor. Not all studies reported remission rates. ***Quality of SR:*** Good ***Comments:*** Studies were small in size, lacked control grps, were funded by the device manufacturer, and included authors w/ financial ties to manufacturer.

Authors and Methods	Methods; Studies and Patients	Main Findings; Authors' Conclusions	Quality of the Evidence; Comments
	mos; 2 studies included pts w/ BD; mean age mid-40's.		
Kalu 2012 SR and MA of tDCS for MDD MEDLINE and Embase. Search dates, January 1998 (date of first published study w/ contemporary parameters) to May 2011.	***Eligible studies:*** Unipolar or bipolar MDD. SR inclusion: Open-label or RCT; published or accepted for publication as journal article or letter (conference abstracts excluded). MA inclusion: RCT, clinician-assessed depression severity measures, data available for calculation of % change in severity. TRD not a study selection criterion. ***Outcomes of interest:*** Depression severity; safety. ***Included studies:*** 10 studies, 6 RCTs (194 pts) and 4 uncontrolled. ***Excluded studies:*** 4 otherwise eligible studies were excluded because of overlapping pt populations. ***Pt characteristics:*** No aggregated data other than prevalence of BD (where reported: all studies, 12%; uncontrolled, 30%; RCTs, 5%).	***Efficacy/Effectiveness:*** *Weighted mean % reduction in depression severity in tDCS grps:* 28.9% (range 14.6%-60%). *Weighted mean response frequency in tDCS grps:* 19.8% of pts (range 0%-80% *Weighted mean remission frequency in tDCS grps:* 8.5% of pts (range 0%-23.8%) *Pooled tDCS-vs-sham effect size based on % change:* 0.74 (95% CI, 0.21-1.27, P=0.006; heterogeneity, P=0.017; tests for publication bias NS). Based on 6 RCTs (2 subgrps analyzed separately in 1 of 6 RCTs). Heterogeneity: Cochrane's Q, P=0.017; I^2=61.4. Heterogeneity not explained in metaregression by BL severity, concurrent AD use, current strength of tDCS, or total # sessions. Publication bias: NS according to 2 tests. ***Safety:*** Generally very minor adverse events except for 3 cases of hypomanic episode during active tDCS tx. Headache, itchiness, and redness at site of stimulation occurred in active and sham arms (4 RCTs; Boggio 2008, Loo 2010, Palm 2011, Loo 2012). No adverse events observed in 5 studies (Fregni 2006a, Verrucci 2009, Brunoni 2011b, Dell'Osso 2011, Martin 2011). 3 cases of hypomanic episode occurred during tx, 1 in each of 3 studies (Loo 2010, Loo 2012, Martin 2011). ***Authors' conclusion:*** Effect size was medium to large. tDCS has the potential to be an effective clinical tx for MDD, but larger studies are needed for identification of optimal tx parameters and studies w/ longer f/u times are needed to assess whether benefits are lasting or maintenance tx is necessary.	***Quality of included studies:*** No systematic quality assessment reported by SR authors; authors mentioned blinding success was reported by only 3 of 6 RCTs and by none of the RCTs w/ sig effect sizes. ***Quality of SR:*** Fair-good (no explicit consideration of study quality). ***Comments:*** TRD was not an inclusion criterion for SR; where reported in individual studies, pts had failed multiple AD trials and/or were selected because of TRD. Outcome measure used for calculation of effect size was a continuous variable rather than a standard dichotomous definition of response.
Berlim 2013a SR and MA of tDCS	***Study inclusion criteria:*** Randomized, double-blind,	***Efficacy/Effectiveness:*** No effect overall.	***Quality of included studies:*** Authors

Authors and Methods	Methods; Studies and Patients	Main Findings; Authors' Conclusions	Quality of the Evidence; Comments
for MDD MEDLINE, Embase, PsycINFO, Cochrane, SCOPUS, and Web of Science's Citations Index Expanded. July 1998 to August 2012.	sham-controlled, parallel or crossover, ≥5 pts, English language, reported response and remission data. ***Study exclusion criteria:*** Narrow dx (e.g., postpartum depression) or secondary MDD (e.g., vascular); tDCS started at same time as new AD. ***Definition of TRD:*** NR (but see pt characteristics). ***Outcomes of interest:*** Response (>50% improvement on HAM-D or MADRS at study endpoint) and remission (≤7 on HAM-D_{17}, ≤8 on HAM-D_{21}, or ≤6 on MADRS). ***Included studies:*** 6 RCTs (5 in common w/ Kalu review) (200 pts). ***Excluded studies:*** 2 otherwise eligible RCTs because of only 1 tDCS session or tDCS initiated at same time as a new AD. ***Pt characteristics:*** Mean age 49.9±4.07 yrs; 49% women; mean # failed AD 2.36±1.19, range 1.3-2.9 (NR in Fregni 2006a and 2006b).	Weighted pooled OR for response, 1.97; CI, 0.85-4.56; no heterogeneity). Weighted pooled OR for remission, 2.13; CI, 0.64-7.06; no heterogeneity). Strong effect in studies (2 RCTs) w/ monotx (switch strategy). Weighted pooled OR for response, 7.54 (CI, 1.63-34.8, *P*=0.01; sig heterogeneity). No differential effect according to # sessions or 1 vs 2 mA electrical current. ***Safety:*** Not evaluated. ***Authors' conclusion*:** There is insufficient evidence to determine whether tDCS is superior to placebo in achieving response and remission in pts w/ MDD. Studies w/ larger sample sizes are needed to evaluate differential efficacy according to MDD subtypes. Research comparing tDCS w/ other neuromodulation techniques is also needed. Identification of optimal stimulation parameters is needed.	asserted that use of stringent inclusion criteria eliminated poor-quality studies ***Quality of SR:*** Good ***Comments:*** TRD was not an inclusion criterion; however, the mean # failed ADs (2.36±1.19) suggests on average, pts met an accepted definition of TRD. Authors conducted this SR/MA because of perceived limitations in the Kalu review.