NEURONETICS FINAL STUDY REPORT
STUDY NO. 44-01101
14 April 2006

"A Randomized, Parallel-Group, Sham-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of the Neuronetics Model 2100 CRS Repetitive Transcranial Magnetic Stimulation (rTMS) System in Patients with Major Depression"

Table of Contents

Table of Contents (cont.)

Table of Contents (cont.)

Appendices

1.0 INTRODUCTION

The clinical development program for the Neuronetics TMS System consisted of three integrated clinical protocols as displayed in Figure 1.

In brief, the efficacy of the Neuronetics TMS System was established in adult outpatients in a 9-week, randomized, placebo-controlled clinical trial, Study 44-01101.

Patients who failed to receive benefit from their randomized assignment in Study 44-01101 were eligible to enter a 9-week, open-label cross-over study with the Neuronetics TMS System in Study 44-01102.

The maintenance of an acute clinical response to the Neuronetics TMS System in either Study 44-01101 or Study 44-01102 was established in a 24 week, open-label continuation clinical trial, Study 44-01103.

The design, objectives and summary results obtained for studies 44-01101, 44-01102 and 44-01103 are summarized in Table 1.

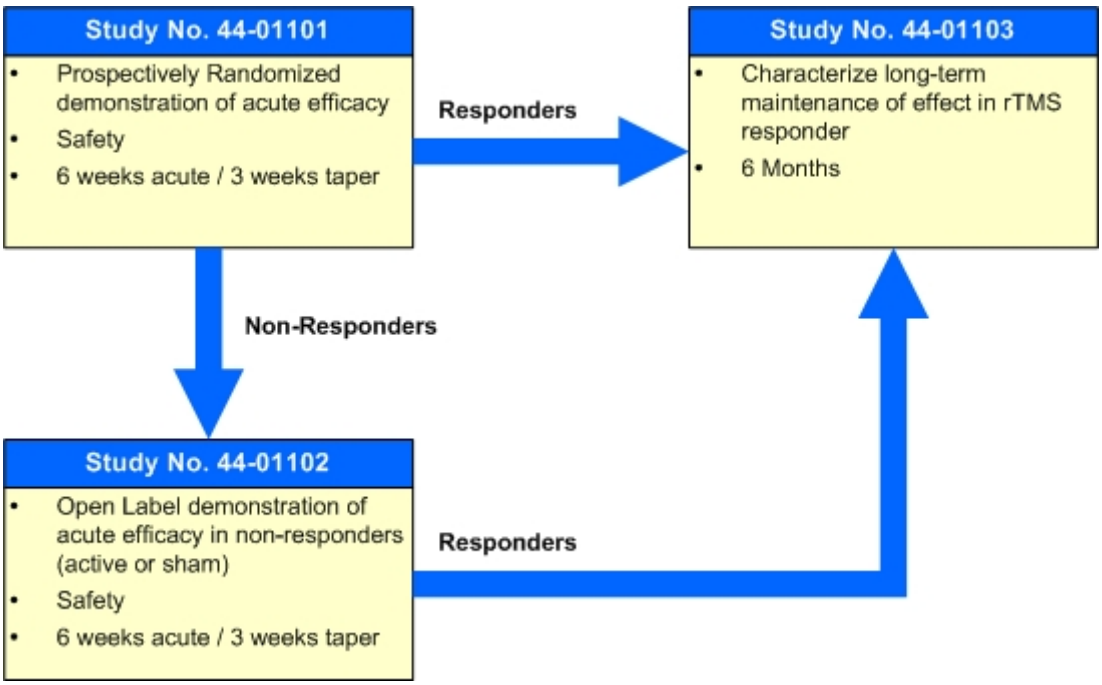


Figure 1. Neuronetics' Clinical Studies and Patient Allocation

Table 1. Summary of Neuronetics Clinical Studies 44-01101, 44-01102 and 44-01103

Study No.	Study Summary	Study Objective
44-01101	A randomized, parallel-group, sham-controlled clinical trial designed to test the efficacy of TMS treatment for patients diagnosed with DSM-IV defined major depression who have not benefited from prior adequate treatment with oral antidepressants. The study design was comprised of three phases: a one week, no-treatment screening phase, a six week acute treatment phase, and a 3 week rTMS taper phase. During the taper phase, as TMS was tapered, monotherapy with oral antidepressant medications was initiated. At the conclusion of Study 44-01101, or at any time after 4 weeks of participation in the acute phase of that study, patients were considered for enrollment in either of the two open-label, uncontrolled extension studies.	The primary objective was to evaluate the antidepressant effect [using the last post-treatment total symptom score on the MADRS] of a specified treatment course of TMS 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. Only patients meeting diagnostic criteria for Major Depression were included in this study. Personnel at the study sites were blind to the choice of primary efficacy measure and to the point of declaration of the efficacy outcome. Secondary outcome measures were HAMD17 and 24 item total symptom score, and response and remission rates for MADRS, HAMD17 and 24. Additional physician and patient rates scale were administered and evaluated as secondary outcome measures. Safety was assessed by adverse event reports, and by targeted safety evaluation of air-conduction auditory threshold. Cognitive function.was assessed with the Mini Mental Status Examination, the Buschke Selective Reminding Test, and the Autobiographical Memory Inventory-Short Form.
44-01102	An open-label, uncontrolled clinical trial for patients who do did not meet pre-defined criteria for response in Study 44-01101. This protocol was otherwise identical in design and treatment sequence to Study 44-01101.	The primary objective was to describe the symptom changes [using the last post-treatment total symptom score on the MADRS] observed with up to 6 weeks of open-label TMS treatment in patients in patients meeting DSM-IV criteria for Major Depressive Episode, single or recurrent episode, who had not shown an acute clinical response to daily dose active of sham rTMS administered for up to 6 weeks. Personnel at the study sites were blind to the choice of primary efficacy measure and the point of declaration of the efficacy outcome.
44-01103	An open-label, uncontrolled clinical trial providing six months of oral antidepressant monotherapy to patients who met pre-defined criteria for response upon exit from Study 44-01101. Study 44-01103 also permitted open-label access, on a defined treatment schedule, to TMS treatment in the event of symptom recurrence despite adequate oral antidepressant treatment.	The primary objective was to evaluate the efficacy of maintenance pharmacotherapy in patients meeting DSM-IV criteria for Major Depressive Episode, single or recurrent episode, who showed an adequate clinical response to daily dose TMS administered for up to 6 weeks by examining the time to first symptom recurrence. To minimize study bias, the Investigator was blinded to the definition of response.

Protocol 44-01101 was conducted under Neuronetics' IDE No. G030185 that was initially approved by the FDA on 10 October 2003, with final approval being granted on 24 May 2004.

A list of investigators participating in Study 44-01101 is provided in Appendix 1. The study protocol and informed consent document for Study No. 44-01101 is provided in Appendix 2. All referenced data tables are provided in Appendix 3. A sample case report form for Study 44-01101 is provided in Appendix 4. SAE vignettes for patients experiencing Serious Adverse Events are provided in Appendix 5.

2.0 PROTOCOL SUMMARY

Protocol 44-01101 was a randomized, parallel-group, sham-controlled clinical trial designed to test the efficacy and safety of TMS treatment using the Neuronetics TMS System for the treatment of patients diagnosed with DSM-IV defined major depressive disorder. The patients in this study included those with DSM-IV-defined MDD who have not benefited from prior adequate treatment with antidepressant pharmacotherapy.

Three hundred and twenty-five (N=325) patients with MDD participated in Study 44-01101 which was conducted at 25 investigational sites. Three sites were non-U.S. sites, two in Australia and one in Canada, which enrolled 25 patients. The non-U.S. studies were conducted under an Investigational Testing Applications (Canada) or Clinical Trial Notifications (Australia) approved by the regulatory authorities in the countries of clinical testing.

The design for this Study 44-01101 was organized into three experimental phases:

- a one-week, no treatment screening phase,
- a six week acute treatment phase, and
- a three week taper phase.

During the acute treatment phase, TMS sessions using the Neuronetics TMS System were scheduled in five-day contiguous treatment blocks, generally scheduled on Monday through Friday, for a maximum possible number of 30 treatment sessions. During the taper phase, all patients were placed onto open-label antidepressant pharmacotherapy and simultaneously cross-tapered off TMS treatment in a schedule of gradually less frequent treatment sessions (3 times per week, twice per week and then once per week).

Clinical evaluations for safety and efficacy occurred at approximately two-week intervals during the acute treatment phase, and weekly during the taper phase of the study.

The design of protocol 44-01101 was structured to address two questions:

1) Is TMS treatment using the Neuronetics TMS System a safe and effective acute antidepressant when administered as monotherapy?

2) Can the acute effect of TMS treatment using the Neuronetics TMS System be sustained in a clinically meaningful manner for a clinically appropriate duration subsequent to completion of an acute treatment course?

The key conclusions drawn from the results of Study 44-01101 in answer to the above questions were:

- TMS treatment using the Neuronetics TMS System was shown to be a clinically and statistically effective antidepressant monotherapy for the treatment of patients with Major Depressive Disorder, with single or recurrent episode, who had not previously been shown to receive adequate benefit from at least one but no more than four antidepressant medications during the qualifying episode.
- The acute clinical response to TMS treatment using the Neuronetics TMS System was successfully maintained over the course of a three week transition to maintenance of effect antidepressant pharmacotherapy.
- TMS treatment using the Neuronetics TMS System was demonstrated to have an adverse event profile consistent with previous exploratory studies and clinical case reports, and was notably absent of suicides, seizures, or of any effect on cognitive function or auditory threshold (with earplug use during TMS treatment) during the course of six weeks of acute treatment.
- TMS treatment using the Neuronetics TMS System was well tolerated by patients as evidenced by a low discontinuation rate during the acute treatment phase.

3.0 METHODS OF DATA COLLECTION AND ANALYSIS

3.1. Clinical Assessment Instruments

A comprehensive set of efficacy instruments was used in the Neuronetics studies to confirm the diagnosis and illness severity of the patient population, and to define the symptomatic and functional response to acute treatment with the Neuronetics TMS System. All instruments used are well-accepted and psychometrically valid psychiatric assessments, and are summarized in Table 2, and include both clinician-rated and patient-reported outcome measures.

Table 2. Diagnostic, Symptom Assessment, Functional Status and Quality of Life Instruments Used in Protocols 44-01101, 44-01102 and 44-01103

Assessment Tool	Description
Psychiatric Diagnostic Interview - Structured Clinical Interview for the DSM-IV (SCID-IV)	- The SCID-IV is a semi-structured diagnostic interview used to confirm the clinical diagnosis according to diagnostic criteria for Major Depressive Disorder consistent with the Diagnostic and Statistical Manual of Mental Disorders, 4th edition
Treatment History - Antidepressant Treatment History Form (ATHF)	- The ATHF is a semi-structured inventory used to rigorously characterize antidepressant treatment in terms of dosing adequacy, treatment duration, patient compliance and outcome. It has been shown to demonstrate predictive validity for the outcome of somatic treatments for depression, and hence is a valid alternative to a prospective treatment trial to establish antidepressant treatment resistance.
Clinician-Rated Symptom Assessments - Montgomery-Asberg Depression Rating Scale (MADRS) - Hamilton Depression Rating Scale (HAMD), 24-item and 17-item versions - Clinician Global Impressions – Severity of Illness (CGI-S)	- The MADRS is a well-recognized, observer-administered disease-specific rating scale that measures core symptoms of major depression on 10 items, with an emphasis on vegetative signs. Each item is scored on an integer scale from 0 to 6. - The HAMD is a standardized, observer-administered disease-specific rating scale that assesses up to 24 items characteristically associated with major depression. Each item is variably anchored with up to 5 integer scores, and item-specific anchor verbatim descriptions. It is reported as the first 17-items (HAMD17) or the full 24-items (HAMD24). - The CGI-S is an accepted, observer-administered, global illness rating scale that measures disease severity on a 7-point Likert scale.

Assessment Tool	Description
Patient-Reported Symptom, Quality of Life, and Functional Status Assessments - Inventory of Depressive Symptoms – Self Report version (IDS-SR) - Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (Q-LES-Q) - Medical Outcomes Study Short Form – 36 Item Questionnaire, version 1 (MOS SF-36) - Patient Global Impressions – Improvement of Illness Scale (PGI-I)	- The IDS-SR is a self-administered, 30-item rating scale that asks patients to identify symptoms characteristically associated with major depression, and rate the severity of each of these symptoms on a 4-point scale. - The Q-LES-Q short form is a self-administered quality of life instrument that asks patients to identify their overall level of satisfaction in 14 different areas of life function and 2 questions about global life satisfaction on a 5-point scale with 1 = Very Poor and 5 = Very Good. - The MOS SF-36 is a well-validated, self-administered questionnaire that measures a patient's functional health status. It has eight subscales that measure physical, social and role functioning, mental health, pain, and general health perceptions. This scale is a criterion standard for health-related quality of life. - The PGI-I is a well-recognized, self-administered, global rating scale that measures disease improvement on a 7-point Likert scale.
Patient-Reported Health Care Resource Utilization and Work Productivity Assessment - Health Resource Utilization Questionnaire (HRQ)	- The HRQ is a multi-item self-reported questionnaire which assesses health care utilization, work status and productivity, and caregiver burden.

Safety was assessed at each study visit by review of spontaneously reported adverse events, and separate reporting of all serious adverse events. All adverse events were initially c ics' contracted vendor for electronic data capture (EDC) using the current version of the Medical Dictionary for edDRA). All coding runs were reviewed and verified by Neuronetics clinical staff prior to final approval. Independent of coding, all adverse events were categorized by the investigative site staff that recorded the event, by severity and by relatedness to the device, i.e., the Neuronetics TMS System.

Additional targeted safety assessments included assessment of cognitive function and auditory threshold. Auditory threshold was examined since animal and human studies have suggested that prolonged exposure to the sound of the magnetic pulses during a TMS treatment course may be associated with short-term changes in auditory threshold. Cognitive function was a specific area of interest because of the known propensity for the relevant predicate device, namely electroconvulsive therapy (ECT) devices, to disrupt critical areas of general cognitive function and memory. The specific cognitive instruments were selected because they were similar or identical to instruments used in studies of cognitive function in patients receiving ECT treatment. These specific measures are shown in Table 3.

Table 3. Cognitive Function Testing Instruments for Neuronetics Studies 44-01101, 44-01102, 44-01103

Assessment Tool	Description
Modified Mini Mental Status Examination (MMSE)	This instrument assesses global cognitive function in several major neuropsychological domains
Buschke Selective Reminding Test (BSRT	This test evaluates short-term memory using immediate and delayed recall of common word lists
Autobiographical Memory Inventory-Short Form (AMI-SF)	This interview assesses the integrity of long-term memory functions by examining the ability to recall basic autobiographical information at post-treatment time-points that were obtained prior to the start of treatment

As commonly done in studies assessing cognitive effects, multiple versions of the MMSE and BSRT were used to allow repeat administrations and to deter potential learning effects.

3.2. Schedule of Events

A detailed discussion of the study protocol and procedures is included in Protocol 44-01101, Appendix 2, of this report. A synopsis of the study procedures is provided here, and the schedule of study events is outlined in Table 4.

The study procedures and foreseeable risks of the protocol and use of the study device were explained to all patients and informed consent was obtained prior to any study procedures.

Table 4. Schedule of Study Events for Protocol 44-01101

Phase	1-Week Prestudy		6-Week Acute Treatment						3-Week Post-Treatment Taper		
Week	Wk –2 to -1[a] (Screening)	Wk 0[a] (Baseline)	Wk 1[b]	Wk 2[b]	Wk 3[b]	Wk 4[b]	Wk 5	Wk 6[b,c]	Wk 1	Wk 2	Wk 3
Day(s)	-7[a]	0[a]	1-5	8-12	15-19	22-26	29-35	36-42	43-49	50-56	57-63
Informed Consent	X										
Medical History	X										
Antidepressant Treatment History Form (ATHF)	X										
Motor Threshold Determination	X		X	X[h]	X	X	X	X			
Structured Diagnostic Interview (SCID)	X										
Efficacy Assessments											
HAM-D_{24}	X	X		X[d]		X[d]		X[d]	X	X	X
MADRS	X	X		X[d]		X[d]		X[d]	X	X	X
CGI-S	X	X		X[d]		X[d]		X[d]	X	X	X
PGI-I		X		X[d]		X[d]		X[d]	X	X	X
IDS-SR		X		X[d]		X[d]		X[d]	X	X	X
Health Outcome Assessments											
SF-36		X				X[d]		X[d]			
Q-LES-Q		X				X[d]		X[d]			
Health Resource Questionnaires		X									
Neuropsychological Assessments											
Mini Mental Status Exam		X				X[d]		X[d]			
Buschke Selective Reminding Task		X				X[d]		X[d]			
Autobiographical Memory Interview		X				X[d]		X[d]			
Safety Assessments											
Physical examination	X										
Laboratory determinations[e]	X										
Pregnancy test[f]	X	X									
Urine drug screen	X										
Audiometry assessment	X					X		X			
ECG[i]	X										
Adverse Events[g]	X ----------					-X					
Prior/Concomitant Treatment	X ----------					-X					
rTMS Treatment Session (daily × 5 weekdays/week)			X---X	X---X	X---X	X---X	X---X	X---X			
Post-Treatment Taper rTMS Session(s) (3X/Wk1, 2X/Wk 2, 1X/Wk 3)									X	X	X

a. A minimum of 7 days may elapse between the screening and baseline visits; a maximum of 5 days may elapse between the baseline visit and the first treatment day of Week 1.

b. The first visit during each week of treatment should occur on a Monday, with daily treatment sessions occurring on Monday through Friday of each week.

c. Patients who prematurely discontinue should complete all Week 6 procedures within 2 days after their last rTMS treatment session.

d. Efficacy and neuropsychological assessments to be performed after last rTMS treatment session on last day of each treatment week block.

e. Laboratory determinations to include standard hematology, blood chemistry, and urinalysis tests.

f. If patient is a female of childbearing potential, a serum β-Human chorionic gonadotropin (β-HCG) test will be performed at screening and a urine pregnancy test will be performed at baseline.

g. Adverse events occurring prior to randomization will be recorded as part of each patient's medical history. Those AEs occurring following the first rTMS treatment session through 30 days after last rTMS treatment session will be collected.

h. In addition to the indicated days, motor threshold may be repeated at any time during the course of the active rTMS treatment sessions based on clinical assessment of the supervising physician.

i. ECG results from within 6 months can be used and another ECG is not required.

4.0 INVESTIGATIVE SITES FOR NEURONETICS STUDY 44-01101

4.1. Investigative Sites and Subjects Per Investigative Site

A listing of the clinical study investigators whose sites were qualified to conduct Study 44-01101 as assessed by Neuronetics staff per standard operating procedure and who participated in study 44-01101 is provided in Appendix 1 of this report. All investigators who participated in the conduct of Study 44-01101 also participated in Neuronetics' continuation studies 44-01102 and 44-01103. Enrollment into protocol 44-01101 for each site and the number of patients who transitioned from protocol 44-01101 into the other two protocols, 44-01102 and 44-01103, is also shown in the listing provided in Appendix 1.

Three hundred and twenty-five (N=325) patients with MDD participated in Study 44-01101 which was conducted at 25 investigational sites. A strict site closure policy was followed in this study, and two sites were closed due to lack of enrollment during their first 3 months of operation. Twenty-three sites contributed patients to the overall clinical development program.

Three sites were non-U.S. sites, two in Australia and one in Canada that, in aggregate enrolled 25 patients. The non-U.S. studies were conducted under an Investigational Testing Application (Canada) or Clinical Trial Notifications (Australia) approved by the regulatory authorities in the countries of clinical testing.

All sites underwent a site-specific study initiation meeting, and all staff were trained in protocol procedures and device use as described below.

4.2. Site Selection Procedures, Training Methods and Follow-Up Procedures for Study Device Operation

All study sites were assessed with an on site visit and interview of potential staff, using established standard operating procedures at Neuronetics. Qualified study sites were provided an extensive training sequence prior to being permitted to utilize the Neuronetics TMS System in the study protocol.

In November 2003, an investigator meeting held prior to the start of the protocol. During this meeting, study site personnel were provided a series of lectures that included a detailed review of the biophysics of magnetic stimulation, safety considerations and currently accepted safety practices, and a review of the safety procedures required for this study. For approximately half of one day, personnel participated in several hands-on didactic training stations that were set up with live demonstrations of the device equipment. All study staff were provided with written materials to review.

Subsequent to the initial training meeting, individual study site initiation visits were scheduled for each site. At these individual visits, all personnel who were expected to be using the Neuronetics TMS system during the trial were required to attend.

No personnel were permitted to use the Neuronetics TMS System unless they obtained specific training conducted and documented by Neuronetics and demonstrated evidence of competence in the use of the device.

The individual study site initiation visits were generally arranged in two meetings separated by approximately a two-week interval to permit the study site to practice use of the device. During the first session, a didactic review of the study specific procedures was held, with all personnel present. Individual two-hour, hands-on training sessions were then conducted with each staff member and a Neuronetics staff trainer, using a live subject. During these sessions, the staff member was individually trained in the technique of obtaining a motor threshold, and then trained in the specific method of treatment session procedures. All aspects of the protocol procedure were standardized to minimize operator-specific error as much as possible.

After the site training, the staff was given approximately a two-week interval to practice the study-specific techniques, after which time an oral examination was held. At these examinations, a Neuronetics trainer observed the staff member performing a live motor threshold uninterrupted, after which the staff member was required to verbally review the specific procedural requirements for the management of a TMS treatment session using the Neuronetics TMS System. All personnel were required to demonstrate facility with each element of the use of the Neuronetics TMS System. Evidence of these sessions was documented for each site member and is contained in the study master files at Neuronetics.

Following these training sessions, within-study follow up occurred in two ways. Neuronetics personnel were present at the first patient's baseline visit and first treatment at each study site. During these visits, Neuronetics staff members were able to observe continued adherence to protocol technique as taught in the training sessions. In addition, Neuronetics staff returned on at least two different occasions within the duration of the study to review procedural technique with all study sites. Any evidence of training deficiency was noted and remediated by the Neuronetics trainer during these visits.

4.3. Training Methods and Follow-Up Procedures for Clinician-Rated Assessments

The HAMD and MADRS were assessed by clinical raters using a semi-structured interview developed for this study by Drs. Harold Sackeim, Judith Kiersky and Mark Demitrack, and modeled after the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D) developed by Dr. Janet Williams at Columbia University (1988). This interview guide provides a verbatim leading question and a series of follow up questions designed to sequentially probe the symptom domains covered in the HAMD and MADRS interview, and permitted simultaneous scoring of the relevant items from both scales.

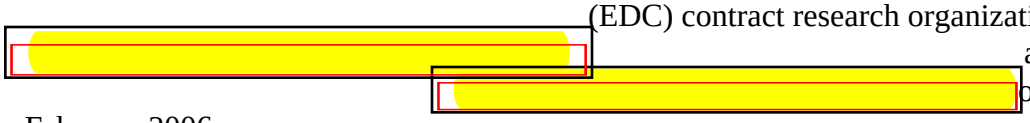
Rater quality and reliability on the use of this interview was assessed in two ways. All prospective raters were required to independently view and score a series of 5

videotapes of different patients interviewed using this structured guide. These tapes were prepared specifically for this study by staff of the Department of Biological Psychiatry at Columbia University and included patients with a broad range of relevant clinical symptomatology. Each rater's scores were compared to a pooled expert score for each tape, and a minimum threshold intraclass correlation statistic was required to be achieved prior to permitting the rater to participate in the study. Once the study ratings began, all patient HAMD/MADRS rating interviews for baseline, week 4 and week 6 assessments were videotaped, and a selected subset of these ratings for each rater were independently reviewed, and quantitatively scored for rater technique by an experienced rater at the New York State Psychiatric Institute. Any deficiencies in rater technique were identified, and if required, the rater was removed from the active rater pool. Details of the rater training program and documentation of the initial rater certification and the follow-up videotaped interviews is contained in the study master files at Neuronetics.

4.4. Case Report Forms and Methods of Data Management

Data was entered from source data records into a web-based electronic case report form database, or electronic data capture (EDC) system, at all participating clinical sites. Only site staff who were trained in data entry using this EDC system were authorized to enter the data.

Study monitoring was conducted by Neuronetics staff and contract research associates from MedSource, Inc., for all Neuronetics US and CA clinical study sites. The Australian sites were monitored by Quintiles, Inc. Both MedSource and Quintiles are qualified, contract research organizations. Neuronetics clinical study monitors verified entered data against source data records and queried all investigative site staff when needed for logical clarification of data or for missing data. The complete dataset for Study 44-01101 was locked on 31 January 2006, and final data was



(EDC) contract research organization analysis on 06 February 2006.

5.0 INCLUSION AND EXCLUSION CRITERIA

Detailed discussion of the inclusion and exclusion criteria and the procedures for their implementation is contained in the original protocol for study 44-01101 that is provided in Appendix 2. A summary of the major features of the inclusion and exclusion criteria is provided here.

The study procedures and foreseeable risks of the protocol and the study device were explained to all patients and informed consent was obtained prior to any study procedures. A copy of the informed consent document is provided in Appendix 2.

- Patients were eligible to participate in this study if they were outpatients ages 18 to 70, who met DSM-IV criteria for Major Depressive Disorder (MDD), single episode or recurrent, with a current illness duration of 3 years or less. The clinical diagnosis was confirmed by structured psychiatric interview with the Structured Clinical Interview for the DSM-IV (SCID-IV).
- At initial screening, patients were required to have a Clinical Global Impressions Severity of Illness (CGI-S) score of at least 4, and a minimum symptom severity as reflected by a total score of at least 20 on the 17-item Hamilton Depression Rating Scale (HAMD17), and an Item 1 score of at least 2. In addition, all patients had to demonstrate sustained symptom severity after the one week no-treatment lead-in period, as reflected by a HAMD17 total score of at least 18, and $\leq$ 25% decrease in score from that observed at the screening assessment.
- All patients were evaluated using the Antidepressant Treatment History Form (ATHF), a structured methodology that characterized their treatment history. To be eligible for study entry, patients must have failed to receive benefit from at least 1 but no more than 4 adequate trials of an antidepressant in the current or a past episode. For purposes of this study, adequacy of treatment was defined as an ATHF antidepressant resistance potency of at least Level 3 for the specific antidepressant.
- Any patient currently receiving treatment with psychotropic medication was required to washout from these medications prior to completion of the screening process.
- Exclusionary criteria for study entry included a history of psychosis, bipolar disorder, or obsessive compulsive disorder. Post-traumatic stress disorder and eating disorders were excluded only if active in the past year.
- Patients who had failed to receive benefit from an adequate trial of electroconvulsive therapy at any point in their lifetime were excluded.
- Patients who had been previously treated with experimental TMS or had received a vagus nerve stimulator implant were excluded from study.
- Patients who had 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 were excluded.
- Pregnancy, or women of reproductive age who were not using a medically accepted form of contraception during intercourse were not permitted to enroll.

- A history of seizure disorder or any neurologic disease or medication therapy known to alter seizure threshold was not permitted.
- The presence of ferromagnetic material anywhere in or in close proximity to the head precluded study entry.

A patient's medical history, physical examination, laboratory studies, including a urine toxicology screen, and electrocardiogram were performed at study entry to ensure that a patient was medically stable and that no excluded psychotropics such as benzodiazepines were being taken.

6.0 STUDY POPULATIONS AND STATISTICAL ANALYSIS

6.1. Study Populations

The *all-randomized study population* (N=325) was defined as those individuals who signed an informed consent, and were subsequently randomized to a treatment condition (this population includes both the evaluable (N=301) and the non-evaluable (N=24) patient samples).

The *modified intent-to-treat study population (N=301, (also known as the evaluable patient sample)* was defined as all subjects who signed an informed consent document, were randomized to a treatment condition and received at least one treatment (whether partial or complete), and for whom a completed post-randomization observation was available for analysis.

The evaluable study population served as the primary population of interest for analysis on all *a priori*-defined primary and secondary outcome measures.

All patients who signed an informed consent document, and received at least one randomized treatment, constituted the *safety population* (N=323). Two patients signed an informed consent document, were randomized to a treatment condition, but were unable to proceed to treatment due to the inability to determine a motor threshold, and were therefore excluded from summary in the safety population tables.

Serious adverse events were reported for all patients who signed an informed consent document.

6.2. Statistical Analysis

6.2.1. Sample Size Justification and Power Analysis

The sample size was arrived at by requiring 90% power and a two-sided 5% test, and is based on the standard t-test method. A standardized effect size (difference in LV means divided by the standard deviation of the score) of d =0.4 was targeted in this study. As stated in Protocol 44-01101, an interim analysis for futility was to be conducted *a priori* when a total sample of approximately N=100 patients were enrolled. Stopping for futility at a conditional power of 20% increases the nominal type II error rate by less than a factor of 10/8 = 1.25. To guarantee a final 10% type II error rate (90% power), a nominal type II error rate was set at 8% (power = 92%) for a total N=286 (143 per treatment group). This sample size includes evaluable patients only, since the specific number of potential non-evaluable patients in the sample could only be observed as the study was underway. Nevertheless, the sample size calculation is conservative in not taking advantage of the adjustment for baseline scores that should reduce the residual variation and therefore increase the power of the test in actuality. In addition, the method for adjustment of type II error inflation due to futility

monitoring was based on a formula for inflation that is known to be conservative.

6.2.2. Statistical Analysis Methods

Study 44-01101 was a randomized, parallel-group comparison of treatment with the Neuronetics TMS System with a matched Neuronetics TMS System sham control. Twenty-five sites completed site initiation; 23 sites contributed patient data to the final study population. Two sites were closed for non-performance. A strict closure policy was used for sites that did not show early signs of enrollment success, which resulted in the early closure of these two sites.

The study used a permuted block design (block size = 6) to improve balance within sites. As stated above, the primary hypothesis to be tested in this study compared the active treatment with the Neuronetics TMS System and sham treatment groups on the last post-treatment symptom score (LV) measured using the primary efficacy outcome measure (MADRS total symptom score at 4 weeks of acute phase treatment) for each patient. The primary efficacy analysis was performed on the intent-to-treat sample of all evaluable patients, meaning those patients with a baseline and at least one post-baseline observation available for analysis.

In the Protocol 44-01101, an *a priori* consideration was made which stipulated that poorly recruiting sites, defined as those with fewer than 2 randomization blocks (randomization schedule block size = 6), would be pooled into one or more pseudo-sites for purposes of analysis. Prior to breaking of the study blind, review of patient recruitment across sites revealed that the most logical pooling of low enrolling study sites would be accomplished by establishing a single pseudosite of all sites that enrolled less than one complete block size, i.e., less than 6 patients. This produced a single pseudosite of N=11 patients, and was employed as such in the statistical analysis.

For the primary efficacy outcome measure (i.e., MADRS total symptom score observed at 4 weeks of treatment during the acute treatment phase), the null hypothesis was tested in an analysis of covariance of the LV, using baseline score, and ATHF medication resistance level as fixed effect covariates, adjusting for site differences using a random effect. The ATHF medication resistance levels were grouped into two categories in the statistical model, 2 or less in the reference episode (current or past) or 3-4 in the reference episode (current or past). All tests were two-sided, with a conventional level of statistical significance set at the 5% level.

As described above, key secondary efficacy outcomes were tested as supportive indices of clinical efficacy of the Neuronetics TMS System and included other continuous measures, and within-patient dichotomous variables. For these secondary analyses, the treatment effect null hypothesis was tested by logistic regression of treatment group assignment with adjustment for site and ATHF

medication resistance level. In addition, the longitudinal symptom scores were analyzed with a repeated measures general linear model, adjusting for baseline scores and ATHF medication resistance level (using Proc Mixed in SAS Version 8.2 or higher). The model included the covariates of baseline score and ATHF medication resistance level as fixed effect covariates, treatment effect, and site differences using a random effect. Time was included in the model as a repeated measure. Additionally, the treatment by time interaction was included in the model. The inclusion of this interaction term allowed for an assessment of the treatment effect at each time point. An unstructured covariance matrix was used in the analysis.

7.0 STUDY PERIOD AND EVALUABLE PATIENTS

The first site initiation for protocol 44-01101 occurred on 18 December 2003, and first patient was enrolled on 26 January 2004. At the closure of study enrollment, 801 patients had been consented for study participation, while 325 patients had been randomized to a treatment condition.

The sample size estimation for the protocol was N=286 evaluable patients as discussed in Protocol 44-01101. Because the exact final number of evaluable patients could only be determined at the conclusion of enrollment, a careful tracking process was instituted to ensure that a sufficient evaluable patient sample would be included in the final study population. Estimates of the attrition of patients from the time of signing of informed consent to the point of randomization to treatment were followed closely during enrollment. Based on these estimates, it was anticipated that a minimum evaluable patient sample would be achieved by closure of further patient consent on 05 August 2005. At this point, 264 evaluable patients were present in the study population. All patients who were consented as of that date were permitted to complete their screening process, and were not denied study enrollment if they met appropriate inclusion criteria. At the same time, a date was declared for all study sites for last patient randomization of 06 September 2005. To be randomized, all patients must have completed their screening procedures prior to that date.

The final enrolled patient population was 325 patients. This ex approved enrollment of 286 patients. This protocol deviation was filed t (Ser. No. 031 dated 04 October 2005) and approved by the FDA on 18 N

Among the all-randomized study population, there were 24 patients who were non-evaluable according to the operational criteria stipulated in the protocol, N=14 were allocated to sham treatment, and N=10 were allocated to active TMS treatment. Patient identification, treatment arm allocation, and reason for discontinuation for all of these patients are listed in Table 5.

Table 5. Summary Patient ID, Treatment Arm Allocation, and Reason for Discontinuation Among Non-Evaluable Patient Sample

Patient ID	Treatment Arm Allocation	Reason for Discontinuation
01-097	Sham	SAE (Suicidal ideation)
03-004	Sham	Protocol violation (use of excluded medication)
05-010	Sham	Adverse event (various somatic symptoms reported)
05-025	Sham	Patient request (due to work schedule interference)
05-037	Active	Adverse event (unable to tolerate treatment)
06-014	Active	Adverse event (worsening depression)
06-027	Sham	Failed to return
08-003	Sham	Motor threshold > 80%
10-040	Active	Other (could not withhold sleep medications)
11-017	Sham	Protocol violation (positive urine drug screen)
11-030	Active	Adverse event (pain at treatment site)
11-046	Active	Failed to return
12-034	Active	Failed to return

Patient ID	Treatment Arm Allocation	Reason for Discontinuation
13-012	Sham	Failed to return
13-017	Active	Adverse event (tension headaches, nausea)
15-005	Sham	Failed to return
15-018	Sham	SAE (suicidal ideation)
15-021	Active	SAE (unsatisfactory response, suicidal ideation)
16-030	Active	Adverse event (use of excluded medication)
17-033	Sham	Protocol violation (age > 70)
21-002	Sham	Other (changed his mind)
23-019	Sham	Protocol violation (positive urine drug screen)
24-001	Active	Patient request (withdrew consent)
24-003	Sham	Could not detect Motor Threshold

8.0 PATIENT DEMOGRAPHICS AND BASELINE ILLNESS CHARACTERISTICS OF THE 44-01101 STUDY POPULATION

The all-randomized study population included 325 patients. Demographic and clinical variables for this population are described Section 7.1. Baseline illness characteristics are described in Section 7.2.

8.1. Patient Demographic and Clinical Variables

A complete description of the demographic features for the all-randomized study population (N=325) are described in Appendix 3, Table 3.1 and for the intent-to-treat, evaluable study population (N=301) in Appendix 3, Table 3.2.

A complete description of the clinical variables for the all-randomized study population (N=325) and for the intent-to-treat, evaluable study population (N=301), that were obtained at screening, are shown in Appendix 3, Tables 3.3 and 3.4, respectively.

A brief summary of key observations from the demographic features and baseline clinical variables are shown in Table 6 for the intent-to-treat, evaluable study population. Please see Tables 3.1-3.4 in the Appendix for further detail. A comparison of the evaluable and the non-evaluable patients on these descriptive features is also provided in Appendix 3 Tables 3.5 and 3.6 to demonstrate that no substantial differences were observed in these two subsets of the all-randomized patient population.

Table 6. Summary of Key Demographic and Clinical Variables Observed at Screening in the Intent-To-Treat, Evaluable Study Population

Variable Name	Treatment Group		P-Value
	Sham (N=146)	Active (N=155)	
Gender N(%) -Male -Female	 72 (49.3) 74 (50.7)	 69 (44.5) 86 (55.5)	 .421
Age [yrs, mean (SD)]	48.7 (10.6)	47.9 (11.0)	.509
Ethnic Origin N(%) -Caucasian -African-American -Asian -Hispanic -Native American -Other ---------- -Caucasian -All other groups combined	 131 (89.7) 3 (2.1) 1 (0.7) 8 (5.5) 0 3 (2.1) 131 (89.7) 15 (10.3)	 146 (94.2) 3 (1.9) 1 (0.6) 3 (1.9) 1 (0.6) 1 (0.6) 146 (94.2) 9 (5.8)	 .394 .201
Motor Threshold	57.0 (9.97)	55.2 (9.67)	.101

8.1.1. Conclusions Regarding Patient Demographics and Clinical Variables

- There were no statistically significant differences between the patient groups allocated to active TMS treatment using the Neuronetics TMS System or sham treatment on any demographic variables.
- The average age of patients was in their 5th decade of life, consistent with expectations for a more treatment-resistant population.
- There was a relatively equivalent representation of men and women in the study population.
- There were no clinically meaningful differences on other clinical variables at study entry.
- Patterns of demographic and clinical variables at screening showed no differences when contrasted between the all-randomized study population and the intent-to-treat, evaluable study population, and when contrasted between the evaluable and the non-evaluable study population, indicating that the efficacy conclusions drawn from the intent-to-treat, evaluable study population are likely to be generalizable across these various population subsets.

8.2. Baseline Illness Characteristics

A summary of *illness history, characterization of treatment resistance history, and baseline symptom severity* is included in Table 7 for the intent-to-treat, evaluable study population. A more complete description for this study population and a similar tabular summary for the all-randomized study population are provided in Appendix 3, Tables 3.7 and 3.8 and shows a similar distribution of illness descriptive variables.

Table 7. Key Observations for Illness History, Characterization of Treatment Resistance History and Baseline Symptom Severity for the Intent-To-Treat, Evaluable Study Population

Variable Name	Treatment Group		P-Value
	Sham (N=146)	Active (N=155)	
Depression History			
- Single episode	9 (6.2)	7 (4.5)	
- Recurrent episodes	136 (93.8)	149 (95.5)	.611
Duration of current episode			
- Length [mean (SD)]	13.2 (9.5)	13.6 (9.9)	.728
- < 24 months N(%)	123 (84.2)	119 (76.8)	
- ≥24 months N(%)	23 (15.8)	36 (23.2)	.112
Secondary Diagnoses N(%)			
- None	104 (71.2)	96 (61.9)	
- Any Other Anxiety Disorder	42 (28.8)	59 (38.1)	.112
ATHF Rating Summary (# of Level 3 Exposures)			
- 1	76 (52.1)	88 (56.8)	
- 2	50 (34.2)	45 (29.0)	
- 3	15 (10.3)	15 (9.7)	
- 4	5 (3.4)	6 (3.9)	
- >4	--	1 (0.6)	.816
Mean # of ATHF Level 3 Exposures	1.6	1.6	
MADRS Total Score [mean (SD)]	32.9 (5.6)	32.6 (5.3)	.476
HAMD24 Total Score [mean (SD)]	30.6 (4.3)	30.7 (3.9)	.803
HAMD17 Total Score [mean (SD)]	22.9 (3.1)	22.6 (2.3)	.325
CGI-Severity Score [mean (SD)]	4.7 (0.7)	4.7 (0.6)	.871
IDS-SR Total Score [mean (SD)]	43.4 (9.9)	42.0 (9.4)	.197

8.2.1. Baseline Illness Characteristics Conclusions

- The overall pattern of illness history in the subject patient population is consistent with a more severe treatment-resistant sample as reflected by the predominance of recurrent depression, and an ATHF assessment which yielded an average Level 3 resistance rating for 1.6 medications in both the active TMS and sham TMS treatment groups in the qualifiying episode.
- Baseline clinical symptom severity was consistent with this illness history as evidenced by the average scores at baseline on the HAMD24, HAMD17, MADRS, IDS-SR and CGI-Severity ratings, which suggest a moderate to severe clinical presentation in the current episode

9.0 HEALTH RESOURCE UTILIZATON AND FUNCTIONAL STATUS

Functional status, work productivity, health resource utilization and quality of life satisfaction were appraised by patient-rated questionnaires at study entry in the all-randomized study population. A summary of key observations obtained from the Work Productivity and Health Resource Utilization Questionnaire is shown in Table 8. A complete, detailed tabular summary of all data measured for functional status, quality of life and health resource utilization is included in Tables 3.9 in Appendix 3.

Table 8. Work/Productivity and Health Resource Utilization in the All-Randomized Study Population at Study Entry

Variable Name	Treatment Group	
	Sham (N=160)	**Active (N=165)**
Productivity/Work Loss due to Illness		
- Work Status N(%)		
o Full time	45 (28.3)	58 (35.6)
o Part time	31 (19.5)	27 (16.6)
o Not working	83 (52.2)	78 (47.9)
- Disability payments		
o Yes	31 (34.1)	28 (32.9)
o No	60 (65.9)	57 (67.1)
Health Utilization and Cost of Illness		
- # visits to HCP for depression in last 3 mos (median)	3.0	3.0
- # visits to HCP for medical problem in last 3 mos (median)	2.0	2.0
Caregiver Support		
- Assisted by a caregiver? N(%)		
o Yes	20 (12.7)	23 (14.3)
o No	137 (87.3)	139 (85.8)
- # hours assisted each week by caregiver (median)	8.0	12.0

9.1. Health Resource Utilization and Functional Status Conclusions

- The pattern of health resource utilization and work productivity impairment indicate a pattern of morbidity consistent with a more difficult to treat history; for example approximately half of the population in each treatment group were currently not working, with nearly 75% of each group reporting that this was due to depression; nearly 15% of each treatment group were receiving the assistance of a caregiver at home for daily tasks.
- On measures of functional health status, patients entering study 44-01101 showed a degree of functional morbidity consistent with their general illness history, presenting symptom severity and degree of treatment resistance.

10.0 PATIENT DISPOSITION

Subsequent to randomization, there were two discrete phases in Protocol 44-01101, the *acute treatment phase* and the *post-treatment taper phase*.

Treatment through Week 4 of the acute treatment phase constituted the *a priori*-defined study period for the primary efficacy analysis.

At week 4 or at later time points, patients were permitted access to the open-label crossover protocol 44-01102 if they or their study site Principal Investigator elected to discontinue their participation in Protocol 44-01101 and if they were otherwise clinically eligible to enroll in Protocol 44-01102 per protocol criteria.

For those patients continuing on their randomized treatment assignment beyond week 4, the time period between week 4 and week 6 served as an *a priori*-defined secondary analysis time point, and provides supportive information whether additional treatment sessions may confer added clinical benefit.

Subsequent to the conclusion of the acute treatment phase, durability of the acute effect of TMS treatment using the Neuronetics TMS System was examined in the patients who proceeded on their randomized treatment assignment into the 3-week, post-treatment taper phase.

The overall pattern of patient disposition across these various study phases is described in Figure 2. The reasons for termination as recorded by the study investigator at the time of patient discontinuation are listed for each critical time point in the study. Per investigator request, two patients were permitted to exit the acute treatment phase at the end of acute treatment week 4, and directly transition to the taper phase and so are not counted in the week 6 totals.

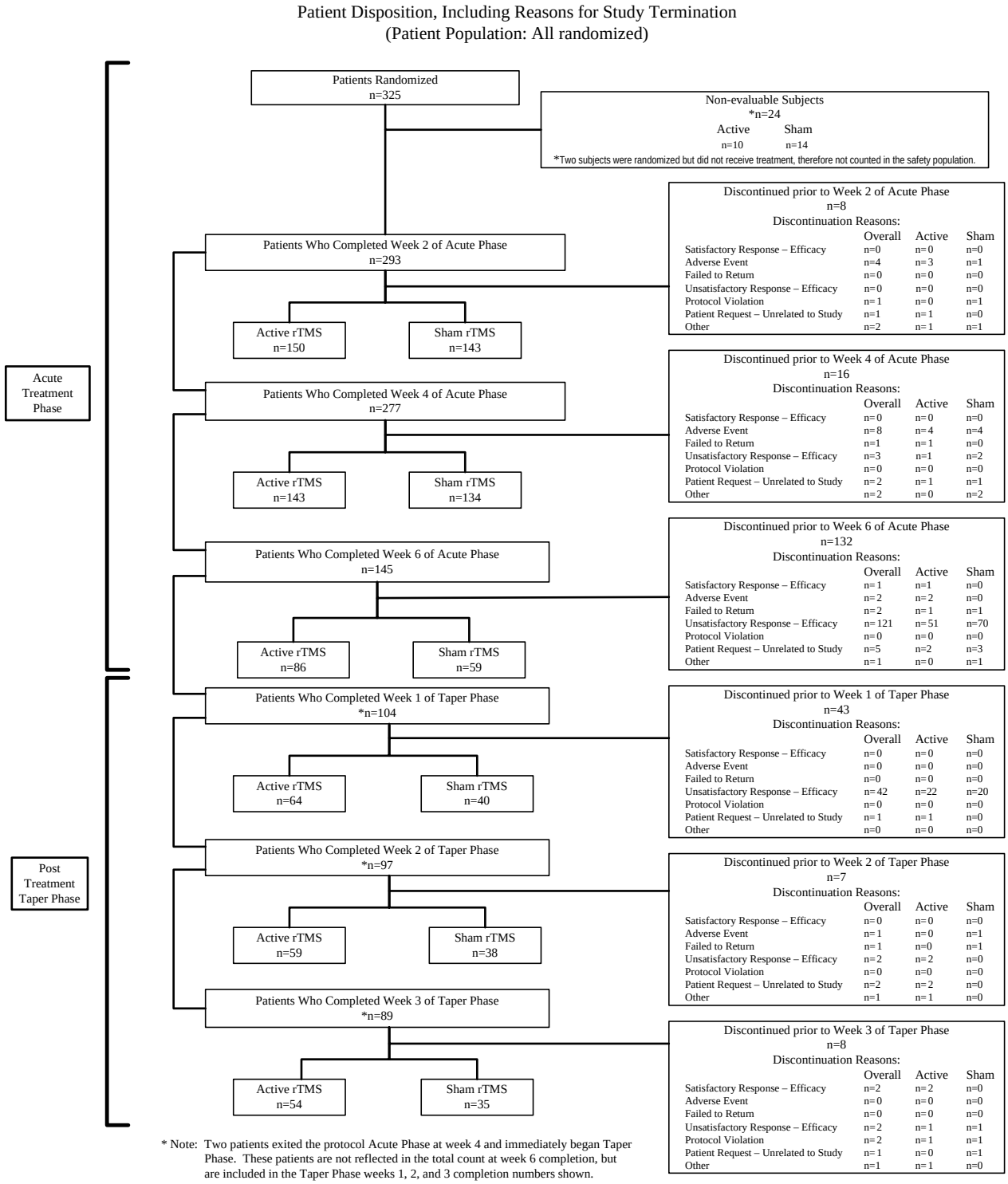


Figure 2. (Corrected) Diagram of Patient Disposition Across Study Phases in Protocol 44-01101

10.1. Patient Disposition Conclusions

- The overall adherence rate through week 4 of the acute treatment phase (the primary efficacy endpoint) was 92%.
- Discontinuation due to adverse events through week 4 of the acute treatment phase was 4.5% for patients allocated to active TMS treatment, and 3.4% for patients allocated to sham TMS treatment.
- By week 6, a greater percentage of patients allocated to sham TMS treatment elected to discontinue due to lack of efficacy (92/146 = 63.0%), compared to those patients allocated to active TMS treatment who elected to leave early due to lack of efficacy (74/155 = 47.7%).

11.0 STUDY DEVICE AND TREATMENT RANDOMIZATION

11.1. Study Device: Neuronetics Model 2100 TMS System

All TMS treatments were delivered using the Neuronetics Model 2100 TMS System. The system is described in detail in

In brief, the Neuronetics Model 2100 TMS System is an electromechanical instrument that non-invasively produces and delivers brief duration (~200 µsec) rapidly alternating, or pulsed, magnetic fields to the patient's head leading to the induction of electrical currents at spatially discrete regions of the cerebral cortex.

This method of cortical stimulation by application of brief magnetic pulses to the head is known as Transcranial Magnetic Stimulation or TMS. The peak magnetic field strength achieved with each pulse is approximately 0.5 Tesla in the cortex.

Study 44-01101 is intended to test the safety and efficacy of TMS as delivered by the Neuronetics Model 2100 TMS System for the treatment of Major Depressive Disorder (MDD). For treatment of MDD, TMS stimulation is directed to the left prefrontal cortex, a discrete region of the brain involved in mood regulation.

In commercial application, the Neuronetics TMS System will be provided on an out-patient basis by a licensed medical professional (i.e., psychiatrists and their staff) and by prescription only.

The Model 2100 TMS System consists of various hardware components, accessories and consumable supplies. The key components are the console which contains the controlling electronics of the system, the ferromagnetic coil that delivers the magnetic field to the patient's head and the E-Shield, which is a disposable circuit placed on the surface of the coil to decrease the induced electric field in the scalp in order to enhance patient tolerability.

Further details regarding the design of the Model 2100 TMS System may be found in

11.2. Treatment Randomization

Three separate "coded" magnetic coils were provided to each site for this study. All coils were identical in weight, external appearance and acoustic properties when actively pulsed. One coil was not blinded, and was used as a known active coil to determine motor thresholds (coil labeled 'MT Active'), and for use in the open-label portions of the clinical development program (i.e., Protocols 44-01102 and 44-01103). The remaining two coils were distinguishable only by external labels as 'coil B' or 'coil C'. One coil was an active treatment coil while the other was a sham coil. Deta and performance characteristics of the sham coil were provided in

Blinded, randomized coil assignment to each patient was indicated by the electronic information previously recorded on flash memory embedded on the unique treatment card assigned to that patient, and was based on the pre-study randomization sequence. When inserted into the console, the operator was prompted to attach the specific coil defined by the randomized treatment assignment, displayed on the console by the text: "Attach Coil B" or "Attach Coil C". The site staff then manually connected the appropriate coil prior to proceeding with each treatment session.

A multi stage process was used at Neuronetics to ensure that the randomization schedule was correctly applied during the programming of the patient treatment cards and that the labeling of the coded treatment coils was correctly allocated to sham or active coil. This process included:

- Patient treatment card programming according to the randomization schedule with verification by the programmer and separate Quality Control verification by the contract manufacturer.
- Manufacture of active and sham coils, labeling with a coil type label (i.e., "B" or "C") and testing for compliance to specification; coil performance and coil type retested at the system integration step before shipment.
- Third-party audit of manufacturing records to confirm correct patient treatment card allocation to active or sham, verification of coil type (active or sham) by review of labeling and test records including review of magnetic field output data, polarity and interlock connector tests for each coil.
- Neuronetics audit of manufacturer's records to verify adequacy of patient treatment card and coil manufacturing and randomization process and review of third-party audits.

These procedural approaches and methods of coil blinding were intended to ensure appropriate, blinded randomization for patient treatment and identical appearance, placement and acoustic properties of the magnetic coils in both active and sham treatment conditions. Additional discussion of measures taken to adequately ensure the integrity of the study blind are discussed in Section 21.2.

12.0 TMS TREATMENT SCHEDULE, TMS TREATMENT PARAMETERS AND COMPLIANCE

TMS treatment sessions were conducted using the Neuronetics Model 2100 TMS System in sequential five-day treatment blocks, generally administered Monday through Friday, during the acute treatment phase. Six additional treatments were administered across the 3 week post-treatment taper phase. A maximum of 36 treatments could have been given to any patient who completed all assigned treatment sessions in this study.

Treatment parameters were standardized for each treatment session using a magnetic field intensity of 120% of the patient's observed motor threshold, at a repetition rate of ten magnetic pulses per second. During the first week of the acute phase only, treatment intensity could be adjusted to 110% of observed motor threshold if clinically indicated for tolerability. Pulses were grouped in 30 second cycles with a stimulation on-time of 4 seconds, and an off-time of 26 seconds. A treatment session lasted for 37.5 minutes for a total number of 3000 magnetic pulses per session.

Motor threshold was determined weekly during the acute treatment phase by visual observation of thumb or finger movement using MT Assist, a standardized mathematical algorithm that provided an iterated estimate of the motor threshold across four estimations (MT1 through MT4). The final motor threshold was computed as the average of the four iterations (Recommended MT).

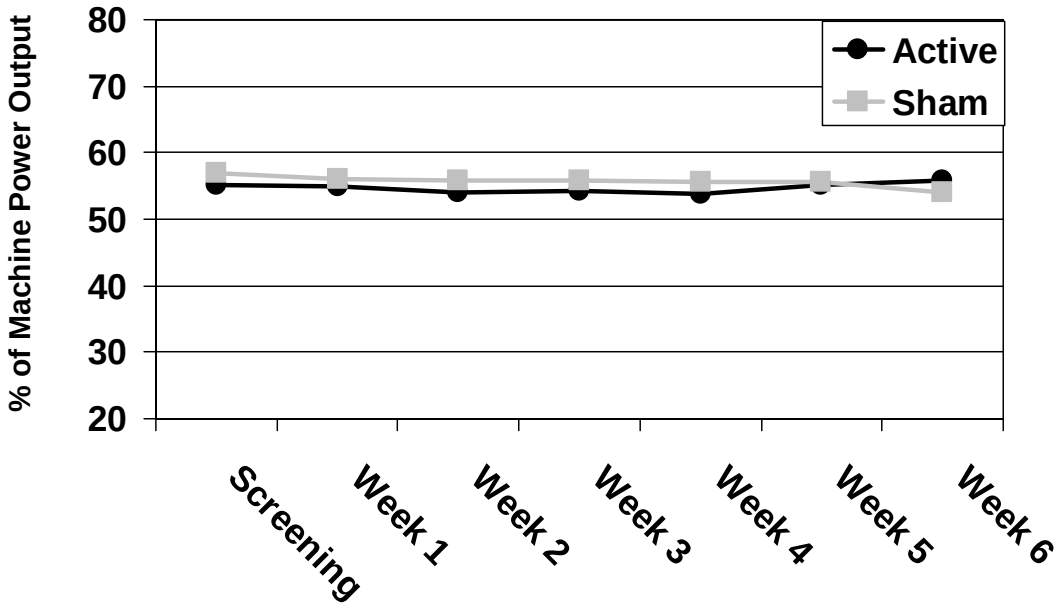
The standardized treatment location was operationally defined in the protocol over the left prefrontal cortex, determined by a standard convention of movement of the TMS coil 5 cm anterior to the motor threshold location along a left superior oblique plane, with a rotation point about the subject's nose. Spatial coordinates of this position were recorded to allow precise placement of the coil in the same position for the next treatment session. Coordinates were reset weekly with each repeat motor threshold. Coil movement within a treatment session was permitted in a limited, pre-defined sequence for comfort as needed, to limit variability in placement.

All patients were assessed for compliance with the intended treatment schedule during the acute treatment phase. Compliance was defined as missing less than 3 treatments in daily sequence, or missing less than 20% of the total number of treatment sessions as outlined in the schedule of events to be administered during the acute treatment phase for that patient.

Detailed tabular summaries of the weekly information obtained for all relevant treatment variables are contained in Tables 3.10 and 3.11 in Appendix 3. The mean number of patient treatment sessions conducted and treatment compliance are summarized in Table 9. The pattern of weekly recommended motor thresholds obtained during the study is shown in Figure 3.

Table 9. No. of Patient Treatments and Compliance in Protocol 44-01101

Treatment Characteristic	Active Treatment Group (N=146)	Sham Treatment Group (N=155)	P-Value
# of Sessions Administered During the Acute Treatment Phase (mean [SD])	24.2 (6.9)	22.8 (6.1)	0.067
Treatment Session Compliance			
• Missed > 2 consecutive sessions N(%)	18 (11.6)	12 (8.2)	0.343
• Missed > 20% of total intended sessions N(%)	6 (3.9)	0	0.03



Note: Average value for recommended MT at the indicated time point is shown based on MT Assist algorithm

Figure 2. Weekly Motor Thresholds Observed During the Acute Treatment Phase of Protocol 44-01101

12.1. Patient TMS Treatment and Compliance Conclusions

- Overall compliance with the scheduled treatment parameters was excellent (6/301 missed > 20% of the intended number of treatment sessions = 98% compliance).
- Motor thresholds demonstrated a stable pattern across the acute treatment phase, and were not clinically meaningfully different between the two treatment groups at any time point.

13.0 CONCOMMITANT MEDICATION USE

Psychotropic medication use during the study was strictly limited. All patients were free of antidepressants or other psychotropic medications directed at treatment of their study diagnosis. Patients 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 interests of patient care and so as not to unduly influence the efficacy and safety assessments in the study.

Table 10 summarizes the frequency of anxiolytic and hypnotic use during the acute treatment phase. As shown, ~30% of patients had some anxiolytic use in both active and sham TMS treatment groups.

Table 10. Frequency of Protocol-Approved Anxiolytic or Hypnotic Medication Use During the Acute Treatment Phase

Medication Name Preferred Term	Sham TMS (N=146) N (%)	Active TMS (N=155) N (%)
Subjects With At Least One Anxiolytic/Hypnotic Medication	44 (30.1)	44 (28.4)
• Chloral Hydrate (no brand name)	1 (0.7)	0
• Clonazepam (Klonopin)	1 (0.7)	1 (0.6)
• Diphenhydramine (Sominex, Benadryl)	0	1 (0.6)
• Eszopiclone (Lunesta)	1 (0.7)	0
• Lorazepam (Ativan)	21 (14.4)	26 (16.8)
• Zaleplon (Sonata)	1 (0.7)	2 (1.3)
• Zolpidem (Ambien)	28 (19.2)	21 (13.5)
• Zopiclone (Immovane)	0	2 (1.3)
• Temazepam	1 (0.7)	2 (1.3)
• Thiopenthyl	0	1 (0.6)

During the post-treatment taper phase, oral antidepressant medication was initiated. The choice of medication was limited to a monotherapy selected from among a protocol-approved list, and also was limited to a medication for which the patient had not previously been shown to have failed to receive benefit. A summary of the antidepressant medications chosen for use during the post-treatment taper phase are listed in Table 11.

The pattern of use of these medications did not differ substantially between treatment groups.

Because of a history of medication intolerance, 8 patients were approved to proceed through the post-treatment taper phase, but were not initiated on antidepressant medication. These patients were not, therefore, eligible to continue into Protocol 44-01103.

Table 11. Antidepressant Medications Used During the Post-Treatment Taper Phase

Antidepressant Medication	Drug Name	Sham TMS (N=40) N (%)	Active TMS (N=64) N (%)
Selective Serotonin Reuptake Inhibitors	Citalopram (Celexa)	1 (2.5)	0
	Escitalopram (Lexapro)	5 (12.5)	12 (18.8)
	Fluoxetine (Prozac)	2 (5.0)	3 (4.7)
	Fluvoxamine (Luvox)	1 (2.5)	1 (1.6)
	Paroxetine (Paxil)	3 (7.5)	0
	Sertraline (Zoloft)	1 (2.5)	3 (4.7)
Serotonin/Norepinephrine Reuptake Inhibitors	Duloxetine (Cymbalta)	15 (37.5)	15 (23.4)
	Venlafaxine (Effexor)	9 (22.5)	15 (23.4)
Other Antidepressants	Bupropion (Wellbutrin)	7 (17.5)	12 (18.8)
	Mirtazapine (Remeron)	1 (2.5)	4 (6.3)
	Trazodone (Desyrel)	0	2 (3.1)

Notes: In a few instances, patients were permitted to change medications after less then 2 days of dosing if immediate tolerability issues emerged, therefore the medication totals in each column exceed the total number of patients in each treatment arm.

14.0 EFFICACY OUTCOMES

The primary and secondary outcome measures used in the analyses for Study 44-01101 and the order of their sequential testing are listed in Table 12 and are also described in the original protocol provided in Appendix 2.

In all analyses, the primary study population of interest was declared as the *intent-to-treat population*, defined as including all subjects who signed an informed consent, were randomized to a treatment condition and received at least one treatment (whether partial or complete), and for whom at least one completed post-randomization observation was available for analysis.

Table 12. Primary Outcome Measure and Secondary Outcome Measures in Protocol 44-01101 and Their Sequential Order of Importance in Testing

Measurement	Evaluation
Primary Outcome Measure	Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using the last post-treatment total symptom score on the Montgomery-Asberg Depression Rating Scale (MADRS) through week 4 of the acute treatment phase 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.
Secondary Outcome Measures	1) Evaluate the antidepressant effect of TMS treatment with the Neuronetics TMS System, using 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, of a specified course of active treatment when compared to sham treatment 2) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using 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, of a specified course of active treatment when compared to sham treatment 3) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using the total symptom score on the MADRS for the last post-treatment value observed through week 6 of the acute treatment phase, of a specified course of active treatment when compared to sham treatment 4) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using 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), of a specified course of active treatment when compared to sham treatment 5) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using 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, of a specified course of active treatment when compared to sham treatment 6) Evaluate the antidepressant effect of treatment with the Neuronetics TMS Sys-

Measurement	Evaluation
	tem, using categorical outcome of remission/recovery (percent of patients achieving HAMD17 total symptom score < 8, HAMD24 total symptom score < 11, and MADRS total symptom score < 10 at the last post-treatment visit through week 4 and week 6, of a specified course of active treatment when compared to sham treatment 7) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using 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, of a specified course of active treatment when compared to sham treatment 8) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using 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, of a specified course of active treatment when compared to sham treatment 9) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using the Clinical Global Impressions – Severity (CGI-S) score, using last post-treatment value through week 4 and week 6, of a specified course of active treatment when compared to sham treatment 10) Evaluate the antidepressant effect of treatment with the Neuronetics TMS System, using the Patient Global Impressions – Improvement (PGI-I) score, using last post-treatment value through week 4 and week 6, of a specified course of active treatment when compared to sham treatment

14.1. Primary Efficacy Outcome – Acute Phase

The *a priori*-defined primary outcome measure in Study 44-01101 was based on the last post-treatment total symptom score on the Montgomery-Asberg Depression Rating Scale (MADRS) through week 4 of the acute treatment phase. This was to be conducted on the intent-to-treat, evaluable study population as defined above. The results of this analysis are shown in Table 13 and in Figure 4.

As shown, the P values for MADRS total symptom score showed a strong statistical trend at p=.057 and p=.058 at 4 and 6 weeks, respectively.

A statistically significant baseline imbalance was observed in the total score on the MADRS between the active TMS and sham TMS treatment groups (LS mean for active TMS = 32.4 [SD 5.99], LS mean for sham TMS = 33.7 [SD 5.69], p = .036). This unexpected outcome arose because of the nature of the study design itself, whereby the baseline screening measure used (i.e., the HAMD17) had a minimum numerical threshold for entry, while the primary outcome measure (i.e., the MADRS) did not. As a result, while no baseline imbalance was detected for the HAMD 17 and 24 item measures (p>0.05), a small (N=6), but nevertheless statistically influential proportion of patients, who had unusually low scores at entry on the MADRS, were over-represented in the active TMS study population (N=4 pa-

tients allocated to active TMS, N=2 patients allocated to sham TMS). This influence was evident predominantly upon the outcome seen on the MADRS total score as a *continuous* measure, i.e., total score.

A MADRS total score less than 20 has been shown to correspond to mild depression (http://www.ids-qids.org, Table 4). In order to characterize the specific influence of the baseline imbalance observed on MADRS scores, a supplementary analysis was conducted of the overall intent-to-treat evaluable study population with this small subset of patients removed from the analysis. Based on the discussion above, as may be expected, the statistical consequence of this truncated analysis is the elimination of the statistical significance of the baseline imbalance in MADRS total score. This analysis also resulted in a statistically significant outcome for MADRS total score, which is consistent with the other two major efficacy outcome measures, namely the HAMD24 and the HAMD17.

It is important to note that the statistically significant outcome on the *a priori*-stated categorical outcome measures seen in the full dataset at the week 4 time point, namely the responder rates, for all three rating scales remains unaffected by the removal of these patients. These analyses are summarized in Table 14, and contrasted with the P-values obtained in the *a-priori* stated analyses at the primary outcome time point, namely week 4. The detailed supporting ANCOVA analyses and logistic regression output for these measures are included in Tables 3.40-3.48 of Appendix 3.

Table 13. Primary Outcome Measure (MADRS Total Score) Last-Observation Carried Forward Analysis

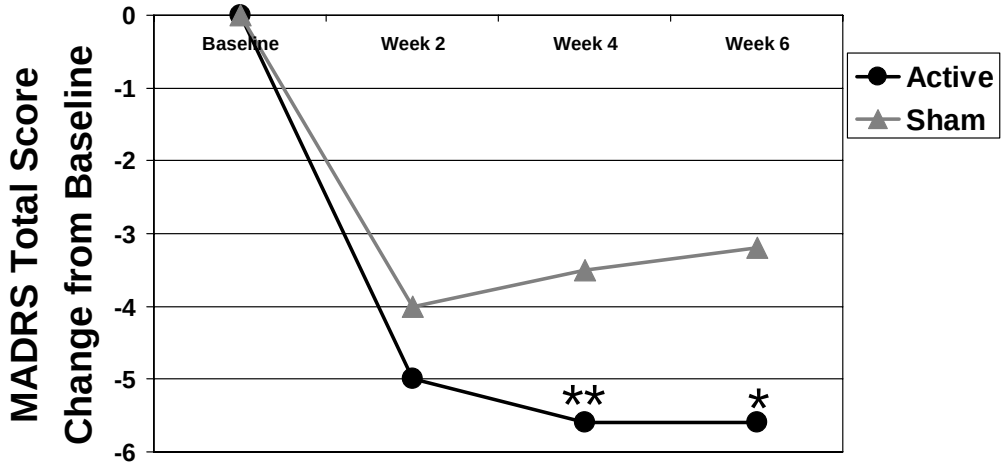
		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Total Score	N	146	146	146	146	155	155	155	155
	Mean	33.9	29.5	29.8	30	32.8	27.7	27	26.8
	LS Mean	33.7	29.3	29.5	29.8	32.4	27.3	26.5	26.4
	SD	5.69	8.55	10.11	10.77	5.99	8.83	11.06	12.78
	Median	34	30.5	32	33	33	28	28	30
	Min	19	3	0	0	14	0	0	0
	Max	46	46	48	48	50	47	51	51
	P-Value1					0.036			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-4.3	-4.1	-3.9		-5.1	-5.8	-6
	LS Mean		-4	-3.5	-3.2		-5	-5.6	-5.6
	SD		7.12	9.08	10.16		7.3	10.21	11.97
	Median		-3.5	-3	-1.5		-4	-4	-2
	Min		-25	-30	-44		-34	-35	-38
	Max		12	15	15		16	16	14
	P-Value2						0.191	0.057	0.058
	P-Value3		0	0	0		0	0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MADRS total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MADRS, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis



** P = .057 vs sham, * P = .058 vs sham, LOCF analysis, LS means

Figure 3. Primary Outcome Measure (MADRS Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

Table 14. Comparison of Week 4 Contrasts Between Active TMS and Sham TMS for Continuous and Categorical Measures on the MADRS, HAMD24 and HAMD17 with Patients Having Baseline MADRS < 20 Excluded

Variable	Week 4 Contrast P-Value	
	Full Dataset (N=301)	Truncated Dataset (N=295)
Primary outcome		
• MADRS Total Score	.057	.0384
Secondary outcomes		
• HAMD24 Total Score	.012	.0064
• HAMD17 Total Score	.006	.0038
• MADRS Responder Rate	.045	.0231
• HAMD24 Responder Rate	.030	.0157
• HAMD17 Responder Rate	.018	.0090

14.2. Secondary Efficacy Outcomes – Acute Phase

Secondary efficacy outcomes are listed in Table 15 in the sequential order of testing as outlined in the original protocol. All analyses are presented for the intent-to-treat, evaluable study population as defined above, and represent a last observation carried forward analysis (LOCF).

Table 15 lists all secondary outcome measures, the number of the data table within this section for each measure and the P value for the primary statistical contrast between the active TMS and sham TMS treatment groups on the specified measure at the time point indicated. All P values that fall below the statistical convention of P <.05 are highlighted. Key outcome measure results show that active TMS treatment was statistically significantly superior to sham TMS treatment for HAMD17 and 24 Item total scores at 4 and 6 weeks, for HAMD17, HAMD24 and MADRS categorical outcomes (>50% reduction in baseline score) scores at 4 and 6 weeks and for CGI-Severity scores at 4 and 6 weeks. Remission was shown to be significantly superior to sham treatment at 6 weeks as shown by HAMD24 and MADRS scores.

Tabular results for all secondary outcomes measures in their *a-priori*-defined order of priority testing are shown from Tables 16 through 41 below. Graphical outcome of the baseline to endpoint change on the HAMD24 and the HAMD17 are displayed in Figures 5 and 6, respectively. Graphical outcome of the responder and remission rates for the MADRS, HAMD24 and HAMD17 are displayed in Figures 7, 8, and 9.

Additional Tables 3.12 and 3.13 are included in Appendix 3 and summarize the individual item change scores for the MADRS and HAMD across the acute treatment phase of the study.

Table 15. Summary of A Priori-Defined Secondary Outcome Measures – Results

Secondary Efficacy Outcome Measures	Table No.	P value Week 4	P Value Week 6
HAMD 24 Total Score	16	.012	.015
HAMD17 Total Score	17	.006	.005
MADRS Responder Rate	18	.045	.007
HAMD 24 Responder Rate	19	.030	.042
HAMD17 Responder Rate	20	.018	.015
SF-36 Item Questionnaire v1: Physical Functioning	21	.299	.229
SF-36 Item Questionnaire v1: Role Physical Score	22	.361	.221
SF-36 Item Questionnaire v1: Bodily Pain Score	23	.520	.301
SF-36 Item Questionnaire v1: General Health Score	24	.049	.047
SF-36 Item Questionnaire v1: Vitality Score	25	.179	.081

Secondary Efficacy Outcome Measures	Table No.	P value Week 4	P Value Week 6
SF-36 Item Questionnaire v1: Social Functioning Score	26	.183	.386
SF-36 Item Questionnaire v1: Role Emotional Score	27	.105	.044
SF-36 Item Questionnaire v1: Mental Health Score	28	.006	.015
Q-LES-Q Total Score	29	.124	.035
MADRS Remission Rate	30	.633	.011
HAMD24 Remission Rate	31	.644	.012
HAMD17 Remission Rate	32	.705	.065
HAMD Anxiety/Somatization Factor Score	33	.025	.023
HAMD Core Depression Factor Score	34	.012	.008
HAMD Maier Factor Score	35	.003	.003
HAMD Gibbons Factor Score	36	.007	.006
HAMD Retardation Factor Score	37	.007	.003
HAMD Sleep Factor Score	38	.211	.109
IDS-SR Total Score	39	.058	.053
Clinician Global Impressions-Severity (CGI-S) Total Score	40	.009	.012
Patient Global Impression-Improvement (PGI-I) Total Score	41	.181	.107

Table 16. Secondary Outcome Measure (HAMD24 Total Score) Last-Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Total Score	N	146	146	146	146	155	155	155	155
	Mean	30.5	25.7	25.9	26	30.1	24	23.4	23.2
	LS Mean	30.2	25.5	25.7	25.9	29.9	23.9	23.1	23.1
	SD	4.85	7.28	8.81	9.39	5.04	7.57	8.93	10.62
	Median	30	26.5	27	28	29	25	24	26
	Min	21	6	2	0	18	0	0	0
	Max	45	42	44	44	46	41	42	49
	P-Value1					0.568			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-4.8	-4.6	-4.4		-6	-6.7	-6.9
	LS Mean		-4.3	-4.1	-3.8		-5.7	-6.5	-6.4
	SD		6.35	8.49	9.31		6.6	8.36	9.77
	Median		-4	-3	-2		-5	-6	-4
	Min		-24	-28	-33		-31	-28	-35
	Max		12	13	13		13	9	11
	P-Value2						0.051	0.012	0.015
	P-Value3		0	0	0		0	0	0

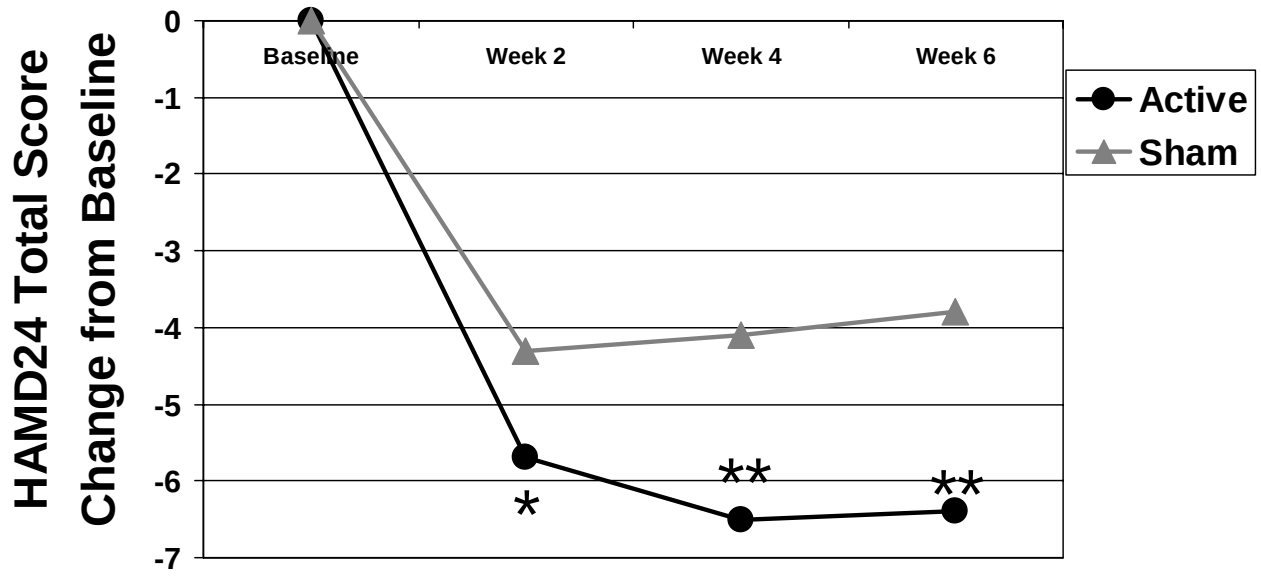
Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MADRS total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MADRS, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

HAMD24 Baseline to Endpoint Change Score



* P = .051 vs sham, ** P < .05 vs sham, LOCF analysis, LS means

Figure 4. Secondary Outcome Measure (HAMD24 Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

Table 17. Secondary Outcome Measure (HAMD17 Total Score) Last-Observation Carried Forward Analysis

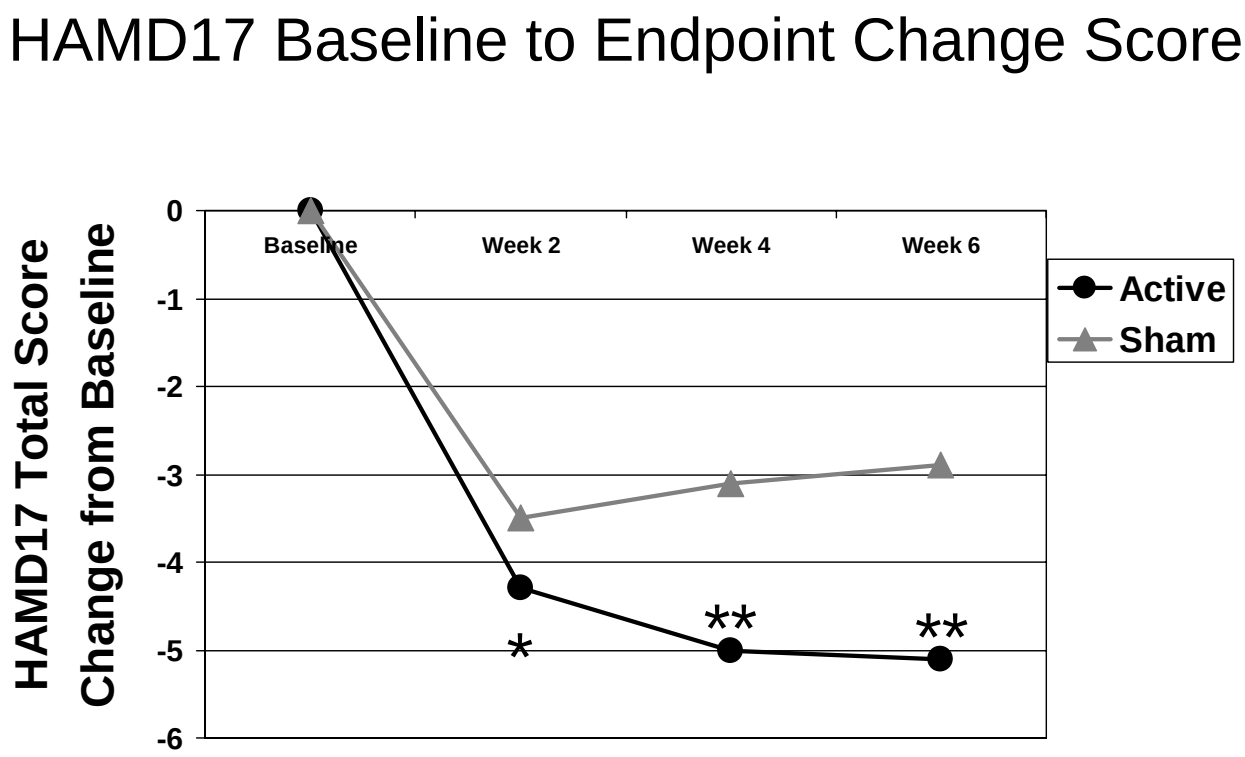
		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Total Score	N	146	146	146	146	155	155	155	155
	Mean	22.9	19	19.4	19.6	22.6	17.9	17.4	17.1
	LS Mean	22.8	18.9	19.3	19.5	22.5	17.8	17.3	17.1
	SD	3.54	5.28	6.51	6.95	3.3	5.42	6.49	7.67
	Median	22	20	21	21	22	18	18	19
	Min	16	6	2	0	13	0	0	0
	Max	34	30	34	34	32	30	31	34
	P-Value1					0.508			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-3.9	-3.4	-3.3		-4.7	-5.2	-5.4
	LS Mean		-3.5	-3.1	-2.9		-4.3	-5	-5.1
	SD		4.49	6.08	6.65		4.95	6.28	7.29
	Median		-3	-2	-1		-5	-4	-4
	Min		-14	-19	-22		-23	-22	-24
	Max		10	10	10		6	5	6
	P-Value2						0.098	0.006	0.005
	P-Value3		0	0	0		0	0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MADRS total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MADRS, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis



* P < .10 vs sham, ** P < .01 vs sham, LOCF analysis, LS means

Figure 5. Secondary Outcome Measure (HAMD17 Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

Table 18. Secondary Outcome Measure (MADRS Responders) Last-Observation Carried Forward Analysis

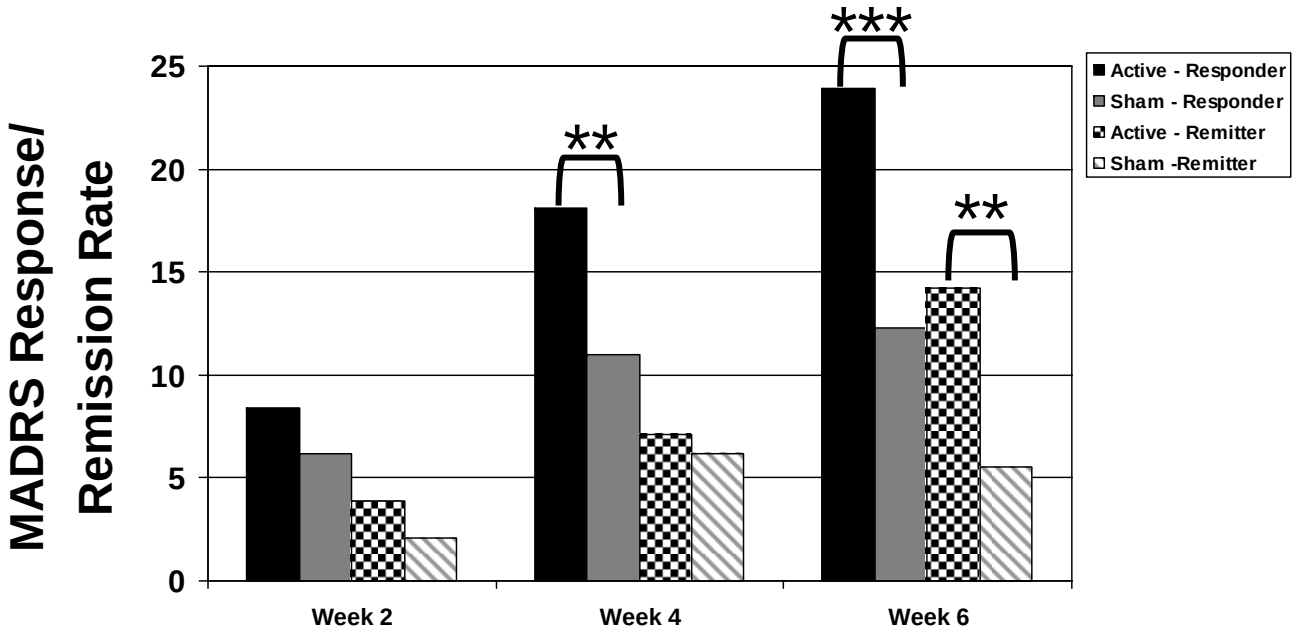
Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Responder	N (%)	9 (6.2)	13 (8.4)	
		Non-Responder	N (%)	137 (93.8)	142 (91.6)	0.384
	Week 4	Responder	N (%)	16 (11.0)	28 (18.1)	
		Non-Responder	N (%)	130 (89.0)	127 (81.9)	0.045
	Week 6	Responder	N (%)	18 (12.3)	37 (23.9)	
		Non-Responder	N (%)	128 (87.7)	118 (76.1)	0.007

Notes: Responder is defined as $\geq$ 50% reduction in total score compared to baseline assessment total score

P value calculated using a logistic regression model: Responder = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

MADRS Categorical Clinical Outcomes



** P < .05 vs sham, *** P < .01 vs sham, LOCF analysis

Figure 6. Secondary Outcome Measures (MADRS Responder and Remission Rates) Last-Observation Carried Forward Analysis

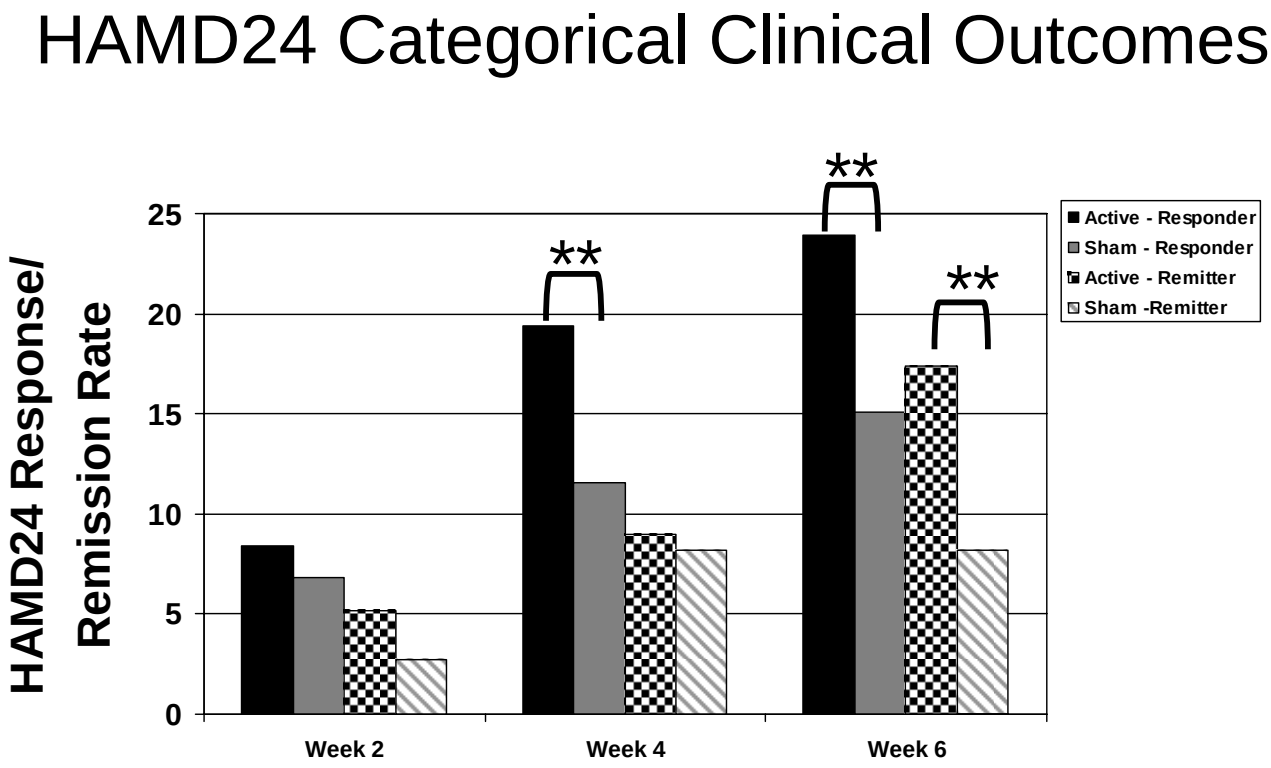
Table 19. Secondary Outcome Measure (HAMD24 Responders) Last-Observation Carried Forward Analysis

Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Responder	N (%)	10 (6.8)	13 (8.4)	
		Non-Responder	N (%)	136 (93.2)	142 (91.6)	0.601
	Week 4	Responder	N (%)	17 (11.6)	30 (19.4)	
		Non-Responder	N (%)	129 (88.4)	125 (80.6)	0.03
	Week 6	Responder	N (%)	22 (15.1)	37 (23.9)	
		Non-Responder	N (%)	124 (84.9)	118 (76.1)	0.042

Notes: Responder is defined as $\geq$ 50% reduction in total score compared to baseline assessment total score

P value calculated using a logistic regression model: Responder = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis



** P < .05 vs sham, LOCF analysis

Figure 7. Secondary Outcome Measures (HAMD24 Responder and Remission Rates) Last-Observation Carried Forward Analysis

Table 20. Secondary Outcome Measure (HAMD17 Responders) Last-Observation Carried Forward Analysis

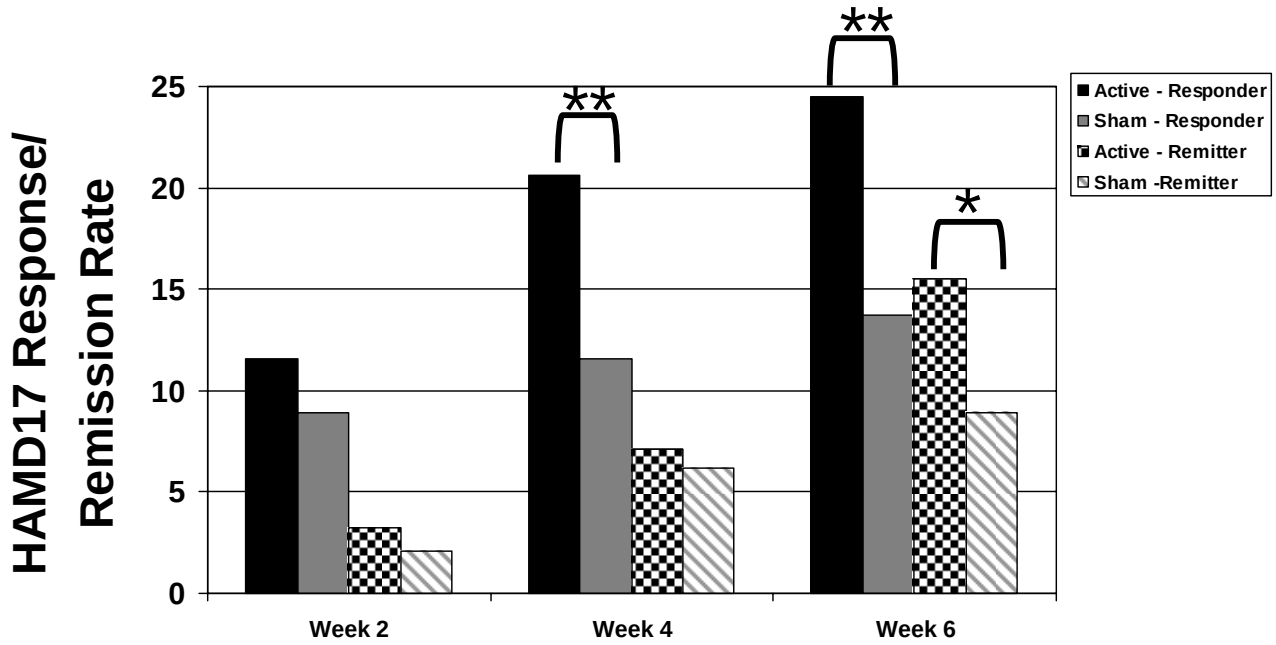
Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Responder	N (%)	13 (8.9)	18 (11.6)	
		Non-Responder	N (%)	133 (91.1)	137 (88.4)	0.451
	Week 4	Responder	N (%)	17 (11.6)	32 (20.6)	
		Non-Responder	N (%)	129 (88.4)	123 (79.4)	0.018
	Week 6	Responder	N (%)	20 (13.7)	38 (24.5)	
		Non-Responder	N (%)	126 (86.3)	117 (75.5)	0.015

Notes: Responder is defined as $\geq$ 50% reduction in total score compared to baseline assessment total score

P value calculated using a logistic regression model: Responder = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

HAMD17 Categorical Clinical Outcomes



* P < .10 vs sham, ** P < .05 vs sham, LOCF analysis

Figure 8. Secondary Outcome Measures (HAMD17 Responder and Remission Rates) Last-Observation Carried Forward Analysis

Table 21. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Physical Functioning Score

		ShamTMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Physical Functioning	N	145	145	145	155	155	155
	Mean	43.2	44.1	44.4	45.9	46.8	47.3
	LS Mean	43.3	44.4	44.6	46.1	47.2	47.5
	SD	11.29	10.36	10.46	10.47	10	9.63
	Median	44.6	46.7	46.7	50.9	48.8	50.9
	Min	15.2	17.3	17.3	17.3	17.3	17.3
	Max	57.1	57.1	57.1	57.1	57.1	57.1
	P-Value1				0.024		
Change from Baseline	N		145	145		155	155
	Mean		0.9	1.2		0.9	1.4
	LS Mean		0.4	0.3		1.3	1.2
	SD		7.39	7.16		7.23	7.12
	Median		0	0		0	0
	Min		-31.5	-21		-23.1	-21
	Max		23.1	27.3		27.3	25.2
	P-Value2					0.299	0.229
	P-Value3		0.146	0.043		0.11	0.019

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 22. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Role Physical Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Role Physical	N	145	145	145	155	155	155
	Mean	38.8	39.9	40	40.3	41.5	42.1
	LS Mean	39.2	39.2	39.7	40.8	41	41.9
	SD	11.64	12.16	12.31	12.53	12.07	12.29
	Median	35	35	35	35	42.1	42.1
	Min	28	28	28	28	28	28
	Max	56.2	56.2	56.2	56.2	56.2	56.2
	P-Value1				0.252		
Change from Baseline	N		145	145		155	155
	Mean		1	1.2		1.2	1.9
	LS Mean		-0.2	0.2		1	1.8
	SD		13.9	13.31		13.16	13.51
	Median		0	0		0	0
	Min		-28.3	-28.3		-28.3	-28.3
	Max		28.3	28.3		28.3	28.3
	P-Value2					0.361	0.221
	P-Value3		0.376	0.291		0.264	0.087

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 23. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Bodily Pain Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Bodily Pain	N	146	146	146	155	155	155
	Mean	43.8	44.4	44.8	43.5	44.7	45.5
	LS Mean	44	44.8	45	43.6	45.1	45.6
	SD	9.06	9.18	8.85	9.47	9.31	9.16
	Median	45.6	45.6	45.6	41.3	45.6	45.6
	Min	19.9	19.9	19.9	19.9	19.9	19.9
	Max	58.5	58.5	58.5	58.5	58.5	58.5
	P-Value1				0.58		
Change from Baseline	N		146	146		155	155
	Mean		0.7	1		1.2	2
	LS Mean		1	0.6		1.4	1.4
	SD		7.37	7.79		7.4	7.75
	Median		0	0		0	0
	Min		-21.4	-21.4		-17.1	-17.1
	Max		38.5	38.5		30	34.3
	P-Value2					0.52	0.301
	P-Value3		0.271	0.124		0.038	0.002

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center
P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment
P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST
All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 24. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: General Health Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
General Health	N	146	146	146	155	155	155
	Mean	40.9	40.6	40.9	41.1	42.4	42.6
	LS Mean	41.2	41.3	41.4	41.4	43.1	43.2
	SD	9.46	10.45	10.12	9.78	9.68	10.12
	Median	40.6	40.6	40.6	40.6	42.9	42.9
	Min	21.9	19.5	19.5	19.5	21.9	21.9
	Max	64	64	64	61.7	64	64
	P-Value1				0.686		
Change from Baseline	N		146	146		155	155
	Mean		-0.3	-0.1		1.3	1.5
	LS Mean		-0.3	-0.2		1.3	1.5
	SD		7.04	6.97		7.8	8.31
	Median		0	0		0	0
	Min		-16.4	-16.4		-18.7	-18.7
	Max		23.4	28.1		28.1	30.4
	P-Value2					0.049	0.047
	P-Value3		0.583	0.89		0.042	0.027

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 25. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Vitality Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Vitality	N	145	145	145	155	155	155
	Mean	29.9	32.6	33	31.8	35.1	36.2
	LS Mean	29.9	32.8	33	31.8	35.3	36.2
	SD	5.88	8.53	9.41	6.84	9.41	11.22
	Median	27.8	30.1	30.1	30.1	32.5	32.5
	Min	23	23	23	23	23	23
	Max	46.7	60.9	65.6	63.3	65.6	68
	P-Value1				0.011		
Change from Baseline	N		145	145		155	155
	Mean		2.7	3.1		3.3	4.4
	LS Mean		2.1	2		3.3	4
	SD		7.86	8.58		8.18	10.71
	Median		0	0		2.4	0
	Min		-11.8	-11.8		-18.9	-18.9
	Max		37.9	42.6		28.4	42.6
	P-Value2					0.179	0.081
	P-Value3		0	0		0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 26. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Social Functioning Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Social Functioning	N	146	146	146	155	155	155
	Mean	26.2	28.9	30.3	27.5	31	31.9
	LS Mean	26.4	29.9	31	27.7	32	32.6
	SD	9.82	10.92	11.6	8.94	10.78	11.51
	Median	24.6	30.2	30.2	30.2	30.2	30.2
	Min	8.1	13.7	13.7	13.7	13.7	13.7
	Max	57.1	57.1	57.1	51.9	57.1	57.1
	P-Value1				0.034		
Change from Baseline	N		146	146		155	155
	Mean		2.7	4.1		3.5	4.4
	LS Mean		1.8	2.7		3.2	3.7
	SD		9.83	10.36		10.53	11.76
	Median		0	0		0	0
	Min		-21.7	-21.7		-21.7	-21.7
	Max		32.6	32.6		37.8	43.4
	P-Value2					0.183	0.386
	P-Value3		0.001	0		0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 27. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Role Emotional Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Role Emotional	N	146	146	146	155	155	155
	Mean	26.5	28.3	28.9	25.9	29.6	30.9
	LS Mean	26.7	28.7	29.5	26.1	30.1	31.5
	SD	7.13	9.01	9.53	4.89	9.69	10.66
	Median	23.7	23.7	23.7	23.7	23.7	23.7
	Min	23.7	23.7	23.7	23.7	23.7	23.7
	Max	55.3	55.3	55.3	44.9	55.3	55.3
	P-Value1				0.381		
Change from Baseline	N		146	146		155	155
	Mean		1.8	2.4		3.7	5
	LS Mean		1.9	2.4		3.6	4.7
	SD		8.07	8.74		10.45	11.54
	Median		0	0		0	0
	Min		-21.2	-31.6		-21.2	-21.2
	Max		31.6	31.6		31.6	31.6
	P-Value2					0.105	0.044
	P-Value3		0.008	0.001		0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 28. Secondary Outcome Measure Medical Outcomes Study Short Form-36 Item Questionnaire v1: Mental Health Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week4**	**Week6**	**Baseline**	**Week4**	**Week6**
Mental Health	N	145	145	145	155	155	155
	Mean	24.6	26	27.1	25.1	29.3	30.5
	LS Mean	25.2	26.3	27.2	25.8	29.8	30.8
	SD	7.8	10.47	11.68	8.74	11.28	13.03
	Median	25.5	25.5	25.5	25.5	27.7	27.7
	Min	7.3	7.3	7.3	9.6	7.3	7.3
	Max	48.2	61.8	64.1	55	57.3	59.5
	P-Value1				0.218		
Change from Baseline	N		145	145		155	155
	Mean		1.4	2.5		4.2	5.4
	LS Mean		0.6	1.4		3.7	4.5
	SD		10.62	11.22		9.56	11.64
	Median		0	0		2.3	2.3
	Min		-20.5	-20.5		-15.9	-22.7
	Max		40.9	43.2		34.1	43.2
	P-Value2					0.006	0.015
	P-Value3		0.104	0.008		0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, MOS SF-36 component score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline MOS SF-36 component score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 29. Secondary Outcome Measure Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q) Total Score

		Sham TMS (146)			Active TMS (155)		
Values	**Statistics**	**Baseline**	**Week 4**	**Week 6**	**Baseline**	**Week 4**	**Week 6**
Total Score	N	146	146	146	155	155	155
	Mean	36.5	39.2	39.3	37.6	41.3	42.4
	LS Mean	36.5	39	39	37.8	41.4	42.2
	SD	7.87	9.78	10.15	8.23	10.32	12.28
	Median	37	38	38	38	41	42
	Min	17	19	19	15	18	18
	Max	56	69	73	64	69	73
	P-Value1				0.173		
Change from Baseline	N		146	146		155	155
	Mean		2.7	2.8		3.7	4.7
	LS Mean		2	1.3		3.5	3.8
	SD		9.24	9.85		9.19	11.58
	Median		2	2		2	2
	Min		-23	-23		-17	-17
	Max		36	42		32	44
	P-Value2					0.124	0.035
	P-Value3		0.001	0.001		0	0

Notes: P value1 represents the between treatment group comparison calculated using ANOVA model, Q-LES-Q total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline Q-LES-Q total score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 30. Secondary Outcome Measure (MADRS Remission Rate) Last Observation Carried Forward Analysis

Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Remission	N (%)	3 (2.1)	6 (3.9)	
		Non-Remission	N (%)	143 (97.9)	149 (96.1)	0.311
	Week 4	Remission	N (%)	9 (6.2)	11 (7.1)	
		Non-Remission	N (%)	137 (93.8)	144 (92.9)	0.633
	Week 6	Remission	N (%)	8 (5.5)	22 (14.2)	
		Non-Remission	N (%)	138 (94.5)	133 (85.8)	0.011

Notes: Remission is defined as a MADRS total score < 10

P value calculated using a logistic regression model: Remission = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 31. Secondary Outcome Measure (HAMD24 Remission Rate) Last Observation Carried Forward Analysis

Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Remission	N (%)	4 (2.7)	8 (5.2)	
		Non-Remission	N (%)	142 (97.3)	147 (94.8)	0.257
	Week 4	Remission	N (%)	12 (8.2)	14 (9.0)	
		Non-Remission	N (%)	134 (91.8)	141 (91.0)	0.644
	Week 6	Remission	N (%)	12 (8.2)	27 (17.4)	
		Non-Remission	N (%)	134 (91.8)	128 (82.6)	0.012

Notes: Remission is defined as a HAMD24 total score < 11

P value calculated using a logistic regression model: Remission = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 32. Secondary Outcome Measure (HAMD17 Remission Rate) Last Observation Carried Forward Analysis

Phase	Time Point	Response	Statistics	Sham TMS (146)	Active TMS (155)	P-value
Acute	Week 2	Remission	N (%)	3 (2.1)	5 (3.2)	
		Non-Remission	N (%)	143 (97.9)	150 (96.8)	0.418
	Week 4	Remission	N (%)	9 (6.2)	11 (7.1)	
		Non-Remission	N (%)	137 (93.8)	144 (92.9)	0.705
	Week 6	Remission	N (%)	13 (8.9)	24 (15.5)	
		Non-Remission	N (%)	133 (91.1)	131 (84.5)	0.065

Notes: Remission is defined as a HAMD17 total score < 8

P value calculated using a logistic regression model: Remission = ATHF group, center, treatment

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 33. Secondary Outcome Measure HAMD Anxiety/Somatization Factor Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week 2**	**Week 4**	**Week 6**	**Baseline**	**Week 2**	**Week 4**	**Week 6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	7.4	6.1	6.2	6.3	7.2	5.7	5.6	5.5
	LS Mean	7.3	6	6.2	6.2	7.1	5.7	5.5	5.5
	SD	1.97	2.13	2.53	2.56	1.92	2.27	2.52	2.78
	Median	7	6	6	6	7	6	5	5
	Min	2	1	0	0	3	0	0	0
	Max	14	11	12	12	13	12	13	13
	P-Value1					0.456			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-1.3	-1.1	-1.1		-1.4	-1.6	-1.7
	LS Mean		-1.2	-1	-1		-1.4	-1.6	-1.7
	SD		2.07	2.34	2.53		2.21	2.52	2.88
	Median		-1	-1	-1		-1	-1	-1
	Min		-7	-8	-9		-7	-8	-10
	Max		7	6	6		4	4	4
	P-Value2						0.3	0.025	0.023
	P-Value3		0	0	0		0	0	0

Notes: HAMD Anxiety/Somatization Factor = HAMD Items 10, 11, 12, 13, 15 and 17

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 34. Secondary Outcome Measure HAMD Core Depression Factor Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week 2**	**Week 4**	**Week 6**	**Baseline**	**Week 2**	**Week 4**	**Week 6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	9.4	8	8.3	8.4	9.3	7.6	7.4	7.3
	LS Mean	9.4	8	8.3	8.4	9.2	7.5	7.3	7.2
	SD	1.84	2.59	3.11	3.47	2.13	2.8	3.3	3.77
	Median	10	8	9	9	9	8	8	8
	Min	5	0	0	0	4	0	0	0
	Max	14	14	16	16	15	13	14	14
	P-Value1					0.437			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-1.4	-1.1	-1		-1.7	-1.9	-2
	LS Mean		-1.2	-1	-0.8		-1.5	-1.9	-1.8
	SD		2.37	2.95	3.3		2.4	3.04	3.52
	Median		-1	-1	-0.5		-1	-1	-1
	Min		-8	-9	-11		-11	-10	-11
	Max		4	6	6		6	6	6
	P-Value2						0.19	0.012	0.008
	P-Value3		0	0	0		0	0	0

Notes: HAMD Core Depression Factor = HAMD Items 1, 2, 3, 7 and 8

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 35. Secondary Outcome Measure HAMD Maier Factor Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week 2**	**Week 4**	**Week 6**	**Baseline**	**Week 2**	**Week 4**	**Week 6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	11.3	9.5	9.9	9.9	11.2	9.1	8.8	8.6
	LS Mean	11.3	9.5	9.8	9.9	11.2	9.1	8.7	8.6
	SD	1.91	2.88	3.31	3.69	1.9	2.96	3.59	4.13
	Median	11	10	10	11	11	10	9	10
	Min	6	1	0	0	6	0	0	0
	Max	17	17	17	17	16	16	16	16
	P-Value1					0.831			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-1.8	-1.4	-1.4		-2.1	-2.5	-2.6
	LS Mean		-1.5	-1.4	-1.1		-1.8	-2.5	-2.4
	SD		2.67	3.23	3.7		2.66	3.41	3.9
	Median		-1	-1	-1		-2	-2	-2
	Min		-9	-10	-13		-12	-11	-12
	Max		7	7	7		6	6	6
	P-Value2						0.276	0.003	0.003
	P-Value3		0	0	0		0	0	0

Notes: HAMD Maier Factor = HAMD Items 1, 2, 7, 8, 9 and 10

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 36. Secondary Outcome Measure HAMD Gibbons Factor Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	14.9	12.7	13.1	13.1	14.6	12	11.7	11.4
	LS Mean	14.8	12.6	13	13	14.5	11.9	11.5	11.4
	SD	2.33	3.68	4.25	4.68	2.51	3.75	4.57	5.27
	Median	15	13	14	14	14	12	12	13
	Min	10	2	1	0	8	0	0	0
	Max	22	22	22	22	23	21	21	23
	P-Value1					0.275			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-2.2	-1.8	-1.8		-2.6	-2.9	-3.2
	LS Mean		-2	-1.8	-1.6		-2.5	-3	-3
	SD		3.3	4.04	4.6		3.26	4.25	4.85
	Median		-2	-1	-1		-2	-2	-2
	Min		-11	-12	-17		-16	-15	-15
	Max		8	8	8		5	8	5
	P-Value2						0.152	0.007	0.006
	P-Value3		0	0	0		0	0	0

Notes: HAMD Gibbons Factor = HAMD Items 1, 2, 3, 7, 9, 10, 11 and 14

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 37. Secondary Outcome Measure HAMD Retardation Factor Score Last Observation Carried Forward Analysis

		ShamTMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	8	6.8	7	7.1	7.9	6.4	6.2	6.1
	LS Mean	8	6.8	7	7.1	7.9	6.4	6.2	6.1
	SD	1.63	2.16	2.59	2.81	1.68	2.34	2.63	3.08
	Median	8	7	7	7	8	6	6	7
	Min	4	0	0	0	3	0	0	0
	Max	12	11	12	12	12	11	11	12
	P-Value1					0.542			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-1.1	-0.9	-0.9		-1.5	-1.6	-1.7
	LS Mean		-0.9	-0.9	-0.7		-1.3	-1.6	-1.6
	SD		1.86	2.43	2.65		2	2.53	3.01
	Median		-1	-1	0		-1	-1	-1
	Min		-6	-7	-9		-8	-8	-9
	Max		2	4	4		3	4	4
	P-Value2						0.057	0.007	0.003
	P-Value3		0	0	0		0	0	0

Notes: HAMD Retardation Factor = HAMD Items 1, 7, 8 and 14

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 38. Secondary Outcome Measure HAMD Sleep Factor Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Factor Score	N	146	146	146	146	155	155	155	155
	Mean	3.7	2.8	2.9	2.9	3.8	2.7	2.7	2.6
	LS Mean	3.7	2.8	2.9	2.9	3.8	2.7	2.7	2.6
	SD	1.71	1.81	1.89	1.86	1.68	1.83	1.82	1.83
	Median	4	3	3	3	4	2	2	2
	Min	0	0	0	0	0	0	0	0
	Max	6	6	6	6	6	6	6	6
	P-Value1					0.617			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-0.9	-0.8	-0.8		-1.1	-1.1	-1.2
	LS Mean		-0.8	-0.6	-0.8		-1	-0.9	-1.1
	SD		1.59	1.71	1.69		1.84	1.99	2.06
	Median		-1	-1	-1		-1	-1	-1
	Min		-5	-6	-6		-6	-6	-6
	Max		3	4	3		3	6	6
	P-Value2						0.388	0.211	0.109
	P-Value3		0	0	0		0	0	0

Notes: HAMD Sleep Factor = HAMD Items 4, 5 and 6

P value1 represents the between treatment group comparison calculated using ANOVA model, HAMD factor score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline HAMD factor score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 39. Secondary Outcome Measure Inventory of Depressive Symptoms-Self Report (IDS-SR) Total Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week2**	**Week4**	**Week6**	**Baseline**	**Week2**	**Week4**	**Week6**
Total Score	N	146	146	146	146	155	155	155	155
	Mean	43.4	37.9	37.5	37.3	42	35.3	34.2	33.4
	LS Mean	42.8	37.4	37.1	37	41.3	34.9	33.6	33
	SD	9.89	11.82	13.35	14.23	9.4	11.25	13.3	15.37
	Median	43	38	38	39	42	36	34	35
	Min	18	2	5	1	21	8	5	1
	Max	69	69	64	64	64	66	74	74
	P-Value1					0.197			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-5.5	-5.8	-6.1		-6.7	-7.9	-8.6
	LS Mean		-4.4	-5.2	-4.7		-5.9	-7.7	-7.7
	SD		9.5	11.84	13.26		9.29	11.88	14.5
	Median		-4	-4	-4		-5	-6	-4
	Min		-44	-47	-56		-33	-42	-54
	Max		16	28	28		20	14	19
	P-Value2						0.142	0.058	0.053
	P-Value3		0	0	0		0	0	0

Notes: IDS-SR total score = Sum of 30 items

P value1 represents the between treatment group comparison calculated using ANOVA model, IDS-SR total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline IDS SR total score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 40. Secondary Outcome Measure Clinician Global Impressions-Severity (CGI-S) Total Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week 2**	**Week 4**	**Week 6**	**Base-line**	**Week 2**	**Week 4**	**Week 6**
CGI-S Score	N	146	146	146	146	155	155	155	155
	Mean	4.7	4.5	4.4	4.4	4.7	4.3	4.1	4
	LS Mean	4.7	4.5	4.4	4.4	4.7	4.3	4.1	4
	SD	0.72	0.81	1.09	1.23	0.62	0.91	1.1	1.41
	Median	5	4	4.5	4	5	4	4	4
	Min	3	3	1	1	4	1	1	1
	Max	6	6	7	7	6	6	6	6
	P-Value1					0.594			
Change from Baseline	N		146	146	146		155	155	155
	Mean		-0.2	-0.3	-0.4		-0.4	-0.6	-0.7
	LS Mean		-0.2	-0.2	-0.2		-0.4	-0.6	-0.6
	SD		0.81	1.07	1.21		0.77	1.05	1.34
	Median		0	0	0		0	0	0
	Min		-2	-5	-5		-3	-4	-4
	Max		2	2	2		2	2	2
	P-Value2						0.047	0.009	0.012
	P-Value3		0	0.001	0		0	0	0

Notes: CGI-S total score = integer score on 7-point scale

P value1 represents the between treatment group comparison calculated using ANOVA model, CGI-S total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline CGI-S total score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Table 41. Secondary Outcome Measure Patient Global Impressions-Improvement (PGI-I) Total Score Last Observation Carried Forward Analysis

		Sham TMS (146)				Active TMS (155)			
Values	**Statistics**	**Baseline**	**Week 2**	**Week 4**	**Week 6**	**Baseline**	**Week 2**	**Week 4**	**Week 6**
PGI-I Score	N	141	146	146	146	155	155	155	155
	Mean	4.4	3.8	3.9	4	4.3	3.7	3.7	3.7
	LS Mean	4.4	3.8	3.9	4	4.3	3.7	3.7	3.7
	SD	0.99	1.13	1.43	1.48	0.86	1.09	1.37	1.56
	Median	4	4	4	4	4	4	4	4
	Min	1	1	1	1	2	1	1	1
	Max	7	7	7	7	7	7	7	7
	P-Value1					0.695			
Change from Baseline	N		141	141	141		155	155	155
	Mean		-0.6	-0.4	-0.4		-0.7	-0.6	-0.6
	LS Mean		-0.5	-0.3	-0.2		-0.6	-0.6	-0.5
	SD		1.39	1.72	1.8		1.3	1.61	1.72
	Median		0	0	0		-1	-1	-1
	Min		-5	-5	-5		-4	-5	-5
	Max		3	4	4		4	3	3
	P-Value2						0.527	0.181	0.107
	P-Value3		0	0.002	0.011		0	0	0

Notes: PGI-I total score = integer score on 7-point scale

P value1 represents the between treatment group comparison calculated using ANOVA model, PGI-I total score = treatment center

P value2 represents the change from baseline between treatment group comparison using ANCOVA model, change from baseline = baseline PGI-I total score, ATHF group, center, and treatment

P value3 represents the change from baseline within treatment group comparison calculated using a paired T-TEST

All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

14.3. Overall Efficacy Conclusions Based on the *A Priori*-Defined Efficacy Outcome Measures

Clinician-rated efficacy outcomes for study 44-01101 are summarized in Tables 42 and 43 and patient-rate efficacy outcomes are summarized in Table 44.

Table 42. Clinician-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS

Variable Name	Week 2	Week 4	Week 6
MADRS Total Score	--	.057	.058
MADRS (baseline adjustment)		.038	.051
HAMD24 Total Score	.051	.012	.015
HAMD17 Total Score	.098	.006	.005
Response Rate (>50% reduction from baseline)			
• MADRS	--	.045	.007
• HAMD24	--	.030	.042
• HAMD17	--	.018	.015
Remission Rate			
• MADRS (Total score <10)	--	--	.011
• HAMD24 (Total score <11)	--	--	.012
• HAMD17 (Total score <8)	--	--	.065
CGI-S Total Score	.047	.009	.012

-- = p>.10

Table 43. Clinician-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS

Variable Name	Week 2	Week 4	Week 6
HAMD Factor Scores			
• Anxiety/Somatization Factor	--	.025	.023
• Core Depression Factor	--	.012	.008
• Maier Factor	--	.003	.003
• Gibbons Factor	--	.007	.006
• Retardation Factor	.057	.007	.003
• Sleep Factor	--	--	--

-- = p>.10

Table 44. Patient-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS

Variable Name	Week 2	Week 4	Week 6
MOS Short Form 36-Item			
• Physical Functioning	N/A	--	--
• Role-Physical	N/A	--	--
• Bodily Pain	N/A	--	--
• General Health	N/A	.049	.047
• Vitality	N/A	--	.081
• Social Functioning	N/A	--	--
• Role Emotional	N/A	--	.044
• Mental Health	N/A	.006	.015
Q-LES-Q	N/A	--	.035
IDS-Self Report	--	.058	.053
PGI-Improvement Total Score	--	--	--

-- = p>.10; N/A = scale not obtained at that time point

14.3.1. Primary Outcome Measure:

MADRS Total Score

- *After 4 weeks*, active treatment using the Neuronetics TMS System showed a *strong statistical trend for superiority compared to sham treatment* on the MADRS total score (p=.057)
- *See section 13.1 for a discussion of the baseline imbalance observed for this outcome measure. Recomputed analysis using a baseline cut-off for MADRS total score of 20 showed that active TMS treatment using the Neuronetics TMS systems was statistically significantly superior compared to sham treatment on the MADRS total score at 4 weeks (p=.038).*

14.3.2. Secondary Outcome Measures:

HAMD24, HAMD17 (Weeks 4 & 6) and MADRS Total Score (Week 6)

- *After 2 weeks*, active treatment with the Neuronetics TMS Systems showed a *strong statistical trend for superiority compared to sham treatment* on the HAMD24 total score (p=.051) and the HAMD17 total score (p=.098)
- *After 4 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the total score on the HAMD24 (p=.012) and HAMD17 (p=.006)
- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the total score on the HAMD24 (p=.015) and HAMD17 (p=.005) and continued to show a *strong statistical trend for superiority compared to sham treatment*

on the MADRS total score (p=.058) and when adjusted for baseline imbalance (p=0.051).

HAMD24, HAMD17, and MADRS Response Rate (Weeks 4 and 6)

- *After 4 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by categorical response rate (≥50% reduction in score from baseline) on all measures, the MADRS, (p=.045), the HAMD24 (p=.030), the HAMD17 (p=.017)
- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by categorical response rate on all measures, the MADRS (p=.007), the HAMD24 (p=.042), and the HAMD17 (p=.015)

Functional Status Outcome (MOS SF-36 and Q-LES-Q) (Weeks 4 and 6)

- *After 4 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the SF-36 General Health (p=.049), and Mental Health (p=.006) subscales
- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the SF-36 General Health (p=.047), Role-Emotional (p=.044) and Mental Health (p=.015) subscales, and showed a *strong statistical trend for superiority compared to sham treatment* on the Vitality (p=.081) subscale
- *After 6 weeks*, active treatment with the Neuronetics TMS System *was statistically significantly superior to sham treatment* as measured by the Quality of Life Enjoyment and Satisfaction Questionnaire (p=.035)

HAMD24, HAMD17, and MADRS Remission Rate (Weeks 4 and 6)

- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham* as measured by categorical remission rate on the MADRS (p=.011), and the HAMD24 (p=.012), and showed a strong statistical trend for superiority on the HAMD17 (p=.065)

HAMD Factor Scores (Weeks 4 and 6)

- *After 2 weeks*, active treatment with the Neuronetics TMS System showed a *strong statistical trend for superiority compared to sham treatment as* measured by the HAMD Retardation Factor (p=.057)
- *After 4 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by all but one of the HAMD Factor Scores, including the Anxiety/Somatization Factor (p=.025), Core Depression Factor (p=.012), Maier Factor (p=.003), Gibbons Factor (p=.007), and Retardation Factor (p=.007)

- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by all but one of the HAMD Factor Scores, including the Anxiety/Somatization Factor (p=.023), Core Depression Factor (p=.008), Maier Factor (p=.003), Gibbons Factor (p=.006), and Retardation Factor (p=.003)

Other Efficacy Measures (IDS-SR, CGI-Severity, PGI-Improvement) (Weeks 4 and 6)

- *After 2 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by all the CGI-Severity score (p=.047)
- *After 4 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the CGI-Severity score (p=.009) and showed a *strong statistical trend for superiority compared to sham treatment* as measured by the IDS-SR Total Score (p=.058)
- *After 6 weeks*, active treatment with the Neuronetics TMS System was *statistically significantly superior to sham treatment* as measured by the CGI-Severity score (p=.012) and showed a *strong statistical trend for superiority compared to sham treatment* as measured by the IDS-SR Total Score (p=.053)

14.3.3. Overall Efficacy Conclusions

These results establish that TMS therapy delivered by the Neuronetics Model 2100 TMS System is *statistically significantly superior to sham treatment at 4 and 6 weeks for key physician-rated depression measures* (HAMD 17, HAMD24 and CGI-S) and for the MADRS total score when corrected for baseline imbalance (see section 14.1).

Additionally, all three key depression measures (HAMD17, HAMD24 and MADRS) were *statistically superior to sham treatment at 4 and 6 weeks for the clinically-meaningful categorical outcomes for >50% reduction in baseline score*.

This data indicate that the TMS therapy delivered by the Neuronetics Model 2100 TMS System is *effective in the treatment of major depressive disorder*.

15.0 SUBSET ANALYSES

An exploratory descriptive analysis that was not previously stipulated in the protocol-defined statistical plan was conducted on specific demographic and illness severity measures. These analyses were intended to determine the if the study results could be generalized across the broad population of patients with major depression regardless of fixed population characteristics (e.g., gender and age), and whether the observed treatment effect when analyzed by baseline severity is also broadly generalized within the overall treatment population.

Specifically, continuous outcome on the total score for the 3 principal disease-specific efficacy instruments, the MADRS, the 24-item HAMD, and the 17-item HAMD was examined for 3 specific patient subsets: gender, age (< 55 or > 55 years), and baseline HAMD17 severity (using a median split of the observed baseline score = 22).

These results are shown in Tables 45, 46, and 47 and summarized in Section 14.1. No inferential statistical comparisons were performed on these subsets since they are presented as exploratory analyses.

Table 45. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for MADRS Total Score at Week 4 (Primary Efficacy Outcome Time Point) - Page 1 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Gender						
Female	Total Score	N	74	74	86	86
		Mean	34.1	29.7	32.5	26.5
		SD	5.43	10.31	5.06	11.02
		Median	35	31	32	27
		Min	19	0	22	3
		Max	46	48	44	47
	Change from Baseline	N		74		86
		Mean		-4.4		-6
		SD		8.43		9.5
		Median		-3.5		-4
		Min		-29		-29
		Max		10		10
		Treatment Effect				-1.1
Male	Total Score	N	72	72	69	69
		Mean	33.7	29.8	33.1	27.6
		SD	5.98	9.98	7	11.16
		Median	33	33	34	30
		Min	19	7	14	0
		Max	46	47	50	51
	Change from Baseline	N		72		69
		Mean		-3.8		-5.5
		SD		9.75		11.1
		Median		-2		-4
		Min		-30		-35
		Max		15		16
		Treatment Effect				-1.3

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 45. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for MADRS Total Score at Week 4 (Primary Efficacy Outcome Time Point)- Page 2 of 3

Subgroup	Values	Statistics	Sham TMS (146) Baseline	Sham TMS (146) Week4	Active TMS (155) Baseline	Active TMS (155) Week4
Age Group						
Age < 55 Years	Total Score	N	98	98	111	111
		Mean	34.4	30.7	33.1	27.3
		SD	5.68	9.94	6.14	11.38
		Median	34	34	33	30
		Min	19	0	14	0
		Max	46	48	50	51
	Change from Baseline	N		98		111
		Mean		-3.7		-5.8
		SD		9		10.11
		Median		-2.5		-3
		Min		-28		-35
		Max		15		16
		Treatment Effect				-1.9
				0		
Age >= 55 - 70 Years	Total Score	N	48	48	44	44
		Mean	32.9	27.9	31.9	26.1
		SD	5.66	10.33	5.56	10.27
		Median	32.5	29	31	26
		Min	19	2	15	5
		Max	46	47	43	47
	Change from Baseline	N		48		44
		Mean		-5		-5.8
		SD		9.27		10.58
		Median		-3.5		-6
		Min		-30		-31
		Max		9		12
		Treatment Effect				-0.1

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 45. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for MADRS Total Score at Week 4 (Primary Efficacy Outcome Time Point)- Page 3 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Baseline Total Symptom Severity						
HAMD17: Total Score < Median	Total Score	N	59	59	65	65
		Mean	29.9	27	29.8	24.1
		SD	4.24	9.72	5.46	10.4
		Median	31	28	30	25
		Min	19	0	14	0
		Max	38	45	44	47
	Change from Baseline	N		59		65
		Mean		-2.9		-5.7
		SD		8.47		9.96
		Median		-2		-5
		Min		-27		-31
		Max		15		16
		Treatment Effect				-2.2
HAMD17: Total Score >= Median	Total Score	N	87	87	90	90
		Mean	36.6	31.6	34.9	29.1
		SD	4.93	10	5.45	11.11
		Median	37	34	35	30.5
		Min	24	8	22	3
		Max	46	48	50	51
	Change from Baseline	N		87		90
		Mean		-5		-5.8
		SD		9.43		10.44
		Median		-3		-3
		Min		-30		-35
		Max		14		12
		Treatment Effect				-0.8

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 46. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD24 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 1 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Gender						
Female	Total Score	N	74	74	86	86
		Mean	31	26.1	30.5	23.4
		SD	4.69	9.09	4.9	8.59
		Median	30.5	27	30	24
		Min	22	2	21	2
		Max	42	44	44	42
	Change from Baseline	N		74		86
		Mean		-4.9		-7.1
		SD		8.49		8.25
		Median		-5		-5.5
		Min		-28		-27
		Max		13		9
		Treatment Effect				-1.9
Male	Total Score	N	72	72	69	69
		Mean	30	25.7	29.6	23.2
		SD	4.99	8.58	5.21	9.39
		Median	29	27.5	29	25
		Min	21	5	18	0
		Max	45	40	46	39
	Change from Baseline	N		72		69
		Mean		-4.3		-6.3
		SD		8.53		8.53
		Median		-2		-6
		Min		-25		-28
		Max		13		8
		Treatment Effect				-1.4

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 46. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD24 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 2 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Age Group						
Age < 55 Years	Total Score	N	98	98	111	111
		Mean	31.1	27.1	30.7	23.8
		SD	5.07	8.67	5.18	9.28
		Median	31	29	30	25
		Min	21	2	18	0
		Max	45	44	46	42
	Change from Baseline	N		98		111
		Mean		-4		-6.9
		SD		8.71		8.46
		Median		-2		-6
		Min		-28		-27
		Max		13		8
		Treatment Effect				-2.8
Age >= 55 - 70 Years	Total Score	N	48	48	44	44
		Mean	29.3	23.6	28.5	22.3
		SD	4.18	8.74	4.35	7.98
		Median	29	25	28	22.5
		Min	22	3	22	4
		Max	42	38	45	40
	Change from Baseline	N		48		44
		Mean		-5.7		-6.2
		SD		7.98		8.17
		Median		-4		-4.5
		Min		-25		-28
		Max		10		9
		Treatment Effect				0.2

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 46. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD24 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 3 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Baseline Total Symptom Severity						
HAMD17: Total Score < Median	Total Score	N	59	59	65	65
		Mean	26.6	22.8	26.4	20.9
		SD	2.44	7.71	2.72	8.57
		Median	27	23	26	22
		Min	21	2	18	0
		Max	33	38	34	34
	Change from Baseline	N		59		65
		Mean		-3.8		-5.5
		SD		7.59		8.16
		Median		-4		-5
		Min		-28		-25
		Max		13		8
		Treatment Effect				-1.7
HAMD17: Total Score >= Median	Total Score	N	87	87	90	90
		Mean	33.1	28	32.8	25.2
		SD	4.23	8.93	4.62	8.79
		Median	33	30	31.5	26.5
		Min	26	3	22	2
		Max	45	44	46	42
	Change from Baseline	N		87		90
		Mean		-5.1		-7.6
		SD		9.04		8.44
		Median		-3		-6.5
		Min		-25		-28
		Max		13		9
		Treatment Effect				-2.2

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 47. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD17 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 1 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Gender						
Female	Total Score	N	74	74	86	86
		Mean	23.2	19.6	22.9	17.5
		SD	3.67	6.86	3.24	6.31
		Median	23	21	23	18
		Min	18	2	18	1
		Max	34	34	32	31
	Change from Baseline	N		74		86
		Mean		-3.7		-5.4
		SD		6.31		6.35
		Median		-2.5		-5
		Min		-19		-22
		Max		10		5
		Treatment Effect				-1.8
Male	Total Score	N	72	72	69	69
		Mean	22.5	19.3	22.2	17.2
		SD	3.39	6.17	3.35	6.74
		Median	22	21	22	19
		Min	16	4	13	0
		Max	33	30	32	29
	Change from Baseline	N		72		69
		Mean		-3.2		-5
		SD		5.87		6.23
		Median		-1.5		-4
		Min		-19		-21
		Max		10		4
		Treatment Effect				-1.7

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 47. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD17 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 2 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Age Group						
Age < 55 Years	Total Score	N	98	98	111	111
		Mean	23.1	20.1	22.8	17.5
		SD	3.65	6.45	3.43	6.73
		Median	23	21	23	18
		Min	16	2	13	0
		Max	33	34	32	29
	Change from Baseline	N		98		111
		Mean		-3		-5.3
		SD		6.13		6.43
		Median		-1		-5
		Min		-18		-22
		Max		10		5
		Treatment Effect				-3
Age >= 55 - 70 Years	Total Score	N	48	48	44	44
		Mean	22.3	17.9	22	17
		SD	3.29	6.44	2.89	5.89
		Median	22	19	21.5	18.5
		Min	18	3	18	3
		Max	34	29	32	31
	Change from Baseline	N		48		44
		Mean		-4.4		-4.9
		SD		5.93		5.93
		Median		-2		-4
		Min		-19		-19
		Max		6		5
		Treatment Effect				0.3

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

Table 47. Subset Analysis by Gender, Age (< 55, > 55 years), and by Baseline Symptom Severity for HAMD17 Total Score at Week 4 (Primary Efficacy Outcome Time Point) – Page 3 of 3

			Sham TMS (146)		Active TMS (155)	
Subgroup	**Values**	**Statistics**	**Baseline**	**Week4**	**Baseline**	**Week4**
Baseline Total Symptom Severity						
HAMD17: Total Score < Median	Total Score	N	59	59	65	65
		Mean	19.5	16.9	19.5	15.3
		SD	1.24	5.57	1.36	5.97
		Median	20	18	20	16
		Min	16	2	13	0
		Max	21	28	21	24
	Change from Baseline	N		59		65
		Mean		-2.6		-4.2
		SD		5.56		6.02
		Median		-1		-4
		Min		-18		-21
		Max		10		4
		Treatment Effect				-2
HAMD17: Total Score >= Median	Total Score	N	87	87	90	90
		Mean	25.1	21.1	24.8	18.9
		SD	2.71	6.59	2.39	6.47
		Median	24	23	24	20
		Min	22	3	22	1
		Max	34	34	32	31
	Change from Baseline	N		87		90
		Mean		-4		-5.9
		SD		6.38		6.39
		Median		-2		-5
		Min		-19		-22
		Max		10		5
		Treatment Effect				-2.2

Notes: All computations are performed on the intent-to-treat, evaluable study population in a last observation carried forward (LOCF) analysis

Treatment effect was computed as (LS mean for active TMS – LS mean for sham TMS)/(standard error of LS mean)

The median cut point is the median (=22) split of the HAMD17 scores observed at the baseline visit

15.1. Subset Analyses Conclusions

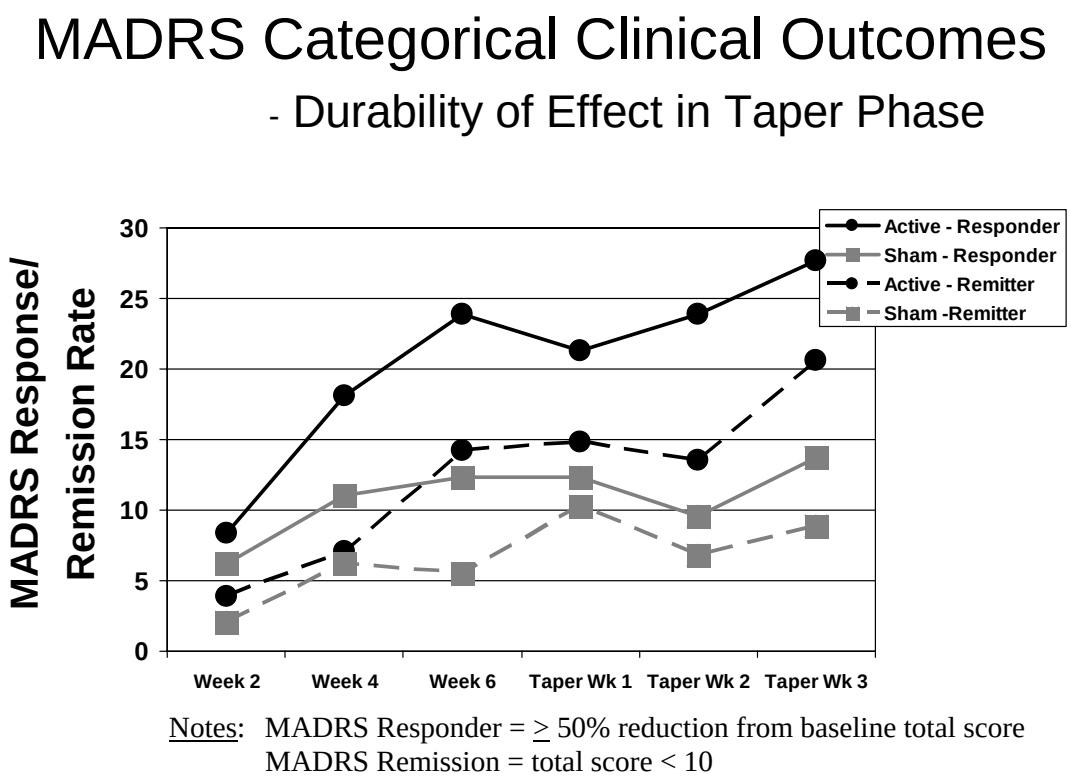
Inspection of the exploratory analyses subset by gender, age and baseline HAMD17 severity do not suggest any clinically meaningfully differential effect of active TMS on any of the observed population features.

16.0 DURABILITY OF EFFECT OF TMS TREATMENT

At the conclusion of the acute treatment phase, all remaining patients were entered into a continuation phase referred to at the *post-treatment taper phase*. During this portion of the study, all patients began a scheduled taper of their blinded treatment assignment across a 3-week schedule. At the same time, *all patients were initiated on open-label pharmacotherapy with a single antidepressant medication* selected from a protocol-defined list. No patient was to be treated with an antidepressant medication for which they had previously been shown to have failed to receive benefit.

Because this phase of the study involved open-label pharmacotherapy, and therefore was uncontrolled, only descriptive statistics are reported for data in this phase of the study as stated *a priori* in the study protocol.

Figures 10, 11, and 12 summarize the categorical responder and remission rates for the primary disease-specific efficacy outcome measures (the MADRS, the HAMD24 and the HAMD17) for all patients continuing into the post-treatment taper phase. Detailed supportive tables for these figures are included in Appendix 3, Tables 3.14-3.19.

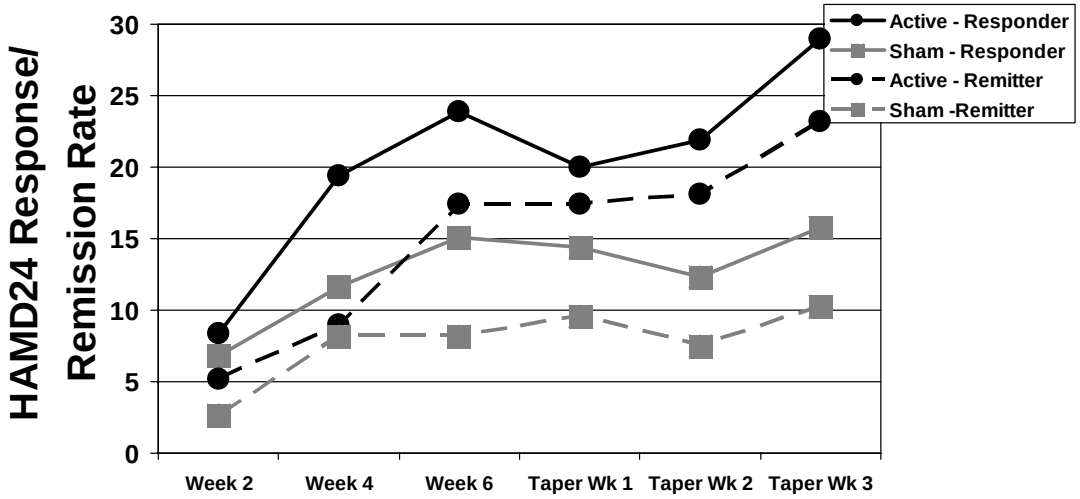


Notes: MADRS Responder = ≥ 50% reduction from baseline total score
MADRS Remission = total score < 10

Figure 9. Responder and Remission Rates for the MADRS for Patients Continuing into the Post-Treatment Taper Phase

HAMD24 Categorical Clinical Outcomes

- Durability of Effect in Taper Phase

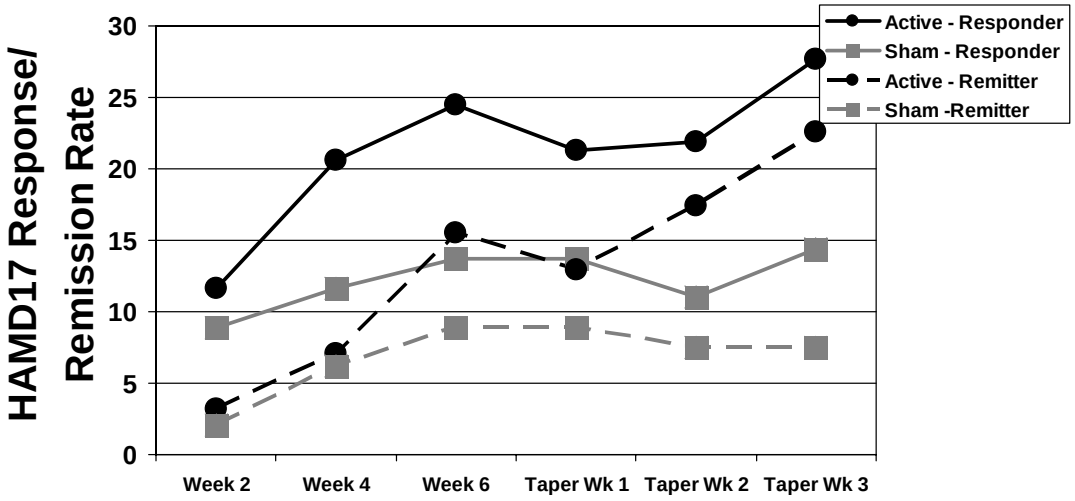


Notes: HAMD24 Responder = ≥ 50% reduction from baseline total score
HAMD24 Remission = total score < 11

Figure 10. Responder and Remission Rates for the HAMD24 for Patients Continuing into the Post-Treatment Taper Phase

HAMD17 Categorical Clinical Outcomes

- Durability of Effect in Taper Phase



Notes: HAMD17 Responder = ≥ 50% reduction from baseline total score
HAMD17 Remission = total score < 8

Figure 11. Responder and Remission Rates for the HAMD17 for Patients Continuing into the Post-Treatment Taper Phase

16.1. Durability of TMS Effect in Taper Phase Conclusions

- The clinical effect of active TMS is sustained during transition to single-drug antidepressant monotherapy (MADRS, HAMD 17 and HAMD mean total score at 6 weeks was maintained through week 3 of taper). This indicates that patients may be appropriately transitioned to clinically relevant continuation treatment without loss of clinical benefit achieved in the acute treatment phase.
- Patients allocated to active TMS showed a greater clinical benefit during this continuation period compared to those patients allocated to sham TMS.
- The remission rate at the end of the 3 week taper phase for active TMS patients was greater than the responder rate seen in the sham TMS group at the same time point.

17.0 SAFETY DATA

17.1. Serious Adverse Events

In addition to the collection of all protocol-emergent adverse events, sites were instructed to collect and document all serious adverse events as defined in the study protocol. Protocol 44-01101 defines a *serious adverse event* (SAE) as an adverse event that:

- Resulted in death,
- Was life threatening,
- Required inpatient hospitalization or prolongation of an existing hospitalization,
- Resulted in permanent impairment of a body function or permanent damage to a body structure,
- Necessitated medical or surgical intervention to preclude such impairment,
- Resulted in a congenital anomaly or birth defect,
- Additionally, *important medical events* that may not have resulted in death, or were not life-threatening, or did not require hospitalization, could have been considered SAEs, based upon appropriate medical judgment of the investigator,
- Seizures, and
- Any malfunction of an investigational device if it was likely to result in death, serious injury or other significant adverse event experience.

Overdose with the Neuronetics device as defined below was considered an adverse event of special interest for reporting purposes of this study. Neuronetics elected to pursue this conservative reporting strategy because the treatment parameters in use in this protocol were higher than previous studied in the TMS literature. This event was asked to be reported in the time frame of a serious adverse event and is reported within the serious adverse event case vignettes below.

17.1.1. Listing of Serious Adverse Events Reported for Study 44-01101

- No deaths or seizures were reported.
- Six (6) events occurred after signing of the informed consent and subsequent to randomization, Sixteen (16) events occurred during the acute treatment phase, and one (1) event occurred in the post-treatment taper phase.
- The types of SAEs or other reportable events are shown in Table 48. The number of SAEs reported and the relationship to study device as determined by the investigator is also provided.

Table 48. Serious Adverse Events Reported for Study No. 44-01101

Serious Adverse Event	Number of SAEs	Relationship to Study Device
Worsening depression	5	Not related (5)
Suicidal ideation	5	Not related (5)
Overdose	5	Not related (5)
Device malfunction/first degree burn	2	Probable (2)
Suicide attempt	1	Not related (1)
Device malfunction/severe pain at treatment site	1	Related (1)
Lower lobe pneumonia	1	Not related (1)
Bowel obstruction	1	Not related (1)
Shortness of breath and increased heart rate	1	Not related (1)

17.1.2. Serious Adverse Event Clinical Case Vignettes for Study 44-01101

Clinical case vignettes for all serious adverse events are provided in Appendix 5 as well as detailed supporting documentation for each vignette, including serious adverse event reporting pages. A CD-rom is provided for the case report forms for the patients for whom an SAE was reported.

17.2. Treatment-Emergent Adverse Events

All investigative sites were trained in the collection of adverse events at every study visit occurring after informed consent was obtained and through 30 days after the last study visit in all Neuronetics clinical protocols.

As defined in the protocol, an *adverse event* was:

- Any untoward, undesired, or unplanned event in the form of signs, symptoms, disease, or laboratory or physiological observations occurring in a person who has received treatment with a Neuronetics device or in a Neuronetics clinical study.

The event need not have been causally related to the Neuronetics device or Neuronetics clinical trial. An adverse event included, but was not limited to:

- Any clinically significant worsening of a pre-existing condition;
- An AE occurring from overdose (i.e., a dose higher than that described in the protocol) of a Neuronetics device, whether accidental or intentional;
- An AE occurring from abuse (e.g., use for non-clinical reasons) of a Neuronetics device;
- An AE that has been associated with the discontinuation of the use of a Neuronetics device

Training in adverse event collection included instruction in proper terminology, as well as methods of assessment of causal relation of the event to study device. Sites recorded all adverse event information in complete form in source data records and on electronic case report forms. Verbatim adverse event terms as recorded by the investigative site staff were coded using the current version of the Medical Dictionary for Regulatory Activities (MedDRA) and reported by MedDRA preferred terms.

Table 49 summarizes adverse events by MedDRA-preferred term that occurred at an incidence of ≥ 2% on active and were greater than the incidence on placebo. Detailed tabular summary of adverse events, including summary of investigator-assigned causal relationship to study device, and clinical severity are contained in Tables 3.20-3.25 in Appendix 3.

Table 49. Summary of MedDRA Preferred Term Adverse Events Occurring with an Incidence on Active TMS of ≥ 2% and Greater Than the Incidence on Sham TMS

Body System (-) Preferred Term	Sham (N=158) N (%)	Active (N=165) N (%)
Ear and labyrinth disorders		
- Ear pain	1 (0.6)	4 (2.4)
- Tinnitus	2 (1.3)	7 (4.2)
Eye disorders		
- Eye pain	3 (1.9)	10 (6.1)
- Lacrimation increased	1 (0.6)	7 (4.2)
- Visual disturbance	2 (1.3)	4 (2.4)
Gastrointestinal disorders		
- Diarrhoea	6 (3.8)	8 (4.8)
- Nausea	10 (6.3)	17 (10.3)
- Toothache	1 (0.6)	12 (7.3)
- Vomiting	3 (1.9)	7 (4.2)
General disorders and site administration conditions		
- Application site discomfort	2 (1.3)	18 (10.9)
- Application site pain	6 (3.8)	59 (35.8)
- Facial pain	5 (3.2)	11 (6.7)
- Pain	3 (1.9)	7 (4.2)
- Pyrexia	1 (0.6)	4 (2.4)
Injury, poisoning and procedural complications		
- Overdose*	0	4 (2.4)
Musculoskeletal and connective tissue disorders		
- Arthralgia	5 (3.2)	10 (6.1)
- Muscle twitching	5 (3.2)	34 (20.6)
- Musculoskeletal stiffness	4 (2.5)	5 (3.0)
- Neck pain	4 (2.5)	8 (4.8)

Body System (-) Preferred Term	Sham (N=158) N (%)	Active (N=165) N (%)
Nervous system disorders		
- Dyskinesia	2 (1.3)	5 (3.0)
- Headache	87 (55.1)	96 (58.2)
- Hypoaesthesia	2 (1.3)	5 (3.0)
- Paraesthesia	4 (2.5)	6 (3.6)
- Tension headache	2 (1.3)	4 (2.4)
Psychiatric disorders		
- Agitation	3 (1.9)	4 (2.4)
- Anxiety	18 (11.4)	19 (11.5)
Reproductive system and breast disorders		
- Dysmenorrhoea	2 (1.3)	5 (3.0)
Respiratory, thoracic and mediastinal disorders		
- Cough	2 (1.3)	4 (2.4)
- Dyspnoea	1 (0.6)	6 (3.6)
Skin and subcutaneous tissue disorders		
- Pain of skin	1 (0.6)	14 (8.5)

Notes: * Overdose refers to events associated with inadvertent smart card operator error resulting in > 75 trains of active or sham TMS delivered to the patient on a single calendar day. Per protocol procedure, all of these events were considered as adverse events to be reported in the time frame and manner of serious adverse events.

17.2.1. Time Course of Common Adverse Events

The most common adverse events experienced by patients were headache (58.2% active TMS treatment vs 55.1% sham TMS treatment) and application site pain (35.8% active TMS treatment vs 3.8% sham TMS treatment). A comparable proportion of patients on active TMS classified their headache severity as 'severe' as compared to sham TMS (active TMS 4.2% vs sham TMS 5.1%). With regard to application site pain, a greater percentage of patients treated with active TMS classified this event as 'severe' compared to sham TMS (active TMS 6.1% vs sham TMS 0%).

Inspection of the investigator-assigned causal relation of the event to the study device revealed that for headache, 27.9% of active TMS treated patients reported their headache as of 'probable' or 'definite' relation to the study device compared to 19.6% of sham TMS treated patients. In the instance of application site pain, all patients in both active and sham TMS treatment groups considered the event of probable or definite relationship to the study device.

In order to determine the time course of incidence of these common adverse events, which were expected to show adaptation and diminishing incidence over time, an exploratory analysis of these symptoms was performed with regard to the time of event within the course of the clinical trial. These data are displayed

in Figures 13 and 14. Supporting data tables for these figures are contained in Appendix 3, Tables 3.26 and 3.27.

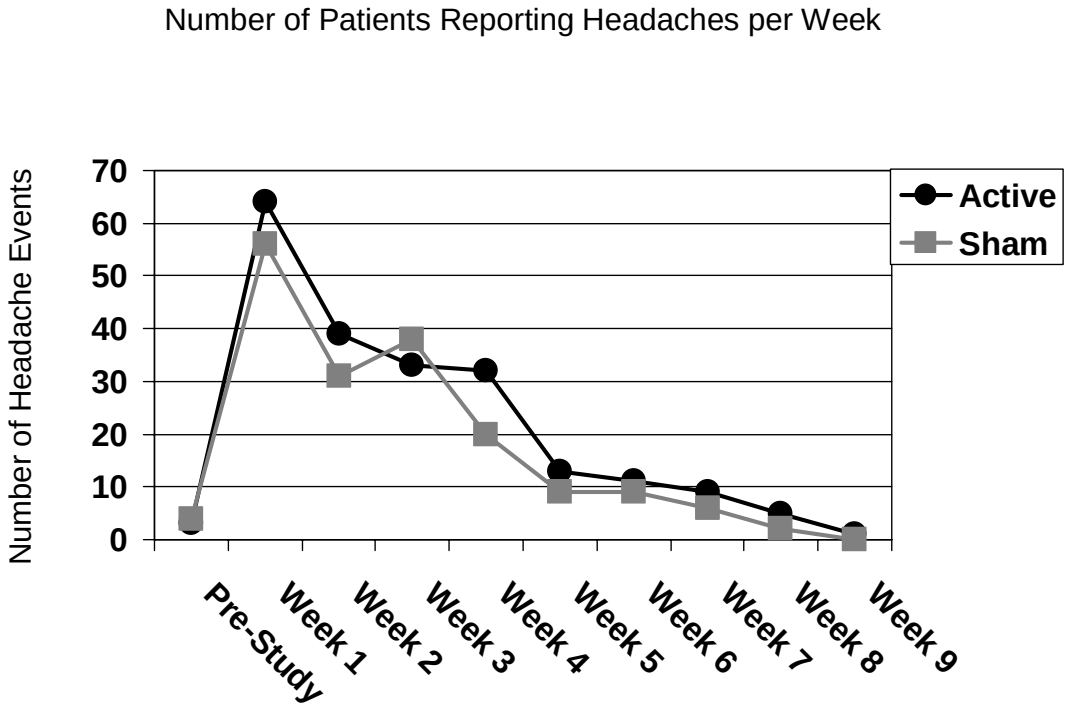


Figure 12. Time Course Incidence of Headache (Any Severity)

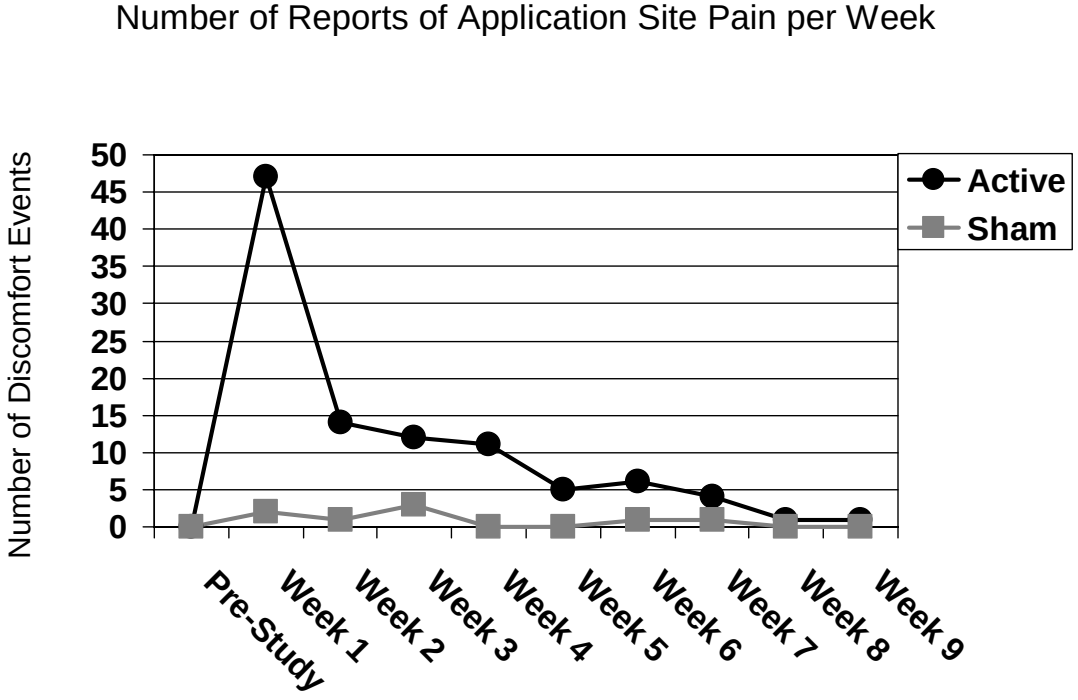


Figure 13. Time Course Incidence of Application Site Pain (Any Severity)

17.2.2. Treatment-Emergent Adverse Events Conclusions

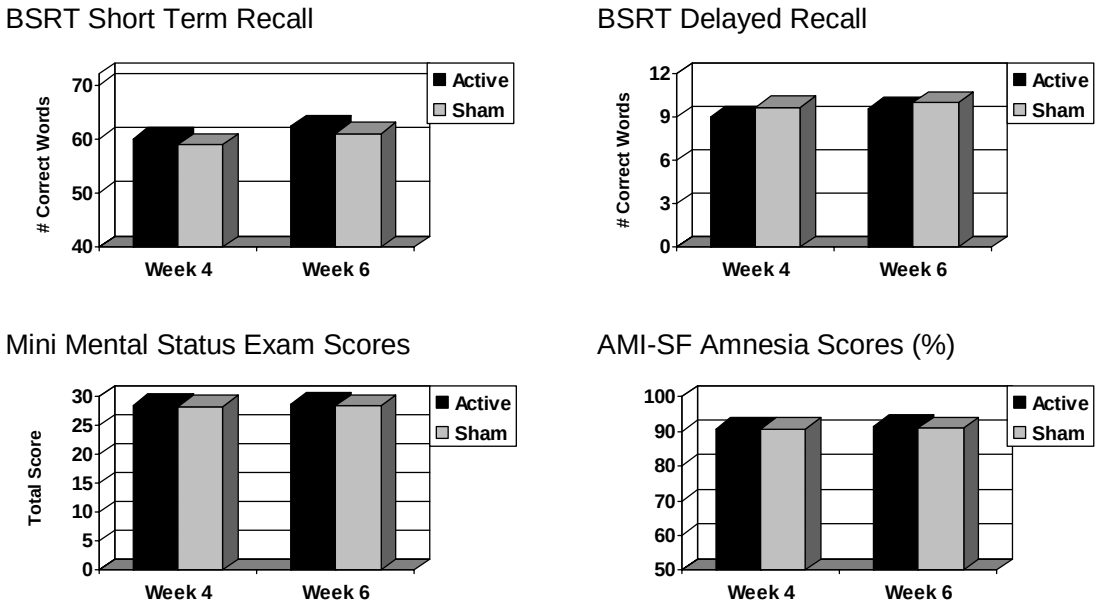
- There were a similar percentage of headaches seen in both the active TMS and sham TMS treatment groups.
- Application site pain was observed in both treatment groups, but the incidence was greater in the active TMS patient group.
- For both headache and application site pain, the greatest incidence was observed during the first week of treatment with a substantial reduction in incidence of these common adverse events after the first week of treatment, consistent with a rapid accommodation to these commonly experienced events. This accommodation effect was more pronounced for application site pain.

17.3. Cognitive Function Testing

Cognitive function was assessed using the modified Mini Mental Status Examination (MMSE), the Buschke Selective Reminding Test (BSRT), and the Autobiographical Memory Inventory-Short Form (AMI-SF) at baseline, week 4 and week 6. Multiple versions of the MMSE and BSRT were used to allow repeat administrations and to deter learning effects.

Results of these tests comparing baseline assessment with 4 and 6 week observations during the acute treatment phase are shown in Figure 15.

Detailed supporting details are provided in Tables 3.28-3.30 in Appendix 3.



All contrasts non-significant, P > .05

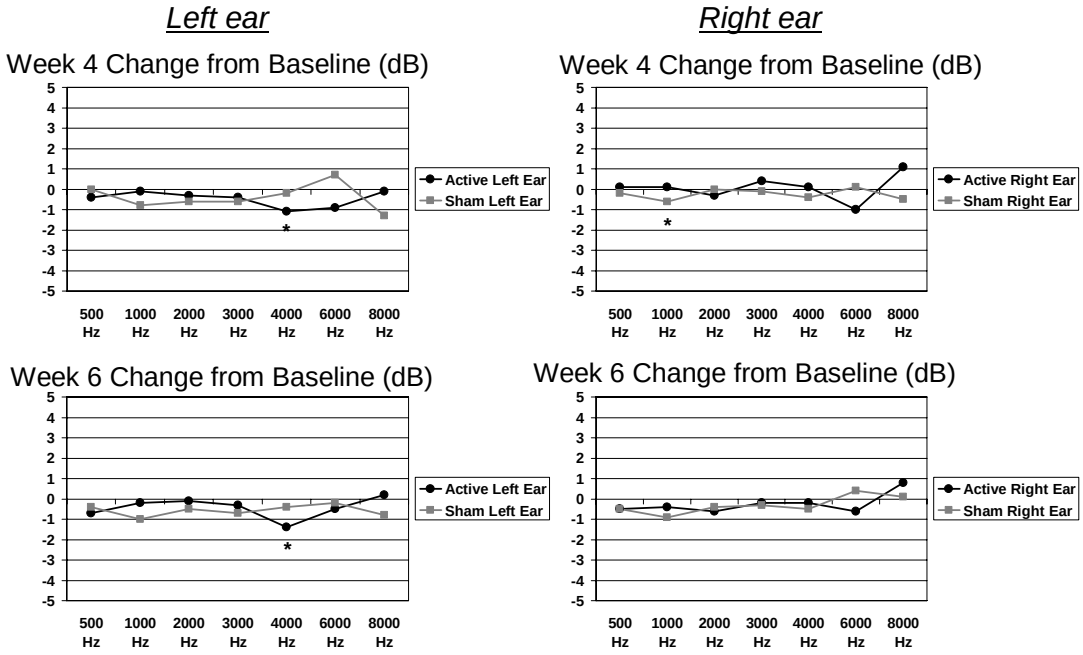
Figure 14. Cognitive Function Testing Results for Study 44-01101

17.3.1. Cognitive Function Testing Conclusions

- There was no evidence of an acute effect of TMS on any measure of cognitive function tested.
- Both TMS active and sham treatment groups showed essentially stable cognitive function on the standard test measures used throughout the acute treatment phase of the study.

17.4. Auditory Threshold Testing

Air-conduction auditory threshold was assessed at baseline, week 4 and week 6. A desktop audiometer (Micro Audiometrics, Inc,) was used, with a standard test sequence that examined the threshold decibel level at which a pure tone signal could be perceived by the patient. Results of these tests are shown in Figure 16. Contrasts within treatment group examining change in decibel level (auditory threshold) are shown for left and right ears. Note that all patients wore ear protection rated at a minimum decibel level reduction of 30 during TMS treatment. Detailed supporting details are provided in Tables 3.21-3.37 in Appendix 3.



All contrasts non-significant, except for * p < .05 showing improvement

Figure 15. Auditory Threshold Testing Results for Study No. 44-01101

17.4.1. Auditory Threshold Testing Conclusions

- There was no evidence of a short-term alteration of auditory threshold with acute treatment with active TMS compared to sham TMS when earplugs (30 db) were worn during TMS treatment.
- Both treatment groups showed essentially stable air conduction auditory threshold throughout the acute treatment phase of the study.

17.5. Emergent Suicidal Ideation

Major depression is a potentially lethal disease. It has been speculated that in some patient populations, antidepressant treatment may be associated with a paradoxical aggravation of the illness, with a resulting abrupt incidence of suicidal ideation. In order to assess if TMS treatment may similarly be associated with a risk for paroxysmal suicidal ideation, an exploratory safety analysis was performed to examine this risk for active TMS treatment.

The Item 3 score on the HAMD (Suicidal Ideation) was examined for incidence of abrupt worsening of this item from a score of 0 or 1 at the baseline assessment to a shift in score to 3 or 4 at any later time point. Results of this analysis are shown in Figure 17, and detailed tabular summary of these results are provided in Appendix 3, Table 3.38.

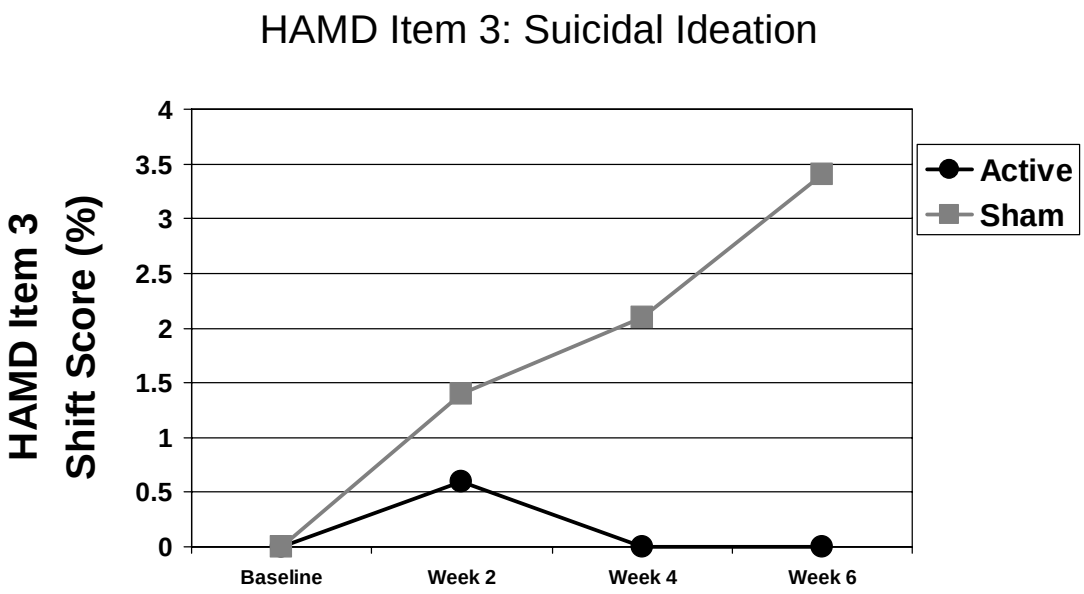


Figure 16. Incidence of Emergent Suicidal Ideation in Study No. 44-01101

17.5.1. Emergent Suicidal Ideation Conclusions

- There was an excess of cases of worsening suicidal ideation in the patients allocated to the sham TMS treatment group.
- There was no evidence that active TMS treatment was associated with worsening of suicidal ideation or emergent suicidal ideation during the acute treatment phase.

17.6. Overall Conclusions Based on the *A Priori*-Defined Safety Outcome Measures

Serious Adverse Events

- There were no deaths or seizures reported in Study 44-01101.
- Serious adverse events related or probably related to TMS treatment, respectively, were confined to a report of severe scalp pain and to device malfunctions of the E-Shield that resulted in minor scalp burns.

Spontaneous Adverse Events During the Acute Treatment Phase

- The adverse event profile associated with acute treatment with the Neuronetics TMS System was similar to the expected profile reported in the scientific TMS literature.
- The most frequently reported events were headache and application site pain. Headache was equally represented in both active and sham TMS groups. Application site pain was more frequently represented in the active TMS group. Both headache and application site pain lessened with time over the TMS treatment course.

Cognitive Function Testing During the Acute Treatment Phase

- There was no evidence of clinically significant cognitive function testing change at either 4 weeks or 6 weeks associated with acute treatment with the Neuronetics TMS System

Auditory Threshold Testing During the Acute Treatment Phase

- There was no evidence of clinically significant auditory threshold change at either 4 weeks or 6 weeks associated with acute treatment with the Neuronetics TMS System (with use of earplugs during TMS treatment).

Emergent Suicidal Ideation During The Acute Treatment Phase

- There was an excess of cases of worsening suicidal ideation in the patients allocated to the sham TMS treatment group.
- There was no evidence that active TMS treatment was associated with worsening of suicidal ideation or emergent suicidal ideation during the acute treatment phase.

18.0 DEVICE FAILURES AND REPLACEMENTS

There were two failure modes that occurred during protocols 44-01101, 44-01102 and 44-01103. The failures involved a malfunction of the clinical trial Model 2100 TMS System console power supply due to a plating defect in the control board and a manufacturing defect of the E-shield that was caused by a shorted trace within the E-shield. The reporting of the failure modes is detailed below and is further defined in Table 50.

18.1. Console Failure

Fifteen console failures at nine clinical sites were reported in Ser. No. 012 on 30 August 2004 and in the IDE Annual Report 2004 (Ser. No. 014). A root cause analysis report for the console failures was submitted as Ser. No. 016 on 19 Oct 2004. The console replacement process concluded on 15 October 2005 with the replacement of all affected consoles.

18.2. E-Shield Failure: first degree scalp burn and E-shield Recall

A single report of overheating of an E-Shield that resulted in a first degree scalp [redacted] rted to the FDA in Ser. No. 009 on 04 June 2004 as stated in the 2004 Annual Report (Ser. No. 014).

As a result of the E-Shield malfunction, a recall of 41 E-Shields was initiated (Ser. No. 009 dated 04 June 2004). The recall was expanded to include an additional 6 E-shields as described in Ser. No. 010 dated 23 July 2004. A root cause analysis was performed and reported in Ser. No. 011 dated 18 August 2004. Unreleased E-shields that met the requirements of the recall were destroyed by the contract manufacturer, DMSI.

A second report of a first degree scalp burn was reported by the Medical University of South Carolina on 26 October 2004 and was reported in Ser. No. 017, dated 05 November 2004. The root cause analysis report for the device malfunction was submitted in Ser. No. 011 on 18 August 2004. The informed consent documents for protocols 44-01101, 44-01102 and 44-01103 had previously been revised to include the risk of scalp burn. They were revised further to indicate that more than one event of scalp burn had occurred (Ser. Nos. 022, 023, 024 dated 07 Feb 2005, 10 Feb 2005, 02 Mar 2005, respectively). The changes to the informed consent documents and the investigational plan were approved in an FDA letter dated 14 April 2005.

One incident of “acute pain” under the treatment coil that was relieved by replacement of the E-Shield occurred on 08 September 2005 at Rush University. The patient’s scalp was examined and there was no evidence of skin irritation, erythema or burn. The event was reported in Ser. No. 031, dated 4 October 2005. The root cause analysis report for the device malfunction was submitted in Ser. No. 033 on

21 October 2005. Based on the findings, the event did not require the alteration of the risk profile of the device or modification of the informed consent documents.

Table 50. Reportable Device Malfunction Event and Regulatory Reporting

Device	Event	Device S/N	Event Date	[redacted] S/N	Report Date
E-Shield	Burn	01979	1 Jun 2004	009, 010, 011	4 Jun 2004
E-Shield	Burn	03645	26 Oct 2004	017	5 Nov 2004
E-Shield	Acute Pain	15021	8 Sep 2005	031, 033	4 Oct 2005
Console	Malfunction	1006	19 Jul 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1015	20 Jul 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1013	21 Jul 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1011	27 Jul 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1005	3 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1007	3 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1009	9 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1012	26 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1008	27 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	1010	30 Aug 2004	012, 014, 016	30 Aug 2004
Console	Malfunction	8006	7 Sep 2004	014, 016	30 Sep 2004
Console	Malfunction	1015	13 Sep 2004	014, 016	30 Sep 2004
Console	Malfunction	8028	21 Sep 2004	014, 016	30 Sep 2004
Console	Malfunction	8025	29 Oct 2004	016	19 Oct 2004
Console	Malfunction	8020	8 Nov 2004	016	19 Oct 2004

19.0 STUDY 44-01101 INDEPENDANT DATA MONITORING COMMITTEE (IDMC) AND MODIFICATIONS TO THE PLANNED STATISTICAL ANALYSIS

19.1. Independent Data Monitoring Committee

An Independent Data Monitoring Committee (IDMC) was organized prior to the first patient enrollment in the study and its function, expertise and independence from Neuronetics was predefined in a charter.

As stipulated in the study protocol, the primary purpose of the IDMC was to conduct an interim analysis for efficacy at an enrollment of approximately 100 patients. An *a priori*-defined statistical threshold was determined such that the study could be halted for futility at that point. The IDMC also was provided with safety information including serious adverse event summaries, and also could have asked to halt the study in the event that a safety concern emerged during the conduct of the trial.

Communications were provided to the IDMC on 02 Aug 2004, 22 Oct 2004, 09 Dec 2004 (interim analysis made available to the IDMC at this point), 03 Mar 2005, 03 Jun 2005. The IDMC was closed on 03 Jan 2006 after the last patient had exited Study 44-01101.

The interim analysis report was performed in a blinded manner, and mailed directly from the contract research organization [redacted] hat conducted that analysis according to N [redacted] the IDMC members. No report of the interim results was provided to Neuronetics staff or to any of the study investigators. The response of the IDMC to the interim analysis was “continue as planned”.

Details of all communications to the IDMC are contained in the study master files at Neuronetics office.

19.2. Modifications to the Planned Statistical Analyses

All *a priori*-defined statistical analyses were conducted as planned. Additional analyses were conducted as follows:

- An analysis was performed of the MADRS, HAMD24 and HAMD17 total scores and the corresponding categorical outcome measures on these instruments, with the subset of patients who showed baseline MADRS scores less than 20 excluded from analysis.
- Subset analyses that were not prospectively defined in the protocol for study 44-01101 were conducted for gender, age and severity to determine if TMS treatment was biased to a demographic subset.

20.0 CONTRAINDICATIONS, WARNINGS AND PRECAUTIONS FOR USE

Safety data obtained from the conduct of Study 44-01101 are provided in the Neuronetics TMS System User Manual. New safety information obtained from study 44-01101 that was not previously included in IDE documents regarding the contraindications, warnings and precautions for use are as follows.

- A listing of adverse events reported with an incidence on active TMS of $\geq$ 2% and greater than the incidence on sham TMS is included in the Neuronetics TMS System User Manual.

21.0 ADDITIONAL CLINICAL INFORMATION FROM THE STUDY

21.1. Deviations to the Protocol and Protocol Violations

During the final database analysis, a summary of the potentially clinically important protocol violations was reviewed and summarized. These are listed in tabular form in Table 3.39 in Appendix 3, as shown for the overall safety study population (N=323).

In general, the pattern of protocol violations was distributed equally across the two treatment assignments. The largest group of observations concerned subjects missing more than two treatment sessions in sequence, or >20% of the number of sessions intended during their study participation as discussed in section 12. These latter violations are determined based on the safety study population, which included the non-evaluable patient sample. Because it was typically difficult to assign a discontinuation date for these patients with as much precision as for evaluable patients, the worst case estimate of latest date was used, resulting in these patients meeting operational criteria for non-compliance.

The overall pattern of protocol violations listed was small relative to the overall sample size, and therefore was not considered to have substantially affected the interpretation of the results, therefore, no post-hoc analyses excluding these data were deemed appropriate.

In addition to these clinically important protocol deviations, a review of the protocol deviation log and of the final data listings used for the development of the data tables was conducted for assessment of potentially clinically non-significant events. This review revealed protocol deviations at the conclusion of the study as shown in Table 51. None of these deviations interfered with patient safety or the risk profile of the device, and none were expected to materially alter the results or interpretation of the study results.

Table 51. Protocol Deviations in Study 44-01101

Protocol Deviations	Number of Deviations
Excluded medications used	26
Regulatory documentation	1
Documentation procedure	92
Protocol procedure	147
Device procedure	115

21.2. Steps Taken to Ensure the Integrity of the Study Blind

Previous exploratory studies of TMS have evaluated methods to assess the effectiveness of study blinding in concert with its potential impact on the interpretation of the study efficacy results (Berman, 2000; Nahas, 2003; Fitzgerald, 2003, 2006; Avery, 2006).

One of the principal concerns has centered on the fact that virtually all previous TMS studies employed a method of blinding that explicitly unblinded the TMS operator, namely the sham TMS treatment used the method of physically tilting the active TMS coil at a specified angle tangential to the surface of the patient's head in order to deflect the magnetic field from the intended treatment area. There are several concerns beyond the gross operator unblinding that have also been discussed in the TMS literature, including the observation that at certain levels of tilt, there remains a considerable amount of magnetic field exposure in the cortex that may alter the profile of apparent efficacy in the sham TMS treatment group (Lisanby, xxxx). For this reason, Neuronetics' clinical studies used a TMS sham coil that reduced the magnetic field to less than 10% of the field delivered by the active TMS coil at typical motor threshold levels.

Issues of the integrity of study blinding should consider the following observations. First, the patient population typically studied in TMS trials represents a more treatment refractory patient subset than is found in studies of antidepressant drugs, since most TMS studies require a pattern of historical treatment resistance. Such TMS patients would be expected to have a low or trivial degree of placebo response compared to the levels typically observed in most non-resistant depression studies. Indeed, this observation is supported by the available clinical trial data where placebo responses to TMS sham treatment is approximately 10% (see for example, Burt, 2002) versus the 30% seen in typical antidepressant drug studies (Khan, 2000; Walsh, 2002).

Second, the degree of active versus placebo differential on adverse event incidence that is seen in many antidepressant drug studies is sizeable. For example, in the adverse event tables contained within the product labels for several commonly used antidepressants, relative adverse event excess on active treatment may range as much as 3 to 8 - fold greater than the incidence observed on placebo: for Effexor XR (Nausea – 31% vs 12%, Dizziness 20% vs 9%, Somnolence 17% vs 8%), for Prozac (Nausea – 21% vs 9%, Anorexia 11% vs 2%), for Cymbalta (Nausea 20% vs 7%), and for Remeron (Somnolence 54% vs 18%, Increased Appetite 17% vs 2%). Given these results, it is generally accepted that integrity of the study blind is rarely complete. It is useful to note that in the majority of typical antidepressant drug study designs, there is rarely a stipulation that the study raters are maintained distinct and isolated from the treatment session, where adverse event data may be collected, as was the case in the Neuronetics TMS studies.

Beyond these general methodologic considerations, several studies have implemented methods of direct query upon study exit of the patient to determine the extent to which the patient felt they could detect to which treatment they had been assigned and the evidence they used to arrive at this guess. Although this method has been used, it is itself problematic because the heightened awareness to guessing may contribute to excess vigilance on the part of all study participants to "look beyond the blind". Nevertheless, the data from these methods are informative, since they generally reveal that while a slightly greater percentage of patients assigned to active TMS appear to guess their assignment correctly, the guess is largely viewed by the patient as being due to the fact that they responded to the treatment condition to which they had been assigned (Berman, 2000; Nahas, 2003; Fitzgerald, 2003, 2006; Avery, 2006).

There were several important design considerations and steps taken in the Neuronetics studies to ensure the integrity of the study and to minimize the likelihood of penetrance of the study blind:

First, this study implemented an important advance in the *sham coil design and randomization*. From a procedural view, the operator was blinded to the sham assignment by use of multiple TMS magnetic coils, all of the same exterior appearance, shape, weight, and physical position when held on the patient's head. In addition, the sham TMS 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 actively pulsing. The randomization of the coils was blinded by the use of coded coils and patient randomization was blinded by use of patient treatment cards that restricted coil use to only the appropriate "B" or "C" coil. Further details regarding randomization of the TMS coils and treatment cards and steps taken to verify their appropriate coding to the randomization schedule are provided in section 12.

Second, the study staff was segregated into *treating staff and rating staff*. In all instances, treaters were never permitted to serve as raters for the patients that they were treating. In virtually all instances, treating staff and rating staff were independent personnel and never served in dual roles. Adverse events were always assessed by treating staff only. Patients were instructed on this segregation of responsibilities, and were instructed to reserve reports of adverse events or questions on treatment session issues for the treating staff only, and to address responses to the symptom questions that were posed to the rater staff only.

The methods of rater training conducted during study 44-01101 and during study follow up are reviewed in Section 20.3. As discussed further below, rater assessment using the HAMD and MADRS outcome scales showed stability of the intraclass correlation coefficient between the on site raters and a site- and time-blind expert reviewer at Columbia University. These results indicate that the rater staff

did not acquire a bias in rater performance that would suggest a systematic pattern of rater unblinding during the time course of the study.

Finally, inspection of *the adverse event profile* in study 44-01101 shows that the two most prominent adverse events were consistent with the expectations from the TMS literature, namely headache and application site pain (see Section 15.2). The most commonly experienced event, headache, was present to a similar degree in both active TMS and sham TMS treatment groups. The next most common adverse event, application site pain, was present in ~1/3 of active TMS patients, and was generally reported as mild or moderate in intensity. At the critical time points of efficacy outcome (4 and 6 weeks), the incidence of these adverse events had fallen to levels substantially less than 50% of the incidence seen during the first week, and differences in incidence between sham TMS and active TMS treated patients were minimal. Based on these observations, the incidence and temporal pattern of these commonly experienced adverse events was unlikely to contribute to penetrance of the study blind at a rate any different than for similarly designed studies for antidepressant medications.

21.3. Rater Training Program

Rater training on patient assessment using MADRS and HAMD instruments was conducted as part of the initial clinical site initiation as described in section 4.3.

A program to assess the adequacy of rater training was used in this study to ensure:

- adequate skill in raters for identification of clinically relevant symptoms on the primary efficacy outcome measures (HAMD and MADRS) as determined from test clinical ratings (up to 5 videotapes of sample patients provided by the sponsor) that were conducted by the rater prior to their first on-study rating, and
- formal videotaped follow up (at baseline, weeks 4 and 6) of rater skill in administering these tests to ensure continued consistency in their administration throughout the study.

During the conduct of Study 44-01101, a subset of the latter videotaped interviews were randomly selected by an expert in conduct of these efficacy measures who was blinded to site- and time-in-study of the videotape in order to allow an unbiased assessment of clinical technique. This assessment also permitted inspection for potential trends in intraclass correlation coefficient drift that may have suggested the development of a systematic bias in rater assessment and unblinding as the study proceeded.

During study 44-01101, the stability of the intraclass correlation coefficient (ICC) between the on site rater, and the site- and time-blind expert reviewer at Columbia University with regard to the HAMD and MADRS interview scores across the study time points indicated that the rater staff did not acquire a bias in rater per-

formance that would suggest a systematic pattern of rater unblinding during the time course of the study.

For both efficacy instruments, a semi-structured interview format was used. In this method, a leading or introductory question for each item was presented, along with a suggested series of clarifying follow-up questions. The specific sequence of items asked was also structured to minimize variability in administration.

One hundred and twenty-two raters underwent initial rater certification and 102 raters were approved to serve as clinical raters in study 44-01101. Of these certified raters, 44 raters actually participated in study clinical ratings. For this group, the average ICC for the HAMD was 0.94 and for the MADRS was 0.95. A total of 98 study videotapes were reviewed during study 44-01101 for continuing certification of raters. The average ICC for the HAMD at baseline, 4 weeks and 6 weeks was 0.87, 0.85 and 0.80, respectively, while for the MADRS at baseline, 4 weeks and 6 weeks was 0.85, 0.88 and 0.78, respectively.

21.4. Post-Data Lock Errata and Data Handling Issues

Data for this clinical study was collected v lectronic data capture system (EDC). Clinical site, monitor, and sponsor personnel were each provided with an individual log in ID. Site personnel entered the data that was collected on patient source documents, patient chart. The data on the EDC was monitored 100% against source document and was additionally reviewed for consistency and clarity. Upon completion of review, patient's data was soft locked by the investigator at the site.

Throughout the study, all adverse events comitant medications were coded via an autoencoder and manually a to MedDRA and WHODrug, respectively. Neuronetics reviewed all as g after each run and for the entire dataset, upon completion.

After the database was complete removed database change access to all users, allowing read only access. The database was then converted to SAS datasets and qual ach field. The dataset was then transferred to for completion of the statistical analysis.

During the analysis, it was determined that the database indicated that two patients had the same randomization code (219). Upon further investigation, it was determined that patient 17-015 was given this treatment card and that patient 17-030 had treatment card and randomization (220). The data was changed post data lock on the s dataset.

For calculation of ANCOVA, the ATHF for the qualifying episode was applied from the ATHF summary record. If current and past episodes were indicated, the past episode was disregarded for this analysis.

Incomplete start dates were provided by the sites for adverse events and concomitant treatments. For missing start months and days with year provided, the worst case scenario was used of January 1 of the year. For missing days with month and year provided, the first of the month was used.

After lock of the database, there were a few patients with adverse events starting in the year 2005, although the patients participation was completed in 2004. We changed the data in the derived dataset for calculation to 2004 without changing the original dataset.