

Transcranial Magnetic Stimulation for Depression



Assessment
Program
Volume 24, No. 5
October 2009

Executive Summary

Background

Depression is a serious psychiatric condition that sometimes does not respond to standard treatments such as medication and/or psychotherapy. Transcranial magnetic stimulation (TMS) has been studied in patients with depression. The U.S. Food and Drug Administration (FDA) recently granted de novo marketing clearance for a specific TMS device for patients who have not responded to one adequate trial of antidepressant medication.

Objective

This Assessment will review the available evidence to determine if TMS therapy is effective for the treatment of depression.

Search Methods

A search of the MEDLINE® database (via PubMed) was completed for the period up through May 2009. The search strategy used the terms “transcranial magnetic” and “depression” as textwords or subject terms. Articles were limited to those published in English language and enrolling human subjects. The MEDLINE® search was supplemented by an examination of article bibliographies and relevant review articles, which were searched for citations.

Study Selection

The search was intended to review sham-controlled studies of TMS. To that end, meta-analyses of such trials were searched for and selected. The 5 most recent meta-analyses were selected for presentation. In light of FDA de novo 510(k) marketing clearance of a specific TMS device, the clinical trial of that device was selected for presentation.

Main Results

The largest and most recent meta-analysis reviewed 30 double-blind sham-controlled trials with a total of 1,164 patients. All the meta-analyses have some common limitations. All of the studies assess outcomes at the time of the end of TMS treatment, which is between 1 and 4 weeks. Only a few of the meta-analyses attempted to synthesize the relatively few studies that assessed outcome beyond the acute treatment period; but these periods tended to be short (i.e., 1 to 2 weeks post-treatment). The studies analyzed within each meta-analysis varied with respect to the anatomic location of treatment, treatment intensity, stimulation frequency, pulses per session, and total number of sessions.

Four of the 5 meta-analyses found statistically significant depression scores between TMS and sham when analyzing the main set of studies. The one meta-analysis that did not find a statistically significant difference had very strict selection criteria and, as a consequence, included only 6 studies with a total of 88 patients. Three of the meta-analyses summarized the individual studies using the standardized mean

NOTICE OF PURPOSE: TEC Assessments are scientific opinions, provided solely for informational purposes. TEC Assessments should not be construed to suggest that the Blue Cross Blue Shield Association, Kaiser Permanente Medical Care Program or the TEC Program recommends, advocates, requires, encourages, or discourages any particular treatment, procedure, or service; any particular course of treatment, procedure, or service; or the payment or non-payment of the technology or technologies evaluated.



An Association
of Independent
Blue Cross and
Blue Shield Plans



difference (SMD). The summary SMD for 2 of these meta-analyses was very similar: 0.39 and 0.35. The SMD allows pooling together studies with different outcome assessment methods, but does not allow an easy translation to measures of clinical response and remission. The one meta-analysis that pooled response and remission rates calculated a clinical response difference of 17% and a clinical remission difference of 14%. However, these differences arise from relatively low response and remission rates for TMS (25% response rate, 17% remission rate).

One meta-analysis compared the effect of TMS in studies evaluating patients with treatment-resistant and nontreatment-resistant depression, and found no difference in effect. In sum, the meta-analyses of sham-controlled studies of TMS are consistent with a short-term antidepressant effect of TMS of uncertain clinical significance. The studies give no information regarding the durability of the effect beyond the treatment period, which was generally between 1 and 4 weeks. The outcome measures used in the meta-analyses make it difficult to assess the clinical significance of the antidepressant effect in terms of response or remission rates.

The randomized trial reviewed in detail here is the largest sham-controlled trial of TMS with 301 evaluable subjects. The primary outcome was evaluated at 4 weeks and showed a 2.1-point difference on the Montgomery and Asberg Depression Rating Scale (MADRS), which was not statistically significant. Other outcomes evaluated at 4 weeks using other depression scales and a comparison of the difference in response rates showed a statistically significant difference. Remission rates did not differ between treatment arms at 4 weeks using any of the depression scales. At 4 weeks, patients who did not have a response to treatment could drop out of the study to enter an open-label treatment extension study. Due to a high proportion of imputed values beyond 4 weeks, results of the study beyond 4 weeks may not be reliable. Sensitivity analyses using imputed values rather than last-value-carried-forward values showed no difference in mean depression scores using any of the scales at 6 weeks.

A subgroup analysis of the clinical trial showed an interaction with amount of prior treatment failure. Treatment effects appeared to be mostly restricted to the group of patients with one prior treatment failure. Clinical response and remission rates were not reported for this subgroup analysis, however.

The open-label extension study treated nonresponders to the randomized trial with 6 weeks of active TMS. Mean change from baseline was greater than the randomized clinical trial, indicating either placebo effects or bias from expectations. Response rates ranged from 26 to 42%, and remission rates ranged from 11 to 27%, depending on original treatment assignment and depression scale. Given the open-label design and lack of a control group for the extension study, it is difficult to make inferences of effectiveness from this study.

In terms of safety, the major adverse effects of TMS are headache and pain or discomfort at the site of application of the device. The clinical trial did not appear to cause significant adverse effects that could be attributed to TMS. Other review articles examining the safety of TMS have not raised concerns about adverse effects beyond headache and pain or discomfort.

Authors' Comments and Conclusions

The randomized clinical trial of TMS does not show definitive evidence of efficacy for its primary endpoint at 4 weeks. Not all outcomes show efficacy, and the analysis is sensitive to alternative methods of analysis. Another limitation of this and other studies of TMS is lack of rigorous evaluation beyond the period of treatment. Although short-term studies are consistent with changes in depression scores due to TMS, the clinical significance and durability of the effect are not well characterized. One meta-analysis indicated no difference in effect between patients with treatment-resistant and nontreatment-resistant depression. The randomized, clinical trial showed a greater effect in patients with only one prior treatment failure, with possibly minimal or no effect in patients with greater than one prior treatment failure.

The indication for which TMS received approval, one prior failure of an adequate antidepressant course, is unusual. A change in antidepressant therapy is usually indicated at this point and has been shown to have a success rate similar to the first course. The current body of evidence cannot determine in a rigorous way whether TMS would be as effective as a second course of antidepressant therapy. Other important gaps in current knowledge include whether TMS is effective as an adjunctive treatment to second-line drug therapy, the durability of TMS treatment, and the effectiveness of retreatment. A clinical trial sponsored by the National Institute of Mental Health has recruited subjects for another clinical trial of TMS. However, this trial also appears to have only a short duration (3 weeks) in which the participants are randomized to TMS or sham before crossovers or alternative treatments are offered.

Based on the available evidence, the Blue Cross and Blue Shield Association Medical Advisory Panel made the following judgments about whether TMS for the treatment of depression meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. The technology must have final approval from the appropriate governmental regulatory bodies.

Devices for transcranial stimulation have received clearance by the U.S. Food and Drug Administration (FDA) for diagnostic uses. One device, NeoPulse (Neuronetics, Atlanta, GA) received approval in Canada and Israel as a therapy for depression. Although initially examined by the U.S. Food and Drug Administration (FDA) under a traditional 510(k) application, the NeoPulse, now known as NeuroStar® TMS, received 510(k) clearance for marketing as a “de novo” device (assessed as low risk, no predicate device) in 2008. NeuroStar® TMS is indicated for the treatment of patients with depression who have failed one 6-week course of antidepressant medication.

2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

An important limitation of the evidence is lack of information beyond the acute period of treatment. Most of the clinical trials evaluate the outcomes at the point of the last TMS treatment, between 1 and 4 weeks. Very few studies evaluated patients beyond this time period. Although meta-analyses are consistent with short-term antidepressant effects, the clinical significance of the effect is uncertain. The large clinical trial of TMS reviewed in this assessment did not unequivocally demonstrate efficacy, as the principal endpoint was not statistically significant at 4 weeks, and some results were sensitive to the methods of analysis. The patients in whom TMS is indicated are usually treated with a second course of antidepressant therapy. The clinical trial, which was sham controlled without active treatment, cannot determine whether TMS would be more or less successful than this standard treatment.

3. The technology must improve the net health outcome; and

4. The technology must be as beneficial as any established alternatives.

The available evidence does not permit conclusions regarding the effect of TMS on health outcomes or compared with alternatives. Comparison to alternatives using other observational studies may not be valid due to unmeasured differences in severity of depression between studies and other differences in studies.

5. The improvement must be attainable outside the investigational settings.

It has not yet been demonstrated whether TMS improves health outcomes in the investigational setting. Therefore, it cannot be demonstrated whether improvement is attainable outside the investigational settings.

For the above reasons, transcranial magnetic stimulation for the treatment of depression does not meet the TEC criteria.

Contents

Assessment Objective	5	Review of Evidence	9
Background	5	Discussion	17
Methods	8	Summary of Application of the Technology Evaluation Criteria	17
Formulation of the Assessment	8	References	19

**Published in cooperation with Kaiser Foundation Health Plan and
Southern California Permanente Medical Group.**

TEC Staff Contributors

Author—David H. Mark, M.D., M.P.H.; **TEC Executive Director**—Naomi Aronson, Ph.D.; **Director, Clinical Science Services**—Kathleen M. Ziegler, Pharm.D.; **Research/Editorial Staff**—Claudia J. Bonnell, B.S.N., M.L.S.; Maxine A. Gere, M.S.

Blue Cross and Blue Shield Association Medical Advisory Panel

Allan M. Korn, M.D., F.A.C.P.—Chairman, *Senior Vice President, Clinical Affairs/Medical Director, Blue Cross and Blue Shield Association*; **Alan M. Garber, M.D., Ph.D.**—Scientific Advisor, *Staff Physician, U.S. Department of Veterans Affairs*; **Henry J. Kaiser, Jr., Professor, and Professor of Medicine, Economics, and Health Research and Policy, Stanford University**; **Steven N. Goodman, M.D., M.H.S., Ph.D.**—Scientific Advisor, *Professor, Johns Hopkins School of Medicine, Department of Oncology, Division of Biostatistics (joint appointments in Epidemiology, Biostatistics, and Pediatrics).*

■ **Panel Members** **Peter C. Albertsen, M.D., Professor, Chief of Urology, and Residency Program Director, University of Connecticut Health Center**; **Sarah T. Corley, M.D., F.A.C.P., Chief Medical Officer, NexGen Healthcare Information Systems, Inc.**—American College of Physicians Appointee; **Helen Darling, M.A. President, National Business Group on Health**; **Josef E. Fischer, M.D., F.A.C.S., William V. McDermott Professor of Surgery, Harvard Medical School**—American College of Surgeons Appointee; **I. Craig Henderson, M.D., Adjunct Professor of Medicine, University of California, San Francisco**; **Mark A. Hlatky, M.D., Professor of Health Research and Policy and of Medicine (Cardiovascular Medicine), Stanford University School of Medicine**; **Saira A. Jan, M.S., Pharm.D., Associate Clinical Professor, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Residency Director and Director of Clinical Programs Pharmacy Management, Horizon Blue Cross and Blue Shield of New Jersey**; **Leslie Levin, M.B., M.D., F.R.C.P.(Lon), F.R.C.P.C., Head, Medical Advisory Secretariat and Senior Medical, Scientific and Health Technology Advisor, Ministry of Health and Long-Term Care, Ontario, Canada**; **Bernard Lo, M.D., Professor of Medicine and Director, Program in Medical Ethics, University of California, San Francisco**; **Barbara J. McNeil, M.D., Ph.D., Ridley Watts Professor and Head of Health Care Policy, Harvard Medical School, Professor of Radiology, Brigham and Women's Hospital**; **William R. Phillips, M.D., M.P.H., Clinical Professor of Family Medicine, University of Washington**—American Academy of Family Physicians' Appointee; **Alan B. Rosenberg, M.D., Vice President, Medical Policy, Technology Assessment and Credentialing Programs, WellPoint, Inc.**; **Maren T. Scheuner, M.D., M.P.H., F.A.C.M.G., Director, Genomics Strategic Program Area, VA HSR&D Center of Excellence for the Study of Healthcare Provider Behavior, VA Greater Los Angeles Healthcare System; Natural Scientist, RAND Corporation; Adjunct Associate Professor, Department of Health Services, UCLA School of Public Health**; **J. Sanford Schwartz, M.D., F.A.C.P., Leon Hess Professor of Medicine and Health Management & Economics, School of Medicine and The Wharton School, University of Pennsylvania**; **Earl P. Steinberg, M.D., M.P.P., President and CEO, Resolution Health, Inc.**; **Robert T. Wanovich, Pharm.D., Vice-President, Pharmacy Affairs, Highmark, Inc.**; **A. Eugene Washington, M.D., M.Sc., Executive Vice Chancellor and Provost, University of California, San Francisco**; **Jed Weissberg, M.D., Senior Vice President, Quality and Care Delivery Excellence, The Permanente Federation.**

CONFIDENTIAL: This document contains proprietary information that is intended solely for Blue Cross and Blue Shield Plans and other subscribers to the TEC Program. The contents of this document are not to be provided in any manner to any other parties without the express written consent of the Blue Cross and Blue Shield Association.

Assessment Objective

Depression is a serious psychiatric condition that sometimes does not respond to standard treatments such as medication and/or psychotherapy. Transcranial magnetic stimulation (TMS) has been studied as a potential treatment of patients with depression.

Background

Depression

Depression is a very common disorder that is most often chronic or recurrent in nature. In the U.S., the lifetime prevalence of major depressive disorder (MDD) is approximately 16%, and the 12-month period prevalence of MDD is approximately 7% (Kessler et al. 2005). Depression is associated with significant morbidity for the patient, patient's family, and society. Among the consequences of depression are functional impairment, impaired family and social relationships, and increased mortality from suicide and comorbid medical disorders.

Although there are many effective treatments for depression, individual response to treatment is variable. Several studies conducted in recent years have attempted to better elucidate rates of response to antidepressant treatments. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study was designed to assess effectiveness of treatments in generalizable samples and to study effectiveness of alternative treatments (Rush et al. 2004). In this study, patients who failed to achieve a remission of depression using a standard antidepressant medication were offered further participation in other treatment alternatives (often randomized to acceptable alternatives).

Definitions of treatment-resistant depression vary, but in general, the term refers to patients who have not responded to adequate trials of one or more treatment strategies. Some authors suggest using different terms for differing levels of poor response to treatment. For example, Thase and Rush (2000) would use the term "treatment-refractory" depression to describe a situation in which the depression has not achieved a response to at least two adequate trials of medication from two different medication classes. Methods for assessing outcomes of depression vary, as well. In general, treatment "response" is usually defined as a 50% reduction in depressive symptoms as measured

by one of the standard scales for measuring depression such as the Hamilton Depression Rating Scale (HAMD). "Remission" is a more stringent measure of treatment effect and is usually defined as achieving a specific threshold low score on the HAMD or other scale.

Cadieux (1998) outlined 5 strategies for treating partial response or nonresponse to antidepressant therapy: 1) optimizing current therapy by dosage and duration; 2) substitution with different classes of pharmaceutical agents; 3) combining drugs; 4) electroconvulsive therapy; and 5) augmentation with drugs not routinely regarded as antidepressants, such as lithium, thyroid hormone, or pindolol.

In the first treatment step of the STAR*D trial (Trivedi et al. 2006), patients were treated for up to 12 weeks with citalopram at sufficient doses. Remission rates were 28–33% based on differing depression scales and the response rate was 47%. Thus, slightly more than one-half of patients with major depression did not respond to an adequate trial of a single antidepressant treatment. In the next level of the STAR*D trial, patients who did not achieve remission were assigned or randomized to alternative treatments such as augmentation of citalopram with another antidepressant or switching to another medication. The overall remission rate for level 2 treatments was 30.6% (Rush et al. 2006). Succeeding levels of sequenced treatment had lower remission rates (13.7% for the third step, 13.0% for the fourth step).

A simple estimation of cumulative remission based on the remission rate of each step and assuming no drop-outs was calculated to be 56% after 2 steps and 67% after 4 steps. Thus, it appears that over half of the patients entering treatment with major depression can achieve successful remission with two trials of adequate antidepressant treatment. It should be noted that in the original entry criteria for the STAR*D trial, only patients with a clear history of nonresponse or intolerance to any of the treatments offered in the first two treatment steps in the current depressive episode were excluded. A detailed treatment history of the current depressive episode was apparently not recorded, but it was known if the patient had received treatment during the current episode. Most patients had received prior treatment during the current depressive episode, but it is not noted if it was considered an adequate trial of treatment.

The results from the STAR*D trial indicate that response and remission rates of antidepressant therapy depend very much on the prior treatment history of the patients, and that progressive levels of treatment resistance (as assessed in this trial by failing successive treatments) predict lower remission rates.

Outcome Assessment in Depression

There are several reasons to measure depressive symptoms in clinical practice or research. Most important for the purposes of this Assessment is to assess treatment outcome. In the research evaluating depression treatments, the 4 most common instruments used are shown in Table 1. Except for the Clinical Global Impression (CGI) scale, the other instruments all involve answering specific inquiries regarding patient symptoms such as mood, affect, energy, appetite, sleep, and suicidal or paranoid ideation. Scale scores have been calibrated with clinical evaluation to correlate to severity levels of depression and changes in scores consistent with good treatment response or remission of depression. A typical threshold for categorizing clinically significant improvement or treatment response is a 50% reduction from baseline score for any of the scales.

The other method for categorizing treatment is for the final value of the test to be below a particular value, indicating very few depressive symptoms. A score below a particular value is often declared to be remission of depression. Remission is generally a more stringent definition of treatment outcome. Studies have shown reasonable concordance between the Inventory of Depressive Symptomatology (IDS) and the HAMD in categorizing patients as treatment responders (Rush et al. 2004).

Except for the self-administered version of the Inventory of Depressive Symptomatology (IDS-SR), the depression rating scales discussed here are designed to be administered by a health care clinician. The CGI scale is the most inherently subjective scale, as it is simply a categorical 7-point scale asking whether, in the judgment of the clinician, the patient is very much improved (CGI=1) to very much worse (CGI=7). The CGI also requires an in-depth knowledge of the patient over the course of the treatment period.

Transcranial Magnetic Stimulation

Transcranial magnetic stimulation (TMS) was first introduced in 1985 as a new method

of noninvasive stimulation of the brain. The technique involves placement of a small coil over the scalp; a rapidly alternating current is passed through the coil wire, producing a magnetic field that passes unimpeded through the scalp and bone, resulting in electrical stimulation of the cortex. TMS was initially used to investigate nerve conduction; for example, TMS over the motor cortex will produce a contralateral muscular-evoked potential. This “motor threshold” (MT), which is the minimum intensity of stimulation required to induce a motor response, is empirically determined for each individual by gradually increasing the intensity of stimulation.

Interest in the use of TMS as a treatment for depression was augmented by the development of a device that could deliver rapid, repetitive stimulation. Imaging studies had showed a decrease in activity of the left dorsolateral prefrontal cortex (DLPFC) in depressed patients, and early studies suggested that high-frequency (e.g., 5–10 Hz) TMS of the left DLPFC had antidepressant effects. Low-frequency (1–2 Hz) stimulation of the right DLPFC has also been investigated. The rationale for low frequency TMS is inhibition of right frontal cortical activity to correct the interhemispheric imbalance. A combination approach (bilateral stimulation) is also being explored. TMS is also being tested as a treatment for other disorders including schizophrenia, migraine, spinal cord injury, tinnitus, and fibromyalgia. In contrast to electroconvulsive therapy, TMS does not require anesthesia and does not induce a convulsion.

In clinical trials of TMS placing the device at the left DLPFC, essential aspects of the treatment that have varied include the following characteristics:

- stimulation frequency
- stimulation intensity, measured as percent of motor threshold
- total pulses per session
- total number of sessions (usually 5 days per week, 1-week increments)
- concurrent medication: no, yes, or introduction of new medication

FDA Status. Devices for transcranial stimulation have received clearance by the U.S. Food and Drug Administration (FDA) for diagnostic uses. One device, NeoPulse (Neuronetics, Atlanta, GA) received approval in Canada and Israel as a therapy for depression. Although initially examined by the U.S. Food and Drug

Table 1. Depression Rating Scales

Scale	Brief Description	Common Thresholds for Treatment Response	Common Thresholds for Remission
Hamilton Rating Scale for Depression (HAM-D)	Observer-rated scoring of 28, 24, 21, or 17 two- to four-point items	50% reduction from baseline	HAM-D 17<8 HAM-D 24<11
Clinical Global Impression (CGI)	Observer-rated 7-point scale 1) Severity of Illness (CGI-S) 2) Global Improvement (CGI-I)	Global Improvement=1 or 2	
Montgomery and Asberg Depression Rating Scale (MADRS)	Observer-rated scoring of 10 six-point items	50% reduction from baseline	Score <10
Inventory of Depressive Symptomatology (IDS)	Observer-rated (IDS-C) or self-rated (IDS-SR) 30 three-point items	50% reduction from baseline	

Administration (FDA) under a traditional 510(k) application, the NeoPulse, now known as NeuroStar® TMS, received 510(k) clearance for marketing as a “de novo” device (assessed as low risk, no predicate device) in 2008. NeuroStar® TMS is indicated for the treatment of patients with depression who have failed one 6-week course of antidepressant medication.

An FDA advisory panel met in January 2007 to determine if the risk-to-benefit profile for the NeoPulse was comparable to the risk-to-benefit profile of electroconvulsive therapy devices. The panel was not asked for a recommendation regarding the regulatory determination of substantial equivalence for this 510(k) submission.

Methods

Search Methods

A search of the MEDLINE® database (via PubMed) was completed for the period up through May 2009. The search strategy used the terms “transcranial magnetic” and “depression” as textwords or subject terms. Articles were limited to those published in English language and enrolling human subjects. The MEDLINE® search was supplemented by an examination of article bibliographies and relevant review articles, which were searched for citations.

Study Selection

In light of the large body of clinical trial evidence on this particular subject, much of which does not reflect the method in which the available FDA-cleared device is intended to be used, the following data sources were selected for inclusion. There have been at least 8 meta-analyses of TMS published. The most recent meta-analyses, published between 2003 and 2008 will be presented and discussed. The meta-analyses published in 2009 by Schutter and in 2008 by Lam et al. contain a considerably larger number of studies than the other meta-analyses and will be presented in more detail.

The other study to be presented in detail is the study by O’Reardon et al. (2007), which is the clinical trial of the Neuronetics device cleared by the FDA, which is the only study that employs TMS in the manner currently intended for use. Further data from this study are published in studies by Avery et al. (2008), Lisanby et al. (2009), and the FDA Executive Summary.

Medical Advisory Panel Review

Current Assessment. This Assessment was reviewed by the Blue Cross and Blue Shield Association Medical Advisory Panel (MAP) on June 25, 2009. In order to maintain the timeliness of the scientific information in this Assessment, literature searches were performed subsequent to the Panel’s review (see “Search Methods”). If the search updates identified any additional studies that met the criteria for detailed review, the results of these studies were included in the tables and text where appropriate. There were no studies that would change the conclusions of this Assessment.

Formulation of the Assessment

Patient Indications

The FDA labeling for the device indicates that it is indicated for patients who have failed one course of antidepressant therapy. This determination was based on a subgroup analysis of a clinical trial showing that benefit of the treatment was restricted to that particular subgroup. This is an unusual indication, as a more standard definition of treatment-resistant or treatment-refractory depression is patients who have failed to respond to two adequate courses of antidepressant therapy. As noted previously, in the STAR*D trial, patients who had failed one course of antidepressant therapy had an overall remission rate of 30.6% in their second course of therapy, a rate almost as high as the first course of treatment. The clinical trial reviewed here enrolled both patients who had failed to respond to one course and patients who had failed to respond to more than one course of therapy.

Technologies to Be Compared

TMS has been compared to sham placebo. TMS has not been evaluated or compared to a course of antidepressant therapy in a published randomized, clinical trial.

TMS has been evaluated in sham-controlled trials with no other concurrent medications, stable concurrent medications, and as an adjunct to initiation of medication. If used as an adjunct to antidepressant therapy, it should be compared to antidepressant therapy plus sham TMS. In the randomized, clinical trial that was reviewed by the FDA, no concurrent antidepressant therapy was used. However, antidepressants were introduced upon tapering TMS

treatments at 6 weeks. There is no established standard course of treatment for TMS; thus the treatment as given in the reviewed randomized, controlled trial shall be assumed to be the one evaluated in this TEC Assessment.

Although TMS shares some characteristics of electroconvulsive therapy (ECT), ECT is not indicated for patients who have failed a single adequate trial of antidepressant therapy. Given that the FDA clearance is for patients who have failed a single adequate trial of antidepressant therapy, ECT is not an appropriate comparator treatment for these patients.

Health Outcomes

Potential Benefits. The primary outcome to be evaluated is relief of depressive symptoms. This outcome can usually be assessed by any one of many different depression symptom rating scales. A 50% reduction from baseline score is considered to be a standard measure of treatment response. Reduction of symptoms to a specific absolute score on a specific depression instrument is also considered a standard measure of remission.

Potential Harms. Harmful outcomes of TMS would include any morbidity, complications, or discomfort associated with the treatment.

Specific Assessment Questions

1. What is the effect of TMS therapy on treatment-resistant depression?
2. What are the adverse effects of TMS therapy?

Review of Evidence

Meta-analyses

We selected the 5 most recently published meta-analyses for presentation and review (Table 2). Besides differing in year of publication and thus availability of different studies, the meta-analyses varied in selection criteria and analytic technique. However, there are certain characteristics of all the studies that are important to note.

The principal outcomes of all the meta-analyses assess the clinical endpoints at the conclusion of TMS treatment only (i.e., 1 to 4 weeks). Only the meta-analyses by Martin et al. (2003) and Lam et al. (2008) analyze any outcomes beyond the acute treatment period, but there were relatively few such studies. Some studies included in the meta-analyses were cross-over

studies that preclude any analysis of outcome beyond the acute treatment period; for such studies, only the initial period of treatment was included in the meta-analysis.

The included studies also varied with regard to the use of concurrent antidepressant therapy. In the studies included by the meta-analysis by Lam et al. (2008), most studies maintained patients on the same medications, some studies kept patients off all medications for a limited time or for the whole duration of treatment, and a few studies initiated treatment concurrent with the initiation of TMS. None of the meta-analyses considered this variation in use of medication in subgroup or meta-regression analyses.

The largest and most recent meta-analysis by Schutter (2009) included 30 double-blind sham-controlled trials with 1,164 patients. The clinical trial by O'Reardon et al. (2007), to be reviewed in detail in this Assessment, is included in this meta-analysis and contributes over one-fourth of the total number of patients. Studies enrolling patients with major depression, employing high frequency (>5 Hz) TMS over the left DLPFC, a minimum of 5 treatment sessions, and measuring pretreatment and post-treatment depression scores using either the HAMD or MADRS were selected. In addition to calculating an overall treatment effect, they performed additional analyses comparing effects between studies that treated patients with medication-resistant depression only (n=17) and patients with non-medication resistant depression (n=8), and studies that used lower-intensity MT intensity (<100% MT, n=14) versus studies that used higher intensity (100–120% MT, n=16).

The overall weighted mean effect size for treatment was 0.39 (95% CI: 0.25–0.54, $p < 0.0001$). An analysis comparing studies with patients with treatment-resistant depression versus nontreatment-resistant depression found no difference in effect size. Treatment intensity, comparing studies using high versus low MT intensity, also did not show a difference in effect size. The study by O'Reardon et al. (2007) appears to have a comparable, but just slightly lower effect size than the overall meta-analytic effect size, as estimated by visual inspection. A problem of assessing outcome in a study using the weighted effect size is that it is not easy to translate this outcome to a clinically meaningful outcome such as a response rate or remission rate difference or relative risk.

Table 2. Summary of Meta-Analyses of Sham-Controlled Trials of TMS

Author	Number of Trials	Year Range of Trials	Total Sample Size (range of studies n)	Main Selection Criteria	Range of Treatment Parameters	Main Results	Additional Analyses/Notes
Schutter 2009	30	1997–2007	1,164 (6–301)	Double-blind >5 Hz Left DLPFC location >80% Motor threshold (MT) HAMD or MADRS outcome on continuous scale	Left DLPFC 5–20 Hz 80–120% MT 1–4 weeks	Weighted effect size 0.39, p<0.0001	Medication resistance vs. non-medication resistance, NS difference High vs. low MT intensity, NS difference
Lam et al. 2008	24	1999–2007	1,092 (10–301)	Treatment-resistant depression Clinical response outcome (threshold % response)	Left or right DLPFC R: 1 Hz; L: 10–20 Hz 80–120% MT 1–4 weeks	Risk difference for response 17% (CI: 10–23) Risk difference for remission 14% (CI: 6–21) Pooled response rate TMS 25% Sham 9% Pooled remission rate TMS 17% Sham 6% Standardized mean difference (SMD) 0.48 (CI: 0.28–0.69)	Risk difference in studies with strict definition for treatment resistance, 16% Risk difference in high quality studies, 15% Risk difference in high-intensity left TMS, 19% 4 studies with follow-up beyond treatment, SMD=1.04 (CI: 0.08–2.01)
Gross et al. 2007	5	2005–2006	274 (27–99)	Any location or frequency Continuous outcome measure	Left, right, or bilateral R: 1 Hz; L: 10–20 Hz 90–110% MT 2–3 weeks	Standardized mean difference 0.76 (CI: 0.51–1.01)	Compared to prior published meta-analysis of Martin et al. 2003, which had SMD of 0.35

Table 2. Summary of Meta-Analyses of Sham-Controlled Trials of TMS (cont'd)

Author	Number of Trials	Year Range of Trials	Total Sample Size (range of studies n)	Main Selection Criteria	Range of Treatment Parameters	Main Results	Additional Analyses/Notes
Couturier 2005	6	1996–2003	91 (6–30)	>10 Hz Duration 5-10 days >80% MT HAMD-21 outcome only Major depressive episode No initiation of antidepressant treatment	Left DLPFC 10–20 Hz 80–110% MT 1–2 weeks	Weighted mean difference 1.1 (CI 4.5 to -2.3; not standardized)	Outcome not standardized, in units of HAMD-21
Martin et al. 2003	14	1997–2001	324 (6–70)	No restrictions other than sham-controlled	Left DLPFC (except 1 study) 10–20 Hz 80–110% MT 1–2 weeks	Standardized mean difference HAMD, 2 weeks treatment (n=9) 0.35 (CI 0.04–0.66, p=0.03) Standardized mean difference HAMD, 1 week treatment (n=5) 0.18 (CI: -0.27–0.64, p=0.5) Standardized mean difference, HAMD, 1 week post-treatment (n=2) 0.08 (p=0.8) Standardized mean difference HAMD, 2 weeks post-treatment (n=3) 0.33 (p=0.2)	Beck depression inventory also analyzed; all results show no difference between TMS and sham

Abbreviations: DLPFC: dorsolateral prefrontal cortex; HAMD: Hamilton Rating Scale for Depression; MT: motor threshold

The next largest meta-analysis by Lam et al. (2008) examined only studies with patients with treatment-resistant depression. They identified 24 trials enrolling 1,092 patients. The study by O'Reardon et al. (2007) is also included in this meta-analysis. It was noted that the studies used various criteria for the definition of treatment-resistant depression. Nine studies used the criteria of failing one or more trials of antidepressants. The rest of the studies used a definition of failing 2 or more trials of antidepressants.

This meta-analysis synthesized the reported response and remission rates of the studies and calculated a summary risk difference. The summary risk difference for clinical response was 17% (95% CI: 10–23%, n=22 studies) and the summary risk difference for clinical remission was 14% (95% CI: 6–21%, n=16 studies). The weighted standardized mean effect size was 0.48 (95% CI: 0.28–0.69, n=21 studies). The pooled response rates were 25% for TMS versus 9% for sham, and the pooled remission rates were 17% for TMS versus 6% for sham. The study by O'Reardon et al. (2007) had a smaller risk difference for response (7.1%) and remission (0.9%) than the summary risk differences.

They found similar summary risk difference estimates in the subsets of studies with the more strict definition (2 or more failures) of treatment resistance (n=14 studies), higher-quality studies (n=10 studies), and the studies using high-intensity stimulation (n=15 studies). Eight studies were found that followed patients beyond the acute treatment period, but only 4 of them could be quantitatively synthesized. The period of follow-up was only 1 to 2 weeks in these 4 studies, but the standardized mean effect size was rather large at 1.04 (95% CI: 0.08–2.01). However, they noted one study that had an extremely large effect of 2.67. When excluding that study, the standardized mean effect size was still significant at 0.58 (95% CI: 0.13–1.02). They note that there were two studies with slightly longer follow-up at 6 weeks and 9 weeks, and in these two studies there no differences in depression scores at these latest follow-up times.

The other recent meta-analyses reviewed here include many fewer studies than those outlines above, either because of particular selection criteria or the limited amount of studies available at that time. Gross et al. (2007) selected only the most recent studies available at the time (from December 2005 to November 2006),

in order to compare to a prior meta-analysis by Martin et al. (2003). They found 5 such studies, but most of them were not included in the two meta-analyses discussed above, because they either did not use left DLPFC as the treatment location or did not study patients with treatment-resistant depression. The studies included in this meta-analysis, other than their sharing a close publication date, are very heterogeneous in most other respects. The standardized mean difference was 0.76 (95% CI: 0.51–1.01). The authors note that this mean difference is larger than the mean difference from the meta-analysis of Martin et al. (2003), suggesting that recent TMS trials show larger antidepressant effects than earlier studies. However, certain aspects of these studies are quite different than most other studies of TMS. Three of the 5 studies used the right DLPFC as a treatment site (2 bilateral, 1 right only).

It may not be valid to directly compare the results of this meta-analysis to the earlier one by Martin et al. (2003) due to differences in analytic pooling techniques. Martin et al. (2003) state that imbalances in the baseline severity of depression between treatment and sham groups were noted in several of the studies (the active treatment groups having greater severity of depression). Analyzing the studies using the difference between baseline and final scores would have biased the analysis in favor of active treatment because of regression to the mean. Thus, Martin et al. pooled only the final values of depression severity between active and sham groups.

Couturier (2005) limited the meta-analysis to studies using high frequency (>10 Hz) left DLPFC treatment site and a treatment duration of 5–10 days. The principal criterion that limited the number of studies in this meta-analysis was that the study was required to report the outcome using the HAMD 21 scale. Based on these selection criteria, 6 studies with a total of 88 patients were included in the meta-analysis. The summary weighted mean difference was 1.1 (95% CI: 4.5 to -2.5; not standardized) which was not statistically significant. The study has less statistical power than the other meta-analyses due to its particular selection criteria, but the magnitude of the effect, a 1.1-point difference on the HAMD 21 is not consistent with a clinically important effect of TMS.

Martin et al. (2003) used broad inclusion criteria for their review of TMS, and included a total of

14 clinical trials. Studies using TMS at any frequency and any localization compared to a sham control were analyzed. In evaluating the quality of the included studies, they noted that only 2 studies undertook an intention-to-treat analysis using the last-observation-carried-forward method to impute values for withdrawn patients.

Pooled analyses were carried out for subsets of the studies that shared common characteristics. In nine studies that used left DLPFC placed stimulation for 2 weeks, the standardized mean difference was 0.35 (95% CI: 0.04–0.66, $p=0.05$), showing a difference in favor of TMS. In 5 studies using left DLPFC stimulation for 1 week, the standardized mean difference was 0.18 (95% CI: -0.27–0.64, $p=0.4$). For assessment of post-treatment effects, standardized mean differences at 1 and 2 weeks were not statistically significant, but very few studies contributed patients to this analysis. Seven studies used the Beck Depression Inventory as an outcome measure. None of the standardized mean differences were significant for a treatment effect.

In sum, these set of recent published meta-analyses are mostly consistent for a statistically significant difference in improvement in depression in TMS patients compared to sham patients for outcomes measured at the last TMS treatment (generally 1–2 weeks). The improvement appears modest overall, and may not have changed significantly over time. The effect size calculated by Schutter (2009) of 0.39 is very similar to the 0.35 calculated by Martin et al. (2003). The recent study by O'Reardon et al. (2007), employing the FDA-cleared device using 4–6 weeks of high-frequency, high-intensity treatment, does not appear to produce greater effects than the other studies, as its effect size and risk difference estimates are actually lower than the pooled means of the meta-analyses in which it was included (Schutter 2009; Lam et al. 2008).

The metric of standardized mean difference as used in most of these meta-analyses is difficult to translate to more clinically meaningful outcome measures such as the response rate or the remission rate. The one meta-analysis pooling data to estimate the risk difference (Lam et al. 2008) shows reasonably sized risk differences (17% risk difference in response), but in the context of overall fairly low response rates. The greatest shortcoming of the meta-analyses is the lack of sufficient studies

examining patient outcomes beyond the time period of acute treatment with TMS.

Randomized Clinical Trial

The study by O'Reardon et al. (2007) is the largest clinical trial of TMS, and was the clinical trial on which the FDA decision to clear TMS for marketing was based. The study was a multicenter study conducted at 23 sites. Patients were enrolled who had uncomplicated major depression meeting severity criteria. They were required to have failed at least one but no more than four adequate antidepressant treatments in the current or most recent episode of depression.

After a 1-week washout, patients were scheduled to have 6 weeks of TMS, 5 sessions per week. Characteristics of the treatment given were left DLPFC treatment location, 120% motor threshold field intensity, 10 Hz pulse frequency, and 3000 pulses per treatment. Patients were not given antidepressant medication during the treatment period. Three-hundred twenty-five patients were randomized, and the analysis presented in the published study is based on 301 patients who had at least one post-baseline assessment.

The study had adequate randomization and assembly of comparable groups into the two treatment arms, and blinding of treating physicians and outcome assessors. However, there were some nonevaluable patients, and a last-value-carried-forward technique was used for missing values. The design of the trial allowed unblinding and cross-over after 4 weeks for nonresponders; thus the trial was no longer a randomized trial after 4 weeks.

After 4 weeks of treatment with active or sham TMS, if the patient showed no or minimal improvement in depression symptoms (e.g., less than 25% reduction in baseline symptoms on the HAMD 17), they could cross over to an open-label, acute treatment-extension study. Thus although treatment was planned for 6 weeks, drop out could occur at 4 weeks according to protocol. At 6 weeks, TMS treatment was tapered over a period of 3 weeks, and treatment with an antidepressant medication was initiated. FDA documents indicate that large numbers of patients dropped out after 4 weeks to enter the open-label trial. Since the analysis used a last-observation-carried-forward technique, many of the 6-week observation values reflect 4-week values.

Table 3 shows the results of the trial's 4-week outcomes as reported in the published paper and the principal analysis of the FDA Executive Summary (U.S. Food and Drug Administration 2007). The mean difference in 4-week change in MADRS was declared a priori to be the primary outcome of the study. The TMS patients improved 5.6 points on the MADRS, and the sham group improved 3.5 points, leading to a mean difference of 2.1 points which was not statistically significant ($p=0.057$). On other outcomes measured by mean changes from baseline, the HAMD 24 had a mean difference 2.4 ($p=0.012$), the HAMD 17 had a mean difference of 1.9 ($p=0.006$) and the CGI-S had a mean difference of 0.4 ($p=0.009$).

Results were also presented in terms of response and remission rates. As Table 3 shows, response rates were generally low, but depending on the scale used, the differences in response rates ranged from 7.1 to 9%, and were statistically significant. The remission rates were even lower, between 7.1% and 9.0%, and were not significant when compared to sham-treated patients.

In the FDA executive summary report of this trial (U.S. Food and Drug Administration 2007), the FDA requested reporting of alternative analyses using multiple imputation rather than last-observation-carried-forward, as an initial response followed by a later decline could bias the results in favor of TMS. There is only limited reporting of these analyses presented in the report, however. For the primary outcome of mean difference in MADRS change at 4 weeks, the multiple imputation method applied to all

325 randomized patients produced a p value of 0.090. The other analyses comparing HAMD 17 and HAMD 24 scores still had significant p values at 4 weeks. There is an analysis reporting 4-week outcomes of 6-week completers only, but this analysis should be discounted because nonresponders from both treatment groups selectively withdrew from the study to enter the open label trial.

The published paper also reports 6-week results for all outcomes, all of which consistently show greater efficacy of TMS at this timepoint than at the 4-week timepoint. For example, the MADRS remission rates increase to 14.2% in the TMS group and decrease to 5.5% in the sham group, now producing a difference of 8.7% which is statistically significant. However, these 6-week results should be viewed with caution, as they are based on 49% (148/301) values imputed from prior visits. However, in the FDA Executive Summary, the FDA requested an analysis of 6-week outcomes using multiple imputation method for missing values rather than the last observation carried forward. When comparing the mean changes from baseline for outcomes estimated at 6 weeks, the p values for MADRS, HAMD 24, and HAMD 17 are not significant ($p=0.276$, 0.73, 0.078, respectively).

Subgroup Analyses of the Clinical Trial.

Subgroup analysis of this clinical trial is reported in the FDA Executive Summary and the study by Lisanby et al. (2009) (Table 4). An essential finding of this subgroup analysis is the results stratified by the history of prior adequate trials of antidepressant treatment. The

Table 3. Outcomes of TMS Clinical Trial at 4 Weeks, Published Results

Outcome	TMS	Sham	p value
MADRS final total score	27	29.8	0.057
HAMD17 final total score	17.4	19.4	0.006
HAMD24 final total score	23.4	25.9	0.012
MADRS response rate (50% impr)	18.1%	11.0%	<0.05
HAMD17 response rate	20.6%	11.6%	<0.05
HAMD24 response rate	19.4%	11.6%	<0.05
MADRS remission rate (score<10)	7.1	6.2	>0.10
HAMD17 remission rate (score<8)	7.2	6.2	>0.10
HAMD24 remission rate (score<11)	9.0	8.2	>0.10

mean number of adequate trials of therapy in the trial was 1.6, with 54% (164/301) of subjects having failed just one adequate trial.

The results stratified by number of prior treatments are consistent with an effect of treatment only in those patients with one prior treatment failure. The study by Lisanby et al. (2009) showed a decrease in MADRS depression score of 7 points in the TMS group versus 2 points in the sham group in those who had only one prior treatment failure. The FDA Executive Summary reports only standardized effect sizes for this comparison. For this same comparison, the standardized effect size was 0.94 ($p=0.001$). For the group with more than one prior treatment failure, the decrease in MADRS scores for both groups appears to be very similar at approximately 3.8 by visual estimation. The p value for interaction between number of treatment failures (1 versus more than 1) and treatment group (TMS versus sham) was significant at $p=0.021$. The FDA Executive Summary reports results separately by the number of prior treatment failures for patients with more than 1 treatment failure. Only the group with 4 treatment failures, that had a small sample size of 12 patients, showed an indication of a treatment effect of TMS.

Neither the Lisanby et al. (2009) study nor the FDA Executive Summary reports these subgroup results in terms of response or remission rates.

Open-Label Extension of the TMS Randomized, Clinical Trial

Results of the open label extension of the clinical trial of O'Reardon et al. (2007) are reported by Avery et al. (2008) and in the FDA Executive

Summary (U.S. Food and Drug Administration 2007) (Table 5). In this study, patients who failed to achieve a greater than 25% reduction in HAMD 17 scores at 4 weeks were eligible to receive TMS for 6 weeks on an open-label basis. Thus, nonresponders from the TMS treatment group would then receive up to 12 weeks of active TMS, and nonresponders from the sham treatment group would cross over and receive up to 6 weeks of TMS. Patients and investigators were blinded to their original randomized treatment assignment. At 6 weeks, TMS treatment was tapered and antidepressant therapy initiated. During the 6-week open-label period, antidepressant medications were not allowed, and hypnotics or anxiolytics were only allowed for urgent problems.

It is not specifically stated how outcomes were calculated for patients who missed their 6-week visits, but it appears that for categorical outcomes, these patients were counted as not having responded or remitted.

For both groups of patients, patients improved their depression scores and many patients responded or had remission of depression. The changes were much greater than in the randomized part of the clinical trial. For example, in the clinical trial, the TMS group improved by 5.6 points on the MADRS at 4 weeks. In the open-label extension trial, the sham-to-TMS group, which is getting its first course of active TMS treatment, had a mean improvement in the MADRS of 17.0 points. Also, the proportion of responders and remitters in the sham-to-TMS group at 6 weeks was much higher than in the TMS group in the randomized, clinical trial.

Table 4. Outcomes of TMS Clinical Trial at 4 Weeks, Stratified by History of Prior Adequate Trials of Antidepressant Therapy

Prior Treatment Failure Stratum	Group Mean Difference in MADRS between TMS and Sham*	p value	Standardized Effect Size between TMS and Sham**	p value
1 treatment failure (n=164)	5 (visual estimation)	0.0006	0.94	0.001
>1 treatment failure (n=137)	0.2 (visual estimation)	0.923	—	—
2 treatment failures (n=95)	—	—	-0.16	0.710
3 treatment failures (n=30)	—	—	-0.55	0.588
4 treatment failures (n=12)	—	—	5.21	0.022

*Group mean differences calculated by visual estimation in published paper of Lisanby et al. (2009)

**Standardized effect size from FDA Executive Summary

Given the unblinded, open-label nature of this extension study, and the lack of an internal control group, it is uncertain how to interpret these findings. The greater changes in depression scores than the randomized trial indicates some effect due to unblinding, either due to patient and/or provider expectations.

“Study 3” Interim Results, Extension Study of Responders

An observational study of various groups of responders in either the randomized trial or open-label extension is only briefly described in the FDA Executive Summary (U.S. Food and Drug Administration 2007). In this study, which is described as having only interim results, all groups of patients that ultimately responded to TMS or sham in the original randomized trial, or responded to TMS in the open-label extension (regardless of initial assignment in the trial) were followed for recurrence of depression and/or need for reintroduction of TMS. The methods and definitions of outcome are incompletely described in the Executive Summary, and a minority of the eligible patients enrolled into this study. Although it cannot be determined how many patients were actually followed for 24 weeks beyond TMS treatment, according to the table of results, between 35.3% and 47.8% of patients required reintroduction of TMS treatment, depending on prior trial assignment. The median time to TMS reintroduction was between 6.5 and 11 weeks. Using a definition of relapse of a HAMD 24 greater than 16 or an absolute increase of 10 points from study 3 entry on the HAMD 24, between 9.1% to 12.9% of patients relapsed

within 4 weeks, and 20.5% to 26.3% relapsed within 24 weeks.

Adverse Effects of TMS

TMS has not been associated with major adverse effects. In the randomized, clinical trial, the most common adverse events in the active treatment group were headache, application site pain, muscle twitching, anxiety, application site discomfort, or nausea. Table 6 shows the incidence of adverse events that occurred with greater incidence than 10% or greater. Of those common events, application site pain, muscle twitching, and application site discomfort occurred more frequently in the TMS group than the sham group. Eighteen serious adverse events occurred, and these were equally distributed between TMS and sham groups.

Adverse effects were also assessed during the open-label extension phase of the study. Recall that one of the groups received extended (up to 12 weeks) of TMS. The common adverse events among this group and the group receiving active TMS for the first time were similar to the randomized clinical trial. Serious adverse events were rare, and included a few instances of worsening depression.

A review of other literature has not raised other major concerns about the safety of TMS. Rachid and Bertschy in a 2006 review, found that tension-type headaches were the most common adverse effect, followed by neck pain. They cite other studies in which significant signs of neuropsychological disturbances, changes in auditory thresholds, or EEG abnormalities were not

Table 5. Outcomes of TMS Clinical Trial Open-Label Extension at 6 Weeks, by Original Randomized Trial Assignment

Outcome Measure	TMS-to-continued-TMS Group n=73	Sham-to-TMS Group n=85
MADRS, mean change from baseline	-12.5	-17.0
HAMD24, mean change from baseline	-11.1	-14.5
HAMD17, mean change from baseline	-8.2	-10.8
MADRS Responders, %	26.0	42.4
HAMD24 Responders, %	31.5	42.4
HAMD17 Responders, %	30.1	37.6
MADRS Remitters (score <10), %	11.0	20.0
HAMD24 Remitters (score <11)	16.4	27.1
HAMD17 Remitters (score <8)	15.1	21.2

found after 2- to 4-week courses of TMS. They note that some studies have shown significant improvements in some cognitive domains, but cannot rule out practice effects in such studies. The potential to cause seizures has been a concern with TMS, but the Rachid and Bertsch review found only two published reports of seizures. Case reports have described some instances of mania or hypomania in patients with bipolar disease.

Discussion

The randomized clinical trial of TMS does not show definitive evidence of efficacy for its primary endpoint at 4 weeks. Not all outcomes show efficacy, and the analysis is sensitive to alternative methods of analysis. Another limitation of this and other studies of TMS is lack of rigorous evaluation beyond the period of treatment. Although short-term studies are consistent with changes in depression scores due to TMS, the clinical significance and durability of the effect are not well characterized. One meta-analysis indicated no difference in effect between patients with treatment-resistant and nontreatment-resistant depression. The randomized, clinical trial showed a greater effect in patients with only one prior treatment failure, with possibly minimal or no effect in patients with greater than one prior treatment failure.

The indication for which TMS received approval, one prior failure of an adequate antidepressant course, is unusual. A change in antidepressant therapy is usually indicated at this point, and has been shown to have a success rate similar to the first course. The current body of evidence cannot determine in a rigorous way whether TMS would be as effective as a second course of antidepressant therapy. Other important gaps in current

knowledge include whether TMS is effective as an adjunctive treatment to second-line drug therapy, the durability of TMS treatment, and the effectiveness of retreatment. A clinical trial sponsored by the National Institute of Mental Health has recruited subjects for another clinical trial of TMS (ClinicalTrials.gov NCT00149838). However, this trial also appears to have only a short duration (3 weeks) in which the participants are randomized to TMS or sham before crossovers or alternative treatments are offered.

Summary of Application of the Technology Evaluation Criteria

Based on the available evidence, the Blue Cross and Blue Shield Association Medical Advisory Panel made the following judgments about whether TMS for the treatment of depression meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. The technology must have final approval from the appropriate governmental regulatory bodies.

Devices for transcranial stimulation have received clearance by the U.S. Food and Drug Administration (FDA) for diagnostic uses. One device, NeoPulse (Neuronetics, Atlanta, GA) received approval in Canada and Israel as a therapy for depression. Although initially examined by the U.S. Food and Drug Administration (FDA) under a 510(k) application, the NeoPulse, now known as NeuroStar® TMS, received clearance for marketing as a “de novo” device in 2008. NeuroStar® TMS is indicated for the treatment of patients with depression who have failed one 6-week course of antidepressant medication.

Table 6. Adverse Events with an Incidence 10% in the TMS Randomized Clinical Trial

Adverse event	TMS (n=165) n (%)	Sham (n=158) n (%)
Headache	96 (58.2)	87 (55.1)
Application site pain	49 (35.8)	6 (3.8)
Muscle twitching	34 (20.6)	5 (3.2)
Anxiety	19 (11.5)	18 (11.4)
Application site discomfort	18 (10.9)	2 (1.3)
Nausea	17 (10.3)	10 (6.3)

2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

An important limitation of the evidence is lack of information beyond the acute period of treatment. Most of the clinical trials evaluate the outcomes at the point of the last TMS treatment, between 1 and 4 weeks. Very few studies evaluated patients beyond this time period. Although meta-analyses are consistent with short-term antidepressant effects, the clinical significance of the effect is uncertain. The large clinical trial of TMS reviewed in this assessment did not unequivocally demonstrate efficacy, as the principal endpoint was not statistically significant at 4 weeks, and some results were sensitive to the methods of analysis. The patients in which TMS is indicated are usually treated with a second course of antidepressant therapy. The clinical trial, which was sham controlled without active treatment, cannot determine whether TMS would be more or less successful than this standard treatment.

3. The technology must improve the net health outcome; and

4. The technology must be as beneficial as any established alternatives.

The available evidence does not permit conclusions regarding the effect of TMS on health outcomes or compared with alternatives. Comparison to alternatives using other observational studies may not be valid due to unmeasured differences in severity of depression between studies and other differences in studies.

5. The improvement must be attainable outside the investigational settings.

It has not yet been demonstrated whether TMS improves health outcomes in the investigational setting. Therefore, it cannot be demonstrated whether improvement is attainable outside the investigational settings.

For the above reasons, transcranial magnetic stimulation for the treatment of depression does not meet the TEC criteria.

NOTICE OF PURPOSE: TEC Assessments are scientific opinions, provided solely for informational purposes. TEC Assessments should not be construed to suggest that the Blue Cross Blue Shield Association, Kaiser Permanente Medical Care Program or the TEC Program recommends, advocates, requires, encourages, or discourages any particular treatment, procedure, or service; any particular course of treatment, procedure, or service; or the payment or non-payment of the technology or technologies evaluated.

CONFIDENTIAL: This document contains proprietary information that is intended solely for Blue Cross and Blue Shield Plans and other subscribers to the TEC Program. The contents of this document are not to be provided in any manner to any other parties without the express written consent of the Blue Cross and Blue Shield Association.

References

- Avery DH, Isenberg KE, Sampson SM et al. (2008).** Transcranial magnetic stimulation in the acute treatment of major depressive disorder: clinical response in an open-label extension trial. *J Clin Psychiatry*, 69:441-51.
- Cadieux RJ. (1998).** Practical management of treatment-resistant depression. *Am Fam Physician*, 58:2059-62.
- Couturier JL. (2005).** Efficacy of rapid-rate repetitive transcranial magnetic stimulation in the treatment of depression: a systematic review and meta-analysis. *J Psychiatry Neurosci*, 30:85-90.
- Gross M, Nakamura L, Pascual-Leone A, Fregni F. (2007).** Has repetitive transcranial magnetic stimulation (rTMS) treatment for depression improved? A systematic review and meta-analysis comparing the recent vs. the earlier rTMS studies. *Acta Psychiatr Scand*, 116:165-75.
- Kessler RC, Berglund P, Demler O et al. (2005).** The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*, 289:5095-5105.
- Lam RW, Chan P, Wilkins-Ho M, Yatham LN. (2008).** Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and metaanalysis. *Can J Psychiatry*, 53:621-31.
- Lisanby SH, Husain MM, Rosenquist PB et al. (2009).** Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology*, 34:522-34.
- Martin JL, Barbanj MJ, Schlaepfer TE et al. (2005).** Repetitive transcranial magnetic stimulation for the treatment of depression. Systematic review and meta-analysis. *Br J Psychiatry*, 182:480-91.
- O'Reardon JP, Solvason HB, Janicak PG et al. (2007).** Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry*, 62:1208-16.
- Rachid F, Bertschy G. (2006).** Safety and efficacy of repetitive transcranial magnetic stimulation in the treatment of depression: a critical appraisal of the last 10 years. *Neurophysiol Clin*, 36:157-85.
- Rush AJ, Fava M, Wisniewski SR et al. (2004).** Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*, 25(1):119-42.
- Rush AJ, Trivedi MH, Wisniewski SR et al. (2006).** Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*, 163:1905-17.
- Schutter DJ. (2009).** Antidepressant efficacy of high-frequency transcranial magnetic stimulation over the left dorsolateral prefrontal cortex in double-blind sham-controlled designs: a meta-analysis. *Psychol Med*, 39:65-75.
- Thase ME, Rush AJ. (2000).** Treatment-resistant depression. Available online at <http://www.acnp.org/g4/GN401000105/Default.htm>. Last accessed May 2009.
- Trivedi MH, Rush AJ, Wisniewski SR et al. (2006).** Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*, 163:28-40.
- U.S. Food and Drug Administration. (2007).** Neurological Devices Panel of the Medical Devices Advisory Committee. January 26, 2007. FDA Executive Summary. Available online at http://www.fda.gov/ohrms/dockets/ac/07/briefing/2007-4273b1_01-FDAExecutiveSummary.pdf. Last accessed May 2009.



**Technology
Evaluation
Center**

**Blue Cross and
Blue Shield Association**
225 North Michigan Avenue
Chicago, Illinois 60601-7680
www.bcbs.com/tec