Cost-Effectiveness of Transcranial Magnetic Stimulation (TMS) in the Treatment of Major Depression: A Health Economic Analysis

Ziad Nahas, MD[1], Kit N. Simpson, DrPH[1], William M. McDonald, MD[2], Mary Jane Welch, DNP, APRN, BC[3], F. Andrew Kozel, MD, MSCR[4], and Mark A. Demitrack, MD[5], for the TMS Therapy Study Group

[1] Department of Psychiatry, Medical University of South Carolina, Charleston, SC; [2] Department of Psychiatry, Emory University, Atlanta, GA; [3] Department of Psychiatry, Rush University, Chicago, IL; [4] Department of Psychiatry, University of Texas-Southwestern, Dallas, TX; [5] Neuronetics, Inc., Malvern, PA

Introduction

In the United States, only 3.2 million of the 14 million patients with depression receive adequate pharmacotherapy (Kessler, 2003). Even in those who are treated, the success rate with pharmacological treatments is not high and short-term antidepressant drugs are only moderately effective when compared to placebo. The recently published results of the Sequenced Treatment Alternatives to Relieve Depression trial (STAR*D) illustrate a pattern of diminishing clinical returns: each successive pharmacological treatment failure predicted a worse prognosis of a subsequent trial (Rush, 2006). The STAR*D study showed that after three successive pharmacological treatment strategies the cumulative remission rate is at best only 67%.

Despite changes in treatment practice, the clinical impact of depression itself and of treatment resistance is growing. The World Health Organization (WHO) projects that by the year 2020, unipolar depression alone will be second in medical burden only to ischemic heart disease (Murray & Lopez, 1996). This somber prospect urges us to look for more effective and economically sound antidepressant treatments.

Non-pharmacological neuromodulation therapies (NMTs) have emerged in the last decade as potentially efficacious treatments for TRD. NMTs sub-convulsively modulate discrete networks with repetitive electrical stimulation. In contrast the therapeutic efficacy of electroconvulsive therapy (ECT) is dependent on the patient having an adequate major motor seizure. Transcranial magnetic stimulation (TMS) is a type of NMT whereby electrodeless stimulation is produced by a rapid oscillation in electrical energy which is converted to magnetic energy. If activated over the skull, anatomically and functionally related brain regions are stimulated through cortical-subcortical neuronal circuits. Prefrontal TMS, repeated over several weeks, has a clinically significant acute antidepressant effect.

TMS has been shown to be safe and effective in the treatment of patients with major depression (O'Reardon et al, 2007; Janicak, et al, 2008; Avery, et al, 2008). Two reports (Kozel, et al, 2004; Knapp, et al, 2008) have already discussed the health economic impact of TMS in patients with pharmacoresistant depression. Both analyses were conducted using either historical estimates of TMS efficacy, or based on the results of a limited, fixed-dose treatment paradigm conducted in a small sample. Since those reports, we have published the results of the first large, multisite, randomized controlled study of TMS used as monotherapy in patients with pharmacoresistant depression. This clinical development program also included two open-label extension studies that provided information on acute efficacy and long-term durability of acute response to TMS.

The present report aims to estimate the incremental cost effectiveness ratio for the NeuroStar TMS Therapy system compared to sham treatment and to current standard of care, using a decision analysis modeling approach. The model is structured to accommodate the clinical data obtained in the recently completed Neuronetics (Malvern, PA) trials referred to here as Studies 101, 102 and 103, and combines these data with cost and utility weights derived from published data, and mean costs from large archival billing data bases for patients with depression.

Quality-adjusted year of life gained (QALY) quantifies the impact of a medical treatment both in terms of the quality and the quantity of life lived. There is no defined cost-effectiveness threshold value for the US, but an ICER below $50,000 is generally considered highly cost-effective; it has been proposed that a de facto value for pharmaceuticals lies somewhere between $50,000 and $100,000 per QALY. The World Health Organization (WHO) suggests a limit for cost-effectiveness of three times a country's gross domestic product (GDP) per capita, which would set the upper limit at $140,000 for the United States. However, acceptable ICER values may vary with time and depend on the burden of the condition being treated, the size of the patient population affected, and health equity considerations. In this economic modeling study, we tested the incremental cost effectiveness ratio (ICER) for TMS compared to sham or compared to standard of care treatments. We hypothesized the incremental cost of TMS treatment would be lower than the societal willingness-to-pay threshold.

Methods

Clinical Development Program Overview

The overall clinical development program was composed of three separate clinical protocols that were related in temporal sequence to one another (Figure 1). A description of the study designs and outcomes for these protocols has been reported elsewhere.

Motor Threshold Assessment, Treatment Location and TMS Treatment Parameters

All TMS sessions were delivered using the NeuroStar TMS Therapy System investigational device.

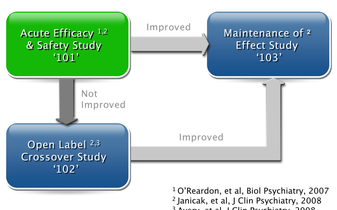
Treatment parameters were fixed per protocol and included stimulation at 120% of observed motor threshold applied at a pulse frequency of 10 pulses per second (cycle time: 4 seconds on and 26 seconds off) for a total of 3000 pulses per treatment session.

Motor threshold was determined using an iterated mathematical algorithm (MT Assist®) to ensure standardization across study sites.

Treatment location was by convention located at a site 5 cm anterior to the motor threshold location oriented on a line extending in a left superior oblique angle with a rotation point about the patient's nose.

During the acute efficacy phases of Study 101 and Study 102, TMS was administered 5 days per week for up to 6 weeks. During Study 103, TMS as an add-on rescue treatment was initiated at 2 days per week for the first 2 weeks, and then 5 days per week for the next 4 weeks.

Figure 1. NeuroStar Clinical Development Program



[1] O'Reardon, et al, Biol Psychiatry, 2007
[2] Janicak, et al, J Clin Psychiatry, 2008
[3] Avery, et al, J Clin Psychiatry, 2008

Decision Analysis Model Structure

In this study, we used a decision tree as the structural framework for organizing the efficacy outcome data for the six week acute treatment and the three week treatment taper phases. The tree segregated patients into groups who finished the six weeks of treatment in study 101 or 102 and those who did not finish. It then categorized patients into those who participated in the planned 3-week tapering phase and those who did not. Within each of these groups we characterized the distribution of patients using outcome criteria defined according to MADRS total score as follows: 1) '**MADRS 0**', no depression (total score 0 to 9); 2) '**MADRS 1**', mild depression, (total score 10 to 17); 3) '**MADRS 2**', moderate depression, (total score 18 to 27); and 4) '**MADRS 3**', severe depression; (total score > 27). At the end of the 9-week treatment period patients who completed the study period with an outcome in either of MADRS 0, MADRS 1 or MADRS 2, i.e., no depression, mild depression, or moderate depression, respectively, moved into a follow up Markov model which estimated the outcomes during the remaining year, based on the likelihood of progression by health state at the end of acute treatment.

Methods

Decision Analysis Model Structure, CONT'D

Patients who were classified as MADRS 3, severe depression at the end of acute treatment were moved into a separate Markov model where they were presumed to be treated with a new drug regimen combination consisting of a new antidepressant, a mood stabilizer, and an atypical antipsychotic. Note that the patients in the MADRS 3 group can be presumed to have demonstrated failure to benefit from at least two prior antidepressant treatments (i.e., ATHF-verified failure of antidepressant pharmacotherapy prior to study entry, and subsequent prospectively demonstrated failure to benefit from TMS during the TMS study itself). Therefore, the estimates of the potential efficacy of future antidepressant treatment for the MADRS 3 group in the Markov model were based on the results reported for Levels 3 and 4 in the STAR*D Study, ie, the groups in that study who had also failed to benefit from at least two antidepressant treatment exposures. The decision analysis model was programmed in Excel (Microsoft Corp., Seattle, WA).

Data Analysis Methods and Model Parameters

An analytical economic dataset and the specific model parameters specifying acute treatment outcome and severity-specific relapse rates estimated over one year of follow up that were used in the decision analysis model were derived from the actual raw clinical efficacy outcome data from the Neuronetics Studies 101, 102 and 103.

Health resource utilization was obtained for all patients in the clinical trial at entry into Study 101 and again at their point of exit from Study 103. This utilization information was obtained using a self-report questionnaire designed specifically for use in this study. The information covered three major domains of information (Work Productivity/Work Loss [9 items], Health Care Utilization and Cost [18 items], and Caregiver Support [4 items]). These costs were applied to both treatment arms in the model (for the Neuronetics study and for the STAR*D trial outcomes).

The overall clinical and health outcomes information was then combined with standard cost weights drawn from large national databases (H-CUP 2004; http://www.hcup-us.ahrq.gov). Detailed cost weights were estimated for subgroups of patients with similar clinical severity, where the clinical severity was defined by the MADRS criteria described above in the Decision Analysis Model. These subgroups were used to define the outcomes for patients at the end of the acute efficacy treatment in the respective clinical trial (ie, either the Neuronetics studies or the STAR*D study), and to link the trial results to cost and quality of life weights reported in the literature. The results of the analysis of the clinical trial data were linked to health state-specific relapse rates estimated over a one year period following their acute treatment, using the actual follow-up data from Study 103 to estimate this one year interval.

The perspective for costing in the model is that of the US health system using 2007 standard cost estimates based on mean study reported resource uses. As noted above, the mean utilization values estimated from the study data were combined with unit cost weights derived from the analysis of a large sample of 2004 Medicaid billing data for patients with depression (South Carolina Medicaid billing dataset; http://www.ors.state.sc.us) The 2004 mean costs (not charges) were then inflated to equivalent 2007 cost weights using the medical care consumer price index over that interval of time. In the model outputs reported here that include the cost of lost productivity, we have assumed 2 hours of lost time per treatment received. As a Base Case, each TMS session was estimated to cost the healthcare system US$300. The incremental cost effectiveness ratio (I-CER) for all reported model estimates was calculated as the difference in the one-year cost of treating 100 patients with TMS minus the difference in the cost expected for treating 100 patients in the control group (sham TMS or pharmacotherapy, depending upon the specific model being considered), divided by the difference in the QALYs produced by the two treatments over one year. The model parameters and cost weights are provided in Table 1. The costs of the various pharmacotherapy regimens considered for the models are summarized in Tables 1 & 2. Final analyses utilized costs from antidepressant treatment regimen 2 for the re-treatment part of the model.

Table 1. Model Parameters and Cost Weights

Parameter/Variables	Base Model Value	Range for Sensitivity Analysis	Data Sources
MADRS 0 Utility weight	.83	.80 - .86	Revicki, 1995; NICE 2006
MADRS 1 Utility weight	.73	.70 - .76	Revicki, 1995 NICE 2006
MADRS 2 Utility weight	.63	.60 - .66	Revicki, 1995 NICE 2006
MADRS 3 Utility weight	.30	.27 - .33	Revicki, 1995 NICE 2006
In Hospital failure	.09	.06 - .12	Kamlet, 1995 NICE 2006
Decision Tree Model:			
TMS treatment (Base case)	US$300	±10%	Neuronetics studies
MADRS 0; Medical care/day	US$2.16	±10%	Neuronetics studies
MADRS 1 Medical care/day	US$2.16	±10%	Neuronetics studies
MADRS 2 Medical care/day	US$3.01	±10%	Neuronetics studies
MADRS 3 Medical care/day	US$3.94	±10%	Neuronetics studies
MADRS 0 Productivity/day	US$33.37	±10%	Neuronetics studies
MADRS 1 Productivity/day	US$43.80	±10%	Neuronetics studies
MADRS 2 Productivity/day	US$99.40	±10%	Neuronetics studies
MADRS 3 Productivity/day	US$128.59	±10%	Neuronetics studies
Median hourly wage for patient during treatment	US$12.00	±10%	Neuronetics studies
Lost wages per treatment	US$509.04	±10%	Neuronetics studies
Markov Model:			
Hospital cost/day	US$880	±10%	Medicaid 2004
ER cost/visit	US$426	±10%	Medicaid 2004
MD office visit	US$129		Medicaid 2004
Antidepressant maintenance drug cost /day	US$1.53	±10%	Red Book 2006
Follow up drug cost to treat failure/day	US$2.20	±10%	Red Book 2006
Retreatment cost for patients in severe health state	US$22.63	See separate table 3	Red Book 2006
Marginal cost of hospital care for suicide	US$852	US$40,000	Medicaid 2004
24 week failure	TMS	Sham	
MADRS 0	30%	33%	Study 103
MADRS 1	23%	50%	Study 103
MADRS 2	33%	50%	Study 103
MADRS 3	33%	50%	Study 103
Efficacy of pharmaceutical treatment regimens used for MADRS 3 patients	14%	14%	Rush et al, 2006

Table 2. Model Cost of Pharmaceutical Treatment Regimens for Patients with Severe Depression

Drug Type	Drug Name	Daily Dose	Cost per Day*	Cost per Week
Regimen 1				
Antidepressant	Fluoxetine	20mg	2.66	
Mood stabilizer	Carbamazepine	100mg	0.43	
Atypical Antipsychotic	Olanzapine	5mg	11.76	
Total Cost of Regimen			[14.85]	[103.95]
Regimen 2				
Antidepressant	Sertraline	150mg	4.21	
Mood stabilizer	Valproic acid	1500mg	6.87	
Atypical Antipsychotic	Aripiprazole	10mg	11.55	
Total Cost of Regimen			[22.63]	[158.41]
Regimen 3				
Antidepressant	Venlafaxine	37.5mg	2.07	
Mood stabilizer	Lamotrigine	200mg	3.83	
Atypical Antipsychotic	Olanzapine	20mg	23.50	
Total Cost of Regimen			[29.40]	[205.80]

* AWP from 2006 Red Book

Methods

The model first estimated outcomes for TMS compared to Sham treatment using the Study 101 overall population data. As the experience in a blinded randomized controlled trial and a comparison to sham are not fully reflective of the expected clinical outcomes with TMS in ordinary clinical practice where other actual treatments options would be considered, we performed three additional model estimates to place these results in context, comparing these model estimates to best estimates of outcomes and costs of pharmacotherapy treatment as usual, using the published STAR*D outcomes for this comparison. These three additional model estimates included:

- a comparison of one year outcomes for the overall patient population treated with TMS in the open-label Study 102 data versus a synthetic comparison group of clinical outcomes observed in the published results from the STAR*D trial for Levels 2 and 3 combined in that study;
- an examination of the impact of prior antidepressant treatment resistance by performing a comparison of the Study 101 active TMS vs Sham using only data from patients who had failed one adequate prior antidepressant treatment in their current episode (ATHF=1). These patients had a median of 4 antidepressant attempts; and finally
- a comparison of the open-label Study 102 clinical outcomes restricted to those patients within the overall study population who experienced one adequate prior antidepressant treatment in their current illness episode (ATHF=1) vs a STAR*D synthetic comparison group restricted to the Level 2 outcomes only.

We tested each model estimate's sensitivity to variations in the assumptions, cost weights and quality of life adjustments used for each estimate. Specifically, we tested the impact of two key model parameters alone and in combination: 1) either excluding or including indirect costs in the model, and 2) varying the estimated cost of a suicide attempt. Clinical outcomes in the STAR*D comparison datasets are reported using the Hamilton Depression Rating Scale (HDRS). We used the information provided on the IDS/QIDS website (http://www.ids-qids.org/index2.html#table4) to establish clinically equivalent rating comparisons of the HDRS data to the MADRS scores used in the Neuronetics dataset. All statistical analyses were performed with SAS version 8.1 (Cary, North Carolina, USA).

Results

Study Population

Demographic and clinical features of the overall study population treated with the NeuroStar TMS System are shown in Table 3.

Decision Analysis Model Outcomes

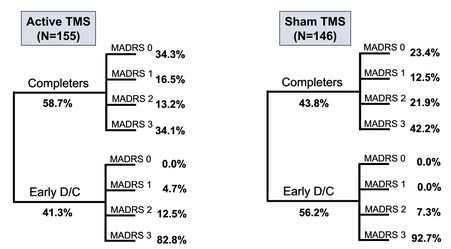
Results of the decision analysis for the acute treatment phase in both active TMS (N=155) and sham TMS (N=146) randomized groups in Study 101 are shown in Figure 2a. Results for the decision analysis for the acute treatment phase of all patients (N=158) in the open-label Study 102 is shown in Figure 2b.

A similar decision analysis model was applied to the outcomes from the STAR*D trial to establish a comparison benchmark for pharmaceutical treatment of patients who had failed initial antidepressant pharmacotherapy. The STAR*D results for Levels 2 and 3 combined were used as a comparison for the overall population in the Neuronetics studies, while the results for the Level 2 STAR*D group alone were used as a comparison for the subgroup of patients in the Neuronetics studies who experienced only one adequate prior antidepressant treatment in their current illness episode (ATHF=1). Results of this analysis are shown in Figure 2c.

Table 3. Demographic and Clinical Features of the NeuroStar Study Population

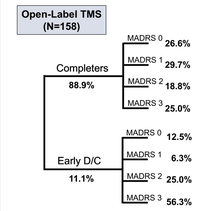
	Active TMS (n=155)	Sham TMS (n=146)	P-Value
Demographic Variables			
n(%) females	86(55.5)	74(50.7)	.421
Age (years ±SD)	47.9±11.0	48.7±10.6	.509
Ethnic Origin, n(%)			
Caucasian	146 (94.2)	131(89.7)	
Other	9(5.8)	15(10.3)	.201
Employment Status, n(%)			
Full Time	58(35.6)	45(28.3)	
Part Time	27(16.6)	31(19.5)	
Unemployed	78(47.9)	83(52.2)	
Receiving Disability Compensation	28(32.9)	31(34.1)	
Disease History			
Recurrent Illness Course (%)	149(95.5)	136(93.8)	.611
Duration of current episode in months, Mean(SD)	13.6(9.9)	13.2(9.5)	.728
n(%) of population with current episode≥2 years	36(23.2)	23(15.8)	.112
Prior Antidepressant Treatment			
Number of antidepressant treatment attempts in current illness episode (mean, SD)	5.5 (3.4)	5.4 (3.6)	.774
Number of dose/duration adequate antidepressant treatments in current episode (mean, SD)	1.6 (0.9)	1.6 (0.8)	.905
Baseline Symptom Scores			
MADRS total score (SD)	32.8(6.0)	33.9(5.7)	.036
HAMD17 total score(SD)	22.6(3.3)	22.9(3.5)	.508
HAMD24 total score(SD)	30.1(5.0)	30.5(4.9)	.568
CGI –Severity (SD)	4.7(0.6)	4.7(0.7)	.197
IDS-SR total score(SD)	42(9.4)	43.4(9.9)	.197

Figure 2a. Decision Tree Outcomes for Acute Phase (six week) Treatment in Study 101: Overall Study Population



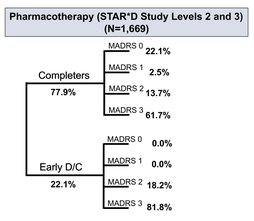
NOTES: MADRS 0 = total score 0-9, MADRS 1 = total score 10-17, MADRS 2 = total score 18-27, MADRS 3 = total score > 27.

Figure 2b. Decision Tree Outcomes for Acute Phase (six Week) Treatment in Study 102: Overall Study Population



NOTES: MADRS 0 = total score 0-9, MADRS 1 = total score 10-17, MADRS 2 = total score 18-27, MADRS 3 = total score > 27.

Figure 2c. Decision Tree Outcomes for Acute Phase Pharmacotherapy Treatment



NOTES: Data adapted from the published results of the STAR*D study. MADRS 0 = total score 0-9, MADRS 1 = total score 10-17, MADRS 2 = total score 18-27, MADRS 3 = total score > 27.

Results

Incremental Cost Effectiveness Ratios (ICERs) or Cost Savings

Data for the incremental cost effectiveness ratios (ICERs) or the economically dominant model estimates of cost savings for the various decision analysis models for the Base Case alone are summarized in Table 4 for both the randomized controlled trial Study 101 and the open-label Study 102.

Cost Savings per Year of Treatment: Summary of Break-Even Time Points

A comprehensive summary of the cost savings (or cost increases) per year of treatment with the NeuroStar TMS System is shown in Figure 3. Outcomes vary depending on the assumed treatment cost. In this display, we show the results for the Open-Label Study 102 with the similar open-label study outcomes for the STAR*D trial. The results are shown for the subgroup of patients at the earliest levels of treatment resistance in both studies (ie, those patients in treated with NeuroStar who had failed to receive benefit from one antidepressant treatment at minimal effective dose and duration in current episode compared to the STAR*D Level 2 outcomes).

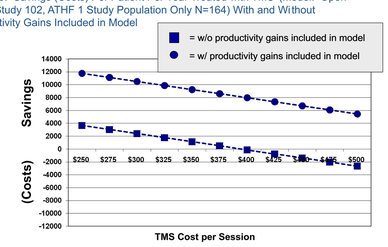
We also estimated the time to break-even cost in this model, for the Base Case model (ie, an estimated TMS per treatment session cost of $300). In this analysis, the mean cost for a TMS patient would be completely offset at 28 weeks, when compared to the expenditure flow expected for patients from Level 2 of the STAR*D trial over 12 months of treatment.

Table 4. Incremental Cost Effectiveness Ratios (incremental cost per QALY gained) or Dominant Cost Savings (per treated patient per year) for the Various Economic Models Examined: Base Case Cost Assumptions

Model Structure	Study Population	Incremental Cost-Effectiveness Ratio (ICER) or [Dominant Cost Savings]	
		With Productivity Costs Included in the Model	Without Productivity Costs Included in the Model
Acute TMS vs Sham (Randomized, Controlled Trial Study 101)	Overall (N=301) ATHF = 1 (N=164)	US$6,667 [US$88]	US$34,999 US$29,556
Acute TMS (Open-Label Study 102) vs Pharmacotherapy Treatment as Usual (STAR*D Study)	Overall (N=301) ATHF = 1 (N=164)	[US$7,621] [US$16,516]	[US$1,123] [US$2,406]

NOTES: Costs in [brackets] represent economically dominant model estimates for TMS, and are reported as cost savings per treated patient per year. All other costs represent incremental cost per QALY gained, or ICER.

Figure 3. Savings (Costs) Per Patient Per Year Treated with TMS (Model: Open-Label Study 102, ATHF 1 Study Population Only N=164) With and Without Productivity Gains Included in Model



Conclusions

- TMS is a safe and effective treatment option for patients who have failed to receive benefit from antidepressant pharmacotherapy
- In patients with pharmacoresistant major depression, significant cost savings may be expected relative to current standard of care pharmacotherapy when used at the earliest stages of treatment resistance
  - The most conservative estimate of the incremental cost-effectiveness ratio (ICER) for TMS is $34,999 per QALY, well below the most stringent willingness-to-pay threshold for a new treatment
  - Treatment with open-label TMS shows an actual cost savings of $10,516 per patient per year when productivity gains are included, and $2,406 with health gains only included in the model estimates
- The mean cost for a patient receiving an acute treatment course of TMS would be offset at 28 weeks ("break-even") when compared to the expenditure for pharmacotherapy over 12 months of treatment

NEURONETICS