Comparative Effectiveness Review
Number 33

Effective Health Care Program

Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults

Executive Summary

Background

Major depressive disorder (MDD) is common and costly. Over the course of a year, between 13.1 million and 14.2 million people will experience MDD. Approximately half of these people seek help for this condition, and only 20 percent of those receive adequate treatment. For those who do initiate treatment for their depression, approximately 50 percent will not adequately respond following acute-phase treatment; this refractory group has considerable clinical and research interest. Patients with only one prior treatment failure are sometimes included in this group, but patients with two or more prior treatment failures are a particularly important and poorly understood group and are considered to have treatment-resistant depression (TRD). These TRD patients represent a complex population with a disease that is difficult to manage.

Patients with TRD incur the highest direct and indirect medical costs among those with MDD. These costs increase with the severity of TRD. Treatment-resistant patients are twice as likely to be hospitalized, and their cost of hospitalization is more than six times the mean total costs of depressed patients who are not treatment resistant. After considering both medical and disability claims from an employer's perspective, one study found that TRD employees cost $14,490 per employee per year, whereas

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm.**

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

the cost for non-TRD employees was $6,665 per employee per year.

Given the burden of TRD generally, the uncertain prognosis of the disorder, and the high costs of therapy, clinicians and patients alike need clear evidence to guide their treatment decisions. The choices are wide ranging, include both pharmacologic and nonpharmacologic interventions, and are fraught with incomplete, potentially conflicting evidence. Somatic treatments, which may involve use of a pharmacologic intervention or a device, are commonly considered for patients with TRD. Antidepressant medications, which are the most commonly used intervention, have decreasing efficacy for producing remission after patients have experienced two treatment failures. Such drugs also often have side effects, sometimes minor but sometimes quite serious. For these reasons, clinicians often look for alternative strategies for their TRD patients.

This review from the RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center (EPC) provides a comprehensive summary of the available data addressing the comparative effectiveness of four nonpharmacologic treatments as therapies for patients with TRD: electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), vagus nerve stimulation (VNS), and cognitive behavioral therapy or interpersonal psychotherapy (CBT or IPT).

The core patient population of interest was patients with MDD who met our definition of TRD: failure to respond following two or more adequate antidepressant treatments. We also included TRD studies in which the patient population could include a "mix" of up to 20 percent of patients with bipolar disorder (i.e., 80 percent or more of patients had only MDD), assuming that this small mix would not substantially alter outcomes seen with MDD-only populations.

We structured our review to maintain our focus on study populations meeting our TRD definition (≥2 antidepressant failures) while not excluding potentially relevant evidence. We identified different tiers of TRD-related studies to use in our analytic strategy:

- **Tier 1** evidence (TRD as defined in this report): studies in which patients specifically had two or more prior treatment failures with medications.
- **Tier 2** evidence: studies in which patients had one or more prior treatment failures.
- **Tier 3** evidence: studies in which the number of prior failed treatments was not specified but the clinical situation suggested a high probability of patients having two or more prior antidepressant treatment failures; these data have probable relevance to TRD. Studies that did not specify the number of failed treatments but noted that all subjects were referred for ECT were included in this tier.

This comparative effectiveness review is intended to help various decisionmakers come to informed choices about the use of nonpharmacologic interventions for TRD in adults. Our principal goal is to summarize comparative data on the efficacy, effectiveness, and harms of ECT, rTMS, VNS, and CBT/IPT in patients with TRD. Comparisons of these nonpharmacologic therapies are our main interest. However, because treatment decisions made by patients with TRD and their clinicians are not limited to nonpharmacologic options, we also compare nonpharmacologic options with pharmacologic ones. We address the following six Key Questions (KQs) as specified by the Agency for Healthcare Research and Quality (AHRQ). "Trials" in these KQs refers to treatment attempts, not experimental studies.

- KQ 1a. For adults with TRD (defined as two or more failed adequate trials of a biologic [i.e., pharmacologic] intervention), do nonpharmacologic interventions such as electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), vagus nerve stimulation (VNS), or demonstrated effective psychotherapy (e.g., cognitive therapy [CBT or IPT]) differ in efficacy or effectiveness in treating acute-phase depressive symptoms (e.g., response and remission), whether as a single treatment or part of a combination treatment?
- KQ 1b. How do these nonpharmacologic treatments compare with pharmacological treatments in efficacy or effectiveness in treating acute-phase depressive symptoms after two or more failed adequate trials?

- KQ 2. For adults with TRD, do nonpharmacologic interventions differ in their efficacy or effectiveness for maintaining response or remission (e.g., preventing relapse or recurrence), whether as a single treatment or part of a combination treatment?
- KQ 3. Do nonpharmacologic interventions (single or combination) differ in their efficacy or effectiveness for treating TRD as a function of particular symptom subtypes (e.g., catatonic [frozen or hyper] or psychotic symptoms)?
- KQ 4. For adults with TRD, do nonpharmacologic interventions differ in safety, adverse events, or adherence? Adverse effects of interest include but are not limited to amnesia, memory loss, headaches, and postoperative complications.
- KQ 5. How do the efficacy, effectiveness, or harms of treatment with nonpharmacologic treatments for TRD differ for the following subpopulations:
  - Elderly or very elderly patients; other demographic groups (defined by age, ethnic or racial groups, and sex)?
  - Patients with medical comorbidities (e.g., seizure history, stroke, diabetes, dementia, perinatal depression, ischemic heart disease, cancer)?
- KQ 6. For adults with TRD, do nonpharmacologic interventions differ in regard to other health-related outcomes (e.g., quality of life)?

We searched MEDLINE, Embase, the Cochrane Library, PsycINFO, and International Pharmaceutical Abstracts. We searched for systematic reviews, clinical controlled trials, meta-analyses, and nonexperimental studies in which the investigator did not assign group allocation. Sources were searched from 1980 through November 18, 2010. AHRQ Scientific Resource Center (SRC) staff contacted device manufacturers and invited them to submit dossiers, including citations. The SRC also provided our EPC with other relevant data that may not have been captured in the literature search.

For efficacy and effectiveness (KQs 1 and 2), we first focused on head-to-head randomized controlled trials (RCTs) comparing one intervention with another. When sufficient head-to-head evidence was unavailable, we evaluated indirect evidence: nonpharmacologic interventions versus placebo- or sham-controlled evidence or "treatment as usual" controls. For KQs 3, 4, 5, and 6, we examined data from both experimental and observational studies (generally prospective cohort studies). We did not formally distinguish efficacy from effectiveness trials.

We rated the quality of individual studies as good, fair, or poor; only good or fair studies are included in these analyses. We evaluated the strength of the various bodies of evidence using principles stated in the AHRQ Methods Guide for Comparative Effectiveness Reviews, which grades strength as high, moderate, low, or insufficient. We evaluated the applicability of the body of evidence using a qualitative assessment of the population, intervention/treatment, comparator, outcomes measured, timing of followup, and setting.

Throughout this report we synthesized the literature qualitatively. If data were sufficient, we conducted meta-analyses of data for comparisons involving trials that were fairly homogenous in study populations, treatment intervention, and outcome assessments. Given our focus on Tier 1 (TRD) studies, for each KQ we first present an overview of the particular comparison, including the strength of evidence findings for the Tier 1 studies. This summary does not present detailed findings from the Tier 2 and Tier 3 studies. The results chapter of the full report presents those data in greater detail.

Results: Overview

From a total of 2,754 citations retrieved, we ultimately identified 79 good-, fair-, or poor-quality articles in this review; they represent 64 studies. Of these studies, there were 17 head-to-head RCTs (19 articles): 7 studies (9 articles) were head-to-head RCTs of a nonpharmacologic intervention versus a nonpharmacologic intervention; 3 were head-to-head RCTS of a nonpharmacologic intervention versus a pharmacologic one; and 7 were head-to-head studies of

a pharmacologic versus pharmacologic intervention. Further, there were 38 additional RCTs (50 articles) that were sham- or placebo-controlled, and 2 observational studies (2 articles). We excluded 8 studies (8 articles) because of poor quality. We present evidence that allows comparison of the four nonpharmacologic treatments of interest (ECT, rTMS, VNS, and psychotherapy) stratified by tiers of evidence.

Comparative clinical research on nonpharmacologic interventions in a TRD population is in its infancy. Many clinical questions about efficacy and effectiveness remain unanswered. The text below presents our principal results; summary tables (A–J) document Tier 1 TRD findings for major comparisons and outcomes for each key question, give the overall strength of evidence for that comparison, and outline key findings. We report first on direct evidence (head-to-head comparisons) and then on indirect evidence (e.g., trials using controls). If a specific comparison did not involve a Tier 1 population but did have trials conducted in a Tier 2 and/or Tier 3 population, we have listed it in this table, noted "No eligible studies identified," and added a footnote indicating the presence of at least one such study.

The greatest volume of evidence is for ECT and rTMS; however, the direct comparative evidence about even these treatments is quite limited. Available indirect evidence primarily involves rTMS; a little information is available on VNS and psychotherapy (chiefly for efficacy and adverse events), and no available indirect evidence involves ECT. Given the limited number of Tier 1 studies incomplete reporting on the number of failed treatment attempts, we were unable to stratify our outcomes by the number of treatment failures within Tier 1.

Table A. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for Key Question 1a, comparative efficacy of nonpharmacologic treatments

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Change in depressive severity	42	Low	1 fair trial: both ECT and rTMS improved symptom but did not differ significantly.
ECT vs. rTMS	Response rate	42	Low	1 fair trial: ECT and rTMS did not differ significantly.
ECT vs. rTMS	Remission rate	42	Low	1 fair trial: ECT and rTMS did not differ significantly.
ECT plus rTMS vs. ECT	Change in depressive severity	22	Low	1 fair trial: both ECT and ECT plus rTMS improved symptom severity but did not differ significantly.
ECT plus rTMS vs. ECT	Response rate	0	NA	No eligible studies identified.[c]
ECT plus rTMS vs. ECT	Remission rate	22	Low	1 fair trial: ECT and ECT plus rTMS did not differ significantly.
ECT vs. sham	Change in depressive severity	0	NA	No eligible studies identified.[c]
ECT vs. sham	Response rate	0	NA	No eligible studies identified.[c]
ECT vs. sham	Remission rate	0	NA	No eligible studies identified.[c]
rTMS vs. sham	Change in depressive severity	497	High	7 trials (3 good, 4 fair): rTMS had a significantly greater decrease in depressive severity than sham. 4 fair trials: rTMS had nonsignificantly greater decrease in depressive severity than sham. 2 fair trials: rTMS had greater decrease than sham but significance NR. 1 fair trial: rTMS did not significantly differ from sham.
rTMS vs. sham	Response rate	471	High	4 trials (3 good, 1 fair): rTMS had a significantly higher response rate than sham. 1 fair trial: rTMS had a nonsignificantly higher response rate than sham. 6 fair trials: rTMS had a higher response rate than sham, but significance NR. 1 fair trial: rTMS did not clearly differ from sham, but significance NR.
rTMS vs. sham	Remission rate	223	Moderate	3 trials (2 good, 1 fair): rTMS had significantly greater remission rate than sham. 1 fair trial: rTMS had a greater remission rate than sham but significance NR.

Table A. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for Key Question 1a, comparative efficacy of nonpharmacologic treatments (continued)

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
VNS vs. sham	Change in severity depressive	235	Low	1 good trial: VNS and sham did not differ significantly.
VNS vs. sham	Response rate	235	Low	1 good trial: VNS and sham did not differ significantly.
Psychotherapy vs. control	Change in depressive severity	0	NA	No eligible studies identified.[c]
Psychotherapy vs. control	Response rate	0	NA	No eligible studies identified.[c]
Psychotherapy vs. control	Remission rate	0	NA	No eligible studies identified.[c]

ECT = electroconvulsive therapy; NA = not applicable; NR = not reported; rTMS = repetitive transcranial magnetic stimulation; VNS = vagus nerve stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 or Tier 3 study addressed this comparison.

Table B. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 1b, comparative efficacy of nonpharmacologic and pharmacologic treatments

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. pharmacotherapy	Change in depressive severity	39	Low	1 fair trial: ECT had significantly greater improvement in symptom severity than pharmacotherapy.
ECT vs. response rates than pharmacotherapy	Response rate	39	Low	1 fair trial: ECT had significantly greater pharmacotherapy.
Psychotherapy vs. pharmacotherapy	Change in depressive severity	0	NA	No eligible studies identified.[c]
Psychotherapy vs. pharmacotherapy	Response rate	0	NA	No eligible studies identified.[c]
Psychotherapy vs. pharmacotherapy	Remission rate	0	NA	No eligible studies identified.[c]

ECT = electroconvulsive therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table C. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 2, comparative efficacy for maintaining remission

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Maintenance of remission	0	NA	No eligible studies identified.[c]
rTMS vs. sham	Maintenance of remission	68	Insufficient	3 fair trials: no significant differences in maintenance of remission; however, small sample sizes in two of the studies and the presence of a co-intervention in the third study make results difficult to interpret.
CBT vs. usual care	Maintenance of remission	0	NA	No eligible studies identified.[c]

CBT = cognitive behavioral therapy; ECT = electroconvulsive therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; vs = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table D. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 3, comparative efficacy for particular symptom subtypes

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Change in depressive severity	0	NA	No eligible studies identified.[c]

ECT = electroconvulsive therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table E. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 4a, impact of nonpharmacologic interventions on cognitive functioning

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Cognitive functioning	72	Insufficient	1 fair trial and 1 fair cohort study: Some evidence suggests no difference between treatments, whereas some evidence suggests ECT may have deleterious impact on cognitive functioning compared with rTMS (1 study: significant effect on 1-week recall; both studies: nonsignificant effect on all other measures).
ECT vs. + rTMS	Cognitive functioning	22	Insufficient	1 fair trial: no significant differences in ECT a single item measure on memory problems.
rTMS vs. sham	Cognitive functioning	161	Insufficient	4 trials (1 good, 3 fair): Some evidence suggests no difference between rTMS and sham, whereas some evidence suggests that rTMS improves cognitive functioning compared to sham (2 trials: significant differences in memory, verbal fluency; all other findings nonsignificant or significance not reported).

ECT = electroconvulsive therapy; rTMS = repetitive transcranial magnetic stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.

Table F. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 4b, specific adverse events

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Adverse events	0	NA	No eligible studies identified.[c]
ECT vs. ECT + rTMS	Adverse events	22	Low	1 fair trial: no significant differences in specific adverse events
rTMS vs. sham	Adverse events	68	Low	1 good trial: rTMS resulted in significantly more scalp pain at the stimulation site than sham.
VNS vs. sham	Adverse events	235	Low	1 fair trial: Some differences in specific adverse events reported (P = NR)

ECT = electroconvulsive therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; VNS = vagus nerve stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table G. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 4c, withdrawals due to adverse event

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Withdrawals	30	Low	1 fair cohort study: no difference in withdrawals between ECT and rTMS groups (P = NR).
ECT vs. sham	Withdrawals	0	NA	No eligible studies identified.[c]
rTMS vs. sham	Withdrawals	337	Insufficient	7 trials (1 good, 6 fair): trials showed mixed results about withdrawals attributed to adverse events.
VNS vs. sham	Withdrawals	235	Low	1 good trial: VNS had greater withdrawals attributed to adverse events than sham (significance NR).
CBT vs. usual care	Withdrawals	0	NA	No eligible studies identified.[c]

CBT = cognitive behavioral therapy; ECT = electroconvulsive therapy; NA = not applicable; NR = not reported; rTMS = repetitive transcranial magnetic stimulation; VNS = vagus nerve stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table H. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 4d, adherence as measured by overall withdrawals

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. rTMS	Overall withdrawals	72	Low	1 fair trial and 1 fair cohort study: studies showed more withdrawals in the ECT group compared with rTMS (P = NR).
ECT vs. sham	Overall withdrawals	0	NA	No eligible studies identified.[c]
rTMS vs. sham	Overall withdrawals	325	Insufficient	8 fair trials: trials showed mixed results about withdrawals.
CBT vs. usual care	Overall withdrawals	0	NA	No eligible studies identified.[c]

CBT = cognitive behavioral therapy; ECT = electroconvulsive therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Table I. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 5, efficacy and harms for selected populations

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
rTMS vs. sham	Changes in depressive severity	34	Low	1 fair trial: rTMS produced better outcome than sham in young adult population (ages 18–37).
rTMS vs. sham	Changes in depressive severity	20	Low	1 fair trial: rTMS produced better outcome than sham in older adults with post-stroke depression.
rTMS vs. sham	Response	34	Low	1 fair trial: rTMS produces better response rates than sham in young adult population (ages 18–37).
rTMS vs. sham	Response	20	Low	1 fair trial: no difference between rTMS and sham for older adults with post-stroke depression.
rTMS vs. sham	Remission	20	Low	1 fair trial: no difference between rTMS and sham in older adults with post-stroke depression.

rTMS = repetitive transcranial magnetic stimulation; vs. = versus
[a]Strength of evidence is based on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.

Table J. Summary of findings on nonpharmacologic treatment of adult treatment-resistant depression (TRD) with strength of evidence for Tier 1 (TRD) for KQ 6, health-related outcomes

Comparison	Outcome	Number of Subjects	Strength of Evidence[a]	Findings[b]
ECT vs. ECT + rTMS	Health-related outcomes	22	Low	1 fair trial: There were no differences between groups in improvements in daily functioning.
rTMS vs. sham	Health-related outcomes	60	Low	1 fair trail: low rTMS had significantly greater improvement in health status and daily functioning than sham, while this relationship approached statistical significance when comparing high rTMS to sham.
VNS vs. sham	Health-related outcomes	214	Low	1 fair trial: VNS and sham groups did not differ significantly in daily functioning.
CBT/DBT vs. control	Health-related outcomes	0	NA	No eligible studies identified.[c]

CBT = cognitive behavioral therapy; DBT = dialectical behavioral therapy; NA = not applicable; rTMS = repetitive transcranial magnetic stimulation; VNS = vagus nerve stimulation; vs. = versus
[a]Strength of evidence is based on the on guidance provided in the AHRQ Methods Guide for Comparative Effectiveness Reviews; see text.
[b]Good and fair designations relate to quality ratings for each study.
[c]At least one Tier 2 and/or Tier 3 study addressed this comparison.

Efficacy of Nonpharmacologic Interventions Against Other Nonpharmacologic Interventions (KQ 1a)

Direct evidence. The available head-to-head literature concerning the efficacy of the nonpharmacologic interventions for Tier 1 TRD is limited to two fair trials (both in MDD-only populations). One compared ECT and rTMS, and the other compared ECT and ECT plus rTMS. They showed, with low strength of evidence, no differences between treatment options for depressive severity, response rates, and remission rates. No trial involved a direct comparison of psychotherapy with another nonpharmacologic intervention.

Indirect evidence. We identified trials that compared a nonpharmacologic intervention, generally rTMS, VNS, or psychotherapy, with a control or sham procedure in Tier 1 populations. We identified no eligible ECT versus control studies. The number of these trials with the same or similar control group was very small, so we could not pool them quantitatively. We could, however, assess the potential benefits of nonpharmacologic interventions versus controls by calculating mean changes in depressive severity, relative risks of response, and relative risks of remission.

rTMS was beneficial relative to controls receiving a sham procedure for all three outcomes (severity of depressive symptoms, response rate, remission rate). rTMS produced a greater decrease in depressive severity (high strength of evidence). Specifically, rTMS averaged a decrease in depressive severity measured by the Hamilton Rating Scale for Depression (HAM-D) of more than 5 points relative to sham control, and this change meets the minimum threshold of the 3-point HAM-D difference that is considered clinically meaningful. Response rates were greater with rTMS than sham (also high strength of evidence); those receiving rTMS were more than three times as likely to achieve a depressive response as patients receiving a sham procedure. Finally, rTMS was also more likely to produce remission than the control procedure (moderate strength of evidence); patients receiving rTMS were more than six times as likely to achieve remission as those receiving the sham.

In the only other Tier 1 comparison, one good-quality VNS versus sham control trial (a mixed MDD/bipolar population) reported no differences between the groups as measured by a change in depressive severity or response rates (low strength of evidence).

Efficacy of Nonpharmacologic Interventions Compared With Antidepressant Pharmacotherapies (KQ 1b)

Direct evidence. The available head-to-head literature concerning the efficacy of the nonpharmacologic interventions compared with pharmacologic treatment (in this case, paroxetine) for Tier 1 trials is limited to one fair trial (a mixed MDD/bipolar population). ECT produced a significantly greater decrease in depressive severity (9 points by HAM-D) and significantly better response rates (71 percent vs. 28 percent) than medications (low strength of evidence).

Indirect evidence. Indirect evidence about procedures or psychotherapy (vs. sham or nonpharmacologic controls) was presented above as part of KQ 1.

We attempted to determine mean changes in depressive severity, relative risks of response, and relative risks of remission for pharmacologic versus control studies to allow a comparison with similar outcomes in the nonpharmacologic versus control trials (KQ 1a, indirect). However, we found no comparable, common control groups (i.e., patients not receiving a mood-related medication) to allow such comparisons.

Instead, we determined mean average outcomes for pharmacologic treatments.

- For switching strategies, mean pharmacologic response rates averaged 39.8 percent (95% CI, 30.7% to 48.9%) and mean remission rates averaged 22.3 percent (95% CI, 16.2% to 28.4%).
- For augmentation, mean response rates averaged 38.1 percent (31.0% to 45.3%) and mean remission rates averaged 27.2 percent (20.4% to 34.0%).
- For maintenance strategies, mean response rates averaged 27.3 percent (19.8% to 34.8%) and mean remission rates averaged 16.8 percent (13.5% to 20.2%).

Although these results provide an idea of the general degree of response seen with next-step pharmacologic treatment in TRD, they serve as an uncontrolled case series and should be compared to nonpharmacologic outcomes only with caution.

Maintenance of Remission or Prevention of Relapse (KQ 2)

Direct evidence. With respect to maintaining remission (or preventing relapse), we had no direct comparisons involving ECT, rTMS, VNS, or CBT.

Indirect evidence. Three fair trials compared rTMS with a sham procedure and found no significant differences. However, too few patients were followed during the relapse prevention phases in two of the three studies, and patients in the third received a co-intervention providing insufficient evidence for a conclusion. We had no eligible studies for ECT, VNS, or psychotherapy.

Efficacy of Nonpharmacologic Interventions for Patients With Different Symptomatology (KQ 3)

Direct evidence. We identified no Tier 1 trials that addressed whether procedure-based treatments differed as a function of symptom subtypes. Also, no comparative evidence was available about psychotherapy in subgroups defined by symptom clusters.

Indirect evidence. We identified no studies testing either procedure-based or psychotherapeutic interventions against sham procedures or other controls.

Safety, Adverse Events, and Adherence (KQ 4)

Direct evidence. In examining safety, adverse events, and adherence, we found some differences across the interventions in the harms and negative side effects to patients. However, the data were insufficient to reach a conclusive result. For just this set of analyses, we examined both clinical trials and cohort studies, and we focus on cognitive functioning, occurrence of specific adverse events, and withdrawals.

Cognitive functioning. For Tier 1 studies on cognitive functioning, some evidence suggests no differences in changes in cognitive functioning between groups, while some evidence suggests ECT may have a deleterious impact on cognitive functioning compared to rTMS (insufficient strength of evidence). No differences between groups on a single-item measure of cognitive functioning were found in a study comparing ECT with ECT and rTMS (insufficient strength of evidence).

Specific adverse events. One Tier 1 study comparing ECT with a combination of ECT and rTMS found no differences in specific adverse events (low strength of evidence).

Withdrawals. We looked at both withdrawals that investigators attributed to adverse events and overall numbers or rates of withdrawals. A single study with a small sample size indicated no difference in withdrawals due to adverse events for the ECT group when compared to rTMS but did not report on the significance of this result (low strength of evidence).

Evidence for ECT compared with rTMS indicated higher rates of overall withdrawals in the ECT compared to the rTMS group (P = NR; low strength of evidence).

Indirect evidence. We attempted to include data from the same types of studies and for the same outcomes as for direct evidence. We identified no studies comparing ECT versus control.

Cognitive functioning. Mixed evidence on cognitive functioning in rTMS versus sham was insufficient evidence to draw a conclusion (insufficient strength of evidence).

Specific adverse events. rTMS groups reported significantly more scalp pain at the stimulation site (low strength of evidence).

Some differences in the frequency of specific adverse events were seen when comparing VNS and sham groups, but the significance of the findings was not reported (P = NR) (low strength of evidence).

Withdrawals. Findings were mixed in Tier 1 studies as to whether rTMS groups had greater rates of withdrawals (overall and due to adverse events) than groups receiving sham procedures (insufficient evidence for both).

Withdrawals attributable to adverse events were higher in the VNS group compared with sham (low strength of evidence).

No Tier 1 studies reported on withdrawals for CBT groups versus those receiving some form of usual care.

Efficacy or Harms of Nonpharmacologic Treatments for Selected Patient Subgroups (KQ 5)

Direct evidence. We found no studies (in any tier) directly comparing nonpharmacologic interventions in selected populations, such as the elderly, those with stroke, or those with other medical comorbidities.

Indirect evidence. Two Tier 1 trials compared rTMS with sham. All findings provided low strength of evidence. For young adults (ages 18–37), one trial found that rTMS produced a greater decrease in depressive severity and a greater response rate than sham. A second trial, conducted in older adults with post-stroke depression, found that rTMS produced a greater decrease in depressive severity and a greater response rate but no difference in remission rates compared with a sham control.

Health-Related Outcomes of Nonpharmacologic Treatments (KQ 6)

Direct evidence. With respect to patient-reported health-related outcomes, we focused on quality of life (various measures) and ability to function in daily life. One Tier 1 study compared ECT with a combination of ECT and rTMS and found no differences between groups in improvement on the Global Assessment of Functioning scale (low strength of evidence).

Indirect evidence. Two trials (both in mixed MDD/bipolar populations) assessed general health status and mental and physical functioning (all health domains related to quality of life). In one fair trial, low rTMS had significantly greater improvement in health status and daily functioning than sham, while this relationship approached statistical significance when comparing high rTMS to sham (as measured by the Global Assessment of Functioning scale; low strength of evidence). In the other fair trial, VNS and sham groups did not differ significantly in daily functioning (as measured by the 36-item Medical Outcomes Study Short Form [MOS SF-36]; low strength of evidence). No studies of psychotherapy were identified.

Applicability

For the limited amount and low strength of evidence available, the data for Tier 1 (TRD) is generally applicable to TRD populations. Populations enrolled in these trials appeared representative of our target population. Studied interventions were comparable to those in routine use, though dose and duration of nonpharmacalogic treatment often varied between studies.

Measured outcomes on the whole reflected the most important clinical outcomes for depression measures, although reporting was inconsistent; outcomes for the other key questions were much more restricted. Followup periods were generally shorter than desirable, but most were sufficient to measure an initial acute-phase treatment response. Study settings were a mixture of inpatient and outpatient, because ECT is generally an inpatient procedure and the others are generally outpatient. Some evidence highlights the importance of patient acceptability of treatment as some patients refuse particular interventions. An individualized balance between a patient's needs and concerns must be taken into account during selection from a range of nonpharmacologic and pharmacologic antidepressant treatment options.

The use of inconsistent definitions of TRD in the trials and the absence of analyses considering the effect of the number of current treatment failures on outcomes hindered interpretation of data, leading to our use of a tiered system for analyses. 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.

Remaining Issues

This area of comparative clinical research is in its infancy. Key areas for future research need primarily to lay more robust foundations for an evidence base that can better inform decisions for clinicians and patients.

The field needs a standard definition of TRD that investigators should use in their clinical trials research. Comparison of any of the potential

interventions in the field, nonpharmacologic or otherwise, is hampered by the variability in TRD definitions. Although these definitions appear to be converging on a single meaning—two or more treatment failures in the current episode—very few studies of TRD have applied it.

Progress in this area of research requires better standardization of this concept, so that future reviews of the evidence do not need to resort to differentiating, as we did, between "Tier 1" studies (i.e., TRD by this definition based on two or more treatment failures) and "Tier 2 or 3" types of studies. The latter do provide information that helps illuminate likely impacts of these interventions on patients with TRD, but that is not the same thing as having robust studies focused clearly on the patient population of greatest interest. The challenge will be to provide a definition that operationalizes TRD to make it feasible for clinicians while at the same time successfully capturing the complexity of treatment resistance.

More clinical trials, as well as other possible study designs, that compare nonpharmacologic interventions with other nonpharmacologic options and with pharmacologic treatments are necessary to inform decisionmaking in TRD. Clinicians, patients, and policymakers need additional relevant data to guide difficult treatment decisions about what to do next: try another medication (and should it be an augmentation, switch, or combination strategy?) or add (or switch to) rTMS, ECT, VNS, or psychotherapy?

Also, given that treatment options for many TRD patients include medications, trials should directly compare nonpharmacologic interventions with each other and with pharmacologic treatments.

The number of treatment failures in the current episode should be delineated carefully. This information, more likely to be accurate than lifetime histories of failures, can help investigators determine whether the particular number of failures, or reaching a particular number of failures in a current episode, can help differentiate between nonpharmacologic treatment choices. For example, for patients with two treatment failures in a current episode, the outcomes may not differ between cognitive therapy and rTMS; however, for patients with a different (higher or lower) number of treatment failures in the current episode, one nonpharmacologic treatment may indeed be better than the other. Currently, we do not know what the proper threshold is for selection of treatment. Clarification of the scientific basis for such a decision would substantially improve decisionmaking.

Clarifying whether responses differ for TRD patients with MDD compared with those with bipolar disorder will help guide future clinical trial design. Our decision to include trials with patient populations including up to 20 percent with bipolar disorder (i.e., the "mixed" populations noted earlier) was guided by clinical experience and common sense but not by data. Testing to see whether outcomes differ between the two groups can yield information about inclusion criteria (should the mix be 0 percent, 10 percent, 20 percent, etc.?) that may be useful to investigators in designing TRD trials and may be important to consider as a potential covariate in analyses involving such mixes.

Greater consideration should be given to the role that the spectrum of depressive severity plays. Using a finer gradation of depressive severity than investigators now typically employ might identify whether particularly severe degrees of depression, most commonly understood currently as a HAM-D17 ≥ 20, may respond differently to the available nonpharmacologic interventions than do less severe levels of depression. These gradations may lead clinicians to a better understanding of severe depression and its role in guiding treatment selection in TRD.

Direct comparisons of treatment strategies, holding consistent any coexisting or concomitant therapies, are imperative. Decisionmakers need to know whether outcomes with nonpharmacologic treatments are better when such a treatment augments the current treatment, replaces the current treatment, or replaces the current treatment in combination with another treatment. When ongoing treatment is uncontrolled and reflects a variety of treatments—e.g., some patients continue with atypical antipsychotics, some with mood stabilizers, some with no psychotropic medications—results of such studies are difficult, if not impossible, to interpret.

Consistent reporting of changes in depressive severity, response rates, and remission rates is crucial. To allow for better comparisons of clinical outcomes in this difficult-to-treat population, all three measures offer useful information for clinicians. Thus, for either clinical trials or observational studies, investigators should attempt to collect data on all three routinely.

Application of consistent, accepted protocols in trials is necessary. Making sure that patients receive equivalent doses of different nonpharmacologic interventions is more difficult than making sure of this for pharmacologic interventions. Nevertheless, investigators designing trials of nonpharmacologic therapies can attempt to do so by implementing standard accepted protocols for their trials. Such "dosing" had been difficult to control when that protocol was in the process of being developed, as with rTMS, but given current treatment parameters, this standardization is a goal well worth trying to reach.

More careful and consistent assessment of adverse events is required. Adverse event reporting is quite limited and tends to cover only a short time span; what reporting does exist is variable and inconsistent. Systematic collection and more consistent reporting of data on harms—that is, adverse events and negative side effects—and information about attrition and withdrawal would provide useful information to help balance information now focused on clinical benefits. Use of the CONSORT statement (available at: http://www.consort-statement.org/home/), which guides proper reporting of study information (including the presentation of adverse events), would strengthen reporting of both harms and other clinical trial findings; it would also aid in the critical appraisal and interpretation of all study results. Further, a more informative assessment of adverse events would require studies to be able to assess long-term and cumulative outcomes.

Including key relevant measures and subgroups in subsequent research is desirable. As indicated by the review, nearly no evidence exists on how the effectiveness of nonpharmacologic treatments differs (or not) as a function of symptom subtypes or for subgroups defined by sociodemographic characteristic (such as age) or coexisting medical conditions (e.g., post-stroke or postmyocardial infarction depression; perinatal depression). Also essentially missing is information about health-related outcomes, especially those reported by patients, that concern their quality of life or levels of functional impairment. Subsequent studies should focus on employing known, reliable, and valid measures of patient-reported outcomes, such as the MOS SF-36, the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), and the EQ-5D.

Including comparisons of newer nonpharmacologic interventions will be important in future research. As new nonpharmacologic treatments are developed and tested, investigators should try to include them as potential comparators. At the time we started this comparative effectiveness review, clinical trial data on some of the developing nonpharmacologic interventions, such as magnetic seizure therapy or deep brain stimulation, were insufficient (from the published literature) for us to try to include them. As the evidence bases grow to support the efficacy of such additional nonpharmacologic interventions, the newer strategies should be included in comparative effectiveness study designs.

Conclusion

Our review suggests that comparative clinical research on nonpharmacologic interventions in a TRD population is early in its infancy, and many clinical questions about efficacy and effectiveness remain unanswered. Interpretation of the data is substantially hindered by varying definitions of TRD and the paucity of relevant studies. The greatest volume of evidence is for ECT and rTMS. However, even for the few comparisons of treatments that are supported by some evidence, the strength of evidence is low for benefits, reflecting low confidence that the evidence reflects the true effect and indicating that further research is likely to change our confidence in these findings. This finding of low strength is most notable in two cases: ECT and rTMS did not produce different clinical outcomes in TRD, and ECT produced better outcomes than pharmacotherapy. No trials directly compared the likelihood of maintaining remission for nonpharmacologic interventions. The few trials addressing adverse events, subpopulations, subtypes, and health-related outcomes

provided low or insufficient evidence of differences between nonpharmacologic interventions. The most urgent next steps for research are to apply a consistent definition of TRD, to conduct more head-to-head clinical trials comparing nonpharmacologic interventions with themselves and with pharmacologic treatments, and to delineate carefully the number of treatment failures following a treatment attempt of adequate dose and duration in the current episode.

Full Report

This executive summary is part of the following document: Gaynes BN, Lux L, Lloyd S, Hansen RA, Gartlehner G, Thieda P, Jonas D, Brode S, Swinson Evans T, Crotty K, Viswanathan M, Lohr KN. 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://www.effectivehealthcare.ahrq.gov/reports/final.cfm.

For More Copies

For more copies of Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults: Executive Summary No. 33 (AHRQ Pub. No. 11-EHC056-1), please call the AHRQ Clearinghouse at 1-800-358-9295 or email ahrqpubs@ahrq.gov.

AHRQ Pub. No. 11-EHC056-1
September 2011