

## **FDA Executive Summary**

## Table of Contents

1.	INTRODUCTION	3
2.	PROPOSED INDICATIONS FOR USE	3
3.	DEVICE DESCRIPTION	4
4.	CLINICAL STUDIES	5
4.1	Study 01: Investigational Plan	5
4.2	Study 01: Results	8
4.2.1	Study 01: Effectiveness Results	9
4.2.2	Study 01: Safety Results	15
4.3	Study 02: Investigational Plan	18
4.4	Study 02: Results	19
4.4.1	Study 02: Effectiveness Results	20
4.4.2	Study 02: Safety Results	21
4.5	Study 03: Investigational Plan	22
4.6	Study 03: Results	23
4.6.1	Study 03: Effectiveness Results	23
4.6.2	Study 03: Safety Results	24
5.	ELECTROCONVULSIVE THERAPY (ECT) TREATMENT	25
	REFERENCES	29
	APPENDIX A	31
	APPENDIX B	32
	APPENDIX C	43
	APPENDIX D	47
	APPENDIX E	50

## 1. **Introduction**

This document is a summary of the information provided in premarket notification (510(k)) submission, K061053, submitted by Neuronetics, Inc. to the Restorative Devices Branch of the Division of General, Restorative and Neurological Devices at the Center for Devices and Radiological Health of the Food and Drug Administration (FDA). Neuronetics, Inc. submitted the 510(k) submission to request marketing clearance for the NeuroStar™ TMS System for the proposed indications for use of the treatment of Major Depressive Disorder (MDD).

When manufacturers submit a 510(k), they must compare their device to a legally-marketed predicate device that does not require review through the premarket approval (PMA) process. Substantial equivalence does not mean the device under review and predicate devices must be identical. A device is determined to be substantially equivalent to a legally-marketed predicate device if it:

- has the same intended use as the predicate **AND** has the same technological characteristics as the predicate;
- OR**
- has the same intended use as the predicate **AND** has different technological characteristics and the information submitted to FDA:
  - does not raise new types of questions of safety and effectiveness; **AND**
  - demonstrates that the device has a comparable risk-to-benefit profile to a legally marketed device.

In previous early discussions between FDA and Neuronetics (e.g., prior to initiation of the clinical trials), the Agency agreed with the sponsor that a premarket notification (or 510(k) submission) using ECT as a predicate would be an appropriate regulatory pathway to pursue. However, in order to use ECT as a predicate device, the sponsor was told that they would need to provide valid scientific evidence, including appropriate clinical performance data, to demonstrate that the risk-to-benefit profile for their device in the treatment of MDD was favorable and was comparable to that of ECT. In particular, the Neuronetics device would not necessarily need to be as effective as ECT if clinical data demonstrated that any reduction in effectiveness was off-set by an improvement in patient safety/risk.

The sponsor designed a clinical study which they believed was adequate and appropriate to demonstrate the device's risk/benefit profile for their desired indication (population). The results of this study, and data for ECT available from the public literature, have been submitted to FDA in support of the substantial equivalence of the NeuroStar TMS System.

Although FDA will not be asking the Advisory Panel (the Panel) any regulatory questions about this submission, we will be looking to the Panel to provide an interpretation of the submitted clinical data and to assist us in gauging the clinical utility in terms of its risk-to-benefit profile for the device and how it compares to that of ECT for the treatment of patients with MDD.

## 2. **Proposed Indications for Use**

The proposed indication for use for the NeuroStar™ TMS System is the treatment of Major Depressive Disorder (MDD). If the Panel believes the NeuroStar™ TMS System has a favorable risk-to-benefit profile they will be asked to comment on the patient population that is most appropriate for treatment with this device.

### 3. Device Description

The NeuroStar™ TMS System is a repetitive transcranial magnetic stimulation (rTMS) therapy system. Note that all technical documentation within the premarket 510(k) notification refers to the Callisto TMS System which has been renamed the NeuroStar™ TMS System. Additionally, a different model of the device, the Model 2100, was used in the clinical investigations. However, the variations in coil output levels between the models are minimal and the sponsor has provided measurement data showing experimental confirmation of magnetic field equivalency, as a function of output level, between the NeuroStar™ TMS System and the Model 2100 that was used in the clinical investigations. The minimal differences are also irrelevant due to the fact that the treatment level is set based on the patient's own motor threshold as determined at each treatment session by the treating clinician.

The NeuroStar™ TMS System (see Figure 1) is a computerized electromechanical device that delivers brief duration, rapidly alternating, or pulsed, magnetic fields to induce electrical currents that are directed to discrete regions of the cerebral cortex. The main system components consist of hardware (a treatment chair and head support, a console and gantry, a touch screen graphic user interface (GUI), system processor and power modules, a magnetic coil (Coil) and gantry, and Coil interface electronics) and software (to provide user interface, control various

The sponsor has hypothesized that the application of a magnetic field by the NeuroStar™ TMS System to the subject's head results in neuronal activation without inducing a seizure.



Figure 1 - The NeuroStar™ TMS System

#### **4. Clinical Studies**

The safety and effectiveness of the NeuroStar™ TMS System, was evaluated in a three interrelated studies that were approved by the FDA under Investigational Device Exemption (IDE) on April 17, 2004. The study phases are identified as numbers 4401101 (Study 01), Study 02) and 4401103 (Study 03). Study 01 was a multi-center, randomized, parallel-group, sham-controlled clinical trial designed to examine the safety and effectiveness of the NeuroStar™ TMS System for subjects diagnosed with DSM-IV defined major depression who have not benefited from prior adequate treatment with oral antidepressants during their current major depressive episode. The data from Study 01 were provided to the agency as the primary basis for determining the substantial equivalence of the NeuroStar™ TMS System to ECT treatment. Study 02 was an open-label clinical trial for subjects from both active treatment and sham treatment groups who did not meet pre-defined success criteria for response to treatment in Study 01. Study 03 was an open-label, uncontrolled clinical trial providing six months of oral antidepressant monotherapy to subjects who met pre-defined success criteria for response to treatment upon exit from Study 01 or Study 02.

##### **4.1 Study 01: Investigational Plan**

###### Objective

To evaluate the antidepressant effect of a specified treatment course with the NeuroStar™ TMS System when compared to sham treatment under the same experimental conditions in subjects meeting the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria for Major Depressive Episode (MDE), single or recurrent episode.

###### Design

A nine-week randomized, parallel-group, sham treatment controlled, multicenter clinical trial. The sham device was identical in appearance and acoustically matched to the active NeuroStar™ TMS System. The subjects, device operators, and evaluators were all masked as to the identity of the device.

###### Primary Effectiveness Endpoint

The primary effectiveness endpoint was the change from baseline in the last post-treatment total symptom score on the Montgomery Åsberg Depression Rating Scale (MADRS) compared to the sham treatment arm at week 4 of the study using the intent-to-treat (ITT) evaluable population. Note that study inclusion was based on the 17-item Hamilton Depression Rating Scale (HAMD<sub>17</sub>) and not the MADRS.

###### Secondary Outcome Measures

The protocol included a listing of secondary outcome measures “in the sequence of intended testing in priority order” (see Table A in Appendix A). Secondary effectiveness outcome measures included both clinician rated and patient rated outcome measures. The following clinician rated measures were included: 24-item Hamilton Depression Rating Scale (HAMD<sub>24</sub>), a HAMD<sub>17</sub> extracted from the hyponymous HAMD<sub>24</sub>, the response and remission rates for the MADRS, the HAMD<sub>24</sub>, the HAMD<sub>17</sub> (a sub-section of the HAMD<sub>24</sub>), and the Clinical Global Impressions - Severity (CGI-S). The following patient rated measures were included: the Medical Outcomes Study Short Form - 36 Item Questionnaire, version 1 (SF-36), the Quality of Life Enjoyment and Satisfaction Questionnaire - Short Form (Q-LES-Q Score), the Inventory of Depressive Symptoms - Self Report version (IDS-SR), and the Patient Global Impressions - Improvement of Illness Scale (PGI-I).

### Standardized Effect Size

The sponsor defined a standardized effect size (d) of 0.4 as the “minimally clinically interesting difference” to be obtained and it was arrived at by requiring 90% power and a two sided 5% test, and is based on the standard t-test method. The standardized effect size is typically defined as the difference in means between sham and active treatment divided by the pooled standard deviation of the scores.

### Key Inclusion Criteria

- Male and female subjects ranging from 18 to 70 years of age;
- Primary diagnosis is by the DSM-IV criteria for Major Depressive Episode (MDE), single episode or recurrent, with a current episode<sup>1</sup> duration of  $\leq 3$  years, as confirmed by the Structured Clinical Interview for DSM-IV (SCID-IV), with the stipulation of a duration of this episode  $\geq 4$  weeks;
- A score CGI-S  $\geq 4$  at initial screening;
- A minimum symptom severity as reflected by a total score of  $\geq 20$  on the HAMD<sub>17</sub>, including an item 1 score<sup>2</sup> of at least 2 and a sustained symptom severity after a one week no-treatment lead-in period as reflected by a HAMD<sub>17</sub> total score of  $\geq 18$  and less than or equal to a 25% decrease in HAMD<sub>17</sub> total score from that observed at the screening assessment;
- Subjects evaluated using the Antidepressant Treatment History Form (ATHF) including failure to receive benefit from at least 1 but not more than 4 adequate trials of an antidepressant in the current or a past episode (adequacy of treatment defined as an ATHF antidepressant resistance potency of at least Level 3 for the specific antidepressant); and
- Have undergone washout of their current psychotropic medication prior to completion of the screening process.

### Key Exclusion Criteria

- Significant acute suicide risk;
- History of substance abuse or dependence within the past year;
- A history of psychosis, bipolar disorder, obsessive compulsive disorder;
- An active history (in past year) of post-traumatic stress disorder or eating disorders;
- ECT treatment within 3 months prior to screening visit;
- Failure to respond to ECT treatment (i.e., consistent with ATHF level 2 or higher) in this or any previous episode;
- Previously treated with experimental rTMS or had received a vagus nerve stimulator implant;
- Have recently (last 3 months) entered or changed psychotherapy or for whom the psychotherapy treatment plan was expected to change during the course of the study;
- Pregnancy or women of reproductive age who were not using a medically accepted form of contraception during intercourse;
- A history of seizure disorder or any neurologic disease or medication therapy known to alter seizure threshold; and

---

<sup>1</sup> The definition of an episode is demarcated by a period  $\geq 2$  months when the patient did not meet the full criteria for the DSM-IV definition of major depressive episode.

<sup>2</sup> Item 1 measures “Depressed Mood” symptoms on a scale of 0 to 4 where 0 is a neutral mood, 2 is when the patient more clearly is concerned with unpleasant experiences, although he still is without helplessness or hopelessness, and 4 includes patient remarks on despondency and helplessness or the non-verbal ones dominate the interview in which the patient cannot be distracted.

- The presence of ferromagnetic material anywhere in or in close proximity to the head.

### Protocol Summary

Randomization assignment was established prior to the start of the study. Subjects were required to washout psychotropic medications prior to the completion of a one-week screening process. Clinical evaluations included all of the evaluations listed above as primary effectiveness endpoint and secondary outcome measures. Safety and tolerability of treatment was assessed by observation of adverse events and serious adverse events. The protocol consisted of three experimental phases:

- Phase I: Subjects were required to washout psychotropic medications prior to the completion of a one-week (7-10 days) screening process. No treatment was applied. A subject needed a sustained symptom severity after the one week screening process (as reflected by a HAMD<sub>17</sub> total score of  $\geq 18$  and a less than or equal to 25% decrease in score from that observed at the screening assessment). This constituted the baseline assessment (Visit 2).
- Phase II: Six weeks of acute treatment with either the active device (see Device Application Protocol below) or the sham device. Treatment was scheduled in 5-day contiguous treatment blocks, generally scheduled on Monday through Friday, for a maximum possible number of 30 treatment sessions. Clinical outcome and safety evaluations occurred at approximately 2-week intervals and the primary efficacy endpoint was at week 4.
- Phase III: A three week taper phase where all subjects were placed onto antidepressant pharmacotherapy and simultaneously tapered off treatment with the NeuroStar™ TMS System (active and sham) in a schedule of gradually less frequent treatment sessions (three times per week, twice per week and then once per week). The blind remained intact. No patient was treated with an antidepressant medication for which they had previously failed to receive benefit. Clinical outcome and safety evaluations occurred weekly.

### Device Application Protocol

The device output stimulus strength, rate, duration of application and location were fixed in the studies as follows: the output stimulus strength was set at a magnetic field intensity of 120% of the patient's observed motor threshold, the rate was 10 pulses/second (grouped in 30 second cycles with a stimulation on-time of 4 seconds and an off-time of 26 seconds), a session lasted 37.5 minutes. Motor threshold was determined weekly by visual observation of thumb or finger movement using the device 'MT Assist' which is a standardized software-based mathematical algorithm that provides an iterated estimate of the motor threshold. The treatment location was over the left prefrontal cortex (determined by movement of the TMS coil 5 cm anterior to the motor threshold location along a left superior oblique angle). The sham device was identical in appearance and acoustically matched to the active NeuroStar™ TMS System. Note that the sponsor did not provide data regarding treatment dose-response characteristics.

### Study Blinding

The operator was blinded to the sham assignment by use of multiple TMS magnetic coils, all with the same exterior appearance, and the physical position on the patient's head was the same for both the sham and active coils. In addition, the sham coil design was constructed with a ferromagnetic core identical to the active TMS coil, but with a method of concealed shielding that prevented clinically meaningful transit of the magnetic field to the cortical surface, but permitted the production of an acoustic artifact that was indistinguishable from the active TMS coil when

both were turned ON. Additionally, the study staff was separated into a treating staff and a rating staff and personnel at the study sites were blind to the choice of primary effectiveness measure and to the point of declaration of the effectiveness outcome.

Concomitant Medication Use

All subjects were free of antidepressants or other psychotropic medications directed at treatment of their study diagnosis. However, subjects were allowed limited use (up to 14 daily doses) of either sedative/hypnotics or daytime anxiolytics for treatment emergent insomnia or anxiety, respectively, subsequent to the initiation of treatment. Any clinical indication for use beyond these limitations required discontinuation from study participation so as not to unduly influence the safety or effectiveness assessments.

**4.2 Study 01: Results**

A total of 325 subjects were enrolled at 23 sites. Two subjects were randomized but did not receive treatment leaving 323 subjects for the safety analyses. Of the 323 subjects, 22 (10 active and 14 sham) were not evaluable because they did not complete any post-randomization follow-up observations, leaving 301 subjects (155 active and 146 sham) for the primary effectiveness analysis (modified ITT population). The protocol defined modified ITT population represents the last observation carried forward (LOCF).

Baseline Screening Demographics

There were no statistically significant differences in the patient baseline demographics at the screening visit between the active and sham treatment groups (see Appendix B, Table B1 for a complete listing of demographic information). Table 1 below depicts the key observations for symptom severity and illness history for the ITT population stratified by ATHF category:

Table 1 – Study 01: Symptom Severity and Clinical Illness Variables for the ITT Study Population, and Stratified for the Separate ATHF Categories (1-4)

<b>Clinical Variable</b>	<b>Overall Study Population (N=301)</b>	<b>ATHF 1 (N=164)</b>	<b>ATHF 2 (N=95)</b>	<b>ATHF 3 (N=30)</b>	<b>ATHF 4 (N=12)</b>
<u>Symptom Severity at Baseline:</u>					
• MADRS Total Score Mean (Standard Deviation (SD))	32.8 (5.4)	32.1 (5.3)	33.2 (5.4)	35.0 (5.8)	33.8 (4.2)
• HAMD24 Total Score Mean (SD)	30.6 (4.1)	30.6 (4.1)	30.1 (3.9)	31.3 (3.9)	33.2 (4.5)
• HAMD17 Total Score Mean (SD)	22.8 (2.7)	22.7 (2.7)	22.6 (2.5)	23.4 (3.2)	24.0 (3.1)
<u>Illness History</u>					
• Recurrent Illness N (%)	285 (94.7)	159 (97.0)	87 (91.6)	27 (90.0)	12(100.0)
• Comorbid Anxiety Disorder Present N (%)	101 (33.6)	52 (31.7)	29 (30.5)	13 (43.3)	7 (58.3)
• Duration of current episode (months)	13.3 (9.6)	12.4 (9.5)	13.8 (9.7)	17.1(10.3)	13.3 (6.9)

Concomitant Medication Use

Approximately, 30% of subjects, in both active and sham treatment groups, used some anxiolytics. There were no statistically significant differences in the use of concomitant

medication use between the active treatment group and the sham treatment group. See Table B2, in Appendix B, for a summary of medication use during the acute treatment phase (Phase II) and Table B3 for a summary of medication use during the taper phase (Phase III).

Protocol Deviations

During Study 01 the following protocol deviations were reported: 147 protocol procedures, 115 device procedures, 92 documentation procedures, 26 excluded medications were used, and 1 regulatory documentation deviation.

**4.2.1 Study 01: Effectiveness Results**

The total protocol defined ITT evaluable population for the primary efficacy endpoint at week 4 was 301 subjects (155 active and 146 sham). A total of 16 subjects (7 active and 9 sham) discontinued prior to the week 4 visit (see Figure B in Appendix B for discontinuation reasons)

**Primary Effectiveness Endpoint**

The primary effectiveness endpoint was the last post-treatment total symptom score on the MADRS through week 4 of the acute treatment phase of the active treatment group as compared to the sham treatment group using the modified ITT population (n = 301). The difference relative to baseline between the treatment group least squares (LS) mean MADRS (-5.6) and the sham treatment group LS mean MADRS (-3.5) was not statistically significant (p=0.057). See Figure 1 for a graphical representation of the MADRS score results during Study 01.

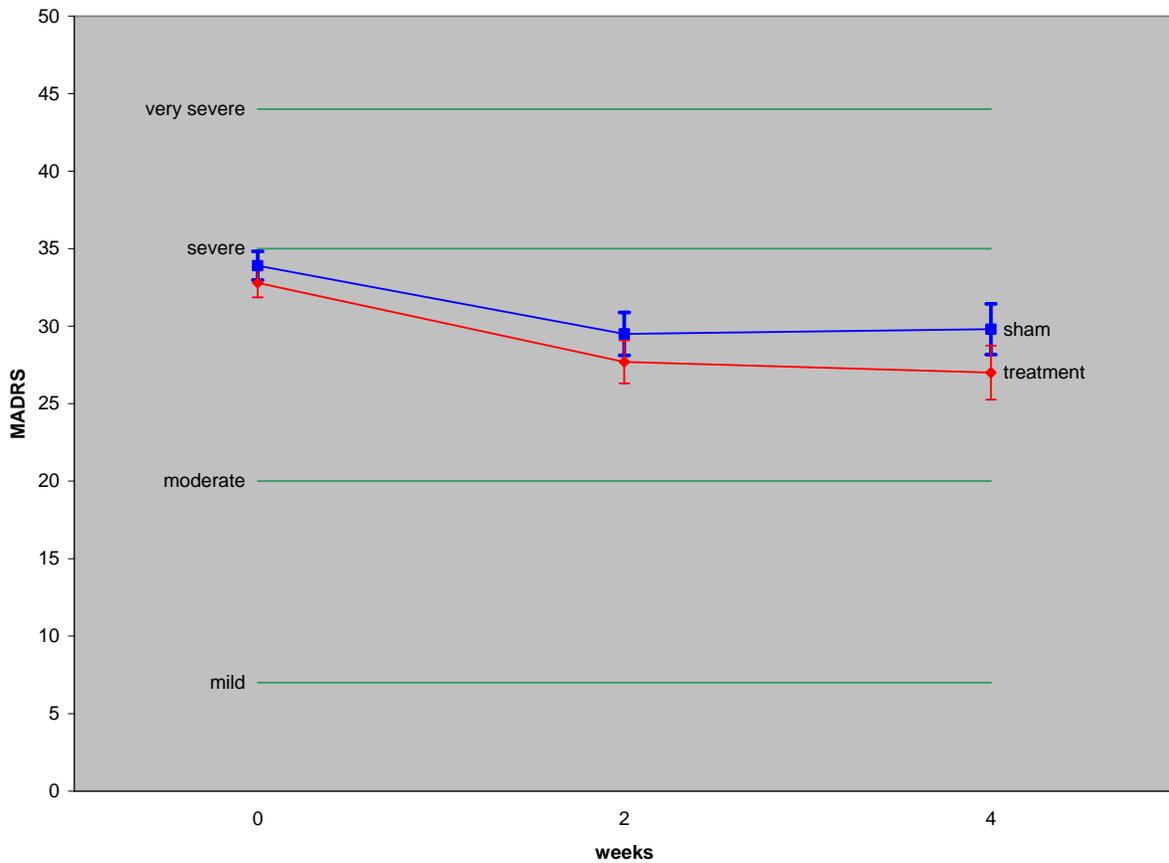


Figure 1 - Mean MADRS score ( $\bar{x}$ ) from study 01 at baseline (week 0), 2 weeks, and 4 weeks with a 95%

confidence interval.<sup>3</sup>

*Sponsor's Post-Hoc Analysis:*

The sponsor states that a statistically significant baseline imbalance ( $p = 0.036$ ) in the total MADRS score is observed between the least squares (LS) means of the active (32.4 [SD=5.99]) and sham (33.7 [SD=5.69]) treatment groups and that this arose because the baseline screening measure for study entry did not include a minimum numerical threshold for a MADRS score. The sponsor performed a post-hoc analysis of subjects by removing all subjects with a MADRS baseline cut-off of less than 20 (the sponsor states this has been shown to correspond to mild depression) in conjunction with meeting the HAMD<sub>17</sub> inclusion threshold. Six subjects (4 active (MADRS scores 14, 15, 18, and 20) and 2 sham (MADRS scores of 19)) were removed from the final study analysis. After removing this set of patients the sponsor calculates that the difference between the LS mean MADRS scores in the active treatment and the sham treatment groups at 4 weeks is statistically significant ( $p=0.038$ ).

The Panel will be asked to comment on the appropriateness of the post-hoc analysis and the results obtained.

Secondary Outcome Measures

The complete results of the secondary analyses are listed in Table B4 in Appendix B. The unadjusted results of the analyses of secondary endpoints specifically targeting depressive symptomatology are listed below in Table 2. Note that the HAMD<sub>17</sub> results were obtained from the HAMD<sub>24</sub> assessment. Additionally it should be noted that the protocol did not pre-specify an adjustment for statistical multiplicity to control for the type I error rate. However, the sponsor's consultant statistician performed four post-hoc multiplicity analyses (Holm, Hochberg, Hommel, and Benjamini-Hochberg) on 13 of the 26 endpoints. The analyses showed that the four methods agree that the primary efficacy endpoint (MADRS at Week 4) had a resultant p-value of greater than 0.05 ( $p > 0.05$ ) and between one and nine, depending on the specific analysis performed, secondary endpoints had an adjusted p-value less than 0.05 ( $p < 0.05$ ). It should be noted that the 13 endpoint scores used for the adjustment by the sponsor's consultant statistician were not chosen by him to be based on the protocol defined listing of secondary outcome measures "in the sequence of intended testing in priority order" (see Table A in Appendix A) but were chosen to include only "scored" variables measuring symptoms of depression." The following endpoint scores were included: MADRS; HAMD<sub>24</sub>; HAMD<sub>17</sub>; SF-36 Mental Health Score; HAMD Anxiety/Somatization, Core Depression, Gibbons, Retardation, and Sleep scores; IDS-SR; CGI-S; and PGI-I. Most notably the response and remission variables were not included. The sponsor states that the rationale for omitting these variables was that the study was not powered to detect differences in binary variables. Scores corresponding to outcomes other than depressive symptomatology were omitted from the analyses based on the intent to focus on the antidepressant effect of the device. In contrast to the post-hoc analyses

---

<sup>3</sup> The data for Figure 1 were taken from Table 13 of the sponsor's final report (see Volume 5, page 44 of the sponsor's submission) that computes missing data using the protocol defined LOCF method. The 95% confidence interval was calculated from the standard deviations ( $\sigma$ ) and number of subjects ( $n$ ) in each group ( $\bar{x} \pm 1.96\sigma / \sqrt{n-1}$ ). It should be noted that the variance at the beginning of the study was low because of patient selection for the study, and this variance increased over time for both groups. Note also that by applying an LOCF method for missing data, the variance by the end of the study would be underestimated. The cut-offs for mild, moderate, severe and very severe depression, were taken from Asperg, Montgomery et al. (1978) and Snaith et al (1986).

described above, FDA notes that, using the most conservative approach, a Bonferroni correction for multiple comparisons, none of the 26 secondary endpoints would have an adjusted p-value less than 0.05 (p<0.05).

The Panel will be asked to comment on the statistical and clinical meaningfulness of analyzing secondary outcomes as a measure of device effectiveness when the primary endpoint was found not statistically significant and an adjustment for statistical multiplicity to control for the type I error rate was not defined *a priori*. They will also be asked to provide a clinical interpretation of the results for the secondary endpoints.

Table 2 – Study 01: Summary Table of Efficacy Outcome Measures Specifically Targeting Depressive Symptomatology at Week 4

Primary Outcome:		LS Mean (SD)			Difference (90% CI)	P-value <sup>1</sup>
		Baseline	Week 4	Week 4 Change from Baseline		
MADRS	Active	32.4 (5.99)	26.5 (11.06)	-5.6 (11.97)	-2.1 (-3.9, -0.3)	0.057
	Sham	33.7 (5.69)	29.5 (10.11)	-3.5 (9.08)		
Clinician Rated Secondary Outcomes:		LS Mean (SD)			Difference (90% CI)	P-value <sup>1</sup>
		Baseline	Week 4	Week 4 Change from Baseline		
HAMD <sub>24</sub>	Active	29.9 (5.04)	23.1 (8.93)	-6.5 (8.36)	-2.4 (-4.0, -0.8)	0.012
	Sham	30.2 (4.85)	25.7 (8.81)	-4.1 (8.49)		
HAMD <sub>17</sub>	Active	22.5 (3.3)	17.3 (6.49)	-5.0 (6.28)	-1.9 (-3.1, -0.7)	0.006
	Sham	22.8 (3.54)	19.3 (6.51)	-3.1 (6.08)		
CGI-S	Active	4.7 (0.62)	4.1 (1.1)	-0.6 (1.05)	-0.4 (-0.6,-0.2)	0.009
	Sham	4.7 (0.72)	4.4 (1.09)	-0.2 (1.07)		
		Number of Patients at Week 4 – N (%)			Difference (90% CI)	P-value <sup>2</sup>
MADRS Responders <sup>3</sup>	Active	28 (18.1%)			7.1% (0.2, 13.9)	0.045
	Sham	16 (11.0%)				
HAMD <sub>24</sub> Responders <sup>3</sup>	Active	30 (19.4%)			7.7% (0.7, 14.7)	0.030
	Sham	17 (11.6%)				
HAMD <sub>17</sub> Responders <sup>3</sup>	Active	32 (20.6%)			9.0% (1.7, 16.1)	0.018
	Sham	17 (11.6%)				
MADRS Remitters <sup>4</sup>	Active	11 (7.1%)			0.9% (-4.2, 5.9)	0.633
	Sham	9 (6.2%)				
HAMD <sub>24</sub> Remitters <sup>5</sup>	Active	14 (9.0%)			0.8% (-4.7, 6.4)	0.644
	Sham	12 (8.2%)				
HAMD <sub>17</sub> Remitters <sup>6</sup>	Active	11 (7.1%)			0.9% (-4.2, 5.9)	0.705
	Sham	9 (6.2%)				
Patient Rated Secondary Outcomes:		LS Mean (SD)			Difference (90% CI)	P-value <sup>1</sup>
		Baseline	Week 4	Week 4 Change from Baseline		
IDS-SR	Active	41.3 (9.4)	33.6 (13.3)	-7.7 (11.88)	-2.5 (-4.8,-0.2)	0.058
	Sham	42.8 (9.89)	37.1 (13.35)	-5.2 (11.84)		
Q-LES-Q	Active	37.8 (8.23)	41.4 (10.32)	3.5 (9.9)	1.5	0.124

	Sham	36.5 (7.87)	39 (9.78)	2.0 (9.24)		
PGI-I	Active	4.3 (0.86)	3.7 (1.37)	-0.6 (1.61)	-0.3 (-0.6, .02)	0.181
	Sham	4.4 (0.99)	3.9 (1.43)	-0.3 (1.72)		

<sup>1</sup> P-value is from the following ANCOVA model: Change from baseline = Baseline Score, ATHF group, center, and treatment

<sup>2</sup> P-value is from the following logistic regression model: Responder = ATHF group, center, and treatment

<sup>3</sup> A responder has a > 50% change from baseline score

<sup>4</sup> Total Score < 10

<sup>5</sup> Total Score < 11

<sup>6</sup> Total Score < 8

#### Post-Hoc Sensitivity Analyses:

The protocol *a priori* defined method of imputation for this missing data, LOCF, could have biased the results in favor of the treatment if there had been an early response followed by an unobserved later decline. Thus, the FDA requested that the sponsor also perform post-hoc sensitivity analyses including an analysis of completers only and an analysis using multiple imputation methods. The results are presented in Table 3 below:

Table 3 - Sensitivity Analysis for the Contrast between Active and Sham Treatment Mean Changes from Baseline at the Primary Outcome Time point of Week 4

Outcome Variable:	P-Values for Contrast Between Active and Sham Treatment Means Changes from Baseline: Primary Effectiveness Time point (Week 4)		
	A-Priori LOCF Analysis (N=301)	Week 6 Completers Only Analysis (N=153)	Multiple Imputation Analysis <sup>1</sup> (N=325)
MADRS Total Score	0.057	0.819	0.090
HAMD <sub>24</sub> Total Score	0.012	0.700	0.008
HAMD <sub>17</sub> Total Score	0.006	0.929	0.004

<sup>1</sup> All-Randomized ITT Population

#### Standardized Effect Sizes

The sponsor *a priori* defined in their investigational plan a standardized effect size, based on the MADRS, of 0.4 (d=0.4) as the “minimally clinically interesting difference” to be obtained in the investigation. Effect sizes have loosely been defined, in absolute value, as “small, d = .2,” “medium, d = .5,” and “large, d = .8” (Cohen, 1988). The resultant effect sizes reported by the sponsor are -0.386, -0.482, and -0.547 for the MADRS, HAMD<sub>24</sub> and HAMD<sub>17</sub> respectively (See Appendix E1). The standardized effect sizes calculated by the FDA are -0.355<sup>4</sup>, -0.481<sup>4</sup>, and -0.556<sup>4</sup> for the MADRS, HAMD<sub>24</sub> and HAMD<sub>17</sub>

<sup>4</sup> FDA calculated the effect sizes in the Study 01 by taking the difference between the LS Means for the change from baseline divided by the pooled baseline standard deviation for the “Total Score”. These values were taken from Final Study Report, Study 44-01101, Tables 13, 16 and 17 for the MADRS, HAMD<sub>24</sub> and HAMD<sub>17</sub> respectively. These calculations differ from the sponsor’s calculations because the sponsor appeared to use a GLM model to derive the pooled baseline standard deviation. This model consisted of the total score as the dependent variable and treatment group and study center as the independent variables. FDA feels that the sponsor’s model based approach may underestimate the baseline standard deviation as the independent variables may be accounting for variation that would otherwise normally be present in the measuring instrument. Additionally, it should be noted that the effect

respectively.

Post-Hoc Evaluation of Response by Treatment Resistance

The FDA requested a post-hoc evaluation of the response by the level of treatment resistance as measured by the ATHF. The ATHF was developed to organize information from various sources about the treatment history of patients with major depression in order to rate the adequacy of the medication trials. The ATHF focuses on several key aspects of the medication trial being assessed: adequacy of dosage, duration of trial, compliance, outcome, and confidence in the overall adequacy of the trial. Thus, the ATHF rating is based on the number of adequate exposures to a drug and does not include medications that were tried but not adequately exposed to the subject. The average ATHF Level score for the subjects enrolled in this study was 1.6. The results of the FDA requested analyses indicate that ATHF Group 1 showed a better response to treatment with the NeuroStar™ TMS System as compared with ATHF Groups 2 and 3. However, it should be noted that the number of subjects enrolled in ATHF Groups 2, 3, and 4 were small compared to those in ATHF Group 1. Table 4 below summarizes the standardized effect sizes and associated p-values for the primary effectiveness outcome (MADRS) and the secondary outcomes HAMD<sub>24</sub>, HAMD<sub>17</sub>, and the IDS-SR at Week 4 (see Table B5 in Appendix B for a responder rate breakdown analyses).

Table 4 - Study 01: Standardized Effect Sizes and Associated P-Values for Primary and Secondary Outcome Measures at Week 4

<b>Primary Effectiveness Outcome Measure</b>	<b>Active Treatment N</b>	<b>Sham Treatment N</b>	<b>Standardized Effect Size Week 4</b>	<b>p-Value Week 4</b>
MADRS Total Score (Overall Sample)	155	146	0.39	0.057
• ATHF Group 1	88 (57%)	76 (52%)	0.94	0.001
• ATHF Group 2	45 (29%)	50 (34%)	-0.16	0.710
• ATHF Group 3	15 (10%)	15 (10%)	-0.55	0.588
• ATHF Group 4	7 (5%)	5 (3%)	5.21	0.022
<b>Secondary Effectiveness Outcome Measures</b>				
HAMD <sub>24</sub> Total Score (Overall Sample)	155	146	0.48	0.012
• ATHF Group 1	88 (57%)	76 (52%)	0.83	0.001
• ATHF Group 2	45 (29%)	50 (34%)	0.03	0.933
• ATHF Group 3	15 (10%)	15 (10%)	0.44	0.577
• ATHF Group 4	7 (5%)	5 (3%)	2.41	0.077
HAMD <sub>17</sub> Total Score (Overall Sample)	155	146	0.55	0.006
• ATHF Group 1	88 (57%)	76 (52%)	0.83	0.001
• ATHF Group 2	45 (29%)	50 (34%)	0.13	0.762
• ATHF Group 3	15 (10%)	15 (10%)	0.81	0.440
• ATHF Group 4	7 (5%)	5 (3%)	2.26	0.089
IDS-SR Total Score (Overall Sample)	155	146	0.27	0.059
• ATHF Group 1	88 (57%)	76 (52%)	0.57	0.002
• ATHF Group 2	45 (29%)	50 (34%)	0.10	0.710

size analysis was a *post-hoc* analysis requested by the FDA and thus, there was no *a priori* specification of a model. The sponsor chose to use the GLM as employed for the primary analysis. Therefore, FDA reported results obtained by using the pooled baseline standard deviations of the “Total Score” as found in Tables 13, 16, and 17 of the Study 01 report.

• ATHF Group 3	15 (10%)	15 (10%)	0.29	0.706
• ATHF Group 4	7 (5%)	5 (3%)	1.85	0.269

The FDA also requested that the sponsor provide a table listing the number of subjects in each group by CGI-S (Table 5 below) and PGI-I (Table B6 in Appendix B) at baseline and week 4 (the primary efficacy endpoint time-point) follow-up.

Table 5 – Study 01: Clinician Global Impressions – Severity Rating: Distribution of Scores By Rating At Baseline and Week 4 of Acute Treatment Phase

CGI-Severity Rating	Baseline		Week 4	
	Active (N=155)	Sham (N=146)	Active (N=155)	Sham (N=146)
1 – Normal, Not at All Ill	0	0	5	3
2 – Borderline Ill	0	0	6	4
3 – Mildly Ill	0	2	24	14
4 – Moderately Ill	56	55	62	52
5 – Markedly Ill	85	67	46	53
6 – Severely Ill	14	22	12	18
7 – Among the Most Extremely Ill	0	0	0	2

If the Panel believes the NeuroStar™ TMS System has a favorable risk-to-benefit profile they will be asked to comment on the patient population that is most appropriate for treatment with this device.

#### **Week 6 Effectiveness Results**

Only 145 patients (48% of the ITT patient population (n=301); 86 Active and 59 Sham) completed the week 6 visit. This was primarily due to subjects who reported a lack of efficacy leaving Study 01 early to enter Study 02. A total of 132 subjects discontinued prior to the 6 week visit, including 119 (50 Active and 69 Sham) for unsatisfactory effectiveness response (see Figure B in Appendix B for a complete accounting of discontinued patients). Two patients exited the acute phase at week 4 and began the taper phase (Phase III) immediately.

Because of the large number of patient’s discontinuing between week 4 and week 6, and the resultant amount of imputed data, the Panel will be asked to comment on whether the clinical outcomes reported during week 6 follow-up visits provide any clinically relevant information.

The ITT patient population, using LOCF, at week 6 was 301 patients. For complete results for week 6 please refer to Table B7 in Appendix B. Using the *a-priori* protocol defined LOCF ITT analysis, the difference between the LS mean MADRS between the active treatment group (LS mean = -5.6) and the sham treatment group (LS mean = -3.2) was not statistically significant (p=0.058) at 6 weeks. Some of the clinician rated secondary outcomes had p-values less than 0.05 (HAMD<sub>24</sub>; HAMD<sub>17</sub>; HAMD<sub>24</sub>, HAMD<sub>17</sub>, and MADRS Responders; MADRS, and HAMD<sub>24</sub> Remitters; and CGI-S) while the HAMD<sub>17</sub> Remitters did not. The patient rated IDS-SR and PGI-I had p-values greater than 0.05 and the Q-LES-Q had a p-value less than 0.05.

#### ***Post-Hoc Sensitivity Analyses:***

The protocol defined method of imputation for this missing data, LOCF, could have biased the results in favor of the treatment if there had been an early response followed by an

unobserved later decline. Therefore, the FDA requested that the sponsor also perform post-hoc sensitivity analyses including an analysis of completers only and an analysis using multiple imputation methods (note that the assumption of missing at random for the multiple imputation analysis may not be applicable here). The results are presented in Table 6 below:

Table 6 - Sensitivity Analysis for the Contrast between Active and Sham Treatment Mean Changes from Baseline at week 6

Outcome Variable:	P-Values for Contrast Between Active and Sham Treatment Means Changes from Baseline: Week 6		
	A-Priori LOCF Analysis (N=301)	Completers Only Analysis (N=153)	Multiple Imputation Analysis <sup>1</sup> (N=325)
MADRS Total Score	0.058	0.881	0.276
HAMD <sub>24</sub> Total Score	0.015	0.984	0.173
HAMD <sub>17</sub> Total Score	0.005	0.615	0.078

<sup>1</sup> All-Randomized ITT Population

### **Phase III (Taper Phase) Effectiveness Results**

Phase III, or the taper phase, of Study 01 consisted of a three week taper where all subjects were placed onto antidepressant pharmacotherapy and simultaneously tapered off treatment with the NeuroStar™ TMS System in a schedule of gradually less frequent treatment sessions (three times per week, twice per week and then once per week). The blind remained intact. Clinical and safety evaluations occurred weekly. One-hundred and four (104) subjects (64 Active, 40 Sham) completed week 1 of the taper phase, 97 subjects (59 Active, 38 Sham) completed week 2 of the taper phase, and 89 subjects (54 Active, 35 Sham) completed week 3 of the taper phase (see Figure B in Appendix B for a table of discontinuation reasons).

Because this phase of the study involved open-label pharmacotherapy, which confounds the device effects with the drug effects, only descriptive statistics were reported by the sponsor as stated *a priori* in the study protocol. The mean total score at week 6 was maintained through week 3 of the taper (week 9 of the study) for the MADRS, HAMD<sub>24</sub>, and HAMD<sub>17</sub> suggesting that any clinical effect of treatment with the NeuroStar™ TMS System may have been sustained. It is also noted that the HAMD<sub>17</sub> Responder ( $\geq 50\%$  reduction from baseline) and Remission Rate (Total Score < 8) were greater in the active treatment group than the sham treatment group throughout the taper phase.

### **4.2.2 Study 01: Safety Results**

#### **Adverse events**

A total of 325 subjects were enrolled at 23 sites. Two subjects were randomized but did not receive treatment leaving 323 subjects (158 Sham, 165 Active) for safety analyses. Investigational sites were queried for adverse event information at each study visit up through 30 days after the last study visit in all clinical protocols. All adverse events, regardless of relationship to stimulation, were recorded. Headache, which was the most common adverse event, was experienced in similar incidence across each treatment arm.

Table B8 in Appendix B summarizes adverse events that occurred at and incidence of  $\geq 2\%$  in the active treatment group and were greater than those experienced by the sham treatment group. Adverse events with an incidence of  $\geq 10\%$ , reported in the order of most common occurrence in the active treatment group, are shown in the table below:

Table 7 - Study 01: Adverse Events with an Incidence  $\geq 10\%$

<b>Adverse Event:</b>	<b>Active (n=165) N (%)</b>	<b>Sham (n=158) N (%)</b>
Headache	96 (58.2)	87 (55.1)
Application site pain	59 (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)

#### Serious Adverse Events

A total of 23 serious adverse events were reported during Study 01. See Table 8 below for a listing of these events.

Table 8 - Study 01: Serious Adverse Events.

<b>Serious Adverse Event</b>	<b>Active (n=165) N</b>	<b>Sham (n=158) N</b>
Worsening of major depression	1	2
Suicidal ideation	1	3
Overdose*	5	0
Device malfunction/first degree burn	1	1
Suicide attempt	0	1
Device malfunction/severe pain at treatment site	1	0
Lower lobe pneumonia	0	1
Bowel obstruction	0	1
<b>All Serious Adverse Events Reported<sup>†</sup></b>	<b>9</b>	<b>9</b>

<sup>†</sup> Five serious adverse events, not depicted in this table, were reported prior to randomization to sham or active treatment groups, including worsening depression (2), suicidal ideation (2), shortness of breath and increased heart rate (1) were reported prior to randomization to treatment groups.

\* Overdose refers to events associated with inadvertent treatment of  $>75$  trains of active TMS delivered to the subject on a single calendar day.

#### Study Blinding Considerations

Application site pain occurred at a much greater frequency in the active treatment group (35.8% or 59/165) than in the sham treatment group (3.8% or 6/158). This potentially created an increased placebo response in the subjects who received active treatment as compared to those who received sham treatment and thus, the FDA requested further analyses to examine the adequacy of the patient blind. The sponsor states that headache and application site pain occurred early in the treatment and at the time points of effectiveness outcome measurements (4 and 6 weeks) the incidence of these adverse events had fallen to levels substantially less than fifty percent of the incidence seen during the first week. Based on these observations, the sponsor believes that the incidence and the temporal pattern of these adverse events was unlikely to contribute to the penetration of the

study blind at a rate any different than for similarly designed studies for antidepressant medications.

The FDA requested that the sponsor perform additional analyses to examine whether the individual types of pain/discomfort reported using the MedDRA preferred adverse events terms (application site pain, eye pain, facial pain, jaw pain, toothache, application site discomfort, and muscle twitching) that occurred in the study had an effect on blinding and thus, final study results (particularly the clinical response using the primary effectiveness endpoint of the MADRS). The sponsor performed two analyses. The sponsor first determined a Spearman correlation coefficient for the relationship between mean change on the MADRS total score and the severity of the adverse event term (i.e., none, mild, moderate, or severe). In the second analysis, an ANCOVA model identical to that used in the *a priori* stipulated ANCOVA analysis was used to examine the difference in mean change from baseline of the MADRS total scores between patients with and without the adverse event. This analysis was done within each treatment group separately and also for the total group.

In the first analysis, the sponsor reported there was no statistically significant observation in the most commonly occurring term, “application site pain”, and only five instances of other events (eye pain, application site discomfort, and aggregate pain). In aggregate (Table 9), the analyses showed a correlation between the presence and severity of any pain/discomfort and mean change in MADRS ( $p = 0.034$ ). This observation suggests that a placebo effect may have occurred during the Study 01. However, there was not a significant relationship between mean change in MADRS score and any occurrence of pain/discomfort ( $p = 0.327$ ). In contrast, when covariate adjusted analyses were performed with inclusion of any pain/discomfort as a covariate, the impact on the observed treatment effects was mixed and two of the assessments were no longer statistically significant. The p-value for the MADRS endpoint went from 0.057 to 0.227. The p-value for HAMD<sub>24</sub> endpoint went from 0.012 to 0.054 and the p-value for HAMD<sub>17</sub> went from 0.006 to 0.020. Based on these findings, the report of any pain in aggregate could account for some of the observed treatment effect. When evaluating application site pain specifically (Table 10), the correlation was not significant in the two treatment groups when examined separately.

Table 9 – Study 01: Relation of Presence and Severity of “Any Pain/Discomfort” Term to Week 4 Effectiveness Outcome

<b>Presence of Adverse Event Term</b>	<b>N</b>	<b>Total Group Mean Change (SD)</b>	<b>N</b>	<b>Active Treatment Mean Change (SD)</b>	<b>N</b>	<b>Sham Treatment Mean Change (SD)</b>
No	220	-4.43 (9.89)	84	-5.02 (11.18)	136	-4.07 (9.03)
Yes	81	-6.46 (9.04)	71	-6.65 (8.92)	10	-5.10 (10.20)
Mild	35	-4.74 (9.48)	27	-4.63 (9.18)	8	-5.13 (11.12)
Moderate	35	-6.69 (7.69)	33	-6.79 (7.77)	2	-5.00 (8.49)
Severe	11	-11.18 (10.55)	11	-11.18 (10.55)	0	--
Correlation Coefficient		-0.12		-0.15		-0.01
P-Value (Spearman correlation)		0.034		0.063		0.908
P-Value (Mean Change		0.327		0.288		0.932

b/w groups [AE present Yes/No]						
--------------------------------	--	--	--	--	--	--

Table 10 - Relation of Presence and Severity of “Application Site Pain” Term to Week 4 Effectiveness Outcome

<b>Presence of Adverse Event Term</b>	<b>N</b>	<b>Total Group Mean Change (SD)</b>	<b>N</b>	<b>Active Treatment Mean Change (SD)</b>	<b>N</b>	<b>Sham Treatment Mean Change (SD)</b>
No	257	-4.85 (9.86)	113	-5.70 (10.70)	144	-4.18 (9.13)
Yes	44	-5.70 (8.77)	42	-5.93 (8.89)	2	-1.00 (4.24)
Mild	15	-1.47 (7.25)	13	-1.54 (7.73)	2	-1.00 (4.24)
Moderate	24	-7.17 (7.83)	24	-7.17 (7.83)	0	--
Severe	5	-11.40 (13.01)	5	-11.40 (13.01)	0	--
Correlation Coefficient		-0.05		-0.05		0.04
P-Value (Spearman correlation)		0.401		0.533		0.663
P-Value (Mean Change b/w groups [AE present Yes/No])		0.734		0.789		0.545

The sponsor also examined potential clinically meaningful relationships between pain/discomfort and clinical outcome, including effectiveness outcomes based on the five subjects observed to have a severe report of any of the MedDRA terms of interest during week 1 of the acute phase. The sponsor believes these additional analyses suggest no statistically or clinically meaningful relationship between the presence or severity of any of the indicated MedDRA-defined adverse event terms and clinical outcome in Study 01.

### **4.3 Study 02: Investigational Plan**

#### Objective

To evaluate the antidepressant effect (using last post-treatment total symptom score on the MADRS) of a specified treatment course of treatment with the NeuroStar™ TMS System in subjects meeting DSM-IV criteria for Major Depressive Episode, single or recurrent episode who did not show an acute clinical response (as defined as a reduction in the baseline total HAMD<sub>17</sub> score greater than or equal to 25%) to daily dose active of sham treatment administered for 4 to 6 weeks in Study 01.

#### Design

A 9 week, uncontrolled open-label, clinical trial for patients who did not meet pre-defined criteria for response in Study 01. Subjects remained blinded to their treatment allocation in Study 01.

#### Primary Effectiveness Endpoint

The primary effectiveness endpoint was a descriptive report of the symptom changes (using the last post-treatment total symptom score on the MADRS) observed over 6 weeks of open-label treatment.

### Secondary Outcome Measures

The protocol included a listing of secondary outcome measures “in the sequence of intended testing in priority order” (see Table A in Appendix A). Secondary effectiveness outcome measures included both clinician rated and patient rated outcome measures. These are the same assessments as used in Study 01.

### Safety Assessments

Safety and tolerability of treatment was assessed by observation of adverse events and serious adverse events.

### Protocol

With the exception that all subjects received stimulation, this protocol was identical in design and treatment sequence to Study 01. Treatment assignment blinding during Study 01 was maintained as patients entered Study 02. Subjects were classified into 2 groups:

- Group A: Subjects who were randomized to active treatment in Study 01, did not respond, and agreed to enter Study 02.
- Group B: Subjects who were randomized to sham treatment in Study 01, did not respond, and agreed to enter Study 02.

Psychotropic medication use during the study was limited.

## **4.4 Study 02: Results**

During Study 02, a total of 166 subjects enrolled at 22 sites. Of the 166 subjects, 8 subjects were non-evaluable (did not complete endpoint observations) leaving 158 subjects in the final evaluable patient study population (i.e., the evaluable population).

### Demographics

There were no statistically significant differences between Group A and Group B on any demographic variables.

### Concomitant Treatment

All subjects were free of antidepressants or other psychotropic medications directed at treatment of their study diagnosis. Psychotropic medication use during the study was limited. Subjects were allowed limited use of either sedative/hypnotics or daytime anxiolytics for treatment emergent insomnia or anxiety, respectively, subsequent to the initiation of treatment. These medications were permitted for up to 14 daily doses (of either or both types of medications) during the acute treatment phase. Any clinical indication for use beyond these limitations required discontinuation from study participation in the interest of subject care and so as not to unduly influence the effectiveness or safety assessments. Approximately, 30% of subjects had some anxiolytics use in both active and sham treatment groups. See Tables C1 and C2 in Appendix C for a complete accounting of medication use.

### Protocol Deviations

During Study 02 the following protocol deviations were reported, including 107 protocol procedures, 52 device procedures, 24 documentation procedures, and 15 excluded medications used.

### Patient Discontinuation

One hundred and fifty-eight (158) subjects enrolled in the study, 144 subjects completed the 4

week visit of the acute phase, 131 subjects completed the 6 week visit, and 114 subjects completed the 3 week taper phase visit. See Table C3 in Appendix C.

**4.4.1 Study 02: Effectiveness Results**

Primary Effectiveness Endpoint

The primary effectiveness endpoint was a descriptive report of the symptom changes (using the last post-treatment total symptom score on the MADRS) observed over 6 weeks of open-label treatment who did not show an acute clinical response to daily dose active or sham treatment administered for 6 weeks. The evaluable population (N total =131; Group A = 73; Group B = 85) was assessed. Subjects previously exposed to active treatment with the NeuroStar™ TMS System (Group A) observed a mean decrease of -12.5 in the MADRS total score at 6 weeks, in comparison to the last assessment obtained in Study 01, prior to entry into Study 02. Subjects previously exposed to sham treatment (Group B) observed a mean decrease of -17 in the MADRS total score at 6 weeks, in comparison to the last assessment obtained in Study 01, prior to entry into Study 02.

As with any trial for MDD, there is a potential for a significant placebo effect and this should be taken into account when interpreting results from this open label trial (Study 02). It is noted that in the Study 01 designated active treatment group, for the MADRS score, the LS mean change from baseline at week 4 of Study 01 (-5.6 (see Table 2)) is much lower than the mean change from baseline observed at week 6 of Study 02 (-12.5 (see Table 11). This may indicate a placebo effect in the open label study.

Considering the open label design of Study 02, the Panel will be asked to discuss any conclusions that can be drawn from the study results.

Secondary Outcome Measures

The secondary outcome measures were the same measures evaluated in Study 01. The results of the analyses of secondary endpoints at 6 weeks are listed below in Table 11.

Table 11 – Study 02: Summary Table of Secondary Outcome Measures at Week 6

Primary Outcome:	LS Mean Change from Baseline <sup>1</sup>	
	Group A <sup>3</sup> N=73	Group B <sup>4</sup> N=85
	MADRS	-12.5
<b>Clinician Rated Secondary Outcomes:</b>		
	LS Mean Change from Baseline <sup>1</sup>	
	Group A	Group B
	HAMD <sub>24</sub>	-11.1
HAMD <sub>17</sub>	-8.2	-10.8
	Percentage (%)	
	Group A	Group B
	MADRS Responders <sup>2</sup>	26.0
HAMD <sub>24</sub> Responders <sup>2</sup>	31.5	42.4
HAMD <sub>17</sub> Responders <sup>2</sup>	30.1	37.6
MADRS Remitters (Total Score < 10)	11.0	20.0

HAMD <sub>24</sub> Remitters (Total Score < 11)	16.4	27.1
HAMD <sub>17</sub> Remitters (Total Score < 8)	15.1	21.2
<b>Patient Rated Secondary Outcome:</b>		
	<b>LS Mean Change from Baseline<sup>1</sup></b>	
	<b>Group A</b>	<b>Group B</b>
IDS-SR	-9.9	-16.8

<sup>1</sup> Mean change from total score observed at baseline upon entry into Study 02

<sup>2</sup> A responder has a > 50% change from baseline score at entry into Study 02

<sup>3</sup> Group A are subjects who were randomized to active treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

<sup>4</sup> Group B are subjects who were randomized to sham treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

#### 4.4.2 Study 02: Safety Results

##### Adverse events

All adverse events regardless of relationship to stimulation were recorded. Table C4 in Appendix C summarizes adverse events that occurred at and incidence of  $\geq 2\%$  in the active treatment group and were greater than those experienced by the sham treatment group. Adverse events with an incidence of  $\geq 10\%$  reported in the order of most common occurrence in Group B are shown in the table below:

Table 12 - Study 02: Adverse Events with an incidence on Active rTMS of  $\geq 10\%$  Incidence on Active Treatment in either Group A or Group B.

<b>Adverse Event:</b>	<b>Group A<sup>1</sup></b> <b>(n=73)</b> <b>N (%)</b>	<b>Group B<sup>2</sup></b> <b>(n=85)</b> <b>N (%)</b>
Headache	35 (47.9)	39 (45.9)
Application site pain	8 (11.0)	27 (31.8)
Insomnia	22 (30.1)	22 (25.9)
Muscle twitching	15 (20.5)	18 (21.2)
Anxiety	11 (15.1)	12 (14.1)
Nausea	10 (13.7)	6 (7.1)

<sup>1</sup> Group A are subjects who were randomized to active treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

<sup>2</sup> Group B are subjects who were randomized to sham treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

##### Serious Adverse Events

See Table 13 below for a listing of serious adverse events:

Table 13 - Study 02: Serious Adverse Events.

<b>Serious Adverse Event</b>	<b>Group A<sup>1</sup></b> <b>(n=73)</b> <b>N</b>	<b>Group B<sup>2</sup></b> <b>(n=85)</b> <b>N</b>
Overdose*	2	2
Atrial fibrillation	0	2

Worsening depression/suicidal ideation	1	1
Left-sided facial numbness	0	1
Suicidal ideation	0	1
Tinnitus	1	0
Worsening of major depression	0	0
All Serious Adverse Events Reported <sup>†</sup>	4	7

<sup>†</sup> One (1) serious adverse event of worsening depression was not identified from Group A.

\* Overdose refers to events associated with inadvertent treatment of >75 trains of active TMS delivered to the subject on a single calendar day.

<sup>1</sup> Group A are subjects who were randomized to active treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

<sup>2</sup> Group B are subjects who were randomized to sham treatment in Study 01, did not respond based on subject self-report, and agreed to enter Study 02.

#### **4.5 Study 03: Investigational Plan**

##### Objective

The primary objective was to evaluate the effectiveness of maintenance pharmacotherapy (antidepressant monotherapy) by assessment at 24 weeks of monotherapy in subjects who met pre-defined criteria for response (i.e., a reduction in baseline total HAMD<sub>17</sub>  $\geq$  25%) upon exit from Study 01 or Study 02. Study 03 also permitted enrollment in the event that a subject had symptom recurrence despite adequate oral antidepressant therapy.

##### Design

This is a 24 week open-label, uncontrolled, clinical trial.

##### Primary Effectiveness Endpoints

The primary effectiveness endpoints were the proportion of subjects remaining relapse-free and the proportion of subjects requiring reintroduction of rTMS with the NeuroStar<sup>TM</sup> TMS System at week 4. Relapse is defined as:

- Recurrence of full criteria for major depression as defined by DSM-IV criteria (confirmed upon two observations over a two week interval of time), or
- Failure of symptom improvement despite administration of a full course of reintroduction of a full course of treatment with the NeuroStar<sup>TM</sup> TMS System.

##### Secondary Outcome Measures

Secondary outcome measures included evaluation of the proportion of subjects who have not experienced the criterion of symptomatic worsening as described in the protocol section below. The longitudinal symptom scores and the change from baseline (i.e., the last assessment prior to entry into Study 03) were also evaluated.

##### Safety Assessments

Safety and tolerability of treatment was assessed by observation of adverse events and serious adverse events.

##### Concomitant Treatment

The pharmacotherapy regimen was limited in order to minimize excessive heterogeneity of medication selection which would confound the final study results. Only monotherapy was permitted and patients were not permitted to be tapered to antidepressant medications to which they had previously not responded. The dose of medication was to be optimized within the

labeled dose range for the specific medication as clinically indicated but no switching of medication was permitted and no augmentation or medication combination regimens were allowed.

#### Protocol

There are four potential routes of entry into Study 03, and they represent the four separate populations contained in the study analysis. The first three groups represent the various paths for active rTMS treated subjects to enter Study 03, while the fourth group represents the sham rTMS responders from Study 01:

- Group 1: Subjects who were randomized to active treatment in Study 01, responded, and agreed to enter Study 03 [Study 01 Active treatment responders].
- Group 2: Subjects who were randomized to active treatment in Study 01, did not respond, and who agreed to enter Study 02, received a course of open-label active rTMS, responded to that course of treatment and then agreed to enter Study 03 [Study 01 Active treatment non-responders/Study 02 responders].
- Group 3: Subjects who were randomized to sham treatment in Study 01, did not respond, agreed to enter Study 02, received a course of open-label active rTMS, and then agreed to enter Study 03 [Study 01 Sham treatment nonresponders/ Study 02 responders].
- Group 4: Subjects who received sham treatment in Study 01, responded to treatment and subsequently agreed to enter Study 03 [Study 01 Sham responders].

In the event that the subject's CGI-S score worsens 1 point or more from the preceding visit, then the subject must be rescheduled for repeat clinical assessment within 1 week. If this symptom change is confirmed at that visit, then the subject is considered to have met criteria for clinical deterioration and open label rTMS treatment with the NeuroStar™ TMS System in conjunction with continuation of pharmacotherapy, for up to 24 sessions across 6 weeks.

#### **4.6 Study 03: Results**

Note that the data reported for Study 03 are interim results since the study was still ongoing at the time of 510(k) submission. During Study 03, at the time of this interim study report, 136 subjects (41.8% of total subjects enrolled in Study 01) had been enrolled at 22 sites (Group 1: N=44; Group 2, N=27; Group3, N=42; Group 4, N=23). At the end of Week 4, only four subjects had discontinued treatment, and one subject was ongoing but had not yet reached the Week 4 time point, leaving N=131 subjects available for analysis.

#### Concomitant Treatment

Concomitant medication use will be provided in a final study report and was not part of the submission.

#### **4.6.1 Study 03: Effectiveness Results**

##### Primary Effectiveness Endpoint

The portion of Study 03 dataset that is of primary interest regarding durability of effect is the first four weeks for patients in Group 1 (Study 01 Active treatment responders). Coupled with the data for this group from the taper phase (Phase III) of Study 01, this dataset allows a descriptive view of this cohort of patients across 7 weeks after exit from the treatment phase. However, given that this is open label study that allowed

antidepressant medications (responders could be similarly responding to the medication), and there was a significant portion of missing data reported in Study 01 leading into Study 03 (only 41.8% of total subjects enrolled in Study 01 were enrolled in Study 03), caution should be taken when interpreting these results. Table 14 below represents the interim study results.

Table 14 – Study 03 Interim Results

	<b>Time Point</b>	<b>Group 1 N=44</b>	<b>Group 2 N=27</b>	<b>Group 3 N=42</b>	<b>Group 4 N=23</b>
<b>Percent of Patients Requiring TMS Reintroduction</b>	24 Weeks	36.4 %	33.3 %	38.1 %	47.8 %
<b>Median Time to 1<sup>st</sup> TMS Reintroduction (weeks)</b>	-	11	7	6.5	10
<b>Relapse Rates<sup>1</sup>:</b>					
• <b>Protocol Defined<sup>2</sup></b>	4 Weeks	2.3 %	0.0 %	7.2 %	4.3 %
	24 weeks	9.1 %	14.8 %	14.4 %	17.3 %
• <b>Literature Defined<sup>3</sup></b>	4 Weeks	9.1 %	11.1 %	9.6 %	12.9 %
	24 weeks	20.5 %	22.2 %	26.3 %	25.8 %

<sup>1</sup> The FDA notes that if one considers the mean change in MADRS score for the active and sham groups, 5.6 and 3.5 respectively, the relapse rates are conservative estimates of relapse.

<sup>2</sup> Discontinuation from the study for all causes.

<sup>3</sup> Literature defined relapse definition of HAMD<sub>24</sub> > 16 on two consecutive visits and an absolute increase of 10 points from the Study 03 entry HAMD<sub>24</sub>.

#### Secondary Outcome Measures

Secondary analyses evaluated the longitudinal symptom scores and the change from baseline (i.e., the last assessment prior to entry into protocol 03). Effectiveness results for Group 1 patients are shown in Table D1 in Appendix D.

#### **4.6.2 Study 03: Safety Results**

##### Adverse events

All adverse events, regardless of relationship to stimulation were recorded (See Table D2 in Appendix D). Adverse events with an incidence of  $\geq 10\%$  reported in the order of most common occurrence are shown in the table below:

Table 15 - Study 03: Adverse Events with an Incidence on Active rTMS of  $\geq 10\%$  Incidence in Any Treatment Group

<b>Adverse Event:</b>	<b>Group 1 (n=44) N (%)</b>	<b>Group 2 (n=27) N (%)</b>	<b>Group 3 (n=42) N (%)</b>	<b>Group 4 (n=23) N (%)</b>
Headache	16 (36.4)	9 (33.3)	13 (31.0)	10 (43.5)
Insomnia	13 (29.5)	10 (37.0)	14 (33.3)	7 (30.4)
Application site discomfort	3 (6.8)	2 (7.4)	2 (4.8)	6 (26.1)
Arthralgia	8 (18.2)	4 (14.8)	8 (19.0)	1 (4.3)
Constipation	0	5 (18.5)	2 (4.8)	0
Muscle twitching	4 (9.1)	1 (3.7)	4 (9.5)	4 (17.4)
Nausea	7 (15.9)	4 (14.8)	3 (7.1)	4 (17.4)
Anxiety	7 (15.9)	2 (7.4)	6 (14.3)	3 (13.0)
Dry mouth	1 (2.3)	4 (14.8)	5 (11.9)	1 (4.3)

Fatigue	2 (4.5)	2 (7.4)	5 (11.9)	3 (13.0)
Diarrhea	5 (11.4)	3 (11.1)	2 (4.8)	1 (4.3)
Back pain	5 (11.4)	2 (7.4)	3 (7.1)	0
Dizziness	5 (11.4)	1 (3.7)	2 (4.8)	1 (4.3)

### Serious Adverse Events

Six serious adverse events were reported. See Table 16 below for a listing of these events:

Table 16 - Study 03: Serious Adverse Events

Serious Adverse Event	Group 1 (n=44) N (%)	Group 2 (n=27) N (%)	Group 3 (n=42) N (%)	Group 4 (n=23) N (%)
Bladder tumor removal	1	0	0	0
Coronary artery surgery	1	0	0	0
Hip pain	0	0	0	1
Overdose*	0	1	0	0
Worsening depression/suicidal ideation	0	0	1	0
Atrial Fibrillation	0	0	1	0
All Serious Adverse Events Reported	2	1	2	1

\* Overdose refers to events associated with inadvertent treatment of >75 trains of active TMS delivered to the subject on a single calendar day.

## **5. Electroconvulsive Therapy (ECT) treatment**

The Advisory Panel will be asked to comment on the risks and benefits associated with treatment with the NeuroStar™ TMS System and the risks and benefits associated with ECT treatment. ECT treatment has been employed for most severe forms of depression since 1938. ECT treatment is intended to treat subjects with severe major depression, treatment resistant depression and depression with psychotic features who may be in immediate danger because of marked physical deterioration or potential suicide. The British ECT guidance, which is strictly evidence-based and was published by the National Health Service (NHS), National Institute of Clinical Excellence (NICE) in 2003 (Guidance on the Use of Electroconvulsive Therapy, Technology Appraisal #59), recommends that "ECT should be used to gain fast and short-term improvement of severe symptoms after all other treatment options have failed, or when the situation is thought to be life-threatening."

The use of ECT treatment has evolved over the years with the use of less electrical charge, more targeted electrode placements, and improved anesthesia and patient management. Effectiveness in the short-term treatment of the severe forms of major depression is reported in the literature. Many health authorities have performed extensive reviews of the literature, including the U.S. National Institutes of Health, the U.S. Surgeon General, the Québec Ministry of Health, the UK ECT Review Group, and the British National Institute for Clinical Excellence. These various governmental findings, as a consensus body of literature, support the effectiveness of ECT treatment for severe major depression in adults.

As background for your consideration and discussion, the following points are presented regarding ECT:

The sponsor made use of the UK ECT Review Group report, which contains comprehensive analyses, of prior ECT studies. The UK report identified six randomized studies and showed that the pooled effect standardized treatment effect of ECT versus sham was -0.91 (range -0.17 to 1.42; 95% CI=-

1.27 to -0.54). The purpose of using standardized effects is to loosely compare treatment effects between studies and studies with different measurement tools. Additionally, effect size estimates are not influenced by sample sizes. Three different standardized treatment effect sizes<sup>5</sup> were calculated by the sponsor for Study 01: MADRS = -0.39, HAMD<sub>24</sub> = -0.48, HAMD<sub>17</sub> = -0.55, and pooled results = -0.47. The sponsor's calculated standardized treatment effect size for the HAMD<sub>17</sub> of -0.55 falls just within the 95% CI range of the UK ECT Group report (-1.27 to -0.54). Note that the summary data presented by the sponsor seem to indicate larger effect sizes than those that were reported for the Wilson 1963, West 1981 and Lambourn 1978 studies which were part of the UK ECT Review Group report. However, the FDA has not examined the source studies to determine whether there is a true discrepancy.

Table 17 - Pooled Standardized Effect Sizes reported by FDA from Study 01 and Six published ECT studies (see also Table E2 in Appendix E).

<b>Study:</b>	<b>Pooled Standardized Effect Sizes</b>
Neuronetics' Study 01	-0.481
Wilson (1963)	-2.244
West (1981)	-1.330
Lambourn (1978)	-0.230
Freeman (1978)	-0.629
Gregory (1985)	-1.418
Johnstone (1980)	-0.739

- Additionally, in the UK analysis, the ECT subjects HAMD<sub>17</sub> scores were reduced by an average of 9.67 (range 0 to 13.9; 95% CI= 5.72 to 13.53) in favor of real ECT (see Table E1 in Appendix E). Note that the largest changes were seen in the smallest sample size studies (n=12 and n=25). In Study 01 the NeuroStar™ TMS System produced a difference between the active treatment and sham treatment of 1.7 on the HAMD<sub>17</sub> scale. However, caution should be used when comparing the literature results to the results of Study 01 with respect to the HAMD<sub>17</sub> because subjects in Study 01 had lower baseline HAMD<sub>17</sub> scores than typical of those reported in the ECT literature report. It should also be noted that the ATHF criteria have been reported in prior ECT literature studies as a range of 0.2 to 1.2. The sponsor reports a mean ATHF of 1.6 in Study 01 which slightly exceeds this range.
- Petrides et al. (2001) found an 85% remission rate (60% decrease in HAMD<sub>24</sub> and score  $\leq 10$ ) with ECT for non-psychotic depression. The Petrides et al. (2001) study may be of particular interest because it analyzed effects in ECT patients without psychotic features (as the NeuroStar trial 01 did). In contrast, Prudic et al. (2004), in a community setting study that included substance abusers, schizoaffective and bipolar disorder patients, found a 46.7% remission rate (HAMD<sub>24</sub>  $\leq 10$ ). Prudic et al. (2004) attributed the lower rate in the community than in clinical trials to comorbid personality disorders, substance abuse, schizoaffective disorder and early termination of ECT treatment by the patient.
- The FDA also performed a literature review on other rTMS studies for the treatment of depression. This information is relevant to the NeuroStar™ TMS System device because the

<sup>5</sup> The standard effect size is typically calculated as the post-treatment score minus the pre-treatment score divide by the pooled baseline standard deviation. Note that the sponsor used a general linear model, which is not a conventional approach, to calculate the pooled standard deviation rather than using the actual values of the treatment and control standard deviations.

technology used in the literature is similar to the NeuroStar™ TMS System, in both the stimulation capabilities and application site, it contains important prior information on the safety and effectiveness of this technology in general, and it shows the deficits in early studies that Neuronetics has addressed in their study. However, caution should also be taken in interpreting the literature study results as compared the NeuroStar™ TMS System study results presented in this Summary since the populations studied, the protocols used, and the stimulation outputs may have varied between the studies. Several randomized sham-controlled studies have been conducted. Results from these trials have varied, and several meta-analyses were performed to discern a treatment effect. Five of the six published meta-analyses found a significant treatment effect of rTMS for depression (Carpenter, 2006). The most prominent was the Cochrane group meta-analysis (Martin et al. 2003). Fourteen previous trials were analyzed, and a favorable treatment effect of rTMS for depression was found at two weeks of treatment with no significant effect after four weeks. The standardized treatment effect was calculated at -0.35 at two weeks with the HAMD<sub>17</sub>. The present Neuronetics trial sought to overcome some of the deficiencies cited in the literature. These included an improved sham treatment, a larger number of subjects, multiple study sites, and a longer study.

The risks associated with ECT treatment are well known. ECT treatment is given with a general anesthetic, a muscle relaxant, and respiratory assistance. The most significant common risks are memory and cognitive changes that occur immediately following each treatment and are dose dependent. ECT treatment specifically can lead to impairment of short-term memory; however, it is not clear for how long this persists. The incidence and degree of long-lasting retrograde amnesia is controversial, but is often cited as a major risk. Other risks include headache, skin burns, residual twitching, bone fractures, mania, and vascular accidents.

The sponsor performed memory and cognitive tests at baseline, 4 weeks, and 6 weeks, and they report that treatment with the NeuroStar™ TMS System produced no significant change over sham rTMS treatment. The memory/cognitive test results were compared to two previous ECT studies (Sackeim et al., 1993 and Sackeim et al., 2000). General cognition was assessed with the Mini Mental Status Exam (MMSE); anterograde amnesia was assessed with the Buschke Selective reminding test; retrograde amnesia was assessed with the Autobiographical Memory Interview – short form and these suggest a favorable profile when compared with the previous ECT publications.

Premarket notification (510(k)) applications are submitted to determine substantial equivalence to a legally-marketed medical device(s), in this case ECT devices. However, the sponsor also comments on the risk benefit ratio of treatment with the NeuroStar™ TMS System in relationship to FDA-approved antidepressant drug treatments. The sponsor compares treatment with the NeuroStar™ TMS System plus monotherapy results (Study 03) to the NIMH-funded, multi-site, open-label clinical trial, the Sequenced Treatment Alternatives to Relieve Depression (STAR\*D) study in which second medications are added after failure for treatment resistant monotherapy. Results from this study have recently been reported in the published literature (STAR\*D Study Team, 2006; Trivedi et al., 2006). The sponsor states that the placebo-adjusted effect sizes in Study 01 are consistent with the magnitude of effects of FDA approved antidepressant drugs. However, the sponsor's report of treatment with antidepressant agents only discusses discontinuation rates due to adverse events but does not look at the specific adverse events or their severity. The sponsor calculated a number needed to harm (NNH) and a number needed to treat (NNT) as a means of demonstrating the risk-to-benefit profile of their device (Altman et al., 1999 and Cook, R. and Sackett, D, 1995). However, their analyses treated all study discontinuation reasons equally and did not distinguish between discontinuation for more minor reasons (e.g., dry mouth) from more serious reasons such as suicidal ideation when determining the NNH. Thus, although the FDA notes that the analyses were performed a detailed discussion by the FDA is not provided.

The sponsor states that treatment with the NeuroStar™ TMS System will expand the range of potentially effective treatment options for patients with MDD. They state that treatment with the NeuroStar™ TMS System should have position in the armamentarium of available antidepressant treatments intermediate between more complex medication regimens on the one hand and ECT on the other.

If the Panel believes the NeuroStar™ TMS System has a positive risk-to-benefit profile they will be asked to comment on the position of treatment with the NeuroStar™ TMS System in the armamentarium of available antidepressant treatments and how this would be conveyed in the product labeling and indications for use.

## References

- Altman, D. and Anderson, P. Calculating the number needed to treat for trials where the outcome is time to event. *BMJ* 1999; 319:1492-1495.
- Åsberg M, Montgomery SA, Perris C, Schalling D, Sedvall G. A comprehensive psychopathological rating scale. *Acta Psychiatr Scand*, 1978, 271:5-27.
- Carpenter L. Neurostimulation in resistant depression. *Journal of Psychopharmacology*, 2006, 20(3) suppl, 35-40 (2006).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cook, R. and Sackett, D. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995; 310:452-454.
- Couturier JL. Effectiveness of rapid-rate repetitive transcranial magnetic stimulation in the treatment of depression: a TMS Systematic review and meta-analysis. *J Psychiatry Neurosci*. 30:83-90, 2005.
- Freeman CP, Basson JV, Crighton A. Double-blind controlled trial of electroconvulsive therapy (ECT) and simulated ECT in depressive illness. *Lancet* 1978;1:738-40.
- Gregory S, Shawcross CR, Gill D. The Nottingham ECT Study: a double-blind comparison of bilateral, unilateral, and simulated ECT in depressive illness. *Br J Psychiatry* 1985;146:520-24.
- Johnstone EC, Deakin JF, Lawler P, et al. The Northwick Park electroconvulsive therapy trial. *Lancet* 1980;2:1317-20.
- Khan A, Warner H, and Brown W. Symptom Reduction and Suicide Risk in Patients Treated With Placebo in Antidepressant Clinical Trials. *Arch Gen Psychiatry* 2000;57:311-317.
- Lambourn J, Gill D. A controlled comparison of simulated and real ECT. *Br J Psychiatry* 1978;133:514-19.
- Martin JLR, Barbanoj MJ, Schlaepfer TE, Thompson E, Pérez V, and Kulisevsky, J. Repetitive transcranial magnetic stimulation for the treatment of depression: TMS Systematic review and meta-analysis. *The British Journal of Psychiatry*, 182:480-491, 2003.
- Petrides G, Fink M, Husain TM, Knapp RG, Rush AJ, Mueller M, Rummans TA, O'Connor KM, Rasmussen KG, Bernstein HJ, Biggs M, Bailin SH, Kellner, CH. ECT remission rates in psychotic versus nonpsychotic depressed patients: A report from CORE. *J ECT* 2001;17:244-253.
- Prudic J, Olfson SC, Marcus SC, Fuller, RB, Sackeim HA. Effectiveness of electroconvulsive therapy in community settings. *Biol Psych* 2004;55:301-312.
- Sackeim HA, Prudic J, Devanand DP, Kiersky JE, Fitzsimons L, Moody BJ, McElhiney MC, Coleman EA, Settembrino JM. Effects of stimulus intensity and electrode placement on the effectiveness and cognitive effects of electroconvulsive therapy. *N Engl J Med*. 328(12):882-3, 1993.

Sackeim HA, Haskett RF, Mulsant BH, et al. Continuation pharmacotherapy in the prevention of relapse following electroconvulsive therapy: a randomized controlled trial. *JAMA* 285:1299-1307, 2001.

Snaith, RP, Harrop FM, Newby DA, Teale C. Grade scores of the Montgomery-Asberg Depression and the Clinical Anxiety Scales. *Br J Psychiatry*, 1986, 148:599-601.

Thase M, et al. Remission Rates Following Antidepressant Therapy With Bupropion or Selective Serotonin Reuptake Inhibitors: A Meta-Analysis of Original Data From 7 Randomized Controlled Trials. *J Clin Psychiatry* 2005;66:974-981.

Trivedi MH, Fava M, Wisniewski SR, Thase ME, Quitkin F, Warden D, Ritz L, Nierenberg AA, Lebowitz BD, Biggs MM, Luther JF, Shores-Wilson K, Rush AJ; STAR\*D Study Team. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*, 354:1243-1252, 2006.

Walsh B, Seidman S, Sysko R, and Gould M. Placebo Response in Studies of Major Depression. *JAMA* 2002;287:1840-1847

West ED. Electric convulsion therapy in depression: a double-blind controlled trial. *BMJ* 1981;282:355-57.

Wilson IC, Vernon JT, Guin T, Sandifer MG. A controlled study of treatments of depression. *J Neuropsychiatry* 1963; 4:331-37.

## Appendix A

Table A - Protocol Defined Outcome Measures

	<b>Outcome Measure Description</b>
<b>Primary Outcome Measure</b>	Evaluate the antidepressant effect of treatment with the NeuroStar System, using the last post-treatment <u>total symptom score on the Montgomery-Asberg Depression Rating Scale (MADRS) through week 4 of the acute treatment phase</u> of a specified course of active treatment when compared to sham treatment given under the same experimental conditions in patients meeting DSM-IV criteria for Major Depressive Episode, single or recurrent episode. The specified data set for this analysis is the intent-to-treat population.
<b>Secondary Outcome Measures</b>	<ol style="list-style-type: none"> <li>1) The last post-treatment total symptom score on the 24- Item Hamilton Depression Rating Scale (HAMD24) through week 4 and week 6 of the acute treatment phase</li> <li>2) The last post-treatment total symptom score on the 17- Item Hamilton Depression Rating Scale (HAMD17) through week 4 and week 6 of the acute treatment phase</li> <li>3) The total symptom score on the MADRS for the last post-treatment value observed through week 6 of the acute treatment phase</li> <li>4) Categorical outcomes of response (percent of patients achieving 50% reduction on each of the MADRS, HAMD24, and HAMD17 total symptom scores at the last post-treatment visit through week 4 and week 6 of the acute phase),</li> <li>5) Health outcomes scores from the Medical Outcomes Study Short Form 36-Item Questionnaire (SF-36, v1) and the Quality of Life, Enjoyment and Satisfaction Questionnaire (Q-LES-Q) at the last post-treatment visit through week 4 and week 6</li> <li>6) Categorical outcome of remission/recovery (percent of patients achieving HAMD17 total symptom score &lt; 8, HAMD24 total symptom score &lt; 11, and MADRS total symptom score &lt; 10 at the last post-treatment visit through week 4 and week 6</li> <li>7) Factor scores derived from the HAMD17 including: Anxiety/Somatization (sum of items 10, 11, 12, 13, 15, 17), Core Factor (sum of items 1, 2, 3, 7, 8), Maier (sum of items 1, 2, 7, 8, 9, 10), Gibbons (sum of items 1, 2, 3, 7, 9, 10, 11, 14), Retardation (sum of items 1, 7, 8, 14), and Sleep (sum of items 4, 5, 6) using the last post-treatment value through week 4 and week 6</li> <li>8) The total score on the Inventory of Depressive Symptoms – Self Report version (IDS-SR), using the last post-treatment value through week 4 and week 6</li> <li>9) The Clinical Global Impressions – Severity (CGI-S) score, using last post-treatment value through week 4 and week 6</li> <li>10) The Patient Global Impressions – Improvement (PGI-I) score, using last post-treatment value through week 4 and week 6</li> </ol>

**Appendix B**

Study 01 Results

Figure B – Patient Disposition

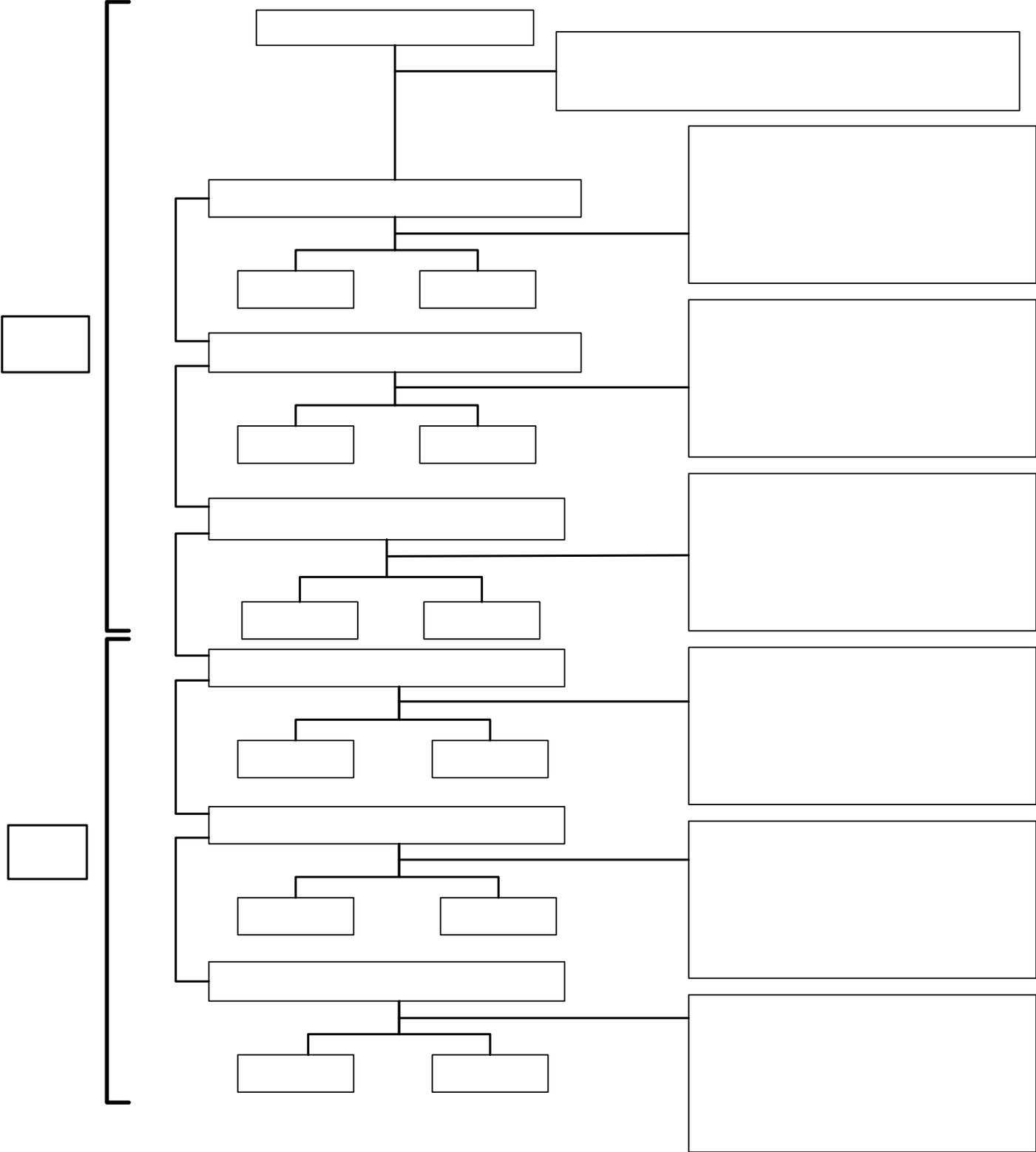


Table B1 - Study 01: Subject Demographics, Active and Sham Treatment.

	Statistic	Sham (N=146)	Active (N=155)	p-value*
<b>Baseline Demographics</b>				
Age [years, mean (SD)]	Mean (S.D.)	48.7(10.6)	47.9(11.0)	0.509
Gender				
Male	N (%)	72(49.3)	69(44.5)	
Female	N (%)	74(50.7)	86(55.5)	0.421
Ethnic Origin				
Caucasian	N (%)	131(89.7)	146(94.2)	
African-American	N (%)	3(2.1)	3(1.9)	
Asian	N (%)	1(0.7)	1(0.6)	
Hispanic	N (%)	8(5.5)	3(1.9)	
Native American	N (%)	0	1(0.6)	0.394
Other	N (%)	3(2.1)	1(0.6)	
<b>Motor Threshold</b>				
Motor Threshold	N	57.0(9.97)	55.2(9.67)	0.101
<b>Depression History</b>				
Single Episode	N (%)	9(6.2)	7(4.5)	
Recurrent	N (%)	136(93.8)	149(95.5)	0.611
<b>Duration of Current Episode</b>				
Length [mean(SD)]	Mean (S.D.)	13.2(9.5)	13.6(9.9)	0.728
<24 months	N (%)	123(84.2)	119(76.8)	
≥24 months	N (%)	23(15.8)	36(23.2)	0.112
<b>Secondary Diagnoses</b>				
None	N(%)	104(71.2)	96(61.9)	
Any Other Anxiety Disorder	N(%)	42(28.8)	59(38.1)	0.112
<b>ATHF Rating Summary</b>				
1	Level3 Exp.**	76(52.1)	88(56.8)	
2	Level3 Exp.**	50(34.2)	45(29.0)	
3	Level3 Exp.**	15(10.3)	15(9.7)	
4	Level3 Exp.**	5(3.4)	6(3.9)	
>4	Level3 Exp.**	-	1(0.6)	
Mean # of ATHF Level 3 Exposures	Mean	1.6	1.6	
<b>Psychiatric Baseline Assessment Scales at Screening (Visit 1)</b>				
MADRS Total Score	Mean (S.D.)	32.9(5.6)	32.6(5.3)	0.476
HAMD <sub>24</sub> Total Score	Mean (S.D.)	30.6(4.3)	30.7(3.9)	0.803
HAMD <sub>17</sub> Total Score	Mean (S.D.)	22.9(3.1)	22.6(2.3)	0.325
CGI Severity Score	Mean (S.D.)	4.7(0.7)	4.7(0.6)	0.871
IDS-SR Total Score	Mean (S.D.)	43.4(9.9)	42.0(9.4)	0.197

\*\*# Level 3 Exposures

Table B2 - Study 01: Frequency of Protocol-Approved Anxiolytics or Hypnotic Medication Use during the Acute Treatment Phase.

<b>Medication Name Preferred Term</b>	<b>Sham rTMS (N=146) N (%)</b>	<b>Active rTMS (N=155) N (%)</b>	<b>P-Value</b>
Subjects With At Least One Anxiolytics/Hypnotic Medication	44 (30.1)	44 (28.4)	0.800
Chloral Hydrate (no brand name)	1 (0.7)	0	0.485
Clonazepam (Klonopin)	1 (0.7)	1 (0.6)	1.000
Diphenhydramine (Sominex, Benadryl)	0	1 (0.6)	1.000
Eszopiclone (Lunesta)	1 (0.7)	0	0.485
Lorazepam (Ativan)	21 (14.4)	26 (16.8)	0.635
Zaleplon (Sonata)	1 (0.7)	2 (1.3)	1.000
Zolpidem (Ambien)	28 (19.2)	21 (13.5)	0.213
Zopiclone (Immovane)	0	2 (1.3)	0.499
Temazepam	1 (0.7)	2 (1.3)	1.000
Thiopenthyll	0	1 (0.6)	1.000
<b>ALL Meds<sup>†</sup></b>	<b>54 (36.9)</b>	<b>56 (36.1)</b>	<b>0.4861</b>

<sup>†</sup> ALL Meds calculated by FDA.

Table B3 - Study 01: Antidepressant Medications Used During the Post-Treatment Taper Phase (Phase III).

<b>Antidepressant Medication</b>	<b>Drug Name</b>	<b>Sham rTMS (N=40) N (%)</b>	<b>Active rTMS (N=64) N (%)</b>	<b>P-Value</b>
<b>Selective Serotonin Reuptake Inhibitors (SSRI)</b>	Citalopram	1 (2.5)	0	0.323
	Escitalopram	5 (12.5)	12 (18.8)	1.000
	Fluoxetine	2 (5.0)	3 (4.7)	0.657
	Fluvoxamine	1 (2.5)	1 (1.6)	0.543
	Paroxetine	3 (7.5)	0	0.032
	Sertraline	1 (2.5)	3 (4.7)	1.000
	ALL (SSRI)	13 (32.5)	19 (29.7)	0.4640
<b>Serotonin/Norepinephrine Reuptake Inhibitors</b>	Duloxetine	15 (37.5)	15 (23.4)	0.024
	Venlafaxine	9 (22.5)	15 (23.4)	0.622
	ALL (S/NRI)	24 (60)	30 (46.9)	0.1353
<b>Other Antidepressants</b>	Bupropion	7 (17.5)	12 (18.8)	0.790
	Mirtazapine	1 (2.5)	4 (6.3)	1.000
	Trazodone	0	2 (3.1)	1.000
	ALL (Other)	8 (20)	18 (28.1)	0.2443
<b>ALL Meds<sup>†</sup></b>		<b>45 (113)</b>	<b>67 (104.7)</b>	<b>#</b>

<sup>#</sup>p-value not calculated. <sup>†</sup> ALL Meds calculated by FDA.

Table B4 – Study 01 Summary Table of Outcome Measures at Week 4

<b>Primary Outcome:</b>	<b>Change from Baseline</b>	<b>Difference (90% CI)</b>	<b>P-value<sup>1</sup></b>
MADRS	-5.6 Active -3.5 Sham	-2.1 (-3.9, -0.3)	0.057
<b>Secondary Outcomes:</b>			
	<b>Change from Baseline</b>	<b>Difference (90% CI)</b>	<b>P-value<sup>1</sup></b>
HAMD24	-6.5 Active -4.1 Sham	-2.4 (-4.0, -0.8)	0.012
HAMD17	-5.0 Active -3.1 Sham	-1.9 (-3.1, -0.7)	0.006
	<b>Proportion of Responders<sup>3</sup></b>	<b>Difference (90% CI)</b>	<b>P-value<sup>2</sup></b>
MADRS Responders	18.1% Active 11.0% Sham	7.1% (0.2%, 13.9%)	0.045
HAMD24 Responders	19.4% Active 11.6% Sham	7.7% (0.7%, 14.7%)	0.030
HAMD17 Responders	20.6% Active 11.6% Sham	9.0% (1.7%, 16.1%)	0.018
	<b>Change from Baseline</b>	<b>Difference (90% CI)</b>	<b>P-value<sup>1</sup></b>
SF 36: Physical Functioning	1.3 Active 0.4 Sham	0.9 (-.05, 2.3)	0.299
SF 36: Role Physical	1.0 Active -0.2 Sham	1.2 (-1.4, 3.8)	0.361
SF 36: Bodily Pain	1.4 Active 1.0 Sham	0.4 (-1.0, 1.8)	0.520
SF 36: General Health	1.3 Active -0.3 Sham	1.6 (0.2, 3.0)	0.049
SF 36: Vitality	3.3 Active 2.1 Sham	1.2 (-0.3, 2.7)	0.179
SF 36: Social Functioning	3.2 Active 1.8 Sham	1.4 (-0.5, 3.3)	0.183
SF 36: Role Emotional	3.6 Active 1.9 Sham	1.7 (-0.1, 3.5)	0.105
SF 36: Mental Health	3.7 Active 0.6 Sham	3.1 (1.2, 5.0)	0.006
Q-LES-Q	3.5 Active 2.0 Sham	1.5 (-0.2, 3.2)	0.124
	<b>Proportion of Remitters<sup>4</sup></b>	<b>Difference (90% CI)</b>	<b>P-value<sup>2</sup></b>
MADRS Remitters	7.1% Active 6.2% Sham	0.9% (-4.2%, 5.9%)	0.633
HAMD24 Remitters	9.0% Active 8.2% Sham	0.8% (-4.7%, 6.4%)	0.644
HAMD17 Remitters	7.1% Active 6.2% Sham	0.9% (-4.2%, 5.9%)	0.705

	<b>Change from Baseline</b>	<b>Difference and 90% CI</b>	<b>P-value<sup>1</sup></b>
HAMD Anxiety/Somatization	-1.6 Active -1 Sham	-0.6 (-1.1, -0.1)	0.025
HAMD Core Depression	-1.9 Active -1 Sham	-0.9 (-1.5, -0.3)	0.012
HAMD Maier	-2.5 Active -1.4 Sham	-1.1 (-1.5, -0.5)	0.003
HAMD Gibbons	-3.0 Active -1.8 Sham	-1.2 (-2.0, -0.4)	0.007
HAMD Retardation	-1.6 Active -0.9 Sham	-0.7 (-1.2, -0.2)	0.007
HAMD Sleep	-0.9 Active -0.6 Sham	-0.3 (-0.7, 0.1)	0.211
IDS-SR	-7.7 Active -5.2 Sham	-2.5 (-4.8,-0.2)	0.058
CGI-S	-0.6 Active -0.2 Sham	-0.4 (-0.6,-0.2)	0.009
PGI-I	-0.6 Active -0.3 Sham	-0.3 (-0.6, 0.02)	0.181

<sup>1</sup> P-value is from the following ANCOVA model: Change from baseline = Baseline Score, ATHF group, center, and treatment

<sup>2</sup> P-value is from the following logistic regression model: Responder = ATHF group, center, and treatment

<sup>3</sup> A responder has a > 50% change from baseline score

<sup>4</sup> MADRS Remission is defined as MADRS total score < 10  
HAMD<sub>24</sub> Remission is defined as HAMD<sub>24</sub> total score < 11  
HAMD<sub>17</sub> Remission is defined as HAMD<sub>17</sub> total score < 8

Table B5 - Study 01: Standardized Effect Sizes and Associated P-Values for Primary and Secondary Outcome Measures at Week 4

<b>Primary Effectiveness Outcome Measure</b>	<b>Active Treatment (N)</b>	<b>Sham Treatment (N)</b>	<b>Standardized Effect Size Week 4</b>	<b>P-Value Week 4</b>
MADRS Total Score (Overall Sample)	155	146	0.39	0.057
• ATHF Group 1	88	76	0.94	0.001
• ATHF Group 2	45	50	-0.16	0.710
• ATHF Group 3	15	15	-0.55	0.588
• ATHF Group 4	7	5	5.21	0.022
<b>Secondary Effectiveness Outcome Measures</b>				
HAMD <sub>24</sub> Total Score (Overall Sample)	155	146	0.48	0.012
• ATHF Group 1	88	76	0.83	0.001
• ATHF Group 2	45	50	0.03	0.933
• ATHF Group 3	15	15	0.44	0.577
• ATHF Group 4	7	5	2.41	0.077
HAMD <sub>17</sub> Total Score (Overall Sample)	155	146	0.55	0.006
• ATHF Group 1	88	76	0.83	0.001
• ATHF Group 2	45	50	0.13	0.762
• ATHF Group 3	15	15	0.81	0.440
• ATHF Group 4	7	5	2.26	0.089
MADRS Responders (Overall Sample)	155	146	0.65	0.045
• ATHF Group 1	88	76	1.23	0.008
• ATHF Group 2	45	50	0.11	0.692
• ATHF Group 3	15	15	1.99	1.000
• ATHF Group 4	7	5	-1.00	1.000
HAMD <sub>24</sub> Responders (Overall Sample)	155	146	0.67	0.030
• ATHF Group 1	88	76	1.35	0.005
• ATHF Group 2	45	50	-0.17	0.747
• ATHF Group 3	15	15	2.99	0.424
• ATHF Group 4	7	5	-0.29	1.000
HAMD <sub>17</sub> Responders (Overall Sample)	155	146	0.78	0.018
• ATHF Group 1	88	76	1.38	0.004
• ATHF Group 2	45	50	-0.21	0.658
• ATHF Group 3	15	15	2.99	0.424
• ATHF Group 4	7	5	-0.29	1.000
IDS-SR Total Score (Overall Sample)	155	146	0.27	0.059
• ATHF Group 1	88	76	0.57	0.002
• ATHF Group 2	45	50	0.10	0.710
• ATHF Group 3	15	15	0.29	0.706
• ATHF Group 4	7	5	1.85	0.269

Table B6 – Study 01: Patient Global Impressions – Improvement Rating: Distribution of Scores By Rating At Baseline and Week 4 of Acute Treatment Phase

PGI-Improvement Rating	Baseline		Acute Treatment (Week 4)	
	Active Treatment (N=155)	Sham Treatment (N=146)	Active Treatment (N=155)	Sham TMS Treatment (N=146)
1 – Very much improved	0	1	4	4
2 – Much improved	1	0	30	20
3 – Minimally improved	9	12	39	36
4 – No change	106	88	34	38
5 – Minimally worse	19	18	33	29
6 – Much worse	16	16	12	11
7 – Very much worse	4	6	3	8

Table B7 – Study 01: Summary Table of Outcome Measures at Week 6

Primary Outcome:	Change from Baseline	P-value <sup>1</sup>
MADRS	Active: -5.6 Sham: -3.2	0.058
<b>Secondary Outcomes* :</b>		
	<b>Change from Baseline</b>	<b>P-value<sup>1</sup></b>
HAMD24	Active:-6.4 Sham:-3.8	0.015
HAMD17	Active:-5.1 Sham:-2.9	0.005
	<b>Proportion of Responders<sup>3</sup></b>	<b>P-value<sup>2</sup></b>
MADRS Responders	Active:23.9% Sham:12.3%	0.007
HAMD24 Responders	Active:23.9% Sham:15.1%	0.042
HAMD17 Responders	Active:24.5% Sham:13.7%	0.015
	<b>LS Mean Change from Baseline</b>	<b>P-value<sup>1</sup></b>
SF 36: Physical Functioning	Active:1.2 Sham:0.3	0.229
SF 36: Role Physical	Active:1.8 Sham:0.2	0.221
SF 36: Bodily Pain	Active:1.4 Sham:0.6	0.301
SF 36: General Health	Active:1.5 Sham:-0.2	0.047
SF 36: Vitality	Active:4 Sham:2	0.081

SF 36: Social Functioning	Active:3.7 Sham:2.7	0.386
SF 36: Role Emotional	Active:4.7 Sham:2.4	0.044
SF 36: Mental Health	Active:4.5 Sham:1.4	0.015
Q-LES-Q	Active:3.8 Sham:1.3	0.035
	<b>Proportion of Remitters<sup>4</sup></b>	<b>P-value<sup>2</sup></b>
MADRS Remitters	Active:14.2% Sham:5.5%	0.011
HAMD24 Remitters	Active:17.4% Sham:8.2%	0.012
HAMD17 Remitters	Active:15.5% Sham:8.9%	0.065
	<b>LS Mean Change from Baseline</b>	<b>P-value<sup>1</sup></b>
HAMD Anxiety/Somatization	Active:-1.7 Sham:-1	0.023
HAMD Core Depression	Active:-1.8 Sham:-0.8	0.008
HAMD Maier	Active:-2.4 Sham:-1.1	0.003
HAMD Gibbons	Active:-3 Sham:-1.6	0.006
HAMD Retardation	Active:-1.6 Sham:-0.7	0.003
HAMD Sleep	Active:-1.1 Sham:-0.8	0.109
IDS-SR	Active:-7.7 Sham:-4.7	0.053
CGI-S	Active:-0.6 Sham:-0.2	0.012
PGI-I	Active:-0.5 Sham:-0.2	0.107

\* The FDA notes that for analysis of the secondary endpoints the protocol did not include an adjustment for statistical multiplicity to control for type 1 error.

<sup>1</sup> P-value is from the following ANCOVA model: Change from baseline = Baseline Score, ATHF group, center, and treatment

<sup>2</sup> P-value is from the following logistic regression model: Responder = ATHF group, center, and treatment

<sup>3</sup> A responder has a > 50% change from baseline score

<sup>4</sup> MADRS Remission is defined as MADRS total score < 10  
HAMD<sub>24</sub> Remission is defined as HAMD<sub>24</sub> total score < 11  
HAMD<sub>17</sub> Remission is defined as HAMD<sub>17</sub> total score < 8

Table B8: Study 01 Adverse Events with an Incidence in the Active Group of  $\geq 2\%$  and Greater than the Incidence Occurring in the Sham Group

<b>Body TMS System (-) Preferred Term</b>	<b>Active (n=165) N (%)</b>	<b>Sham (n=158) N (%)</b>
<b>Ear and labyrinth disorders</b>		
- Ear Pain	4 (2.4)	1 (0.6)
- Tinnitus	7 (4.2)	2 (1.3)
<b>Eye disorders</b>		
- Eye pain	10 (6.1)	3 (1.9)
- Lacrimation increased	7 (4.2)	1 (0.6)
- Visual disturbance	4 (2.4)	2 (1.3)
<b>Gastrointestinal disorders</b>		
- Diarrhea	8 (4.8)	6 (3.8)
- Nausea	17 (10.3)	10 (6.3)
- Toothache	12 (7.3)	1 (0.6)
- Vomiting	7 (4.2)	3 (1.9)
<b>General disorders</b>		
- Application site discomfort	18 (10.9)	2 (1.3)
- Application site pain	59 (35.8)	6 (3.8)
- Facial pain	11 (6.7)	5 (3.2)
- Pain	7 (4.2)	3 (1.9)
- Pyrexia	4 (2.4)	1 (0.6)
<b>Injury, poisoning, complications</b>		
- Overdose*	4 (2.4)	0
<b>Musculoskeletal, Connective tissue disorders</b>		
- Arthralgia	10 (6.1)	5 (3.2)
- Muscle twitching	34 (20.6)	5 (3.2)
- Musculoskeletal stiffness	5 (3.0)	4 (2.5)
- Neck pain	8 (4.8)	4 (2.5)
<b>Nervous TMS System disorders</b>		
- Dyskinesia	5 (3.0)	2 (1.3)
- Headache	96 (58.2)	87 (55.1)
- Hypoaesthesia	5 (3.0)	2 (1.3)
- Paraesthesia	6 (3.6)	4 (2.5)
- Tension headache	4 (2.4)	2 (1.3)
<b>Psychiatric disorders</b>		
- Agitation	4 (2.4)	3 (1.9)
- Anxiety	19 (11.5)	18 (11.4)
<b>Reproductive TMS System and breast disorders</b>		
- Dysmenorrhoea	5 (3.0)	2 (1.3)
<b>Respiratory, thoracic, and mediastinal disorders</b>		
- Cough	4 (2.4)	2 (1.3)
- Dyspnoea	6 (3.6)	1 (0.6)
<b>Skin and subcutaneous tissue disorders</b>		

- Pain of skin	14 (8.5)	1 (0.6)
All Adverse Events Reported	404 (**)	183 (**)

\* Overdose refers to events associated with inadvertent treatment of >75 trains of active TMS delivered to the subject on a single calendar day.

\*\* Not calculated.

**Appendix C**

Study 02 Results

Table C1 - Study 02: Frequency of Protocol-Approved Anxiolytics or Hypnotic Medication Use during the Acute Treatment Phase

Medication Name Preferred Term	Group A-Active (N=73) N (%)	Group B-Sham (N=85) N (%)	P-Value
Subjects With At Least One Anxiolytics/Hypnotic Medication	30 (41.1)	34 (40)	1.000
Aprozolam	0	3 (3.5)	0.250
Lorazepam	20 (27.4)	19 (22.4)	0.579
Zaleplon	1 (1.4)	0	0.462
Zolpidem	15 (20.5)	21 (24.7)	0.573
Zopiclone	1 (1.4)	0	0.462
Temazepam	1 (1.4)	0	1.000
Valium	0	1 (1.2)	1.000
ALL Meds <sup>†</sup>	68 (93.2)	78 (91.8)	0.4921

<sup>†</sup> ALL Meds calculated by FDA.

Table C2 - Study 02: Antidepressant Medications Used During the Post-Treatment Taper Phase.

Anti-depressant Medication	Drug Name	Group A-Active (N=73) N (%)	Group B-Sham (N=85) N (%)	P-Value
Selective Serotonin Reuptake Inhibitors (SSRI)	Citalopram	2 (2.7)	3 (3.5)	1.000
	Escitalopram	4 (5.5)	6 (7.1)	0.753
	Fluoxetine	1 (1.4)	0	0.462
	Fluvoxamine	0	1 (1.2)	1.000
	Paroxetine	1 (1.4)	0	0.462
	Sertraline	3 (4.1)	0	1.000
	ALL (SSRI)	11 (15.1)	10 (11.8)	0.3528
Serotonin/Nor-epinephrine Reuptake Inhibitors (S/NRI)	Duloxetine	17 (23.3)	10 (11.8)	0.060
	Venlafaxine	7 (9.6)	8 (9.4)	1.000
	ALL (S/NRI)	24 (32.9)	18 (21.2)	0.0696
Other Anti-depressants	Clomipramine	1 (1.4)	0	0.462
	Bupropion	11 (27.5)	11 (17.2)	0.819
	Mirtazapine	4 (10.0)	4 (6.3)	1.000
	Nardil	1 (1.4)	0	0.462
	Parnate	1 (1.4)	0	0.462
	Tofranil	1 (1.4)	0	0.462
	Trazodone	1 (1.4)	0	0.462
	ALL (Other)	20 (27.4)	15 (17.6)	0.1005
	ALL Meds <sup>†</sup>	55 (75.3)	43 (50.6)	0.0011

<sup>†</sup> ALL Meds calculated by FDA.

Table C3 – Study 02: Discontinuation Reasons

Discontinuation Reason:	Acute Phase - Through Week 6*		Taper Phase – Through Week 3**	
	A	B	A	B
Adverse Event	0	8	5	0
Failed to Return	1	1	0	1
Satisfactory Effectiveness Response	0	0	0	0
Unsatisfactory Effectiveness Response	4	1	2	3
Patient Request – Unrelated to Study	5	3	1	0
Protocol Violation	0	1	2	1
Other	1	1	2	1
<b>Total</b>	11	15	12	6

\* Includes patients discontinuing prior to week 2, week 4, and week 6.

\*\* Includes patients who completed the week 6 visit but who discontinued prior to the week 1, week 2, and week 3 taper phase visits.

Table C4 - Study 02: Adverse Events with an incidence on Active rTMS of  $\geq 2\%$  Incidence on Active rTMS Treatment in either Group A or Group B.

Body TMS System (-) Preferred Term	Group A (n=73) N (%)	Group B (n=85) N (%)
Gastrointestinal disorders		
- Diarrhea	6 (8.2)	7 (8.2)
- Nausea	10 (13.7)	6 (7.1)
- Toothache	3 (5.4)	1 (1.4)
- Vomiting	5 (6.8)	1 (1.2)
General disorders		
- Application site discomfort	7 (9.6)	8 (9.4)
- Application site pain	8 (11.0)	27 (31.8)
- Facial pain	0	5 (5.9)
- Fatigue	6 (8.2)	5 (5.9)
- Pain	4 (5.5)	3 (3.5)
Infections and infestations		
- Nasopharyngitis	4 (5.5)	2 (2.4)
- Upper respiratory tract infection	4 (5.5)	1 (1.2)
Musculoskeletal, Connective tissue disorders		
- Arthralgia	4 (5.5)	8 (9.4)
- Back pain	5 (6.8)	2 (2.4)
- Muscle twitching	15 (20.5)	18 (21.2)
- Pain in extremity	5 (6.8)	4 (4.7)
Nervous TMS System disorders		
- Dizziness	6 (8.2)	7 (8.2)
- Headache	35 (47.9)	39 (45.9)
- Migraine	4 (5.5)	2 (2.4)
- Paraesthesia	5 (6.8)	4 (4.7)
Psychiatric disorders		
- Anxiety	11 (15.1)	12 (14.1)

- Insomnia	22 (30.1)	22 (25.9)
Skin and subcutaneous tissue disorders		
- Pain of skin	1 (1.4)	5 (5.9)
All Adverse Events Reported	170 (**)	189 (**)

\* Overdose refers to events associated with inadvertent treatment of >75 trains of active TMS delivered to the subject on a single calendar day.

\*\* Not calculated.

**Appendix D**

Study 03 Results

Table D1 – Study 03: Secondary Outcome Effectiveness Results for Group 1 (i.e., Study 01 Active treatment responders), N=44

<b>Effectiveness Outcome Measures</b>	<b>Week 1</b>	<b>Week 2</b>	<b>Week 3</b>	<b>Week 4</b>
MADRS Total Score Mean Change <sup>1</sup>	-20.1	-21.4	-20.3	-21.2
HAMD <sub>24</sub> Total Score Mean Change <sup>1</sup>	-18.0	-19.0	-18.4	-19.6
HAMD <sub>17</sub> Total Score Mean Change <sup>1</sup>	-14.0	-14.4	-13.9	-14.6
MADRS Remission Rate (%) <sup>2</sup> (MADRS Total Score < 10)	50	59.1	52.3	45.5
HAMD <sub>24</sub> Remission Rate (%) <sup>2</sup> (HAMD <sub>24</sub> Total Score < 11)	47.7	54.5	47.7	43.2
HAMD <sub>17</sub> Remission Rate (%) <sup>2</sup> (HAMD <sub>17</sub> Total Score < 8)	50	56.8	43.2	43.2

<sup>1</sup> Baseline is defined as Study 01 Baseline

<sup>2</sup> Remission Rates were calculated using total enrolled sample (N=44)

Table D2 - Study 03: Adverse Events with an incidence on Active rTMS of  $\geq 2\%$  Incidence in Any Treatment Group.

<b>Body TMS System (-) Preferred Term</b>	<b>Group 1 (n=44) N (%)</b>	<b>Group 2 (n=27) N (%)</b>	<b>Group 3 (n=42) N (%)</b>	<b>Group 4 (n=23) N (%)</b>
Gastrointestinal disorders				
- Constipation	0	5 (18.5)	2 (4.8)	0
- Diarrhea	5 (11.4)	3 (11.1)	2 (4.8)	1 (4.3)
- Dry mouth	1 (2.3)	4 (14.8)	5 (11.9)	1 (4.3)
- Nausea	7 (15.9)	4 (14.8)	3 (7.1)	4 (17.4)
- Vomiting	0	1 (3.7)	0	2 (8.7)
General disorders				
- Application site discomfort	3 (6.8)	2 (7.4)	2 (4.8)	6 (26.1)
- Fatigue	2 (4.5)	2 (7.4)	5 (11.9)	3 (13.0)
- Pain	3 (6.8)	0	2 (4.8)	1 (4.3)
Immune TMS System disorders				
- Seasonal allergy	1 (2.3)	0	2 (4.8)	1 (4.3)
Infections and infestations				
- Upper respiratory tract infection	4 (9.1)	1 (3.7)	4 (9.5)	1 (4.3)
Musculoskeletal, connective tissue disorders				
- Arthralgia	8 (18.2)	4 (14.8)	8 (19.0)	1 (4.3)
- Back pain	5 (11.4)	2 (7.4)	3 (7.1)	0
- Muscle twitching	4 (9.1)	1 (3.7)	4 (9.5)	4 (17.4)
- Musculoskeletal stiffness	1 (2.3)	2 (7.4)	0	0
- Myalgia	1 (2.3)	1 (3.7)	5 (11.9)	0
- Pain in extremity	2 (4.5)	0	3 (7.1)	0
Nervous TMS System disorders				

- Dizziness	5 (11.4)	1 (3.7)	2 (4.8)	1 (4.3)
- Headache	16 (36.4)	9 (33.3)	13 (31.0)	10 (43.5)
Psychiatric disorders				
- Agitation	3 (6.8)	0	0	0
- Anxiety	7 (15.9)	2 (7.4)	6 (14.3)	3 (13.0)
- Depressive symptom	0	1 (3.7)	4 (9.5)	2 (8.7)
- Insomnia	13 (29.5)	10 (37.0)	14 (33.3)	7 (30.4)
- Irritability	2 (4.5)	2 (7.4)	2 (4.8)	0
- Libido decreased	4 (9.1)	3 (11.1)	1 (2.4)	0
Respiratory, thoracic and mediastinal disorders				
- Nasal congestion	1 (2.3)	0	1 (2.4)	2 (8.7)
- Sinus congestion	2 (4.5)	0	1 (2.4)	2 (8.7)
Skin, subcutaneous tissue disorders				
- Hyperhidrosis	2 (4.5)	2 (7.4)	0	0
Uncoded verbatim terms				
- Increased frequency of headaches	0	1 (3.7)	0	0
- Menorrhea	0	0	0	1 (4.3)
All Adverse Events Reports	103 (**)	63 (**)	94 (**)	53 (**)

\*\* Not calculated

## **Appendix E**



Table E2 --FDA Summary of Evidence from Study 01 and Randomized, Sham-Controlled Clinical Trials of ECT treatment.<sup>7</sup>

Study Measure	Neuronetics Study 44-01101 (N=301)		Wilson, 1963 (N=12)		West, 1981 (N=25)		Lambourn, 1978 (N=32)		Freeman, 1978 (N=40)		Gregory, 1985 <sup>2</sup> (N=69)		Johnstone, 1980 (N=70)	
	BL	End	BL	End	BL	End	BL	End	BL	End	BL	End	BL	End
<u>Symptom Rating Method</u>														
[Score (SD)]														
- MADRS														
- Active	32.4(6.0)	26.5(11.0)	--	--	--	--	--	--	--	--	--	[24.0]	--	--
- Sham	33.7(5.7)	29.5(10.1)	--	--	--	--	--	--	--	--	--	--	--	--
- 24-Item HAMD														
- Active	29.9(5.0)	23.1(8.9)	--	--	--	--	--	--	--	--	--	--	--	--
- Sham	30.2(4.9)	25.7(8.8)	--	--	--	--	--	--	--	--	--	--	--	--
- 17-Item HAMD														
- Active	22.5(3.3)	17.3(6.5)	26.5(4.9)	5.6(3.5)	26.6(9.3)	10.8(8.6)	24.9(7.4)	11.8(9.7)	28.0(--)	27.0(--)	--	[30.53]	27.5(--)	9(--)
- Sham	22.8(3.5)	19.3(6.5)	28.8(4.1)	18.0(12.2)	24.1(11.6)	22.2(12.6)	27.0(7.4)	15.6(10.8)	28.0(--)	20.5(--)	--	--	25.5(--)	12(--)
- BDI														
- Active	--	--	--	--	--	--	--	--	--	--	--	--	--	--
- Sham	--	--	--	--	--	--	--	--	--	--	--	--	--	--
- Other														
- Active	--	--	--	--	67.9(15.6)	19.5(15.3)	--	--	--	--	--	--	--	--
- Sham	--	--	--	--	70.7(21.6)	63.4(17.9)	--	--	--	--	--	--	--	--
<u>Standardized Effect Size</u>														
MADRS		-0.355		--		--		--		--		--		--
HAMD24		-0.481		--		--		--		--		--		--
HAMD17		-0.556)		-2.244		-1.330		-0.230		-0.629		-1.418		-0.739

<sup>2</sup> Only change from baseline scores provided in published data report.

<sup>3</sup> FDA calculated the effect sizes in the Study 01 by taking the difference between the LS Means for the change from baseline divided by the pooled baseline standard deviation for the “Total Score”. These values were taken from Final Study Report, Study 44-01101, Tables 13, 16 and 17 for the MADRS, HAMD24 and HAMD17 respectively. These calculations differ from Sponsor’s calculations because the Sponsor appeared to use a GLM model to derive the pooled baseline standard deviation. This model consisted of the total score as the dependent variable and treatment group and study center as the independent variables. FDA feels that the sponsor’s model based approach may underestimate the baseline standard deviation as the independent variables may be accounting for variation that would otherwise normally be present in the measuring instrument. Additionally, the GLM model used by Sponsor was not pre-specified. Therefore, FDA reported results obtained by using the pooled baseline standard deviations of the “Total Score” as found in Tables 13, 16, and 17 of the Study 01 report.

<u>Pooled Effect</u> <u>Size [95% CI]</u>	-0.418 (LS Means -0.464)	-1.255 (-6.86 to 4.35)
--	--------------------------	---------------------------