Sponsor Executive Summary

Neuronetics, Inc.

19 December 2006

Neuronetics Premarket 510(k) Notification K061053

NeuroStar™ System for the Treatment of Major Depressive Disorder

Table of Contents

SUMMARY

Neuronetics' premarket 510(k) notification K061053 provides evidence of the safety and efficacy of Transcranial Magnetic Stimulation (TMS) Therapy, delivered by Neuronetics' NeuroStar System in the treatment of patients meeting DSM-IV-defined criteria for Major Depressive Disorder. Specific evidence and analyses in support of this claim includes:

- Results from a randomized, sham-controlled clinical trial, Study No. 44-01101, that provides evidence of safety, acute efficacy, and durability of effect;

- Additional results from the open-label extension Study 44-01102 that provides confirmatory evidence of safety, acute efficacy and durability of response;

- Additional results from the interim report of the open-label maintenance of effect extension Study 44-01103 that provide confirmatory evidence of safety and durability of response over a 24 week observation period;

- Clinical significance of the efficacy results obtained in the acute controlled and extension studies has been supported through demonstration of comparable clinical effect to currently available antidepressants. This is accomplished via an extensive analysis of the Neuronetics data compared to large, definitive data bases of FDA-approved pharmaceutical antidepressants and ECT; and

- Quantitative analysis demonstrating that the risk-to-benefit ratio of the NeuroStar System compares favorably to FDA-approved pharmaceutical antidepressants and to Electroconvulsive Therapy (ECT) devices.

The information presented in this Sponsor Executive Summary summarizes the above evidence that was submitted in 510(k) K061053 and demonstrates that Neuronetics' NeuroStar System is substantially equivalent to ECT devices, the predicate device, for the treatment of major depressive disorder, and therefore meets regulatory requirements for marketing clearance by premarket notification.

Neuronetics, Inc. submitted premarket 510(k) notification K061053 to the FDA on 14 April 2006 for the following intended use:

The NeuroStar System is intended for the treatment of Major Depressive Disorder (MDD).

This Sponsor Executive Summary, as given in the Table of Contents, summarizes the information submitted to the FDA in K061053 regarding the NeuroStar System and the data that support its safety, efficacy, and durability in the treatment of MDD.

An Expert Statistical Opinion Letter that was requested by Neuronetics and provided by Phillip W. Lavori, Ph.D., Professor of Biostatistics, Stanford University, gives an independent review of the clinical data from Study 44-01101 and is provided in Tab 6 of this Panel Package.

Additional information cited in this Sponsor Executive Summary is found in Attachments 1-16 that are provided in electronic (PDF) format on the accompanying CD-ROM.

Summary Basis of the Neuronetics Premarket 510(k) Notification

This Summary highlights several key aspects of the Investigational Plan for the NeuroStar System for the treatment of Major Depressive Disorder that should serve to assist the Advisory Panel in understanding the rationale for the Plan carried out by Neuronetics and its study designs, as well as key considerations in efficacy and safety evaluation and analysis. It also provides a high level review of the principal outcomes of the Neuronetics' randomized, controlled study and how these outcomes support a determination of substantial equivalence of the NeuroStar System to the predicate ECT devices.

1. **There is a strong rationale for pursuing TMS Therapy as a treatment for Major Depressive Disorder (MDD)**

   - Animal and human studies of transcranial magnetic stimulation (TMS) have suggested its potential antidepressant effects. Over 20 controlled clinical trials have been conducted to evaluate active TMS versus sham TMS in the treatment of MDD (See Section 1.3). Meta-analyses of these single-center, controlled trials of TMS have generally concluded that active TMS is more effective than sham TMS but that a large, randomized multicenter trial is needed to verify these results.

   - Recent single-site, randomized, controlled clinical trials have examined the efficacy of TMS in the treatment of MDD and have used the largest sample sizes studied to date, with improved study designs and treatment parameters, built on the knowledge learned from prior work (Avery, 2005; Fitzgerald, 2003, 2006). These studies demonstrated statistically significant superiority of active TMS as compared to sham treatment. Most importantly, these studies have provided the most compelling literature support available for the clinical significance of the antidepressant benefit of TMS. For example, in the report by Avery and colleagues, approximately 30% of patients treated with TMS achieved the categorical criterion of response using the HAMD24, and the overall standardized effect size for this outcome was 0.69.

   - There is clearly a large, unmet need for antidepressant treatments. Of the estimated 14 million Americans with depression in any year, nearly half go untreated and of those treated, nearly 4 million do not receive sufficient benefit from existing therapies. Existing treatment options for this "difficult-to-treat" patient population typically consist of more complex augmentation or combination treatment regimens, which lack a strong empirical database to support their use, and that also provide a range of systemic adverse events. More invasive treatment options are also available for use and include ECT and the implantable treatment with vagus nerve stimulation.

2. **Marketing clearance of the NeuroStar System is sought by premarket 510(k) notification to the FDA, which provides evidence of substantial equivalence of the NeuroStar System to predicate ECT devices as shown by comparison of their risk-to-benefit ratios (See Section 2)**.

   Neuronetics presented to the FDA an Investigational Plan intended to support marketing clearance of the NeuroStar System for the treatment of Major Depressive Disorder. After extensive discussions with the FDA, Neuronetics embarked on t al Plan and study protocols that were approved by the FDA under with the following understandings:

   - The NeuroStar device would be evaluated for marketing clearance by FDA under the premarket 510(k) regulation ***by showing substantial equivalence to the proposed predicate device, ECT devices***.
     - Substantial equivalence would be shown by demonstrating that the ***risk-to-benefit ratio*** of the NeuroStar System is comparable to ECT devices.
     - A head-to-head trial of TMS and ECT was not required due to the substantial limitations of conducting such a study in a blinded manner. These limitations are due to the considerable safety measures that must be taken with ECT but not for TMS. For example, with ECT, extensive safety measures are taken for electrically-induced seizure induction, including cardiac monitoring, the use of anesthesia and muscle relaxants during treatment, and following treatment, the extensive post-treatment care that is required due to the cognitive dysfunction that occurs in all patients immediately following treatment. None of this occurs with TMS which is conducted as an outpatient procedure after which the patient can immediately return to normal activities.
   - The safety and effectiveness of each device compared to its relevant control would be used to assess their comparative risk-to-benefit ratio. Clinical trial data for ECT would be obtained from the peer-reviewed medical literature. ***Safety and effectiveness*** of the NeuroStar System for the ***acute treatment*** of major depression would be evaluated in a randomized, multicenter trial versus sham treatment after 4-6 weeks (20-30 treatments). The rationale and justification for this study design is based in part on the following considerations.
     - Clearance for acute use of TMS is consistent with the regulatory approval requirements for pharmaceutical antidepressants.
     - ECT is also primarily used as an acute treatment for major depression, usually in a treatment “to effect” of up to 20 treatments delivered over several weeks.

- The ***short-term durability of effect*** would be determined for the NeuroStar System.

  - Treatment with ECT is known to have a dramatic loss of its acute effect with up to 50% of patients relapsing within 1 month of treatment in the absence of transition to a regimen of maintenance pharmacotherapy. Even with maintenance pharmacotherapy, there is still ~20-30% relapse at 1 month (see Section 6.4). Therefore, it is important to determine whether the durability of TMS, when observed over a similar time frame, is at least comparable to the durability of effect observed with ECT.

  - Durability was assessed after 6 weeks of TMS Therapy, by transitioning patients over 3 weeks from TMS to maintenance single-medication pharmacotherapy and then evaluated for 1 month after the last TMS treatment.

  - Neuronetics also followed patients who transitioned to maintenance pharmacotherapy over a 6 month period. This study was ongoing at the time of 510(k) submission and interim results at 6 months were provided to assess relapse and NeuroStar re-treatment rates over this time period.

- ***Safety*** of the NeuroStar System was evaluated over the course of each study by collection of spontaneous adverse events. Additional testing on cognitive function and auditory threshold was also conducted.

3. Neuronetics randomized, controlled clinical study, Study 44-01101, that demonstrates the safety and effectiveness of the NeuroStar System in the treatment of patients with major depression, is a *well-designed* study (See Section 4.3.1).

- The design of Study 44-01101 is consistent with standards of practice for contemporary studies of antidepressants. Clinical development of antidepressants is traditionally separated into demonstration of acute efficacy and then followed at a later time point by demonstration of maintenance of effect (or relapse prevention) in separate studies. Demonstration of acute efficacy for an antidepressant is considered sufficient evidence to bring a new antidepressant to market, and is a crucial foundation to any further work, since all known antidepressants that are effective acutely have ultimately been demonstrated to show efficacy in long-term maintenance. On the other hand, the significantly different study design approaches for demonstration of acute and long-term outcomes mandates that these two aspects of antidepressant development are conducted as separate clinical programs.

- Study 44-01101 was conducted as an acute efficacy study with a 3 week taper phase for transition to maintenance antidepressant pharmacotherapy.

- Study 44-01101 was conducted under a pre-specified analysis plan with an *a priori*-determined set of efficacy variables, rank ordered by level of importance in study outcome (see Section 4.2.1, Table 2). This is a traditional practice in the development of antidepressants.

- Primary and high-ranking secondary outcome measures were well-validated, clinician-rated efficacy instruments, the MADRS, HAMD24 and HAMD17.

  - The HAMD and MADRS scales are the most commonly used instruments to measure antidepressant outcome and are well-correlated with each other (See Tab 5 for a summary of depression rating scales and their use in the study of major depression).

  - In addition to assessment of continuous symptom change, traditional categorical outcome endpoints of response and remission on both the HAMD and MADRS scales were reported.

  - Clinician-rated and patient-rated outcomes were both collected. Higher importance was given to the clinician-rated outcomes, as these measures are universally considered to be the primary basis of evidence to establish clinical efficacy in the study of antidepressants.

  - Rigorous methods, including independence of the roles of TMS treater and clinical efficacy rater, were used to protect the integrity of the study blind.

4. Study 44-01101 was executed well and the study performed as expected (See Section 4.3.1).

- The study plan was followed with good protocol compliance.

- The study blind remained intact.

  - Because of the unique methodologic challenges posed by a device-based study, specific care was taken in the design of the trial to address this issue. Additional steps to protect the integrity of the study blind included isolating the role of the TMS treater from the study efficacy Rater. The latter individual was kept blind to the activities that occurred during the treatment session itself, and to the adverse event collection procedures. These additional levels of study blind protection are notably more stringent than typically utilized in pharmaceutical antidepressant studies.

  - Several post-hoc analyses were conducted at the request of the FDA to assess the relationship between device-related adverse events and efficacy outcome. These analyses showed no evidence of study unblinding. In one analysis that used these events as a covariate in the ANCOVA, there was a reduction of significance levels for MADRS and HAMD24 while the

HAMD17 remained statistically significant. As described in Section 4.2.4.2, these results are not meaningful because the method used is not appropriate and will, by definition, always result in a reduction in the statistical strength of the observed effect.

- Patient attrition through Week 4, the primary efficacy time point, was low (8%) which supports the appropriateness of the LOCF analysis at Week 4.
    - Attrition beyond Week 4 to Week 6 was high, as was expected, due to the study design allowance of study exit at or after Week 4 for patients who did not receive benefit and were therefore potentially eligible for open-label Study 44-01102.

5. Efficacy results are statistically significant and clinically meaningful (See Section 4.3.1.5)

- The primary outcome measure, the MADRS mean total score at Week 4, demonstrated ***favorable benefit for active TMS compared to sham TMS, with an observed P=0.057***. That this measure did not reach statistical significance stands in contrast with the statistically significant outcomes of the HAMD and other clinician-rated measures and was likely due to the overall greater variance in the scale performance of the MADRS compared to the HAMD in this study. This difference in scale variance was likely further aggravated by a feature of the study design, the implementation of a floor at baseline for the HAMD but not the MADRS, that allowed a statistically significant imbalance in baseline MADRS scores between active and sham treatment arms ($p = 0.036$). To assess the effect of this imbalance on study outcome, Neuronetics conducted a *post-hoc* subset analysis in which the MADRS baseline imbalance was addressed by using a well-defined "floor" for study entry (MADRS < 20). This subset analysis resulted in a P value of 0.038 on the primary outcome measure. We believe the combination of greater variance of the MADRS, and the baseline imbalance that occurred with the MADRS only, contributed to this outcome falling just above the traditional threshold for statistical significance in contrast to the other clinician-rated measures.

- Clinician-rated secondary outcomes, including response rate as measured by MADRS and HAMD at 4 weeks, and remission rate at 6 weeks, showed ***statistically significant benefit*** for active vs. sham TMS.

- Well-established and *a priori*-defined HAMD factors scores that detail the clinical response on depression and anxiety core symptoms showed strong, consistent and ***statistically significant effects,*** as expected for an antidepressant treatment.

- Patient-rated outcomes also performed as expected, per the *a priori*-defined test plan, showing ***statistical significance for specific measures of emotional***

function and well-being, and in the expected temporal pattern, relative to the timing of change on clinician-rated outcome measures.

- A third-party multiplicity analysis independently conducted by an expert statistician (See Tab 6) used four post-hoc multiplicity analyses (Holm, Hochberg, Hommel, and Benjamini-Hochberg) on 13 of the 26 primary and secondary endpoints using defined criteria. These analyses showed, as expected, that the four methods agree that the primary efficacy endpoint (MADRS at Week 4) had a resultant p-value of greater than 0.05 ($p>0.05$), and between one and nine, depending on the specific analysis performed, secondary endpoints had an adjusted p-value less than 0.05 ($p<0.05$). The conclusion of the statistician was that these multiplicity analyses did not favor the null hypothesis, thus indicating the overall effect of the outcomes was in favor of a significant positive outcome of active TMS versus sham TMS.

- Standardized effect size calculations for key outcome variables at 4 weeks were well within the estimated variance around the protocol-defined targeted effect size of 0.4 (SE = ± .12); MADRS (ES = 0.39), .HAMD24 (ES = 0.48), HAMD17 (ES = 0.55).

- Response and remission rates for TMS Therapy in Study 44-01101 ***compare favorably*** to similar data obtained in registration trials of FDA-approved antidepressants. Remission rates for TMS Therapy in the open-label Study 44-01102 compare favorably with similar data obtained from the recent open-label STAR*D trials of antidepressant use in difficult-to-treat patients with major depression (See Section 5).

- Although a direct comparison is not possible, TMS Therapy with the NeuroStar System appears roughly two-thirds as effective as ECT when comparing outcomes in the Neuronetics study to randomized controlled trials conducted with ECT and simulated ECT. In general, the effect sizes for TMS Therapy (range = 0.39 to 0.55) fall within the large range of effect sizes observed from single-center, controlled trials of ECT versus simulated ECT (range = 0.17 to 1.42).

6. The NeuroStar System is safe and well-tolerated in patients with Major Depressive Disorder (See Section 4.3.1.8).

- The attrition rate in Study 44-01101 was 8% through week 4, few patients exited the study due to adverse events, and this was similar between the two treatment conditions.

- There were no suicides, deaths, or seizures reported. A specific examination of the development of emergent suicidal ideation showed that virtually all of these events were observed in the sham treatment condition, indicating that

active TMS does not worsen underlying depression, as has been reported to occur with pharmaceutical antidepressants in similar analyses.

- The most frequently reported adverse event was headache, which was mild to moderate in reported severity and occurred less frequently as the treatment course progressed. Headache was equally reported in active TMS and sham TMS groups.

- The most frequently-reported device-related adverse event was application site pain that occurred in about 35% of patients with active TMS versus 4% with sham TMS. Application site pain was generally mild to moderate in reported severity and occurred most frequently within the first week of treatment, and then dissipated substantially at later time points.

- Two reports of first degree scalp burns in the area of treatment were due to a manufacturing defect in the single-use coil shield that was corrected during the course of the study.

- There were no negative effects on cognitive function.

- There were no effects on auditory threshold.

7. The risk-to-benefit ratio for the NeuroStar System is favorable as compared to ECT; the NeuroStar System is substantially equivalent to the ECT predicate devices (See Section 6).

- A direct, within-study comparison of ECT and TMS efficacy cannot be performed. Obtaining such data from a direct head-to-head study of TMS and ECT is not methodologically or ethically feasible due to the differences between these treatments with regard to the safety issues associated with ECT, but not TMS treatment. Therefore, as agreed with the FDA, an estimate of the comparative profiles of TMS and ECT was performed by comparing the Study 44-01101 data to data from controlled clinical trials of ECT versus simulated ECT as reported in the literature. This data was obtained from a recent rigorous analysis reported in the UK ECT Review Group report. This report as well as data from two more recent ECT studies, the OPT-ECT study, (Sackeim, 2001) and ECT Community Study (Prudic, 2004) were used as reference ECT databases for efficacy and safety comparison. Due to the differences in the design and conduct of ECT and TMS trials, only general comparisons should be made between these treatments.

- The Neuronetics study population shows clinically meaningful and substantive overlap with the population treated with ECT with regard to clinical diagnosis, demographics, symptom severity, prior treatment failure and the level of treatment resistance. The ECT trials studied patients with

somewhat *less chronicity or treatment resistance* than those studied in the Neuronetics trials (See Section 6.3).

- The standardized effect size for the NeuroStar System in randomized, controlled Study 44-01101 is within the range of effect sizes calculated for randomized, controlled trials of ECT as evaluated in the UK ECT Review Group report; ECT effect size range HAMD17 = 0.17 to 1.42, Study 44-01101 HAMD17 = 0.55.

- The categorical clinical remission rates (HAMD24 < 11) observed after 6 weeks (end of acute phase) and 9 weeks (end of taper phase) in the Neuronetics open-label extension Study 44-01102 was 27.1% and 36.5%, respectively, which reach the lower range of HAMD24 remission rates observed in the open-label Community ECT Study (36.4% to 57.1%).

- The durability of effect at 1 month with pharmacotherapy after acute treatment with the NeuroStar System is comparable to that seen with ECT with pharmacotherapy, using the same definition for relapse (Neuronetics Study 44-01103 = 9.1% relapse, ECT = 4.5%-36% relapse rates). Furthermore, persistence of effect with NeuroStar TMS Therapy at 6 months after treatment, was sustained and compares favorably to that observed with ECT after 6 months (Neuronetics Study 44-01103 ~ 20% relapse, ECT ~ 50-63% relapse) (See Section 6.4).

8. The safety and effectiveness data submitted in Neuronetics premarket 510(k) notification support its proposed labeling and intended use (See Section 7).

- TMS Therapy as delivered by the NeuroStar System has proven efficacy and safety for the proposed indication of "treatment of patients with major depressive disorder" as defined by DSM-IV criteria.

- The safety and efficacy of the NeuroStar System for the treatment of MDD has been proven in a patient population that has the same demographic and disease characteristics as the population that is treated with ECT.

- NeuroStar System labeling, like ECT labeling, does not dictate a minimum level of treatment resistance. Efficacy has been shown in the difficult-to treat MDD population and the strong safety profile of NeuroStar TMS Therapy supports its use in the broader MDD patient population. This would include patients who may be intolerant to or otherwise not be expected to receive benefit from other antidepressant therapies.

- The proposed labeling for the NeuroStar System is for the treatment of MDD and does not include additional ECT indications as follows.

- ECT devices are also indicated for the treatment of patients with Major Depressive Disorder with psychosis or bipolar disorder. At the request of the FDA, these patient populations were excluded from Neuronetics studies so as to focus on the primary indication of Major Depressive Disorder. Therefore, treatment of patients with MDD with psychosis or with bipolar disorder is not part of the proposed indication.
- ECT devices are also indicated for patients with MDD with emergent suicidal symptoms. These patients were excluded from Neuronetics studies for safety reasons for the conduct of a randomized, sham-controlled outpatient trial. Therefore, Neuronetics is not seeking this additional label indication.
- The conditions regarding safe clinical use of the NeuroStar System are provided as part of the product labeling. The NeuroStar User Manual provides the proposed indication for use, warnings, contraindications, and precautions, as well as detailed step-by-step instructions on use of the device. All psychiatrists who will use the NeuroStar System will have training on the procedures and conditions of use which will be provided by Neuronetics.

SECTION 1. MAJOR DEPRESSIVE DISORDER AND TREATMENT OPTIONS

1.1. Major Depressive Disorder

Major Depressive Disorder (MDD) is a common, disabling and potentially lethal condition. It is estimated that by the year 2020, depression will be second only to heart disease in magnitude of disease burden as determined by disability-adjusted life years (Murray and Lopez, 1996). In the most recent epidemiologic estimate in the United States, the National Comorbidity Survey replication study, it was estimated that the lifetime prevalence of formally diagnosed major depression was 16.2% (Kessler, et al, 2003). Over a twelve month interval, in over half of all cases, the clinical significance was independently classified as either severe (38.0%) or very severe (12.9%). Notably, only about 10% were seen as mild.

In addition to its sheer prevalence, major depression rarely occurs as an isolated disease state, but frequently occurs in a comorbid manner with both psychiatric and medical illnesses. The presence of major depression also has an aggravating impact on the morbidity and mortality of a range of other medical conditions, including heart disease (Katon, 2003; Rugulies, 2002), cancer (Fawzy, 2003), HIV infection (Cook, et al, 2002), and diabetes mellitus (Eaton, et al, 1996). Indeed, there is essentially no health condition whose course is not adversely affected by untreated major depression. The impact of this illness is also seen when examining patterns of health resource use and their resulting financial costs, both direct and indirect (Katon, 2003). The concurrence of untreated or unrecognized major depression results in an excess utilization of health care resources among affected individuals, along with a substantial disruption in their productive work life. In many instances, health care visits are for the evaluation or treatment of presumed medical conditions that in fact represent untreated symptoms of the underlying major depression.

In summary, major depressive disorder is a common and potentially lethal disease that is frequently associated with comorbidities and related increases in healthcare costs.

1.2. Treatment Options for Major Depressive Disorder

Treatments for major depression are generally grouped into the biological therapies and the psychological therapies, with the most common approach to treatment combining these two modalities as clinically indicated in the individual patient. Regardless of the approach pursued, clinical outcomes to first line treatment for major depression remain modest at best. For instance, in randomized, controlled clinical trials of antidepressants used in a treatment-naïve or non-refractory patient population, approximately 50-60% of patients may be expected to achieve the symptomatic criterion of response at the end of 4-6 weeks of acute treatment (i.e., a reduction of >50% in total symptom score on a standardized rating scale, compared to the level seen on that scale at baseline).

Only one-third of such treatment-responsive patients will experience complete relief of illness, typically expressed as achievement of an *a priori*-defined remission score on a standardized symptom rating scale. Despite continued attempts with available antidepressant treatments, approximately 15-20% of patients will fail to receive clinical benefit from any currently available intervention, including ECT (Thase and Rush, 1997; Sackeim, 2001). This latter population is sometimes referred to as the treatment-refractory depressed patient population.

Among the biological interventions, antidepressant medication treatment typically serves as the initial step in treatment planning. Selective serotonin reuptake inhibitors such as fluoxetine (Prozac), sertraline (Zoloft), paroxetine (Paxil) or citalopram (Celexa) have largely replaced the older tricyclic antidepressants and monoamine oxidase inhibitors as drugs of first choice in most clinical settings in contemporary practice. In the event that a patient remains symptomatically ill after at least 4-8 weeks of exposure to appropriate doses of their starting antidepressant, the clinician may consider several more options such as changing medications to one of the same or a different chemical class, then to more complex medication regimens consisting of combination treatment or augmentation approaches. These progressively more complex pharmacotherapeutic regimens are also associated with increasing burdens of adverse effects.

Beyond mono-pharmacotherapy or combination pharmacotherapy, the most commonly considered alternative approach is ECT. Aside from clinical situations where suicidal ideation is emergent and immediately life-threatening, or in the setting of catatonic stupor that creates a medical crisis, ECT is often approached with apprehension, if at all due to the invasive nature of the procedure (i.e., requiring anesthesia, respiratory ventilation) and adverse effects, especially cognitive deficits. However, for those patients without other options, ECT can be effective (APA Committee on ECT, 2001).

Electroconvulsive Therapy (ECT) has been used as a therapeutic antidepressant since its introduction to clinical practice in 1938 and is considered the most effective antidepressant available. The ECT procedure involves the direct application of electrical current to the brain through the placement of electrodes on the surface of the head. It is generally accepted by practitioners that upon exposure of the brain to a sufficient amount of electrical energy, clinically meaningful antidepressant activity can be achieved in patients with major depression and in other clinical conditions, including some psychotic illnesses (APA Committee on ECT, 2001). During the application of ECT, an electrical seizure and an accompanying motor convulsion are intentionally produced, although it is generally recognized that the production of the seizure is not in and of itself sufficient for antidepressant effect. The clinical outcome appears to be related to the magnitude of the electrical dose applied above the amount needed to induce a seizure. In fact, the work of Sackeim and colleagues has clearly

established that induction of a seizure alone is itself insufficient to result in therapeutic effect with ECT (Sackeim, HA, et al., 1993).

Since its introduction, a number of important modifications in ECT treatment technique have been established to enhance its safety, including the use of general anesthesia, muscle paralysis, and improvements in cardiovascular monitoring (APA Committee on ECT, 2001). In addition, improvements in the design of ECT devices have permitted the administration of shorter duration electrical pulses. Overall, these modifications have led to significant changes in the practice of ECT, with accompanying reductions in the morbidity of the procedure itself (Sackeim, et al, 2000; APA Committee on ECT, 2001). Nevertheless, few practitioners would dispute the fact that ECT remains the most complex and poorly tolerated of all contemporary antidepressant techniques. In addition, although acutely effective, ECT effectiveness is not persistent and typically requires pharmacologic antidepressant maintenance for retention of acute benefit beyond 1 month. Additionally, for both patients and practitioners alike, there is a significant social stigma associated with ECT "shock therapy" that further limits its use.

Vagal Nerve Stimulation (VNS) has recently become another approved treatment option for patients with MDD (Sackeim, et al, 2001; Rush, et al, 2005). This treatment requires surgical implantation of a vagal nerve stimulator and electrode in the patient's chest and neck, respectively, which stimulates the vagal nerve in an effort to reduce depression symptom burden. VNS therapy is indicated for the adjunctive treatment of major depressive disorder for those patients who have failed to receive benefit from at least 4 failed antidepressant medication trials. Therefore, VNS provides another, albeit invasive, treatment option for the treatment refractory MDD patient.

In summary, the patient with major depressive disorder who has failed to receive benefit from pharmacologic monotherapy is a candidate for treatment with combination pharmacotherapy, ECT or VNS therapy. However, the lack of efficacy, invasiveness and/or poor side effect profile associated with these therapies are significant and many patients and physicians choose not to pursue these treatment options.

1.3. The NeuroStar System as a Therapeutic Option for Patients with MDD

Transcranial magnetic stimulation (TMS) has been developed as an alternative to the treatment options outlined above. TMS Therapy delivered by the NeuroStar System uses magnetic, rather than electrical energy, to induce an electrical field in a region of the brain associated with mood, i.e., the left prefrontal cortex, which results in stimulation of local neurons, leading to relief of depression symptoms (Burt, et al, 2002). The question in the medical community regarding the clinical utility of TMS was whether this localized stimulation was sufficient to cause

clinically-relevant symptom relief in the absence of seizure, as is required with ECT.

To date, a large number of clinical trials have been conducted to evaluate TMS in the treatment of MDD, including over 20 controlled studies and 5 studies comparing TMS and ECT. Most of these studies show antidepressant effects and for many controlled trials, these effects are significantly greater than sham treatment. However, these studies evaluated small to moderate patient numbers and treatment parameters vary considerably from study to study. In addition to these individual studies, there are now six published independent meta-analyses of the published or public TMS antidepressant literature, each differing in the articles included and the statistics used. The majority (4 of the 6 reports) have definitively concluded that active TMS exerts a statistically measurable and clinically notable effect greater than sham TMS. For example, Burt and colleagues examined 23 published comparisons for controlled TMS trials and found that TMS had a combined effect size of 0.67, indicating a moderate antidepressant effect. In the two reports that reported indeterminate conclusions, the authors examined a more restricted subset of published studies, and noted that the limitations in study design, patient selection, duration of treatment and treatment parameter selection limited the ability to draw meaningful conclusions.

More recent single-site randomized controlled clinical trials examining the efficacy of TMS in the treatment of MDD have used the largest sample sizes studied to date, with improved study designs and treatment parameters built on the knowledge learned from prior work (Avery, 2005; Fitzgerald, 2003, 2006). These studies have provided further support for the statistical and clinical significance of the antidepressant benefit of TMS. For example, in the report by Avery and colleagues, approximately 30% of patients treated with TMS achieved the categorical criterion of response using the HAMD24, and the overall standardized effect size for this outcome was 0.69.

It is generally thought that the current body of prior work has established the efficacy of TMS in proof-of-concept, however, the field has continued to lack a definitive, large, confirmatory, multisite, randomized, controlled, clinical trial. Neuronetics conducted a series of 3 clinical studies to determine the efficacy and safety of the NeuroStar System in the treatment of patients with Major Depressive Disorder as described further in Section 4. Unlike previous trials, Neuronetics Study 44-01101 used optimized TMS treatment parameters, as learned from the previous literature, rigorous study design and controls, and due to its large, multicenter design, allowed an evaluation of the generalizability of the safety and efficacy of TMS Therapy as delivered by the NeuroStar System when used as part of the psychiatrist's general clinical practice.

Randomized Controlled Trial:

- Study 44-01101 evaluated the acute safety and efficacy of the Neuronetics TMS System in adult outpatients in a 9-week, randomized, placebo-controlled multicenter clinical trial.

Extension Studies (Open-Label Treatment Conditions):

- Study 44-01102 enrolled patients from Study 44-01101 who failed to receive benefit from their randomized assignment in Study 44-01101 and was a 9-week, open-label study using the same treatment protocol as Study 44-01101.
- Study 44-01103 enrolled patients who met criteria for response in either Study 44-01101 or Study 44-01102 and evaluated maintenance of their acute clinical response over this 24 week, open-label continuation clinical trial.

The results of these studies show that TMS Therapy as delivered by the NeuroStar System is an effective, safe and well-tolerated antidepressant for the treatment for patients with major depressive disorder. TMS Therapy delivered by the NeuroStar System lacks the adverse effects associated with combination pharmacotherapy or ECT treatment or the invasiveness of VNS therapy. Furthermore, the acute response to TMS treatment can be effectively maintained in a clinically meaningful manner during a follow up period of up to 24 weeks.

As described in more detail in Section 5 of this Summary, TMS Therapy as delivered by the NeuroStar System provides clinically significant patient benefit while offering low risk due to its excellent safety profile with a risk-to-benefit ratio that is more favorable than currently available antidepressant treatments. Therefore, the NeuroStar System can serve to fill the large, unmet need for an effective, non-invasive antidepressant therapy with minimal adverse effects. The utility of the NeuroStar System as part of the psychiatrist's patient treatment plan is described in more detail in Section 5.

SECTION 2. REGULATORY CONSIDERATIONS FOR CLEARANCE OF NEURONETICS NEUROSTAR SYSTEM FOR THE TREATMENT OF MAJOR DEPRESSIVE DISORDER

2.1. Regulatory History

The NeuroStar System is a new device that has never been marketed in the United States. There are no devices that deliver transcranial magnetic stimulation (TMS) that are approved for marketing in the United States for the treatment of Major Depressive Disorder.

Neuronetics clinical studies were condu[redacted]s approved by the FDA under Investigational Device Exemption[redacted]The following Investigational Plan and data collection for Neuronetics' NeuroStar System for treatment of MDD was reviewed with FDA's Division of General, Restorative and Neurological Devices, and consisted of the following:

- Conduct of a single, randomized, sham-controlled multicenter clinical trial to demonstrate ***acute safety and effectiveness at 4 weeks*** of treatment with the NeuroStar System in patients suffering from major depressive disorder (Study 44-01101; maximum treatment duration was 6 weeks with primary endpoint at 4 weeks – see Section 4).
- Provision of data that evaluates the ***durability of the acute response***; (1) over a 3 week period (weeks 7, 8 and 9) during taper from TMS treatment to pharmaceutical antidepressant and, (2) after 4 weeks (week 13) without TMS treatment – see Section 4.
- Provision of ***safety data*** over 13 weeks, including data available for patients followed to 3 and 6 months post-TMS treatment – see Section 4.
- Comparison of clinical data from the Neuronetics studies to clinical data for ECT from the medical literature to ***demonstrate substantial equivalence by comparison of risk-to-benefit ratio*** to an FDA-approved predicate device as required under 21 CFR 807.87, and filing to the FDA of a Premarket Notification – see Section 6.

2.2. *A priori*-defined Outcomes Support Regulatory Clearance

Neuronetics' premarket 510(k) notification K061053 contains the above information in support of marketing clearance for the NeuroStar System. The data presented indicates that the NeuroStar System is ***safe and effective*** in the treatment of Major Depressive Disorder.

The primary efficacy outcome measure, the change in MADRS total score at the primary efficacy time point of 4 weeks, showed a trend towards statistical significance at a value of p=0.057. The mean change in value at the same time

point for the HAMD17 and HAMD24 total scores showed statistical significance (p=.006, p=.012, respectively). Importantly, all three key efficacy outcome measures, MADRS, HAMD17 and HAMD24, achieved statistical significance at 4 weeks for the categorical outcome of *response* (p=.045, p=.018, p=.030, respectively).

At the secondary efficacy time point of 6 weeks, a similar pattern of results was observed for mean change on the MADRS total score (p=.058), HAMD17 total score (p=.005), and HAMD24 total score (p=.015). For the categorical outcome of response at 6 weeks, all three scales continued to show statistical significance (MADRS, p=.007, HAMD17 p=.015, HAMD24 p=.042). In addition, the more stringent measure of categorical outcome, i.e., *remission*, was statistically significant for 2 of 3 scales (MADRS, p=.011, HAMD24, p=.012, HAMD17, p=.065).

Overall, the pre-specified analyses in Study 44-01101 display a preponderance of evidence for a statistically significant benefit of active TMS versus sham TMS in patients with major depressive disorder (See Expert Statistical Opinion, Tab 6).

For clinician-rated *a priori*-defined outcome measures, which used well-validated traditional efficacy instruments, 11 of 16 measures showed strong signals of statistical significance at 4 weeks and 13 of 16 measures were significant at 6 weeks. These significant effects, particularly on the categorical outcomes of response and remission, demonstrate that the effects of the NeuroStar System on depression symptoms are ***clinically significant*** for patient outcome.

Further discussion of the study data obtained from Neuronetics' randomized, controlled clinical trial Study 44-01101 in support of Neuronetics' claim of safety and effectiveness of the NeuroStar System in the treatment of MDD is provided in more detail in Section 4. This section also presents the results of the open-label Study 44-01102 which provides ***confirmatory evidence of the efficacy*** observed in the randomized, controlled trial 44-01101. Additional maintenance-of-effect data from the first 4 weeks of Study 44-01103 provide evidence of ***durability*** of acute response after cessation of TMS treatment.

2.3. Post-hoc Efficacy Analyses for Study 44-01101

In addition to the planned analyses described above, several *post-hoc* analyses were conducted to further evaluate the efficacy data.

Neuronetics' *Post-hoc* Analysis

Adjustment for MADRS Baseline Imbalance:

- Analysis of the MADRS primary efficacy endpoint determined that it was substantially influenced by a baseline imbalance in symptom severity between

active and sham TMS groups that was not observed with other outcome measures (HAMD17 and HAMD24). This was due to the use of a study entry floor for HAMD scales that was not instituted for the MADRS. Correction of the baseline imbalance resulted in statistical significance of the MADRS primary outcome measure (p=0.038) while all other secondary clinician-rated outcome measures remained unchanged. This was the only supplementary *post-hoc* analysis provided by Neuronetics as part of the premarket 510(k) notification - See Section 4.3.1.5.1.

FDA-requested *Post-hoc* Analyses

Integrity of the Study Blind:

- FDA requested an evaluation of the integrity of the study blind by determining if a relationship existed between clinical outcome and any adverse events, in particular those events that occurred more frequently with active than sham treatment, i.e., application site pain, that could have led to study unblinding. As described in Section 4, three types of analyses were conducted to examine the relationship between these adverse events and clinical response (See Section 4.2.4.2).

- Comparison of MADRS total score at week 4 versus occurrence of adverse events associated with application site pain and other characteristic terms indicating discomfort. This analysis showed *no relationship* between these adverse events and outcome.

- Comparison of MADRS and HAMD scores with occurrence and severity of the above adverse events during the first week of treatment and mean change at week 4. A total of 48 comparisons were conducted and no *statistically significant observation was found in the most commonly occurring term, "application site pain"*. In only 5 instances were any p-values observed to exceed the $p < 0.05$ threshold which occurred for adverse events that constituted <5% of the sample.

- MADRS and HAMD total scores were examined in two separate methods of covariate analyses. The first method was performed using a covariate analysis with the term "application site pain" occurring during the first week of treatment in the ANCOVA model. The second method used a covariate analysis with a term comprised of any of pain or discomfort terms stipulated by the FDA and occurring during the first week of treatment, as a variable of the ANCOVA model.

- These covariate analyses are statistically problematic. This is due to the fundamental fact that the occurrence of pain or discomfort, even if restricted to that occurring during the first week of treatment, is a post-randomization event, and is not an intrinsic pre-randomization characteristic that would be appropriate for use as an ANCOVA variable. As a result, the *expected* effect

for an adverse event that occurs with outcome would be the reduction in signal.

- Nevertheless, with the first method of covariance analysis with the use of the "application site pain" term, which is the event that was most variable between active and sham treatment groups, the *overall conclusions of the study remain virtually unchanged* including the P values for all outcome measures. In the second model, the effect for the MADRS is eliminated (p=.227), reduced for the HAMD24 (p=.054), and the HAMD17 outcome remains statistically significant (p=.020).

Clinical Outcome by Level of Prior Treatment Failure:

- FDA requested an evaluation of response in Study 44-01101 by level of treatment resistance as measured by the ATHF (Antidepressant Treatment History Form). The standardized effect sizes and associated p-values were provided for the primary efficacy variable, MADRS, for the overall population and for subgroups ATHF Levels 1-4. The standardized effect size for NeuroStar TMS Therapy for the overall sample is 0.39, using MADRS mean total score. Subset analysis by ATHF Level showed an inverse correlation between treatment effect and treatment resistance which is consistent with the large published antidepressant literature where a diminished response is observed in the increasingly treatment-resistant populations – See Section 4.3.1.5.2.1, Table 15. This relationship has been demonstrated for all known antidepressant treatments including pharmaceuticals and ECT, (i.e., see Section 5). For ATHF Level 1 patients, the primary efficacy endpoint, mean change in MADRS total score at 4 weeks, reached statistical significance at p=.001. Additionally, as described in Section 4, the standardized effect size in the ATHF Level 1 group (standardized effect size = 0.94, p=.001) was greater than seen in the overall sample (standardized effect size = 0.39, p=.057) and is comparable to that reported in the literature for the predicate device, ECT (effect size = 0.9 (range 0.17-1.42) – See Section 6.3.

Clinical Outcome by Level of Response:

- FDA requested an evaluation of the level of response in Study 44-01101 in active and sham treatment arms as grouped by <0% (worsening), 0-25%, (partial response), 25-50% (partial response) and >50% (response). The results of this analysis demonstrated that the overall pattern of clinical benefit was consistently evident in favor of active TMS compared to sham TMS when viewed across the full response spectrum from mild improvement to full response. In general, there was a over-representation of active TMS patients compared to sham TMS as the level of response increased, while there was an over-representation of sham TMS compared to active TMS in the category of worsening of symptoms.

Durability of Response:

- FDA requested an evaluation of the durability of response over the course of Study 44-01101 by evaluating the change in mean total score from baseline at weeks 4, 6 and 9. These analyses showed that the benefit obtained at week 4 was maintained over weeks 6 and 9. Additionally, in a within-group comparison of week 4 and week 9 scores, the active TMS group showed statistically significant improvement from week 4 to week 9 (MADRS total mean score, p= 0.016) whereas the sham TMS group did not (p=.915). These data support two key findings: (1) the acute effect associated with active TMS is durable, and (2) the sustained effect in the taper phase cannot be ascribed to medication treatment alone given that the active TMS group had a statistically significantly greater degree of improvement at 9 weeks than was seen in the sham TMS group – See Section 4.3.1.6.1, Table 17.

Clinical Significance:

- FDA requested further discussion of the clinical significance of the clinical efficacy data obtained with the NeuroStar System in Neuronetics clinical studies. As described in Section 5, the NeuroStar efficacy data are consistent with the large, well established registration clinical trial database for pharmaceutical antidepressants. This database is critically important because it defines the standard for clinical significance and forms the basis for FDA approval of all currently marketed antidepressants in the United States. Section 5 also demonstrates that the risk-to-benefit profile of the NeuroStar System is favorable as compared to FDA-approved antidepressant treatments.

2.4. NeuroStar System Safety

The excellent safety profile of the NeuroStar System deserves to be highlighted. Neuronetics clinical studies evaluated the following key safety metrics: type and

frequency of spontaneous adverse events, effects on cognitive function, and the incidence of emergent suicidal ideation.

Spontaneous Adverse Events:

- Treatment tolerability was excellent and patient compliance with the treatment protocol was high. There was a near absence of systemic adverse events. The most frequent adverse event was headache that was equally represented in active TMS and sham TMS groups. Application site pain was the most frequent device-related adverse event. Both of these events were generally mild to moderate in nature and attenuated over time.
- A comparison of adverse event data from the Neuronetics Study 44-01101 with those reported in registration trials for current U.S. marketed pharmaceutical antidepressants, demonstrated markedly reduced levels of systemic side effects with acute TMS Therapy relative to pharmacologic therapy.

Cognitive Function:

- No negative effects on cognitive function were observed.

Emergent Suicidal Ideation:

- There was no increase of emergent suicidal ideation in patients treated with active TMS. Virtually all instances of emergent suicidal ideation were observed in the sham TMS group.

2.5. Risk-Benefit and Substantial Equivalence of the NeuroStar System

Efficacy and safety of the NeuroStar System are important in considering the risk-benefit of TMS Therapy in relationship to FDA-approved antidepressants, including the predicate device, ECT. As stated in Section 6, ECT is an effective treatment for MDD but suffers from substantial treatment-related toxicity. Treatment with the NeuroStar System in this same patient population, offers many patients an effective, clinically-significant treatment without any of the safety concerns of ECT, or any of the systemic side effects associated with pharmaceutical antidepressant treatment. This favorable risk-benefit ratio of the NeuroStar System as compared to current FDA-approved antidepressants, including ECT, is central to its demonstration of substantial equivalence to ECT devices, and is substantiated by the data presented in this premarket 510(k) notification, and therefore, deserves FDA clearance as a therapeutic option for patients with major depressive disorder.

SECTION 3. DESCRIPTION OF THE NEUROSTAR SYSTEM

3.1. General Description of the NeuroStar System

The NeuroStar System is the commercial name that has been applied to Neuronetics commercial TMS system. The clinical version of the NeuroStar System that was used in Neuronetics' clinical studies is called the Model 2100 TMS System and is also referred to in technical documentation as the Callisto TMS System. Equivalency testing verified the comparability of the clinical and commercial devices (See CD-ROM Attachment 7 – 510(k) Section 11: Device Description)]. For ease of review, the NeuroStar System name is used synonymously with the clinical device name throughout this document.

The NeuroStar System is a class III medical device predicated on regulatory clearance for substantial equivalence to class III, ECT devices that are indicated for use in the treatment of major depressive disorder.

The NeuroStar System is designed to be a user-friendly, non-invasive tool for the modulation of cortical neurons and is intended for the treatment of psychiatric disorders. This premarket 510(k) notification seeks clearance of the NeuroStar System for the ***treatment of major depressive disorder*** (MDD).

Patient treatment using the NeuroStar™ System is ***by prescription only***. The NeuroStar System will be operated only by trained, licensed medical professionals in both inpatient and outpatient settings including physician offices and clinics, free-standing psychiatric hospitals, and general medical/surgical hospitals with psychiatric units.

The NeuroStar System is a ***computerized electromechanical instrument for magnetic stimulation*** that produces and delivers brief duration, rapidly alternating, or pulsed, magnetic fields to induce electrical currents directed at spatially discrete regions of the cerebral cortex.

The principles of operation of the NeuroStar System for its intended use of treatment of major depressive disorder are described in detail in the NeuroStar User Manual which is provided in CD-ROM Attachment 15. As discussed there, pulsed magnetic fields are applied to the cortex, inducing electrical currents that produce local neuronal depolarization. Neuronal depolarization is associated with various physiological changes in the brain that are associated with the symptomatic relief of depression.

The safety of the NeuroStar System was evaluated in Neuronetics' clinical trials. In addition, safety considerations relevant to the use of a high intensity magnetic field are discussed below and are prominently presented in the NeuroStar User Manual.

3.2. Physical Description of the NeuroStar System

The NeuroStar System is an integrated system consisting of a combination of hardware, software, and consumable supplies:

- A console and gantry;
- A graphic user interface;
- Software for the console and graphic user interface;
- Ferromagnetic coil for NeuroStar System;
- SenStar™ Treatment Link (referred to as "SuperShield" in 510(k) K061053), a disposable unit placed on the face of the coil to reduce local scalp stimulation and incorporating a contact sensor (added to ensure proper magnet placement against the scalp) and field detect feature (added to detect correct magnetic field pulse).
- A separate Patient Data Management System (PDMS) allows and facilitates recording and retrieval of patient and treatment information and communication of data among multiple NeuroStar System units.

The components of the NeuroStar System and their functions are summarized in Table 1. The NeuroStar System and its components are shown in Figure 1.

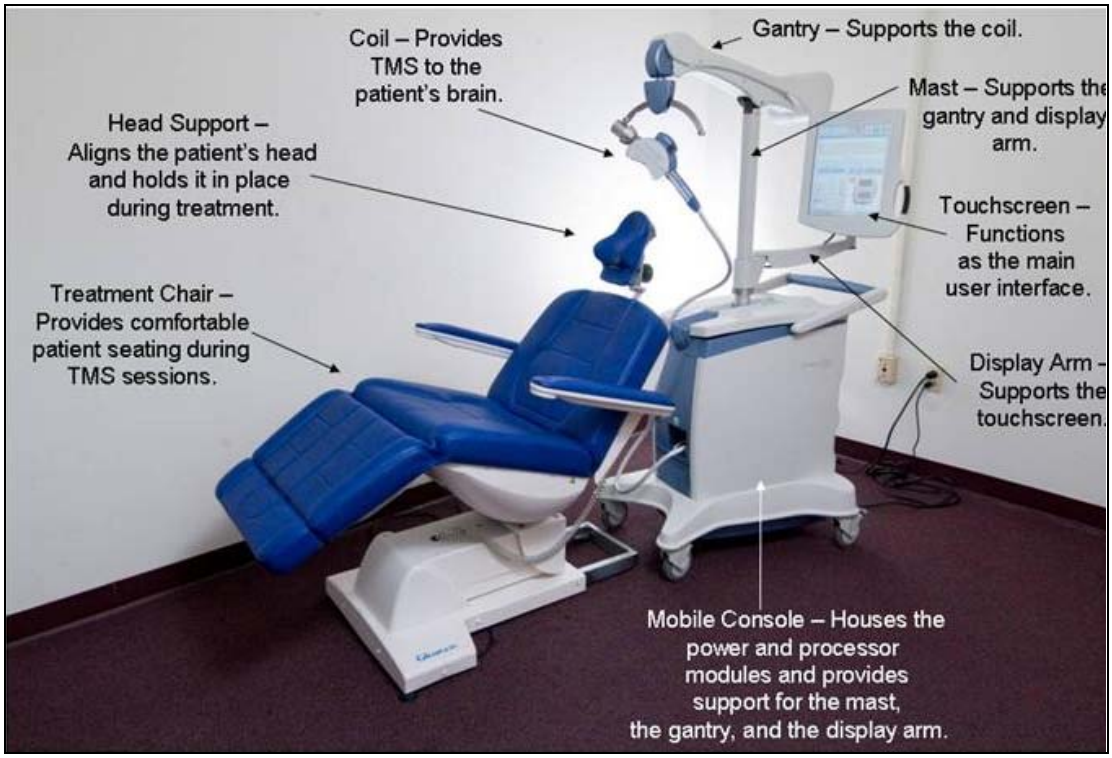


Figure 1. Main Components of the NeuroStar System

Table 1. NeuroStar System Components and Functions

Hardware	• The NeuroStar System wheeled-base console contains and integrates the various subsystems into a single package. • System processor and power modules (electronic subsystems that control the user interface and system diagnostics; and generate TMS pulses). • Magnetic coil (i.e., electromagnet) converts electrical pulses from the power module into magnetic pulses for stimulation of the patient's cortex • Magnetic coil gantry and mast that position and suspend the Coil for treatment delivery. • Coil interface electronics (part of coil assembly) manages all the hardware interfaces between the coil and external components of the mobile console module via a microprocessor-based circuit board that controls components such as brakes and sensors and facilitates manual pulsing and manual single pulse adjustments. • LCD-ROM touch screen display and integrated display arm (attached to the mobile console), provides the primary user interface for the operation of the system. • Treatment chair, head support unit, and positioning pads act together to align the patient's head with the mobile console and coil and support the patient's head during a treatment session.
Proprietary System Software	• Provides user interface. • Controls various subsystems. • Organizes work flow to deliver TMS treatments.
Consumables	• SenStar, the key disposable component, which is a one-time use component that: o Provides contact sensing (between coil and patient's scalp) to aid positioning, magnetic field detection (to verify pulse strength) o Reduces the magnetic field strength at the scalp in an effort to improve treatment comfort. o Must be in-place and valid in order for the NeuroStar System to operate. o Detects magnetic field strength (to verify correct delivery)
Patient Data Management System (PDMS)	• Proprietary software that runs on a standard personal computer that is connected using wireless or cabled Ethernet to one or more NeuroStar System mobile consoles that are running the system software. • Provides capabilities to enter and retrieve patient and treatment information and history, and to manage users and passwords for all connected consoles. • Patient data (demographic data, patient depression scores, patient reports, motor threshold data and treatment histories from all connected NeuroStar System units) can be entered or viewed.

The manufacturing of the NeuroStar System and its components is handled by contract manufacturers or packagers that are qualified under Neuronetics' quality management system. Neuronetics' quality management system complies with FDA's Quality System Regulations (21 CFR Part 820) and is certified for compliance with ISO 13485:2003, "Medical devices - Quality management systems - Requirements for regulatory purposes".

3.3. Device Safety

The clinical safety of the NeuroStar System was evaluated in Neuronetics' clinical studies, the results of which are presented in Section 4. Contraindications, warnings and precautions related to clinical use of the NeuroStar System are provided in the NeuroStar User Manual (CD-ROM Attachment 15). Safety considerations regarding the use of a high intensity magnetic field are as follows:

Magnetic Field Exposure:

- A safety-related consideration is the potential effect of magnetic field exposure. An expert analysis, provided in Attachment 5, was conducted of the potential effects of exposure to TMS magnetic fields including: induced translational forces, interaction with implanted devices, effects on nerve conduction velocities, voltages induced due to flowing blood or moving particles, comparison to power line exposure, chemical reaction rates, proton tunneling in DNA, single photon interactions, cardiac stimulation, T-wave modifications, mutagenic effects, and temperature rise due to power deposition.

- The analysis concluded that there appears to be little or no evidence of harmful effects from magnetic field exposure up to 8 T. Theoretical concerns start as low as 10 T. These and other data led the U.S. Food and Drug Administration to consider static fields below 4 T to be a non-significant risk. The magnetic field of the NeuroStar System is well below these levels in the cortex. Also, the time varying nature of the TMS field is of significantly low frequency that translational forces and tissue heating are negligible. In summary, experimental and observational evidence indicate that exposure to the magnetic fields similar to those produced by the Neuronetics' TMS coil is not associated with long term adverse health effects.

Electromagnetic and Electromechanical Safety:

- Safety Testing has been completed on a fully functional NeuroStar System which has shown compliance with the requirements of the EN60601-1 standard which covers electromagnetic and electromechanical safety.

Effects of the Magnetic Field on Other Devices:

- The magnetic field of the NeuroStar System may cause movement or impact the functioning of implanted medical devices if they are in relative proximity to the TMS coil. Therefore, use of the NeuroStar System is contraindicated in patients implanted with these devices – see Attachment 15. Contraindications, warnings and precautions for patients to be treated with the NeuroStar System are prominently discussed in the NeuroStar User Manual (CD-ROM Attachment 15).

SECTION 4. NEUROSTAR SYSTEM INVESTIGATIONAL PLAN, STUDY DESIGNS AND CLINICAL STUDY RESULTS

4.1. NeuroStar System Investigational Plan

The clinical development program was designed to demonstrate safety and effectiveness of Neuronetics' NeuroStar System in the treatment of Major Depressive Disorder and to provide data in support of marketing clearance as described in Section 2. The investigational plan consists of three integrated clinical protocols as shown in Figure 2.

In brief, the efficacy of the NeuroStar 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 extension study with the NeuroStar System in *Study 44-01102*.

The maintenance of an acute clinical response to the NeuroStar 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 relationship of Studies 44-01101, 44-01102 and 44-01103 and patient allocation are summarized in Figure 2.

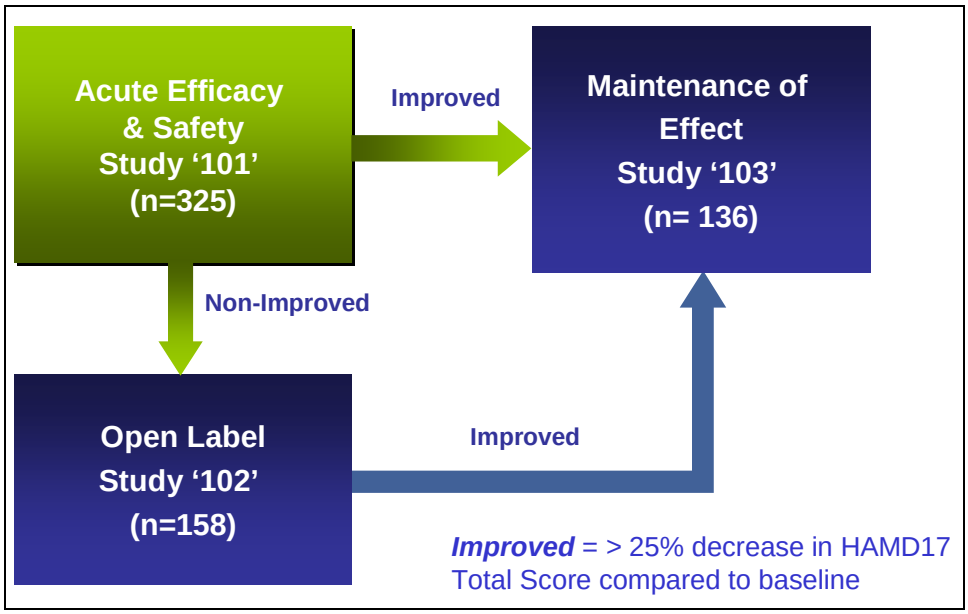


Figure 2. Neuronetics' Clinical Studies and Patient Allocation

4.2. Design Considerations of Study 44-01101

Neuronetics' clinical Study 44-01101 that was conducted in support of premarket 510(k) notification K060153 was designed consistent with current regulatory and clinical research practices used in the design of pharmaceutical antidepressant studies. In particular, the following items were considered critical to the overall study design and its implementation.

- Primary and Secondary Efficacy Outcome Measures
- Diagnostic, Symptom Assessment, Functional Status and Quality of Life Instruments
- Safety Outcome Measures
- Study Blinding

4.2.1. Primary and Secondary Efficacy Outcome Measures

The primary and secondary outcome measures that were collected and analyzed in Neuronetics studies are shown below in Table 2 in their *a priori*-planned prioritized order of testing and relative order of importance in study outcome. All are well-validated instruments that are typically used in clinical studies of pharmaceutical antidepressants (see Section 4.2.2 below). These instruments are described in detail in the study protocols provided in the Final Study Report for Study 44-01101 (CD-ROM Attachment 11).

The *a priori*-declared primary efficacy time point was at 4 weeks and was based on previous literature of TMS treatment outcomes over 1-4 weeks of treatment. Supportive information was obtained from the secondary efficacy time point at 6 weeks.

Table 2. Primary Outcome Measure and Secondary Outcome Measures in Protocol 44-01101 and Their Sequential Order of Importance in Testing

Measure	Description
Primary Outcome Measure	Evaluate the antidepressant effect of treatment with the NeuroStar 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) 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 2) The last post-treatment total symptom score on the 17- Item Hamilton Depression Rating Scale (HAMD17) through week 4 and week 6 of the acute treatment phase

Measure	Description
	3) The total symptom score on the MADRS for the last post-treatment value observed through week 6 of the acute treatment phase 4) Categorical outcomes of response (percent of patients achieving 50% reduction on each of the MADRS, HAMD24, and HAMD17 total symptom scores at the last post-treatment visit through week 4 and week 6 of the acute phase), 5) Health outcomes scores from the Medical Outcomes Study Short Form 36-Item Questionnaire (SF-36, v1) and the Quality of Life, Enjoyment and Satisfaction Questionnaire (Q-LES-Q) at the last post-treatment visit through week 4 and week 6 6) 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 7) Factor scores derived from the HAMD17 including: Anxiety/Somatization (sum of items 10, 11, 12, 13, 15, 17), Core Factor (sum of items 1, 2, 3, 7, 8), Maier (sum of items 1, 2, 7, 8, 9, 10), Gibbons (sum of items 1, 2, 3, 7, 9, 10, 11, 14), Retardation (sum of items 1, 7, 8, 14), and Sleep (sum of items 4, 5, 6) using the last post-treatment value through week 4 and week 6 8) The total score on the Inventory of Depressive Symptoms – Self Report version (IDS-SR), using the last post-treatment value through week 4 and week 6 9) The Clinical Global Impressions – Severity (CGI-S) score, using last post-treatment value through week 4 and week 6 10) The Patient Global Impressions – Improvement (PGI-I) score, using last post-treatment value through week 4 and week 6

4.2.2. Diagnostic, Symptom Assessment, Functional Status and Quality of Life 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 NeuroStar System. All instruments used are well-accepted and psychometrically valid psychiatric assessments, and are summarized in Table 3, and include both clinician-rated and patient-reported outcome measures. A summary of the key features of the MADRS and HAMD depression rating scales, which are the main clinician-rated efficacy instruments used in the Neuronetics studies, follows this Sponsor Executive Summary in Tab 5 of this Panel Package.

Table 3. Diagnostic, Symptom Assessment, Functional Status and Quality of Life Instruments Used in Study 44-01101

Assessment Tool	Description
Psychiatric Diagnostic Interview - Structured Clinical Interview for the DSM-IV (SCID-IV) (First, et al, 2002)	- 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) (Sackeim, 2001)	- 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) (Montgomery and Asberg, 1979) - Hamilton Depression Rating Scale (HAMD), 24-item and 17-item versions (Hamilton, 1960) - Clinician Global Impressions – Severity of Illness (CGI-S) (Guy, ECDEU Assessment Manual, 1976)	- 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.
Patient-Reported Symptom, Quality of Life, and Functional Status Assessments - Inventory of Depressive Symptoms – Self Report version (IDS-SR) (Rush, et al, 1996) - Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (Q-LES-Q) (Endicott, 1993) - Medical Outcomes Study Short Form – 36 Item Questionnaire, version 1 (MOS SF-36) (Ware and Kosinski, 2005) - Patient Global Impressions – Improvement of Illness Scale (PGI-I) (Guy, ECDEU Assessment Manual, 1976)	- 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.

Assessment Tool	Description
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.

4.2.3. Safety Outcome Measures

In all Neuronetics studies, 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 coded by staff at Neuronetics' contracted vendor for electronic data capture (EDC) using the current version of the Medical Dictionary for Regulatory Activities (MedDRA). 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 NeuroStar System.

Additional targeted safety assessments included assessment of cognitive function and auditory threshold. Auditory threshold was assessed using a desktop audiometer with a standard test sequence that examined the threshold decibel level at which a pure tone signal could be perceived by the patient. 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 (Loo, 2001; Pascual-Leone, 1992; Pascual-Leone, 1993). 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 (APA Committee on ECT, 2001; Sackeim, et al, 1993). 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 (Sackeim, et al, 1993). These specific measures are shown in Table 4. 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.

Table 4. Cognitive Function Testing Instruments for Neuronetics Studies 44-01101, 44-01102, 44-01103

Assessment Tool	Description
Modified Mini Mental Status Examination (MMSE) (Folstein, et al, 1975)	This instrument assesses global cognitive function in several major neuropsychological domains
Buschke Selective Reminding Test (BSRT) (Buschke, 1974)	This test evaluates short-term memory using immediate and delayed recall of common word lists
Autobiographical Memory Inventory-Short Form (AMI-SF) (Lisanby, et al 2000)	This interview assesses the integrity of long-term memory functions by examining the ability to recall basic autobiographical information at post-treatment timepoints that were obtained prior to the start of treatment

4.2.4. Study Blinding

4.2.4.1. Measures to Protect the Study Blind

Maintaining the integrity of the study blind is one of the most critical issues in the conduct of a randomized controlled clinical trial. Neuronetics Study 44-01101 was innovative in several important design considerations directed at minimizing the likelihood of penetrance of the study blind. Several of these design considerations and the resulting measures taken to ensure the integrity of the study blind are shown in Table 5 and are described in more detail in the Final Study Report for Study 44-01101 (CD-ROM Attachment 11). As outlined below, the preponderance of evidence suggests that differences between the active TMS and sham TMS treated groups were due to a TMS effect and not due to a failure of the study blind.

Table 5. General Consideration in Study Design and Conduct To Ensure Adequate Integrity of Study Blind

Method	Measures
Method of sham coil design and randomization	▪ Active and Sham coils identical in external appearance and labeled as “B” or “C” coil ▪ Coils acoustically matched ▪ Randomization controlled by electronic “smart cards” that could only be read by the stimulator making treatment assignment blind to site staff
Triple blind: Patient, Treaters, and Independent Raters	▪ Patients were “treatment-naïve” for TMS ▪ Treaters collected adverse events but did not rate patients ▪ Rater was blinded to the entire TMS treatment setting and to adverse event collection

Method	Measures
	▪ A site-blind independent expert demonstrated no evidence of site-rater bias
Blinding of the primary endpojnt and key study design criteria	▪ Primary endpoint variable (MADRS) and timing (Week 4) were concealed to the investigator and staff ▪ Criteria used to permit patient entry into follow-on studies were blinded to investigator and staff
Adverse event profile active vs sham difference	▪ Adverse events were comparable in magnitude to the observed difference seen in pharmaceutical antidepressant studies (see below) ▪ Common adverse events were, in general, mild to moderate in severity, and transient ▪ Low discontinuation due to adverse events ▪ No relationship between adverse event profile and study outcome (data discussed below)

4.2.4.2. Post-hoc Analysis of the Study Blind (As Requested by the FDA)

At FDA's request, *post hoc* analyses were conducted to evaluate the integrity of the study blind in Study 44-01101. These analyses evaluated whether response rates observed for active TMS vs sham TMS groups could be associated with device-related adverse events that could have the potential to unblind the patient or treater.

As shown in Table 5 above, patients were TMS-treatment naïve and did not know what active TMS treatment would encompass. In addition, adverse events were collected by the TMS treater who did not conduct patient ratings. Activities of the TMS treater and the efficacy rater were always kept independent and separate from each other, and patients were instructed not to discuss any adverse events with raters. In addition, the choice of primary efficacy outcome measure (MADRS) and timing (Week 4) were blinded from the investigator and staff. These restrictions served to further protect the study blind and mitigate the impact of any device-related observations by patient or treater, or knowledge of efficacy measurement or timing, on treatment outcome.

Controlled clinical trials of antidepressant medications typically demonstrate adverse event differences between placebo and active treatment groups. 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%) (data obtained from current package labeling). This range of adverse event differential is similar to the magnitude of difference seen in the Neuronetics studies. The stipulation applied in Study 44-01101, which required that study raters remain distinct and isolated from adverse event data collection, is a method that is rarely, if ever, used in registration trial of pharmaceutical antidepressants.

In Study 44-01101, the most frequent device-related adverse event was application site pain which was present in 35.7% of patients receiving active TMS vs 3.8% receiving sham TMS across the entire acute treatment phase. Application site pain was reported as mild or moderate in intensity in the majority of patients. 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 half the incidence seen during the first week, and differences in incidence between sham TMS and active TMS treated patients were minimal (see Section 4 for a listing of adverse events by treatment arm). These observations suggest that 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 pharmaceutical antidepressants.

The FDA requested that Neuronetics evaluate the possibility that the following adverse events could have led to study unblinding and thereby influenced the efficacy outcomes reported in Study 44-01101. These events were:

- Application site pain
- Eye pain
- Facial pain
- Jaw pain
- Toothache
- Application site discomfort
- Muscle twitching

Three types of analyses were conducted to examine the relationship between these adverse events and clinical response using the MADRS total score as follows:

- Analysis of the relationship between the occurrence of these adverse events during the first week of treatment and categorical outcome of clinical response at week 4.

- Analysis of the relationship between the occurrence and severity of these adverse events during the first week of treatment and mean change at week 4.
- Analysis of covariance with the inclusion of these adverse event terms as a covariate in the model.

Analysis #1. Relationship of Incidence of Adverse Event (Week 1) and Categorical Outcome (Week 4):

To perform this analysis, all patients who experienced *any one or more* of the MedDRA noted terms above were identified and then separately, those patients who met the response criterion at week 4 on the MADRS were identified. These patient data were separated by treatment group, and the relationship between the experience of pain or discomfort and the categorical outcome of MADRS response at week 4 was determined. The results, shown in Figures 3 and 4, indicated that there is no relationship between reports of "Pain/Discomfort", as defined using the group of terms cited above, and treatment response (≥50% reduction in MADRS score at week 4) in either TMS treatment arm (Active TMS Treatment, p=1.000, Sham TMS Treatment, p=0.301).

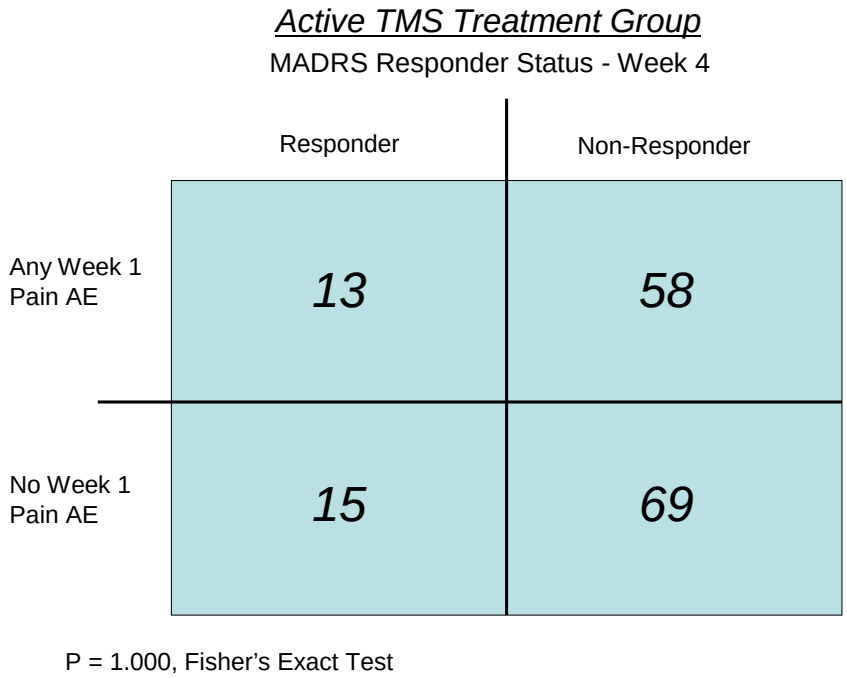


Figure 3. Absence of Relationship Between the Presence of Any Pain-Related AE at Week 1 and MADRS Responder Outcome at Week 4 – Active TMS Group

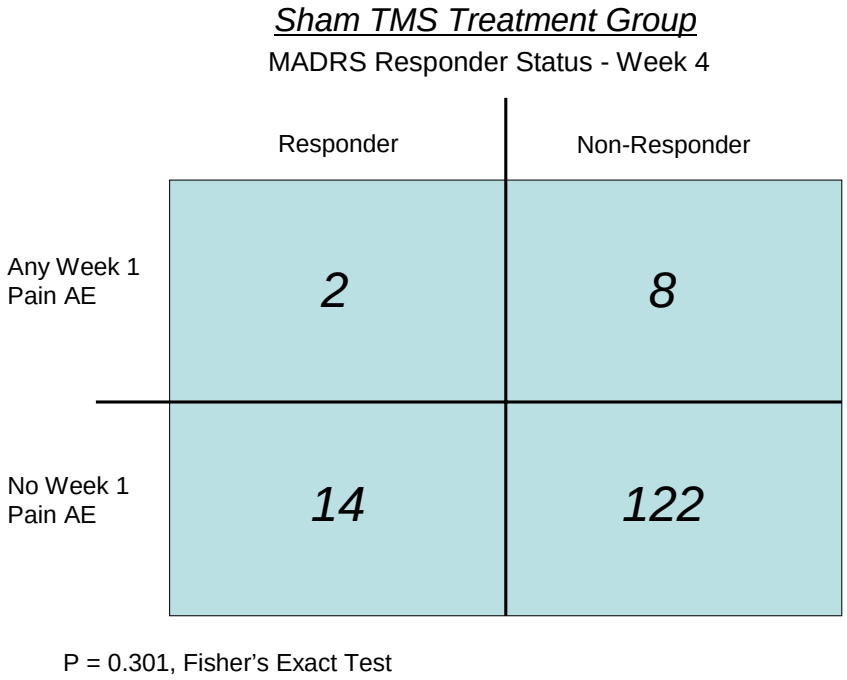


Figure 4. Absence of Relationship Between the Presence of Any Pain-Related AE at Week 1 and MADRS Responder Outcome at Week 4 – Sham TMS Group

Analysis #2 Relationship of Incidence of Adverse Event (Week 1) and Continuous Outcome (Week 4):

Two analyses were conducted. In the first, a Spearman correlation coefficient was determined for the relationship between mean change on the MADRS total score and the severity of the adverse event term (four categories: none, mild, moderate, and severe). In the second analysis, an ANCOVA model identical to that used in the *a priori*-stipulated ANCOVA analysis was used to examine the difference in mean baseline to endpoint change of the MADRS total scores within each treatment group separately and also for the total group. Separate analyses were conducted for the MedDRA-defined terms: "application site pain", "eye pain", "facial pain", "jaw pain", "toothache", "application site discomfort", "muscle twitching", and for the aggregation of any of the above pain/discomfort terms. Tables 6 and 7 below provide the results of these analyses for application site pain, the most common of these adverse events, and for the aggregate of pain/discomfort terms.

Table 6. MedDRA Preferred Term and Week 4 Outcome on MADRS Total Score – Relation of Presence and Severity of Term to Week 4 Efficacy Outcome

Application Site Pain

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

Table 7. MedDRA Preferred Term and Week 4 Outcome on MADRS Total Score – Relation of Presence and Severity of Term to Week 4 Efficacy Outcome

Any Pain/Discomfort Term

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

A total of 48 comparisons were conducted in these analyses. There was ***no statistically significant observation in the most commonly occurring term, "application site pain"***. In only 5 instances were any p-values observed to exceed the $p < 0.05$ threshold. In those instances, the strength of the observed p-value ranges from 0.010 to 0.047. Of these five instances, 3 were observed in association with the term "eye pain" that occurred in less than 3% of the overall sample (n = 9). One of the 5 instances was observed in association with the term "application site discomfort" that occurred in less than 5% of the overall sample (n = 15). In the last instance, only a weak inverse correlation ($r = -0.12$, $p = .034$) was observed when the aggregate terms were used. This correlation was not significant in the two treatment groups when examined separately. Moreover, when the aggregated pain terms were used, there was no statistically significant relationship between the mean change in MADRS total score at the primary outcome time point of week 4 and the presence or absence of pain/discomfort during week 1.

It is worth pointing out that prior clinical studies with TMS have shown that the most commonly occurring adverse event associated with TMS therapy is headache, and is usually over-represented in the active TMS treatment condition. In Study 44-01101, headache was observed with equal frequency in the active TMS and the sham TMS groups, and assessment for causality by the reporting investigators was equally assigned to the active and sham treatments. This suggests that the sham methodology implemented in Study 44-01101 effectively improved on the prior methods of sham TMS that were used in the literature. Furthermore, the equal occurrence of headache in both treatment groups served to further obfuscate the ability of the patient to discriminate active TMS from sham TMS treatment conditions.

Analysis #3 Analysis of Covariance (Pain Included as a Covariate)

MADRS total score, and the key secondary outcome measures, HAMD24 total score and HAMD17 total score were examined in two separate methods of covariate analyses. The first method was performed using a covariate analysis with the term "application site pain" occurring during the first week of treatment in the ANCOVA model. The second method used a covariate analysis with a term comprised of any of the pain or discomfort terms stipulated by the FDA and occurring during the first week of treatment, as a variable of the ANCOVA model.

In the first method of covariance analysis with the use of the "application site pain" term (the most prominent of the pain/discomfort adverse event terms observed in the study), the overall conclusions of the study remain virtually unchanged. Specifically, the strength of the effect on the MADRS is similar to the *a priori*-specified model ($p = .060$). For the HAMD24 and

the HAMD17, the strength of the statistical effect is identical to the *a priori* planned analysis (p=.012, p =.006, respectively).

In the second model, the effect for the MADRS is eliminated (p=.227), while for the HAMD24, there is a statistically notable trend (p=.054), and the HAMD17 outcome remains statistically significant (p=.020).

In contrast to the first two analyses presented above that examine adverse events and their relationship to outcome, these covariate analyses are statistically problematic. They preclude any meaningful interpretation of potential causal relationship between outcome and the presence or absence of pain or discomfort during the first week of treatment. This is due to the fundamental fact that *the occurrence of pain or discomfort, even if restricted to that occurring during the first week of treatment, is clearly a post-randomization event, and is not an intrinsic pre-randomization characteristic that would be appropriate for use as an ANCOVA variable. A post-randomization event, such as an adverse event, cannot be appropriately utilized to answer the question of causality when beneficial clinical effect and the presence of the adverse event are both expected outcomes of allocation to active treatment.* Therefore, the inclusion of such a post-randomization characteristic as a covariate in an analysis of variance can only be *expected to diminish* the strength of the overall treatment group difference

In summary, Study 44-01101 was designed with measures taken to protect the integrity of the study blind. *Post-hoc* analysis of device-related adverse events in relationship to clinical response in TMS active vs TMS sham groups showed no relationship between incidence of these events and treatment outcome. Therefore, the study blind in Study 44-01101 was maintained.

4.3. Clinical Study Results

4.3.1. Study 44-01101

"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."

The Final Study Report for Study 44-01101 is provided in PDF format in CD-ROM Attachment 11.

4.3.1.1. Study Design

Outpatients ages 18 to 70, meeting DSM-IV criteria for MDD, single episode or recurrent, with a current illness duration of 3 years or less and who had never previously been treated with TMS, were enrolled in Study 44-01101. The clinical diagnosis was confirmed by structured psychiatric interview. Patients were required to have a minimum symptom severity as reflected by a total score of at least 20 on the 17-item Hamilton Depression Rating Scale (HAMD17). In addition, patients were evaluated using the Antidepressant Treatment History Form (ATHF) and shown to have failed to receive benefit from definitive and adequate treatment with at least 1 but no more than 4 adequate trials of an antidepressant in their current episode. Patients who had failed to receive benefit from an adequate trial of electroconvulsive therapy at any point in their lifetime were excluded. Patients with psychiatric disorders other than MDD were also excluded. All patients were free of psychotropic medications for at least one week prior to and throughout the trial. The *a priori*-defined evaluable population consisted of 301 patients.

This study design was comprised of three phases: a one week, no-treatment and drug washout lead-in phase, a six week acute treatment phase, and a 3 week taper phase as shown in Figure 5. During the taper phase, treatment with the Neuronetics TMS System was tapered, while the patient was simultaneously tapered onto monotherapy with oral antidepressant medication. At any time after at least 4 weeks of participation in the acute phase of Study 44-01101, patients could be considered for enrollment in Study 44-01102 or 44-01103.

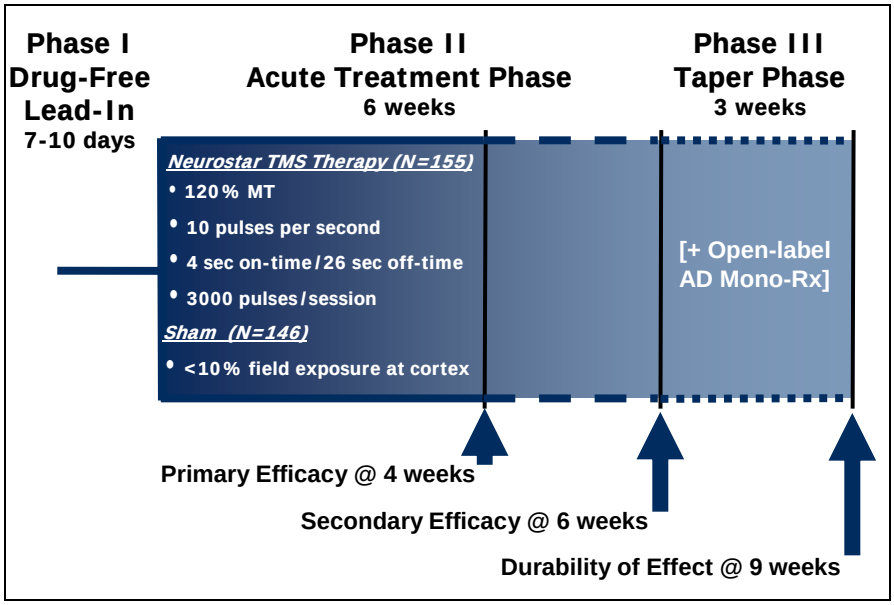


Figure 5. Study 101 Trial Design: Randomized, Double-blind, Sham-Controlled

Patients were randomized to receive either active treatment or placebo treatment. Randomization assignment was established prior to the start of the study, and was electronically recorded on sequentially assigned treatment cards that were used to control the operation of the NeuroStar System in a blinded manner. This blinding method ensured identical appearance, placement and acoustic properties of the NeuroStar System for both active and placebo treatments. Efficacy outcome was assessed by study personnel not included in the treatment session itself as an additional means to ensure the study blind.

Treatment sessions were conducted in sequential five day series for the 6 week acute treatment phase. Six additional treatments were administered across the 3 week taper phase. The maximum number of treatments was 36. 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. 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 by visual observation of thumb or finger movement using MT Assist which is a standardized software-based mathematical algorithm that provides an iterated estimate of the motor threshold. The treatment location was over the left prefrontal cortex, determined by a standard convention of movement of the TMS coil 5 cm anterior to the motor threshold location along a left superior oblique angle.

The primary efficacy outcome was difference between active TMS treatment and sham TMS treatment using the total score of the Montgomery Asberg Depression Rating Scale (MADRS) at week 4 of the acute treatment phase.

Secondary outcome measures included active versus sham treatment outcomes using the total score of the 24-item Hamilton Depression Rating Scale, and categorical outcomes such as 50% response rates and remission rates using both the HAMD and MADRS scores. Functional status and quality of life was measured using the Medical Outcomes Study 36-Item Short Form, and the Quality of Life Enjoyment and Satisfaction Questionnaire. 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.

4.3.1.2. 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. Key characteristics of the population were:

- The average age of patients was ~48 years, consistent with expectations for a more treatment-resistant population (i.e, roughly 10 years older than treatment naïve patients normally present in pharmaceutical antidepressant trials).
- 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.

4.3.1.3. Baseline Illness Characteristics and Functional Status

A summary of illness history, characterization of treatment resistance history, and baseline symptom severity is included in Table 8 for the intent-to-treat, evaluable study population.

Clinical symptom severity describes a ***moderate to severe*** clinical presentation in the current episode as evidenced by the average scores at study entry on the HAMD24, HAMD17, MADRS, IDS-SR and CGI-Severity ratings.

The overall pattern of illness history in the subject patient population is consistent with a more ***severe, difficult-to-treat sample*** as reflected by the predominance of recurrent depression, and an ATHF assessment which yielded an average number of adequate treatment exposures of 1.6 in both the active TMS and sham TMS treatment groups in the qualifying episode.

There is an extensive literature that supports the validity of the ATHF in assessing treatment adequacy. These data have shown that for each ATHF-verified adequate treatment exposure, patients usually have been exposed to ***at least 4 separate antidepressants treatments*** (Prudic, 2004).

Table 8. Key Observations for Illness History, Characterization of Treatment Resistance History and Symptom Severity for the Intent-To-Treat, Evaluable Study Population – Information Obtained at Screening Visit

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

Variable Name	Treatment Group		P-Value
	Sham (N=146)	Active (N=155)	
ATHF Rating Summary (# of adequate treatments in qualifying episode)			
- 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-defined Adequate Treatments in Qualifying Episode	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

The FDA requested a *post hoc* analysis of the type of antidepressant used by ATHF Group for Study 44-01101. This analysis examined the frequency and type of medication failures that occurred at each ATHF level.

This *post hoc* ATHF analysis described a population of subjects who were utilizing, in substantial proportion, antidepressant medications that are typically used in later stages of treatment complexity. For example, nearly half of the sample was using second and third generation medications, while ~15% were using more complex medication combination or augmentation strategies. The proportion of the population using these more complex treatments increased as a proportion of the sample with each progressive ATHF group. This provides validity to the view that each progressive cluster of patients within the ATHF categories reflects a patient subgroup with manifestly more difficult to treat depression.

Additional *a priori* defined evaluations of the study population included functional status, work productivity, health resource utilization and quality of life satisfaction. These 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 9.

As shown, 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.

Table 9. Work/Productivity and Health Resource Utilization in the All-Randomized Study Population at Study Entry - Selected Variables

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

4.3.1.4. Patient Disposition

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.

4.3.1.5. Efficacy Data

In all analyses, the primary study population of interest was declared as the *intent-to-treat, evaluable 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.

4.3.1.5.1. Primary Efficacy Outcomes – 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. The results of this analysis are shown in Table 10 and in Figure 6.

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, but did not meet the conventional level of statistical significance of $p<0.05$.

Table 10. 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

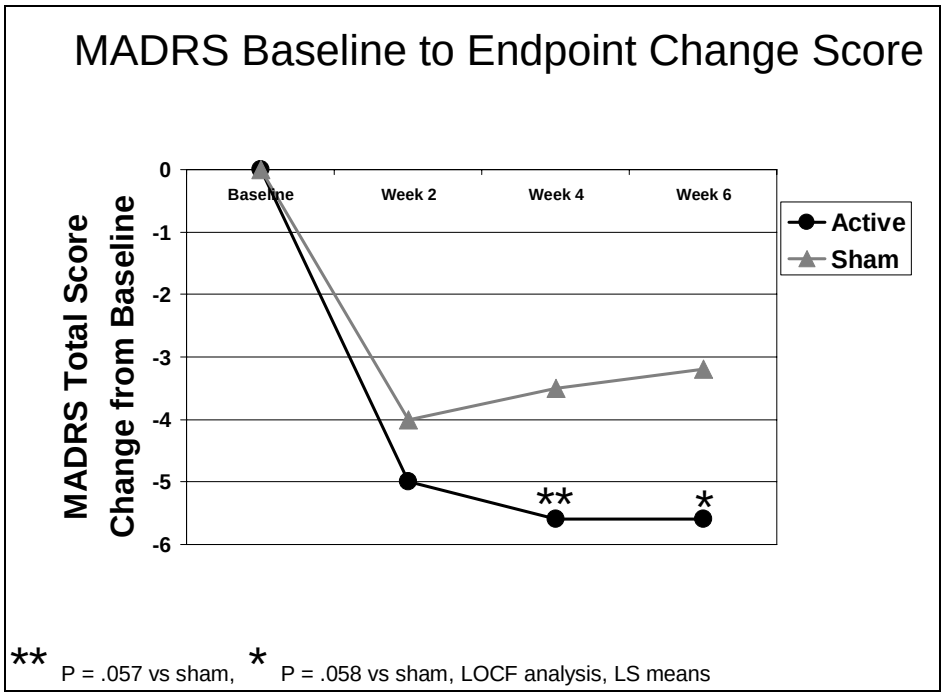


Figure 6. Primary Outcome Measure (MADRS Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

A *post hoc* evaluation of scores showed a statistically significant baseline imbalance 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 study entry, while the primary outcome measure (i.e., the MADRS) did not. As a result, while no baseline imbalance was detected for the HAMD17 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 patients 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 and is a commonly used minimum severity threshold for that scale at study entry (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

(n=6) removed from the analysis. The statistical consequence of this truncated analysis is the elimination of the statistical significance of the baseline imbalance in MADRS total score and a statistically significant outcome for MADRS total score (p=0.038) at the primary endpoint of 4 weeks. This data is consistent with the other two major efficacy outcome measures, the HAMD24 and the HAMD17.

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 the cited patient data. The detailed supporting ANCOVA analyses and logistic regression output for these measures are included in Tables 3.40-3.48 of the Final Study Report for Study 44-01101, (CD-ROM Attachment 11).

4.3.1.5.2. Secondary Efficacy Outcomes – Acute Phase

All secondary efficacy outcomes are shown in the following tables below by clinician-rated or patient-rated outcomes. Clinician-rated outcome measures are provided in Tables 11, 12 and 13. All analyses are presented for the intent-to-treat, evaluable study population as defined above, and represent a last observation carried forward analysis (LOCF). Detailed tables of secondary outcome results are provided in the Final Study Report for Study 44-01101.

Table 11 provides p-values at 2, 4 and 6 weeks and Table 12 provides p-values for the HAMD Factor Scores that are components within the HAMD24 item scale. Table 13 shows patient-rated outcomes at 2, 4, and 6 weeks. In all tables, p-values that fall below the traditional significance value of $p<.05$ are shown in green; those that trend towards statistical significance at $p<.1$ are shown in yellow. In all instances where statistical significance was reached between groups, active TMS treatment was superior to sham TMS treatment.

Table 11. Clinician-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS[1]

Variable Name	Week 2	Week 4	Week 6
MADRS Total Score	.191	.057	.058
MADRS (baseline adjustment)		.038	.052
HAMD24 Total Score	.051	.012	.015
HAMD17 Total Score	.098	.006	.005
Response Rate (>50% reduction from baseline)			
• MADRS	.384	.045	.007
• HAMD24	.601	.030	.042
• HAMD17	.451	.018	.015
Remission Rate			
• MADRS (Total score <10)	.311	.633	.011
• HAMD24 (Total score <11)	.257	.644	.012
• HAMD17 (Total score <8)	.418	.705	.065
CGI-S Total Score	.047	.009	.012

1. All P values are in favor of active TMS over sham TMS

Table 12. Clinician-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS for HAMD Factor Scores[1]

Variable Name	Week 2	Week 4	Week 6
HAMD Factor Scores			
• Anxiety/Somatization Factor	.300	.025	.023
• Core Depression Factor	.190	.012	.008
• Maier Factor	.276	.003	.003
• Gibbons Factor	.152	.007	.006
• Retardation Factor	.057	.007	.003
• Sleep Factor	.388	.211	.109

1. All P values are in favor of active TMS over sham TMS

Table 13. Patient-Rated Efficacy Outcomes: P Values for LOCF contrasts between active TMS vs sham TMS[1]

Variable Name	Week 2	Week 4	Week 6
MOS Short Form 36-Item			
• Physical Functioning	N/A	.299	.229
• Role-Physical	N/A	.361	.221
• Bodily Pain	N/A	.520	.301
• General Health	N/A	.049	.047
• Vitality	N/A	.179	.081
• Social Functioning	N/A	.183	.386
• Role Emotional	N/A	.105	.044
• Mental Health	N/A	.006	.015
Q-LES-Q	N/A	.124	.035
IDS-Self Report	.142	.058	.053
PGI-Improvement Total Score	.527	.181	.107

1. All P values are in favor of active TMS over sham TMS
N/A = scale not obtained at that time point

Three key secondary outcome measure results that provide strong, statistically significant evidence of an active TMS treatment effect superior to sham TMS treatment are:

- for HAMD24 and HAMD17 Item *total scores* at 4 and 6 weeks (Figures 7 and 8, respectively)
- for MADRS, HAMD24 and HAMD17 categorical outcomes for *response* (defined as >50% reduction in baseline score) scores at 4 and 6 weeks (Figures 9, 10 and 11, respectively), and
- for *remission* at 6 weeks as shown by MADRS and HAMD24 scores (Figures 9 and 10, respectively) .

These results are displayed graphically below and show a consistent, time-dependent outcome. As shown, response rates for the 3 outcomes measures were ~20% and ~25% at weeks 4 and 6, respectively, for active TMS and ~13% for sham TMS. The differential between active and control TMS treatment on both continuous and categorical outcome measures is similar to that obtained in controlled, clinical studies supporting the FDA approval of pharmaceutical antidepressants.

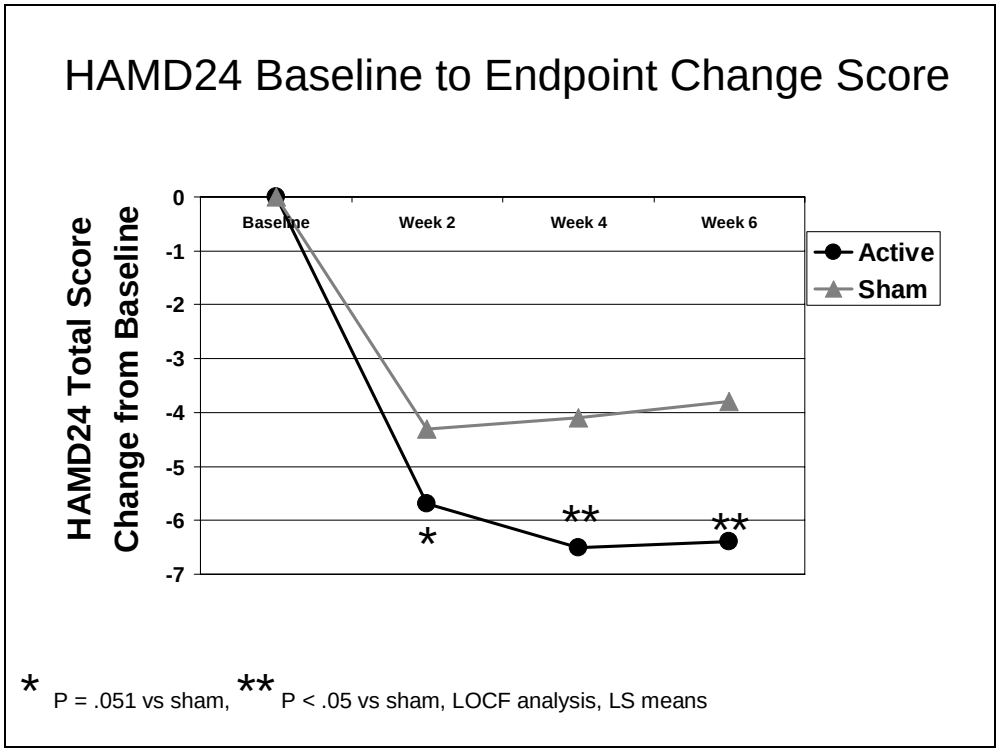


Figure 7. Secondary Outcome Measure (HAMD24 Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

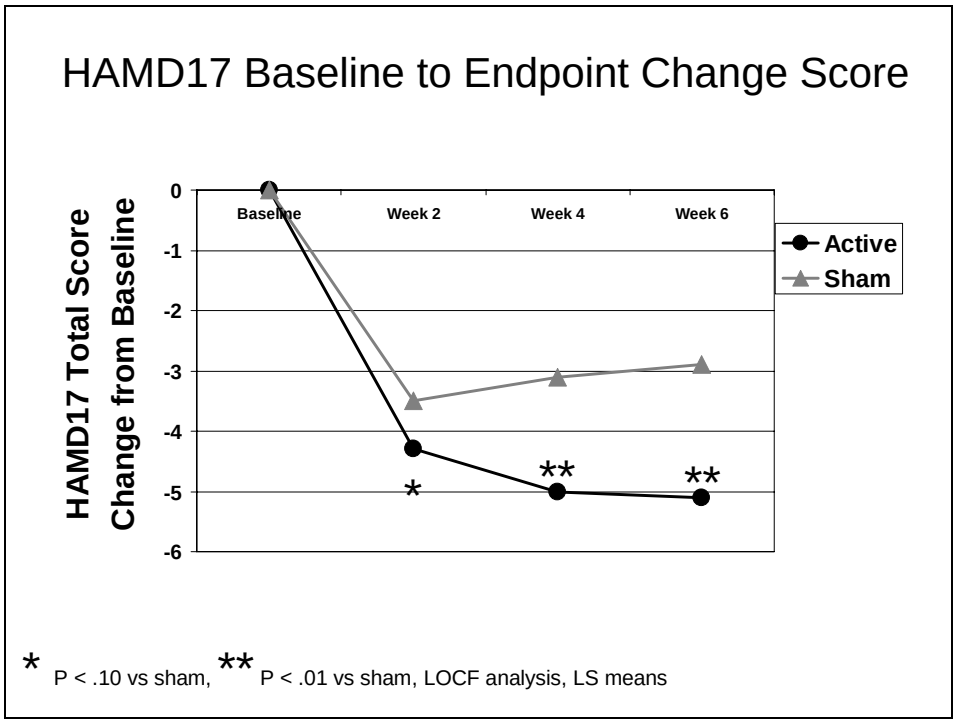


Figure 8. Secondary Outcome Measure (HAMD17 Total Score) Baseline to Endpoint Change Last-Observation Carried Forward Analysis

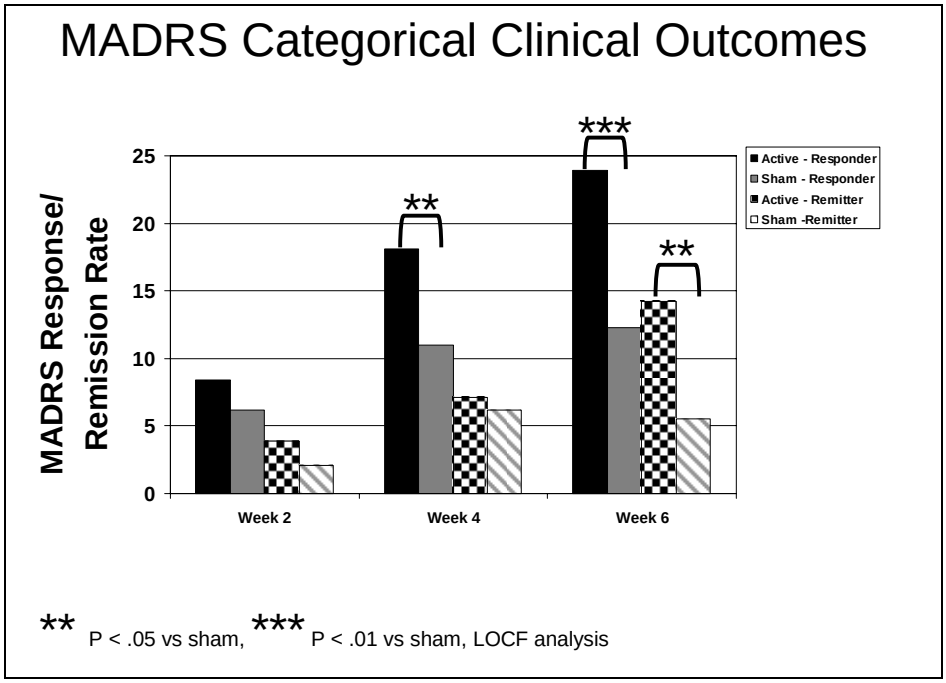


Figure 9. Secondary Outcome Measures (MADRS Responder and Remission Rates) Last-Observation Carried Forward Analysis

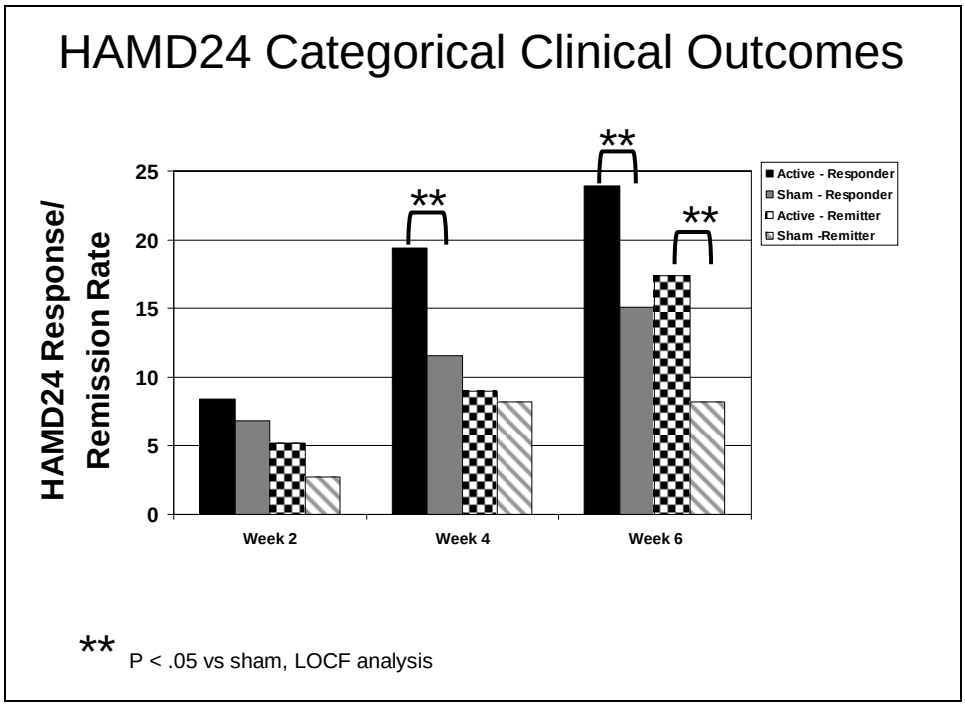


Figure 10. Secondary Outcome Measures (HAMD24 Responder and Remission Rates) Last-Observation Carried Forward Analysis

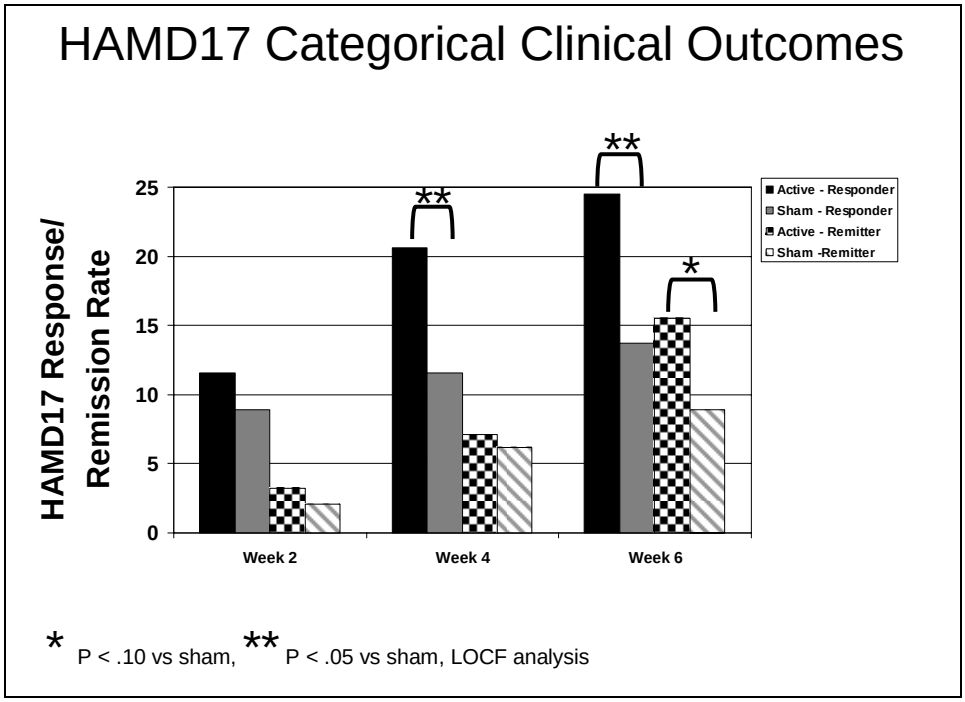


Figure 11. Secondary Outcome Measures (HAMD17 Responder and Remission Rates) Last-Observation Carried Forward Analysis

Neuronetics' randomized, sham-controlled clinical trial, Study 44-01101, showed that, at the *primary efficacy outcome time point of 4 weeks*:

- The primary efficacy outcome measure, the MADRS***, showed a statistical trend*** (p=0.057). The MADRS reached statistical significance (p=0.038) after correction for imbalance in baseline score.
- Active treatment with TMS was ***statistically significantly superior to sham*** TMS treatment for the change in mean symptom score using the HAMD17 and HAMD24 item scales.
- Symptom change was ***statistically significant for the categorical outcomes of response*** (> 50% reduction of baseline scores) on all major efficacy rating scales, namely the MADRS, HAMD17 and HAMD24.
- The clinician-rated CGI-score showed ***statistically significant*** evidence of efficacy as early as 2 weeks, and maintained this effect through the primary efficacy time point.
- The strength of the clinical effect of TMS on the core symptoms of depression and anxiety were shown by ***statistically significant*** changes in the *a priori*-defined sub-factor scores of the HAMD and in individual item analyses of the HAMD and the MADRS.

- Specific outcomes on the patient-rated SF-36 scale that reflected improvements in general and mental health were ***statistically significant*** in favor of active treatment with TMS.

Further evidence of clinical effect was provided by a continuing, and strengthened pattern of benefit in the analysis of the secondary efficacy time point of 6 weeks, in particular:

- Statistically significant efficacy was observed on both the HAMD17 and HAMD24.
- Further improvement in categorical outcomes was observed by sustained evidence of categorical response on all 3 depression rating scales, and also by the achievement of the more stringent threshold of remission on 2 of the 3 major depression rating scales, namely the MADRS and the HAMD24.
- Statistically significant effects on the core symptoms of depression and anxiety were sustained as shown by changes in the *a priori*-defined sub-factor scores of the HAMD and in individual item analyses of the HAMD and the MADRS.
- Further improvement in patient-rated outcomes were observed in statistically significant effects on quality of life enjoyment and satisfaction as shown on the Q-LES-Q total score, and on the general and mental health and role emotional scales of the SF-36.

The FDA questioned why remission did not reach statistical significance at 4 weeks. At 6 weeks, statistical significance was reached for remission for 2 of the 3 major depression rating scales, namely the MADRS and the HAMD24. The absence of reaching this outcome threshold at week 4 in Study 44-01101 was expected, given the significant symptom severity of the patient population at the start of the acute phase. For example, considering that the average MADRS score at baseline was ~32 in Study 44-01101, in order to achieve remission by week 4, the magnitude of clinical change would need to be far in excess of 16 points (i.e., the reduction needed to meet the criteria of "response") in order to meet the stringent remission metric of a MADRS total score < 10. In studies of pharmaceutical antidepressants, the general time course of symptom improvement shows that response occurs prior to the more pronounced effect of remission (Rush, et al, 2006). In instances where response and remission are more tightly coupled, i.e., where the response and remission rates are more nearly equal, this is usually observed to be an artifact of a lower baseline score at entry, making the remission target within easier reach from the starting point. This is not the case in Study 44-01101 with a starting baseline of ~32. Therefore, the achievement of response but not remission at the primary outcome time point of week 4, and the additional

achievement of remission by the secondary efficacy time point of week 6, in the Neuronetics studies are clinically reasonable and meaningful.

It is also worth considering why 2 of the patient-rated outcomes reached statistical significance by weeks 4 or 6 (SF-36 and O-LES-Q) whereas 2 did not (IDS-SR and PGI-S). In protocol 44-01101, the SF-36 and Q-LES-Q outcomes measures were listed earlier in the *a priori* listing of outcomes than the IDS-SR or PGI-I because clinical experience with the SF-36 and Q-LES-Q in clinical studies of major depressive disorder have generally shown them to be more sensitive to clinical change. The results of Study 44-01101 generally followed the *a priori* determined priority of efficacy outcomes.

Another important measure of patient-rated outcomes is obtained from the specific items of the SF-36. SF-36, which measures functional status, showed that patient-rated improvements were not non-specific, but were focused and most pronounced in the areas of mental health and emotional role functioning which reflects an improvement in emotional well being (i.e., General and Mental Health and Role Emotional) as expected for an antidepressant effect. The SF-36 outcomes were further evaluated in a *post hoc* analysis that was requested by the FDA, by an aggregated Physical Component Scale (PCS) total and Mental Component Scale (MCS) total. These analyses showed no effect on the Physical Component Scale at either 4 weeks ($p=0.892$) or 6 weeks ($p=0.682$) and a statistically significant effect on the Mental Component Scale at both 4 weeks ($p=0.019$) and 6 weeks ($p=0.0332$), consistent with the individual subscale results.

A third-party multiplicity analysis independently conducted by an expert statistician (See Tab 6) used four post-hoc multiplicity analyses (Holm, Hochberg, Hommel, and Benjamini-Hochberg) on 13 of the 26 primary and secondary endpoints using defined criteria. These analyses showed, as expected, that the four methods agree that the primary efficacy endpoint (MADRS at Week 4) had a resultant p-value of greater than 0.05 ($p>0.05$), and between one and nine, depending on the specific analysis performed, secondary endpoints had an adjusted p-value less than 0.05 ($p<0.05$). The conclusion of the statistician was that these multiplicity analyses did not favor the null hypothesis, thus indicating the overall effect of the outcomes was in favor of a significant positive outcome of active TMS versus sham TMS.

Observed Standardized Effect Size Compared to Target Effect Size

The targeted effect size for Study 44-01101 was stated in the protocol as 0.4 for the primary outcome measure, the MADRS, which represents a "moderate" effect size (Cohen, J. [1988] Statistical Power Analysis for the

Behavioral Sciences, 2nd Edition, Hillsdale N.J.). An estimate of the standard error of this target effect size is 0.12 (Kraemer, 2006). The standardized effect size calculated for MADRS at the primary efficacy time point of week 4 is 0.39, a result that is well within the expected standard error around the *a priori*-defined target estimate [.28-.52]. Similarly, on the two secondary efficacy symptom measures, the HAMD24 and the HAMD17, the standardized effect sizes were 0.48 and 0.55, respectively. Standardized effect sizes were computed using the between-group difference in means for the change from baseline divided by the pooled baseline population variance. Although no method for effect size calculation was pre-specified in the protocol, this standard method (GLM) was specified *a priori* for the primary outcome measure, and was used to compare Neuronetics study effect sizes with ECT data reported in the UK ECT Group report that used this same statistical method (See Section 6.3). Neuronetics' and the ECT computations utilized the adjusted pooled baseline variance term generated by the ANCOVA model for the primary outcome measure. We believe the FDA computed the standardized effect sizes for Study 44-01101 using the unadjusted raw baseline standard deviations as shown below in Table 14. As can be observed, these methods of calculation do not result in meaningfully different results and all lie within or above the expected effect size range.

Table 14. Standardized Effect Sizes Estimates for Study 44-01101

Outcome Variable	Observed Standardized Effect Size (Neuronetics Report)	Observed Standardized Effect Size (FDA Report)
MADRS	0.39	0.36
HAMD24	0.48	0.48
HAMD17	0.55	0.56

NOTE: Differences in reported effect size estimates between the Sponsor report and the FDA report are due to different derivations of the variance term used to calculate the effect sizes (see text above).

4.3.1.5.2.1 Post-hoc Analysis of Effect Size by ATHF Group (Requested by the FDA)

A comparison of the standardized effect sizes obtained with the MADRS, HAMD17 and HAMD24 efficacy instruments were reported for total score, response and remission, and for the IDS-SR for the overall population sample and distributed by subset of ATHF Levels 1-4 groups.

As shown in Table 15, the standardized effect size for the primary efficacy variable, MADRS total score, is 0.39 for the overall sample (p=0.057). This effect size is consistent with a moderate effect size as defined by Cohen, 1988 (small effect size = <0.20, moderate effect size = 0.30 – 0.80, large effect size >0.80) and lies within the estimated error of the a priori-determined targeted effect size.

Effect size, using the MADRS total score, for the ATHF Level 1 group is 0.95 (p=0.001). ATHF Level 2-4 groups do not have meaningful effect sizes and are not statistically significant between active and sham TMS groups. This result is repeated for all the outcome measures that were evaluated, including the patient-rated IDS-SR. As discussed further in the following section, these subset results are of interest and are notably consistent with expectation, since the strongest result is observed in the subset with the least treatment resistance rating. Nevertheless, these are *post-hoc* explorations, and Study 44-01101 was not powered *a priori* to detect differences between ATHF Level subgroups.

Table 15. Standardized Effect Sizes and Associated P-Values for Primary and Secondary Outcome Measures Observed in Study 44-01101

Primary Efficacy Outcome Measure	Active TMS (N)	Sham TMS (N)	Standardized Effect Size Week 4	P-Value Week 4
MADRS Total Score (Overall Sample)	155	146	0.39	0.057
• ATHF Group 1	88	76	0.94	0.001
• ATHF Group 2	45	50	-0.16	0.710
• ATHF Group 3	15	15	-0.55	0.588
• ATHF Group 4	7	5	5.21	0.022
Secondary Efficacy Outcome Measures				
HAMD24 Total Score (Overall Sample)	155	146	0.48	0.012
• ATHF Group 1	88	76	0.83	0.001
• ATHF Group 2	45	50	0.03	0.933
• ATHF Group 3	15	15	0.44	0.577
• ATHF Group 4	7	5	2.41	0.077
HAMD17 Total Score (Overall Sample)	155	146	0.55	0.006

• ATHF Group 1 • ATHF Group 2 • ATHF Group 3 • ATHF Group 4	88 45 15 7	76 50 15 5	0.83 0.13 0.81 2.26	0.001 0.762 0.440 0.089
MADRS Responder Rate (Overall Sample) • ATHF Group 1 • ATHF Group 2 • ATHF Group 3 • ATHF Group 4	155 88 45 15 7	146 76 50 15 5	0.65 1.23 0.11 1.99 -1.00	0.045 0.008 0.692 1.000 1.000
HAMD24 Responder Rate (Overall Sample) • ATHF Group 1 • ATHF Group 2 • ATHF Group 3 • ATHF Group 4	155 88 45 15 7	146 76 50 15 5	0.67 1.35 -0.17 2.99 -0.29	0.030 0.005 0.747 0.424 1.000
HAMD17 Responder Rate (Overall Sample) • ATHF Group 1 • ATHF Group 2 • ATHF Group 3 • ATHF Group 4	155 88 45 15 7	146 76 50 15 5	0.78 1.38 -0.21 2.99 -0.29	0.018 0.004 0.658 0.424 1.000
IDS-SR Total Score (Overall Sample) • ATHF Group 1 • ATHF Group 2 • ATHF Group 3 • ATHF Group 4	155 88 45 15 7	146 76 50 15 5	0.27 0.57 0.10 0.29 1.85	0.059 0.002 0.710 0.706 0.269

It is evident that the strength of the statistical effect observed in the largest subset of the overall population, namely the ATHF Group 1, is large. These effect sizes are well within the range of those observed in controlled studies of pharmaceutical antidepressants and also are consistent with the historical controlled trial literature of ECT vs simulated ECT that are discussed further below (see also 510(k) Section 12, Substantial Equivalence – CD-ROM Attachment 8). In all instances, including in the patient-rated measure, the IDS-SR, the p-value for the effect size in the ATHF Group 1 is observed to be substantially less than 0.01.

The ATHF Group 1 also demonstrates a superior response to treatment compared with any of the remaining ATHF Groups 2-4. This is consistent with the well-established relationship between failure of prior treatment and subsequent clinical outcome that has been observed for all known effective antidepressant treatments in the literature, both for pharmaceuticals and for ECT (Prudic, 1996; Rush, et al, 2006; Trivedi, et al, 2006; Fava, et al, 2006; Nierenberg, et al, 2006; McGrath, et al, 2006). Specifically, as evidence for failure of prior treatment mounts (e.g., increasing ATHF score), the likelihood of a response to subsequent treatment intervention diminishes. This is precisely the relationship

observed in the Study 44-01101 population. The treatment resistance evident in this study population can be found in the pattern and type of antidepressant use associated with the ATHF resistance ratings across the ATHF groups as was discussed in Section 4.3.1.3 as observed from a *post hoc* analysis that was requested by the FDA. That analysis underscored the fact that, in the more severe ATHF Groups, there was utilization of an increasing number of second and third generation antidepressants, and the report of greater use of more complex combination and augmentation strategies. This data lends substantial validity to the claim that ATHF Groups 2 through 4 represent a clinically more treatment resistant population than the patients represented in ATHF Group 1.

Although the pattern of antidepressant utilization across the ATHF Groups substantiates the clinical complexity of the treatment context with progressively worsening levels of treatment failure, it may still be asked whether the observed effect is due to a substantially less severe illness in those patients identified as ATHF Group 1 by treatment history alone. To assess this, Table 16 provides a summary of symptom severity and illness clinical features for the overall study population, stratified by ATHF Groups. As shown, there is a modest trend for the patients in the more severe ATHF Groups (2 through 4) to have slightly greater symptom severity scores and duration of current episode than observed for the patients in the ATHF Group 1. Nevertheless, across all ATHF Groups, the symptom ratings and general characteristics of the illness history are more similar than different. This indicates that the overall study population, regardless of ATHF Group, represents a moderately to severely ill patient population, who also demonstrate a range of prior treatment failures in the current episode.

Table 16. Symptom Severity and Clinical Illness Variables for the Overall Study Population, and Stratified for the Separate ATHF Groups (1 thru 4)

Clinical Variable	Overall Study Population (N=301)	ATHF Group 1 (N=164)	ATHF Group 2 (N=95)	ATHF Group 3 (N=30)	ATHF Group 4 (N=12)
Symptom Severity at Baseline					
• MADRS Total Score Mean (SD)	32.8 (5.4)	32.1 (5.3)	33.2 (5.4)	35.0 (5.8)	33.8 (4.2)
• HAMD24 Total Score Mean (SD)	30.6 (4.1)	30.6 (4.1)	30.1 (3.9)	31.3 (3.9)	33.2 (4.5)
• HAMD17 Total					

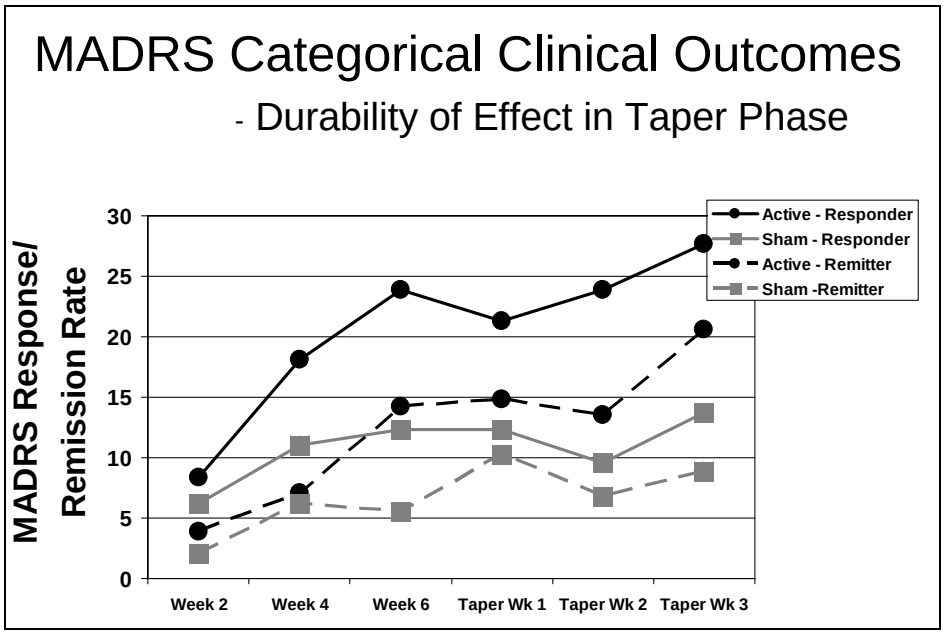
Clinical Variable	Overall Study Population (N=301)	ATHF Group 1 (N=164)	ATHF Group 2 (N=95)	ATHF Group 3 (N=30)	ATHF Group 4 (N=12)
Score Mean (SD)	22.8 (2.7)	22.7 (2.7)	22.6 (2.5)	23.4 (3.2)	24.0 (3.1)
Illness History					
• Recurrent Illness N(%)	285 (94.7)	159 (97.0)	87 (91.6)	27 (90.0)	12(100.0)
• Comorbid Anxiety Disorder Present N(%)	101 (33.6)	52 (31.7)	29 (30.5)	13 (43.3)	7 (58.3)
• Duration of current episode (months)	13.3 (9.6)	12.4 (9.5)	13.8 (9.7)	17.1(10.3)	13.3 (6.9)

4.3.1.6. Durability of Effect

At the conclusion of the acute treatment phase in Study 44-01101, all remaining patients were entered into a continuation phase referred to as 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. The blind remained intact throughout this taper phase. 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.

Figure 12 summarizes the categorical responder and remission rates for the primary disease-specific efficacy outcome measure for the MADRS for all patients continuing into the post-treatment taper phase. Similar results were observed with the HAMD24 and the HAMD17. Detailed supportive tables for these figures are included in the Final Study Report for Study 44-01101, (CD-ROM Attachment 11).



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

Figure 12. Responder and Remission Rates for the MADRS for Patients Continuing into the Post-Treatment Taper Phase

As shown, the clinical effect of active TMS is sustained during transition to single-drug antidepressant monotherapy (MADRS, HAMD17 and HAMD mean total score at 6 weeks was maintained through week 3 of taper). Therefore, 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. In addition, 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.

4.3.1.6.1. Post-hoc Analysis of Durability of Response (As Requested by the FDA)

In response to a request from the FDA for further analysis, the p-values for the between treatment group contrasts of the baseline to endpoint change on the MADRS total score and the HAMD24 total score, and also the categorical outcomes (responder rate and remission rate) on these scales were calculated as provided in Table 17. These analyses were performed in the protocol-specified evaluable study population in a last-observation carried forward manner at the week 4, week 6 and week 9 (end of taper phase) time points. Contrasts are also provided for each of these outcomes between the week 4 and end-of-taper phase time points.

Table 17. Summary of P-Values for Contrast Between Active TMS and Sham TMS at Week 4, Week 6 and Week 9 of Study 44-01101 and Summary of P-Values for Contrasts Within-Group for Week 4 and Week 9 – MADRS and HAMD24 Total Scores, Responder Rates and Remission Rates (Last Observation Carried Forward Analysis)

Primary Efficacy Outcome Measure	P-Value[1] Week 4	P-Value[1] Week 6	P-Value[1] Week 9	P-Value[2] Contrast of Week 4 vs Week 9
MADRS Total Score • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.057	0.058	0.011	0.016 0.915
Secondary Efficacy Outcome Measures	**P-Value[1] Week 4**	**P-Value[1] Week 6**	**P-Value[1] Week 9**	**P-Value[2] Contrast of Week 4 vs Week 9**
HAMD24 Total Score • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.012	0.015	0.003	0.016 0.657
MADRS Responder Rate • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.045	0.007	0.002	0.024 0.298

HAMD24 Responder Rate • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.030	0.042	0.008	0.026 0.140
MADRS Remission Rate • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.633	0.011	0.004	0.001 0.281
HAMD24 Remission Rate • Active TMS Week 4 v Week 9 • Sham TMS Week 4 v Week 9	0.644	0.012	0.001	0.001 0.545

NOTES:
P-Value[1] reflects analysis of change from baseline for treatment group comparison (Active TMS v Sham TMS) calculated using ANCOVA model, where change from baseline = baseline score, ATHF group, center, and treatment.
P-Value[2] reflects comparison of week 4 and week 9 scores within treatment group, calculated using paired T-Test.

As shown, these further analyses continue to demonstrate that active TMS shows a statistically significant superiority to sham TMS. These effects observed during the acute treatment phase are also evident at the end of the taper treatment phase and are statistically significantly more pronounced in the active TMS group as compared to the sham TMS group, when the contrasts between the acute treatment phase and the taper treatment phase are examined.

Taken together, these data are consistent with the *a priori*, protocol-specified statistical analyses demonstrating statistically significant efficacy of active TMS compared to sham TMS, and of the durability of this effect in follow up after discontinuation of acute treatment.

In summary, the acute clinical response to TMS treatment in Study 44-01101 using the NeuroStar System, ***was shown to be durable*** and was successfully maintained over the course of a three week transition to maintenance-of-effect antidepressant pharmacotherapy. During this time, a continuing improvement in both response and remission outcomes was observed on the MADRS, HAMD24 and HAMD17 in the active TMS group.

4.3.1.7. Overall Efficacy Conclusions

The results of Study 44-01101 demonstrate clinically and statistically significant efficacy of the NeuroStar System using well-validated efficacy instruments, and demonstrate that this acute clinical response is durable over the 9 week trial period.

These data address the elements of proof of efficacy as stated in Section 2.0 in support of Neuronetics' premarket 510(k) notification. These results also provide proof of acute durability of response. Additional evidence of durability at 4 weeks past cessation of TMS treatment is provided by the results from Study 44-01103.

4.3.1.8. Safety Data

4.3.1.8.1. 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, including,

- 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 by protocol for exceeding protocol treatment parameters.

Serious adverse events reported in Study 44-01101 are listed in Table 18.

- No deaths were reported
- No seizures were reported.
- No suicides were reported.
- Seven events occurred after signing the Informed Consent and prior to randomization. 15 events occurred after treatment and during the acute treatment phase and 1 event occurred in the post-treatment taper phase.

Table 18. Serious Adverse Events Reported for Study No. 44-01101

Serious Adverse Event[1]	Number of SAEs	Relationship to Study Device
Worsening depression	6	Not related (6)
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)

1. Clinical case vignettes for all serious adverse events are provided in the Final Study Report for Study 44-01101.

Among the overall reported SAE terms that reflected worsening depression or emergence of suicidal ideation or suicide attempt, these events represented 11 unique patients. Five of these patients were allocated to the sham TMS treatment arm, 2 patients were allocated to the active TMS treatment arm, and 4 patients were never randomized.

- Worsening of depression was a specifically reported term in the SAE in 6 patients. Four of these patients were not randomized into the study at the time of the event. Two patients had been allocated to the sham TMS treatment arm. None were allocated to active TMS treatment.
- Report of suicidal ideation was a specifically reported term in the SAE in 5 patients. One of these patients was not randomized into the study at the time of the event. Two patients had been allocated to the sham

TMS treatment arm. Two patients had been allocated to the active TMS treatment arm.

- Suicide attempt was a specifically reported term in the SAE in 1 patient. This patient had been allocated to the sham TMS treatment arm.

First degree burns were reported for 2 patients receiving active TMS. These were due to a defect in the single-use shield that is attached to the face of the NeuroStar TMS coil (see Section 4.3.1.8.2).

4.3.1.8.2. Device Malfunctions

There were two failure modes that occurred with the clinical version of the NeuroStar System during Neuronetics clinical studies. The failures involved (1) a malfunction of the console power supply due to a plating defect in the control board and (2) a manufacturing defect of the single-use shield that was caused by a shorted trace within the shield. These defects were corrected in the clinical devices during the trial and were included as part of the failure mode and effect analysis for implementation of the commercial NeuroStar System.

4.3.1.8.3. Treatment-Emergent Adverse Events

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.

Table 19 summarizes adverse events, by MedDRA-preferred term, that occurred at an incidence of $\geq$ 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 CD-ROM Attachment 11.

Table 19. 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)
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.

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. 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. These data are contained in CD-ROM Attachment 11, Tables 3.26 and 3.27.

4.3.1.8.4. 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 (See Section 4.2.3). Multiple versions of the MMSE and BSRT were used to allow repeat administrations and to deter learning effects.

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. Details are provided in Tables 3.28-3.30 in CD-ROM Attachment 11.

4.3.1.8.5. Auditory Threshold Testing

Auditory threshold testing was conducted to assess any effects from the acoustic noise of the NeuroStar device. Air-conduction auditory threshold was assessed at baseline, week 4 and week 6. A desktop audiometer 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. All patients wore ear protection rated at a minimum decibel level reduction of 30 during TMS treatment. 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.

Supporting details are provided in Tables 3.21-3.37 in CD-ROM Attachment 11.

4.3.1.8.6. Overall Safety Conclusions

The results of Study 44-01101 demonstrate that the NeuroStar System is safely tolerated.

Only 8% of patients did not complete the protocol-required treatment course through the primary efficacy endpoint of Week 4. 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 were generally mild to moderate and lessened with time over the TMS treatment course.

There were no deaths or seizures reported. 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 single-use shield that resulted in minor scalp burns. The device design was corrected to address the failure mode and the failure did not occur during the remainder of the trial. There was no evidence of clinically significant change in cognitive function testing or in auditory threshold at either 4 weeks or 6 weeks. There was no evidence that active TMS treatment was associated with worsening of depression or emergent suicidal ideation during the acute treatment phase.

These data provide evidence of acute safety of the NeuroStar System and address the elements of proof of safety provided to the FDA in support of Neuronetics' premarket 510(k) notification.

4.3.2. Study 44-01102

"A 9-week, Uncontrolled, Open-Label Study to Evaluate the Efficacy and Safety of the Neuronetics Model 2100 Repetitive Transcranial Magnetic Stimulation (rTMS) System in the Treatment of Patients with Major Depression Previously Non-Responsive to Active or Sham rTMS Treatment"

The Final Study Report for Study 44-01102 is provided in PDF format in CD-ROM Attachment 12.

4.3.2.1. Study Design

A total of 158 patient formed the evaluable dataset for Study 44-01102. Patients who participated in Study 44-01101 for at least 4 weeks of acute phase treatment and who failed to receive benefit from their randomized treatment assignment in that study, and who also showed less than a 25% decline in their HAMD17 total score at exit compared to their baseline score, were eligible for enrollment into Study 44-01102. Treatment assignment from Study 44-01101 was not unblinded at the time of enrollment into Study 44-01102. Study 44-01102 was an open-label, uncontrolled clinical trial otherwise identical in design and treatment sequence to Study 44-01101.

Since there were two potential routes of entry into Study 44-01102, depending upon their treatment assignment into Study 44-01101, data were always reported for the two groups separately. The two Groups are:

Group A: Patients who were randomized to active TMS in Study 44-01101, did not respond, and who agreed to enter Study 44-01102

Group B: Patients who were randomized to sham TMS in Study 44-01101, did not respond, agreed to enter Study 44-01102.

4.3.2.2. Efficacy Data

4.3.2.2.1. Efficacy Outcomes – Acute Phase

Efficacy results for Study 44-01102 are summarized below for clinician-rated outcome measures as shown in Table 20. The table shows the mean change in total efficacy assessment scores, and responder and remission rates for patients who were randomized in Study 44-01101 to either active TMS or to sham TMS and were subsequently treated with open-label TMS in Study 44-01102.

Table 20. Open-Label TMS Study 44-01102: *A Priori*-Defined Outcome Measures

Efficacy Outcome Measures	Week 4 Study 101 Active Non-Responder (Group A)	Week 6 Study 101 Active Non-Responder (Group A)	Week 4 Study 101 Sham Non-Responder (Group B)	Week 6 Study 101 Sham Non-Responder (Group B)
MADRS Total Score Mean Change[1]	-10.5	-12.5	-11.9	-17.0
HAMD24 Total Score Mean Change[1]	-9.0	-11.1	-11.0	-14.5
HAMD17 Total Score Mean Change[1]	-6.4	-8.2	-8.2	-10.8
MADRS Responder Rate (%)[2,6]	20.5	26.0	24.7	42.4
HAMD24 Responder Rate (%)[2,6]	21.9	31.5	28.2	42.4
HAMD17 Responder Rate (%)[2,6]	21.9	30.1	27.1	37.6
MADRS Remission Rate (%)[3,6]	5.5	11.0	5.9	20.0
HAMD24 Remission Rate (%)[4, 6]	9.6	16.4	12.9	27.1
HAMD17 Remission Rate (%)[5, 6]	6.8	15.1	10.6	21.2

Group A = Study 101 Active TMS Non-Responders; Group B = Study 101 Sham TMS Non-Responders.

[1] Mean change from total score observed at baseline upon entry to Study 44-01102
[2] Responder is >50% change from baseline score at entry to Study 44-01102
[3] MADRS Remission is defined as MADRS total score <10
[4] HAMD24 Remission is defined as HAMD24 total score <11
[5] HAMD17 Remission is defined as HAMD17 total score <8
[6] Responder and Remission Rates were calculated as a proportion of the total evaluable sample at all time points.

Figure 13 below displays the primary outcome measure (MADRS Total Score), baseline to endpoint change for the evaluable study population. Figure 14 shows the secondary outcome measure (MADRS responder and remission rates) using the Last-Observation Carried Forward analysis which is representative of other measures (HAMD17 and HAMD24).

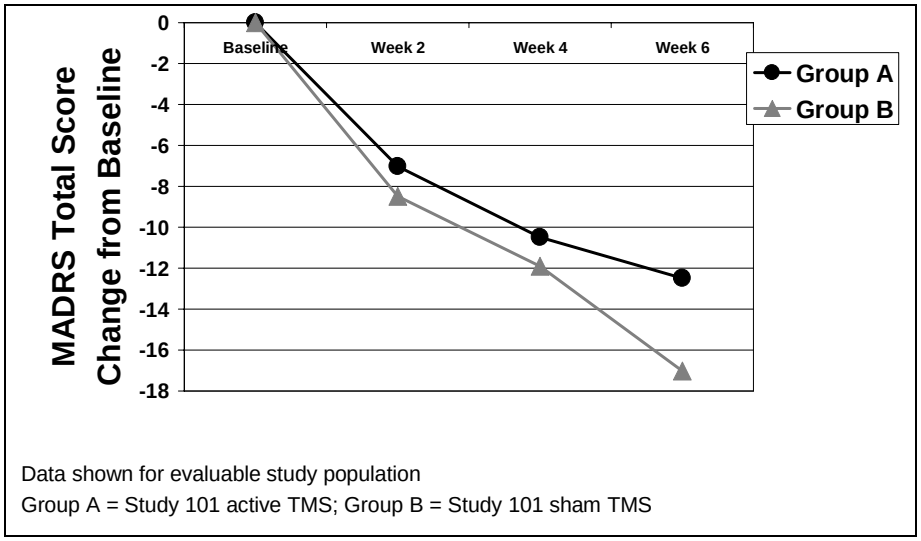


Figure 13. Primary Outcome Measure (MADRS Total Score) Baseline to Endpoint Change for the Evaluable Study Population in Study 44-01102

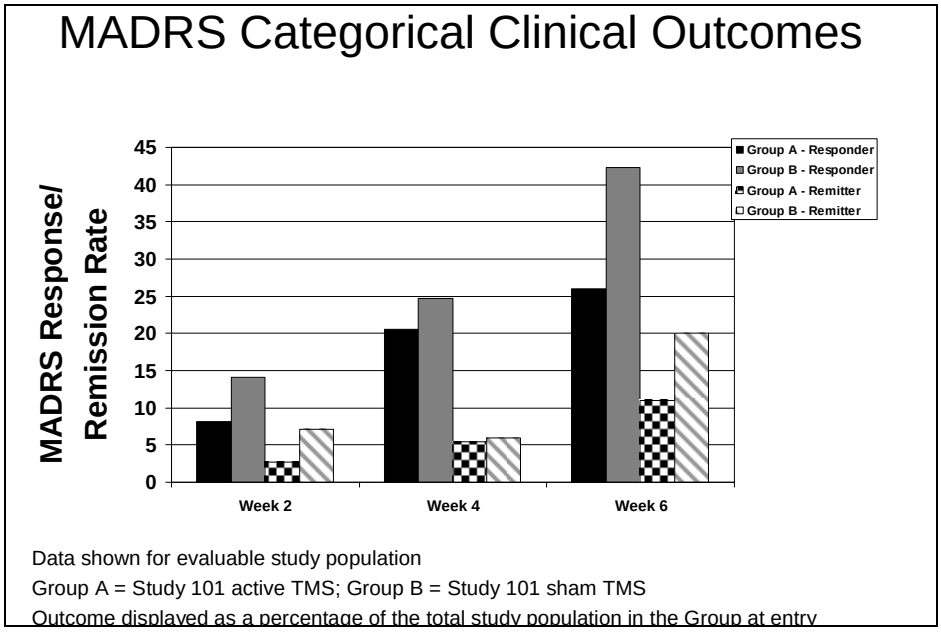


Figure 14. Secondary Outcome Measures (MADRS Responder and Remission Rates) Last-Observation Carried Forward Analysis in Study 44-01102

As shown, patients with major depression who have failed to receive adequate clinical benefit from medication therapy show a clinically meaningful response to open-label treatment with the Neuronetics TMS System:

- After failure to receive benefit from their randomized treatment assignment in Study 44-01101, patients previously assigned to sham TMS show clinical benefit with open-label TMS treatment (at week 6; 42.4% MADRS response rate, 20% MADRS remission rate).
- A clinically meaningful proportion of patients who failed to receive clinical benefit after at least 4 weeks of active TMS, respond successfully to an extended duration of active treatment with TMS (at week 6; 26% MADRS response rate, 11% MADRS remission rate).

4.3.2.3. Durability of Effect

As in Study 44-01101, 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 where they began a scheduled taper of their open-label, active TMS treatment across a 3-week schedule with initiation onto open-label pharmacotherapy. The results showed the patients previously allocated to sham TMS treatment in Study 44-01101 consistently showed a greater clinical benefit during this continuation period in Study 44-01102 as compared to those patients previously allocated to active TMS treatment.

4.3.2.3.1. Post-hoc Analysis of Effect Within and Between Treatment Groups over Time (As requested by the FDA)

At the request of the FDA, analyses were conducted to examine the primary outcome measure, MADRS total score, and the two key secondary outcome measures, namely the HAMD24 total score and the HAMD17 total score, for any statistically significant change in scores from the beginning to the end of the study (i.e., baseline to Taper Phase Week 3). An ANOVA model was used, and p-values were examined for the within-group effect of change from baseline, along with p-values for the between-group effect at each time point (Table 21).

Table 21. Summary of P-Values for Change from Baseline for Primary and Secondary Outcome Measures in Study 44-01102 – Overall Group Results

Primary Efficacy Outcome Measure	Week 2	Week 4	Week 6	Week 7	Week 8	Week 9

MADRS Total Score • Group A (N=73) • Group B (N=85) (P-Values for within group contrast from baseline)	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001
P-Values (Between group contrast)	0.247	0.395	0.032	0.010	0.083	0.287
Secondary Efficacy Outcome Measures	**Week 2**	**Week 4**	**Week 6**	**Week 7**	**Week 8**	**Week 9**
HAMD24 Total Score • Group A (N=73) • Group B (N=85) (P-Values for within group contrast from baseline)	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001
P-Values (Between group contrast)	0.483	0.131	0.040	0.010	0.067	0.086
HAMD17 Total Score • Group A (N=73) • Group B (N=85) (P-Values for within group contrast from baseline)	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001	< 0.001 < 0.001
P-Values (Between group contrast)	0.356	0.069	0.042	0.005	0.028	0.065

NOTES:
Group A = Study 101 Active TMS Non-Responders; Group B = Study 101 Sham TMS Non-Responders.
P-Values for between group contrast performed using ANOVA model, change from baseline = treatment.
P-Values for within group contrast performed using paired T-Test.

In the overall sample, there is a statistically significant change from baseline to the final endpoint of study for all three symptom rating scales (MADRS, HAMD24 and HAMD17) ($p<.001$ for all scales). The effect was robust, appearing as early as the first observation point at two weeks, and was maintained throughout the course of the acute treatment (weeks 2, 4 and 6) and taper treatment (weeks 1, 2 and 3) phases of the study.

There was also a statistically significant difference between the treatment groups that was uniformly evident on all symptom rating scales at weeks 6 and 7. This difference was always in favor of Group B, i.e., those patients

who had previously been allocated to sham TMS treatment in Study 44-01101 and were receiving their first exposure to active TMS treatment in this open-label study. As antidepressant treatment was initiated in all subjects in combination with their open-label active TMS taper during the study Taper Phase, by the end of study, the statistical significance between treatment groups was no longer evident.

An important observation can be made regarding the sustained clinical benefit of active TMS when used in combination with pharmacotherapy. In Study 44-01101, there was a statistically significantly greater benefit during the taper phase of Study 44-01101 in the active TMS group as compared to the sham TMS group (see Section 4.3.1.6). This suggested that an additional clinical benefit was obtained from the combination of active TMS and antidepressant pharmacotherapy during taper.

Overall, it appears that a similar pattern of benefit of active TMS and antidepressant pharmacotherapy during the taper phase is evident in Study 44-01102. Since both groups are being treated with open-label active TMS in Study 44-01102, it is expected that this pattern of benefit should be evident in both treatment groups, as appear to be the case. This effect appears to account for the loss of statistical significance in the between-group comparison as the two groups enter the taper phase of the study.

4.3.2.4. Safety Data

The safety of TMS treatment using the NeuroStar System was evaluated by the collection and evaluation of serious adverse events, spontaneous adverse events, cognitive function testing, auditory threshold testing and emergent suicidal ideation.

There were no deaths or seizures reported in this study. Adverse events and their temporal relationship in Study 44-01102 were similar to that reported in Study 44-01101.

Application site pain was observed in both treatment groups, but the incidence was greater in the patient group that had previously been allocated to sham TMS treatment prior to entry into study 44-01102, suggesting that the prior exposure assisted in accommodation to this effect. For both headache and application site pain, the greatest incidence was observed during the first week of treatment with a substantial reduction in incidence 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.

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

There was no evidence of clinically significant auditory threshold change at either 4 weeks or 6 weeks. There was no evidence that active TMS treatment was associated with worsening of depression or emergent suicidal ideation during the acute treatment phase.

It is important to note that for patients previously allocated to active TMS in Study 44-01101 each received up to 40 treatment sessions with the NeuroStar System totaling 120,000 pulses over approximately 8 weeks. Therefore, Study 44-01102 showed sustained safety over this extended number of TMS treatments.

4.3.2.5. Efficacy and Safety Conclusion

Study 44-01102 provides confirmatory evidence of the safety and efficacy observed for the NeuroStar System in Study 44-01101.

This study confirms that treatment with the NeuroStar System is an effective antidepressant for patients with DSM-IV defined major depression for those patients who had not previously received sufficient clinical benefit from treatment with pharmacotherapy for their illness. Patients previously allocated to sham TMS show substantial and clinically meaningful improvement in symptom scores (42% response rate at 6 weeks). Additionally, in patients previously allocated to active TMS, a clinically meaningful proportion of patients show evidence of late response to treatment with continued active TMS.

4.3.3. Study 44-01103 (Interim Report)

"A 6-month, Open-Label Maintenance Study of Patients with Major Depression Previously Responsive to rTMS Treatment with the Neuronetics Model 2100 CRS Repetitive Transcranial Magnetic Stimulation (rTMS) System."

An Interim Study Report for Study 44-01103 is provided in PDF format in CD-ROM Attachment 13.

4.3.3.1. Study Design

A total of 136 patients were enrolled in Study 44-01103. This study was an open-label, uncontrolled clinical trial providing 24 weeks of continuation oral antidepressant monotherapy to patients who showed a 25% or greater improvement in their HAMD17 total score at the end of participation in

either Study 44-01101 or Study 44-01102, compared to their baseline HAMD17 score in those studies.

There were four potential routes of entry into Study 44-01103, and they represent four separate populations contained in the study analysis. The first three groups represent the various paths for active TMS-treated subjects to enter Study 44-01103, while the Group 4 represents the sham TMS responders from Study 44-01101:

Group 1 (N=44): Patients who were randomized to active TMS in Study 44-01101, responded, and agreed to enter Study 44-01103 [Study 101 Active Responders]

Group 2 (N=27): Patients who were randomized to active TMS in Study 44-01101, did not respond, and who agreed to enter Study 44-01102, received a course of open-label active TMS, responded to that course of treatment and then agreed to enter Study 44-01103 [Study 101 Active Non-responders/Study 102 Responders]

Group 3 (N-42): Patients who were randomized to sham TMS in Study 44-01101, did not respond, agreed to enter Study 44-01102, received a course of open-label active TMS, and then agreed to enter Study 44-01103 [Study 101 Sham Non-responders/Study 102 Responders]

Group 4 (N=23): Patients who received sham TMS in Study 44-01101, responded to treatment and subsequently agreed to enter Study 44-01103 [Study 101 Sham Responders]

During the course of Study 44-01103, patients were permitted to adjust their monotherapy antidepressant medication schedule as clinically indicated, but were not permitted to switch or combine antidepressant regimens. In the event that a patient's clinical symptoms deteriorated as determined by change in the Clinical Global Impressions – Severity of Illness score, observed on at least two sequential study visits, then open-label TMS treatment was reintroduced in conjunction with continuation pharmacotherapy, for up to 24 sessions across six weeks. Treatment with the NeuroStar System was discontinued if symptom resolution occurred. Patients were discontinued from this study if they experienced a recurrence of DSM-IV defined MDD or if they failed to receive benefit from a full course of reintroduction of treatment with the NeuroStar System. Efficacy and safety outcomes were assessed using the same measurement instruments as in Study 44-01101.

4.3.3.2. Interim Efficacy Data

[Note: the data reported for Study 44-01103 are interim in nature. The study was ongoing at the time of 510(k) submission]. All patients enrolled in

Study 44-01103 had completed 4 weeks of the study at this interim analysis and 70% of patients had completed the study across 24 weeks.

4.3.3.2.1. Durability of Effect – Incidence of Relapse During the First 4 Weeks

Interim efficacy data from Study 44-01103 demonstrate that the durability of the acute treatment response to active TMS is maintained over the first four weeks of TMS-free treatment expressed in terms of the incidence of illness relapse. Using the protocol-defined definition of discontinuation for all cause during this time interval, the cumulative incidence of relapse is 2.3% (Figure 15).

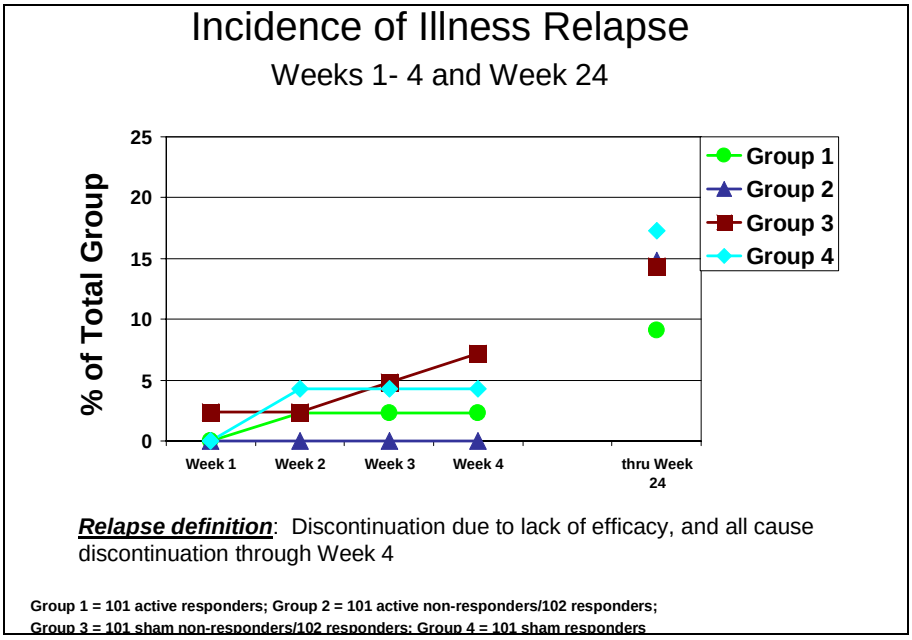


Figure 15. Incidence of Relapse Using the *A Priori*-Defined Protocol Criterion During Weeks 1 through 4 and At Study Endpoint in Study 44-01103

At the request of the FDA, an alternative definition was also applied in an exploratory manner over this same time interval, using a relapse definition derived from the observed change in HAMD24 score using a relapse definition commonly used in the ECT literature. Based on this definition, the cumulative incidence of relapse across the first 4 weeks of TMS-free treatment is 9.1% (Figure 16). These data compare favorably to the expected incidence of relapse in a difficult to treat patient population with major depression, as seen in the published ECT literature. For example, in recent studies where best available pharmacotherapy options are pursued, relapse rates following cessation of an acute course of ECT range from 4.5% to 36% at 4 weeks (Prudic, 2004, Sackeim, 2001).

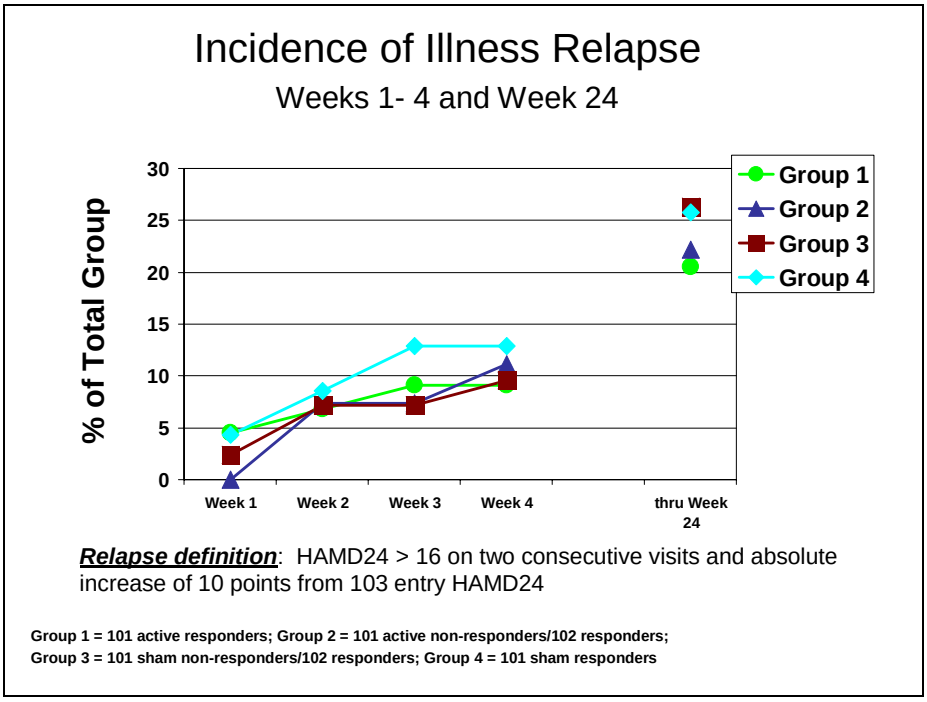


Figure 16. Incidence of Relapse Using the ECT Literature Definition Criterion During Weeks 1 through 4 and At Endpoint in Study 44-01103

4.3.3.2.2. Durability of Effect – Longitudinal Course of Symptom Scores Across the First 4 Weeks

Efficacy results for patients who were responders in the active treatment group in Study 44-01101 and continued directly into Study 44-01103 (Group 1) are summarized in Table 22 that displays the mean score change and remission rates for MADRS, HAMD24 and HAMD17 item scores.

Table 22. *A Priori*-Defined Outcome Measures for Group 1[1] (N=44)

Efficacy Outcome Measures	Week 1	Week 2	Week 3	Week 4
MADRS Total Score Mean Change[2,6]	-20.1	-21.4	-20.3	-21.2
HAMD24 Total Score Mean Change[2,6]	-18.0	-19.0	-18.4	-19.6
HAMD17 Total Score Mean Change[2,6]	-14.0	-14.4	-13.9	-14.6
MADRS Remission Rate (%)[3,6]	50	59.1	52.3	45.5
HAMD24 Remission Rate (%)[4,6]	47.7	54.5	47.7	43.2
HAMD17 Remission Rate (%)[5,6]	50	56.8	43.2	43.2

[1] Group 1 are patients who were responders in the active treatment group in Study 44-01101
[2] Baseline is defined as baseline of Study 44-01101
[3] MADRS Remission is defined as MADRS total score <10
[4] HAMD24 Remission is defined as HAMD24 total score <11
[5] HAMD17 Remission is defined as HAMD17 total score <8
[6] Responder and Remission Rates were calculated using total enrolled sample

Overall, 38.2% of all patients in Study 44-01103 experienced at least one cycle of TMS reintroduction during the 6 month duration of the study. Most treatments occurred subsequent to the first month, with the median time to reintroduction ranging from 6.5 to 11 weeks after enrollment in Study 44-01103**.**

These results suggest that symptomatic change sufficient to require TMS reintroduction occurs in less than half of the patients entering Study 44-01103 overall, and that the time to reintroduction is not immediate, but occurs after approximately 1-3 months.

4.3.3.3. Interim Safety Data

The safety of TMS treatment using the NeuroStar System in Study 44-01103 was evaluated by the collection and evaluation of serious adverse events and spontaneous adverse events for this interim report.

There were no deaths or seizures. Patients who showed an acute response to TMS treatment during either controlled or open-label treatment with the NeuroStar System show a pattern of adverse events during 24 week maintenance treatment with antidepressants that is consistent with the expected profile of adverse events with medication use and with the expected profile of adverse events seen in Neuronetics' Studies 44-01101 and 44-01102.

4.3.3.4. Efficacy and Safety Conclusions

A majority of patients who experienced symptomatic response to acute TMS treatment in Study 44-01101 showed a clinically meaningful and stable pattern of symptomatic response during 4 weeks of maintenance antidepressant pharmacotherapy alone in Study 44-01103. The cumulative incidence of protocol-defined relapse through 4 weeks is 2.3%.

The percentage of subjects who experienced symptomatic worsening and were provided with *reintroduction of active TMS* treatment for at least one cycle observed at this interim report across 24 weeks ranged from 33.3% to 47.8% depending on path of entry into Study 44-01103.

Active TMS was safe and well tolerated when administered in an adjunctive manner with antidepressant pharmacotherapy.

This data from Study 44-01103 provides ***additional proof of evidence of durability of effect*** submitted in support of marketing clearance.

SECTION 5. CLINICAL SIGNIFICANCE OF NEUROSTAR TMS THERAPY IN THE TREATMENT OF MAJOR DEPRESSIVE DISORDER

5.1. Clinical Significance - Outcome Measures

In the clinical development of new antidepressants, a complete interpretation of the study results derives from observation of the outcomes on well-validated depression rating scales that are reported in both a *continuous* and a *categorical* manner. The *continuous* approach uses the mean value of the population as a whole, i.e., it measures the average total symptom score of the active group compared to the average score in the placebo group (i.e., this approach compares values that may vary along a numerical continuum), whereas the *categorical* outcome measures the percentage of patients in each group that have achieved a particular clinical outcome, such as "response" or "remission", each of which have specific definitions (i.e., this approach compares values that occur in a dichotomous fashion: present or not present). Therefore, in the interpretation of clinical study results, although the population mean score (the continuous outcome) is a valuable tool to determine statistical significance of the effects between groups as a whole, the categorical outcomes are critical in determining whether the effect achieved a criterion level reflecting clinically meaningful results such that more patients in the active group than the sham group achieved the defined response or remission criteria. Put simply – did many patients get a little better or did a smaller number of patients get a lot better? It is this distinction that cannot be determined from the continuous outcome alone. The categorical endpoints help to determine the *clinical impact* of the group difference in treatment response.

In discussions with the FDA during the process of designing Neuronetics' clinical studies, FDA emphasized the importance of including the categorical clinical outcomes using the three key observer-administered rating scales, the MADRS, the HAMD24 and the HAMD17 for the overall interpretation of the clinical significance of the study results. All of these outcome measures were collected and determined as part of the *a priori*-stated statistical analysis plan for Study 44-01101.

As stated in Section 4.3.1.5.2, response rates for the 3 outcomes measures in Study 44-01101 were ~20% and ~25% at weeks 4 and 6, respectively, for active TMS and ~13% for sham TMS – a statistically significant difference for all outcome measures (MADRS, HAMD17, and HAMD24). This magnitude of benefit in response rates for active TMS compared to its relevant within-study control are similar to the outcomes obtained in controlled, clinical studies for FDA-approved pharmaceutical antidepressants and demonstrate a statistically significant difference between treatment arms.

In this section, the data from *controlled registration trials* of pharmaceutical antidepressants is compared to data from Study 44-01101 and data from *open-*

label studies of antidepressants are compared to the data from open-label Study 44-01102. These comparisons are important because they demonstrate that the efficacy data from Neuronetics studies are consistent with similar data obtained from clinical trials of FDA-approved pharmaceutical antidepressants. The response rates obtained in Neuronetics clinical trials are similar to those obtained in controlled, clinical studies for FDA-approved pharmaceutical antidepressants and similarly demonstrate a statistically significant difference between treatment arms.

A similar discussion of the comparative efficacy of the Neuronetics studies to controlled and open-label studies with the predicate device, ECT, can also be performed. Despite limitations of differences in study design and outcome assessment, it can reasonably be inferred from this analysis that the clinical efficacy outcome with TMS is within the range of outcomes observed in either controlled or open-label studies of ECT.

The safety and efficacy of the NeuroStar TMS System is also described in this section in terms of its risk-to-benefit ratio as compared to current FDA-approved antidepressant therapies. As described below, the NeuroStar TMS System presents a favorable risk-to-benefit ratio as compared to FDA-approved antidepressants, including ECT.

5.2. Efficacy of the NeuroStar System as Compared to FDA-Approved Antidepressant Treatments

Comparison of Efficacy of TMS Therapy to Pharmaceutical Antidepressants

The extant literature for antidepressant medication registration clinical trials for the general antidepressant treatment-naïve population is the most relevant source for evaluation of antidepressant clinical trial outcomes. The Khan (2000) report is particularly instructive for evaluations using continuous outcome measures because it is a synthesis of this information that is available from the FDA under the Freedom of Information Act and contains the summary basis of approvals for the majority of the currently marketed pharmaceutical antidepressants. This study reported that the mean percentage reduction from baseline in HAMD17 total score across the entire dataset of antidepressants was 40.7% for the active treatment group, compared to a mean percentage change from baseline of 30.9% for placebo treated patients.

On average, these study results represent an overall relative advantage of approximately 9.8% in the reduction of total score from baseline on the HAMD17 when comparing active antidepressant treatment to placebo treatment in this treatment-naïve study population (see Figure 17).

By using the Khan (2000) FDA dataset analysis, it can be seen that a comparison of study results with the appropriate within-study control shows that, on average,

the observed relative advantage of active TMS compared to sham TMS (i.e., ~8% seen in Study 44-01101) is comparable to the relative advantage seen with any of the most commonly used pharmaceutical antidepressants (i.e., ~9.8%, range -10.0% to +33.6%). This is a particularly important observation since pharmaceutical antidepressants represent the predominant method of clinical management for major depression in the United States and therefore describes a benchmark that can be considered to define a clinically meaningful outcome in current practice.

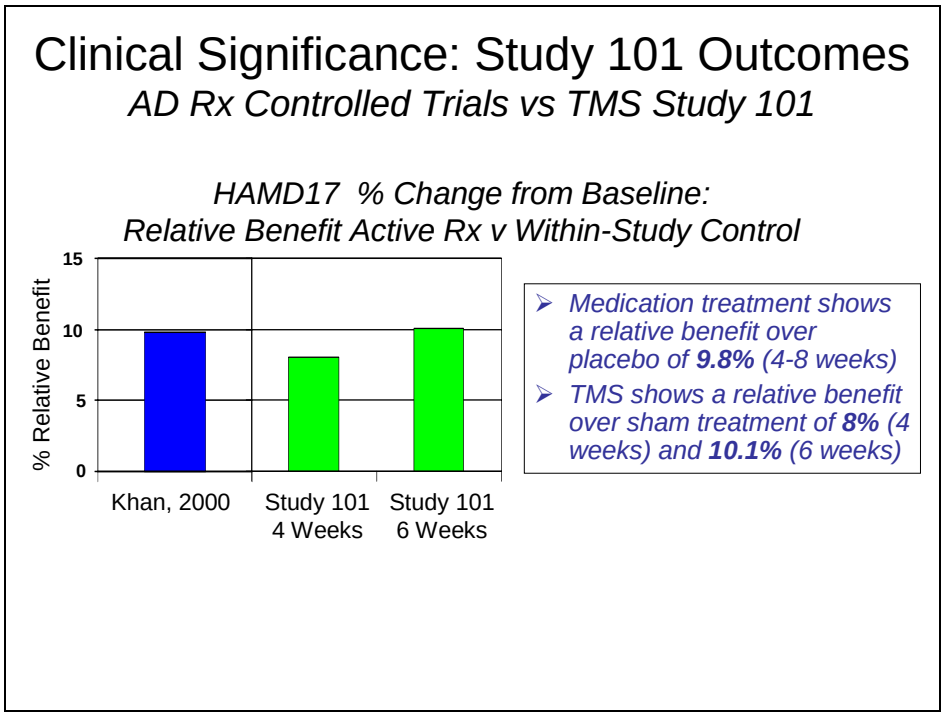


Figure 17. Relative Benefit of Percentage Change from Baseline on the HAMD17 Total Score Between Active and Placebo Conditions

Describing clinical results on continuous outcome measures alone provides only part of the story of effectiveness. The magnitude of relative benefit should also be described with regard to the categorical outcome measures. Khan and colleagues have not provided a similar summary of the FDA dataset on categorical outcome measures. Therefore, other reports that have aggregated data across large pooled samples of antidepressant studies were examined to estimate the clinically meaningful difference on categorical outcomes. The Walsh (2002) dataset is of interest because its size and comprehensive nature provides a thorough summary of the peer-reviewed, published literature on the clinical outcomes of currently marketed pharmaceutical antidepressants. This dataset is based on a summary of published reports of largely positive study results and excludes the observed results of unpublished negative studies, so it may be

reasonably assumed to define the upper bound, or best outcome estimate of clinically meaningful difference in categorical outcomes associated with drug therapy. The report of Thase and colleagues (2005) describes the pooled summary of results from 7 randomized controlled trials (N=1975) comparing the marketed antidepressant bupropion with several other SSRI antidepressants (fluoxetine, sertraline and paroxetine) or placebo, and includes studies with both positive and negative results and therefore serves to provide a reasonable lower bound estimate for response. Walsh (2002) cites that the average percentage of patients achieving the categorical outcome of response in these studies is commonly in the range of 31.6% to 70.4% in the active treatment study population (average active response rate of 50.1%), compared to an average percentage of patients achieving the categorical outcome of response of 12.5% to 51.8% for the placebo treatment condition (average placebo response rate of 29.7%). The Thase (2005) report noted that overall response rate in the active treated patients was 62.8% compared to 50.8% in the placebo treated patients.

These study results represent an overall relative advantage of approximately 1.2-fold (Thase, lower bound estimate) to 1.7-fold (Walsh, upper bound estimate) greater likelihood of responding to treatment when comparing active antidepressant treatment to placebo treatment on the HAMD17 in a relatively treatment-naïve study population **(Figure 18).**

A similar analysis for the HAMD17 data at week 4 in Study 44-01101 shows an overall relative advantage of ~**1.8-fold** for TMS relative to its relevant control. Therefore, the Neuronetics studies show a similar or greater advantage over its control than that seen in registration trials for FDA-approved antidepressants.

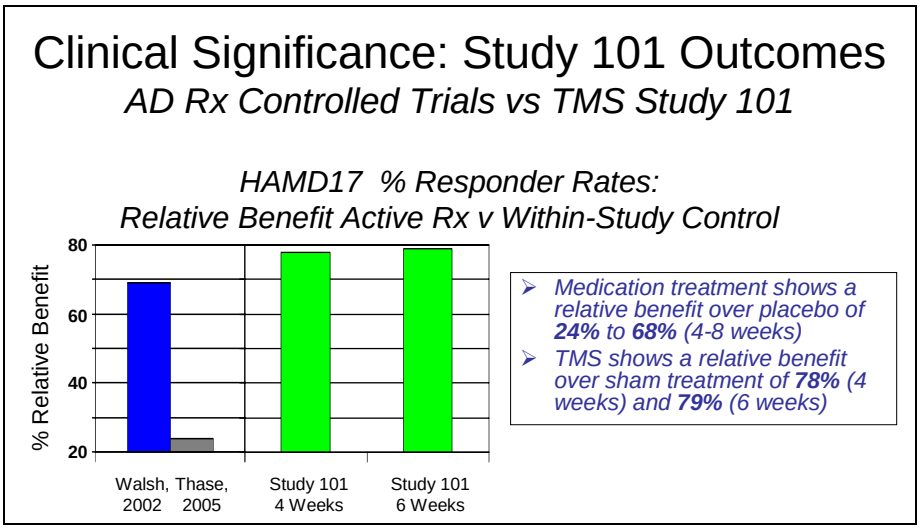


Figure 18. Relative Benefit of *Response Rate on the HAMD17 Total Score* Between Active and Placebo Conditions

In 510(k) K061053, the results from Neuronetics' open-label Study 44-01102 was presented as confirmatory evidence for the clinical significance of the observed treatment outcome. Although Study 44-01102 was an uncontrolled clinical trial and must be interpreted in that context, the pattern of results nevertheless should follow a specific and predictable path if they are to confirm the antidepressant effects of TMS seen in Study 44-01101. It would be expected that the cohort of patients who were previously allocated to active TMS in Study 44-01101 should represent a more treatment-resistant patient group than the group of patients previously allocated to sham TMS in that study. Therefore, the outcome for the patient cohort receiving extended active TMS should always show an inferior outcome to the group receiving first exposure to TMS. The results of Study 44-01102 confirmed this hypothesis showing superiority of the previously sham-allocated arm on all observed outcome measures.

510(k) K061053 also reviewed the outcome of the open-label Study 44-01102 compared to the recently reported results from the large, NIMH-funded, multi-site, open-label clinical trial, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. STAR*D was developed to provide a semi-naturalistic treatment algorithm to assess outcome as close to clinical practice as possible.

The results of Level 1 and Level 2 of that work were recently published (Trivedi, 2006a, 2006b; Rush, 2006). These two treatment Levels deal with the early continuum of the treatment resistance spectrum, largely patients who have persistent disease (most patients were in a recurrent course of illness), but who have minimal to no evidence of failure to receive benefit to date. The treatment history criteria defining entry into Level 1 of STAR*D was having no evidence of having failed to receive benefit from any of the various treatment options offered in the first two Levels of the study, as described below. In other words, the patients entering this study in Level 1 would not have qualified for entry into the Neuronetics studies because of insufficient evidence of resistance to treatment. On the other hand, many of the patients in the Level 2 to Level 3 continuum showed a prospectively demonstrated pattern of resistance to treatment consistent with the range of the treatment resistance history required for entry into the Neuronetics clinical development program. The primary outcome in the STAR*D study was remission, measured using the 17-Item HAMD, with an endpoint total score of less than 8, the same as the 17-Item HAMD remission endpoint used in the Neuronetics studies.

During Level 1, patients presenting with a clinically diagnosed major depression (N=2,876), and had not shown non-response or intolerance to any of the antidepressants to be used in the first two Steps of the STAR*D algorithm were eligible for entry. At entry to Level 1, the majority of the patients were relatively young (mean age = 40.8 years), however, the majority (~75%) had recurrent major depression. Their presenting HAMD17 total score was 21.8 (SD=5.2). In comparison, over 90% of the patients entering the Neuronetics clinical program

had a history of recurrent depression, their average age was nearly a decade greater (48 years of age), and their mean HAMD17 score at study screening was about 1-2 points greater (active TMS mean score of 22.7, SD = 2.37, and sham group mean score of 22.9, SD=3.13) than in the STAR*D cohort. This data points to a less severely ill population in the Level 1 STAR*D cohort than in the Neuronetics studies.

For the STAR*D population, after flexible-dose, open-label treatment with citalopram for up to 14 weeks, the overall remission rate was 27.5% by HAMD17 criterion. Those patients who did not respond to treatment in Level 1, were offered the opportunity to proceed to Level 2 (N=1292), where two additional options were offered in an equipoise statistical design. These included either switching medication (3 different choices: sertraline, bupropion-SR, or venlafaxine-XR), or antidepressant augmentation (2 choices: add-on bupropion-SR, or add-on buspirone). Outcomes in Level 2 revealed a slight decline in the rate of remission with the two treatment options: for the switching strategy, the average rate of remission across the 3 options was 21.2 % (range: 17.6% to 24.8%), while for the two augmentation options, the average remission rate was 29.9% (range: 29.7% and 30.1%). It is worth noting that in Level 2, the average HAMD17 total score at entry for the switching cohort was 18.9 (SD=7.3) and for the augmentation cohort was 15.8 (SD=7.1). While response rates for the HAMD17 were not reported in the study, based on the entry HAMD17 total scores, it may be concluded that these results are consistent with the range of outcomes expected based on the data from Khan and Walsh described above, and are consistent with a gradually diminishing rate of response with progressive treatment resistance.

It should be noted that the baseline HAMD17 total scores in Level 2 are at or below the values observed for patients entering the Neuronetics studies, indicating that a proportion of these patients still precede the clinical severity of the population studied in the Neuronetics sample in terms of symptom severity and treatment resistance.

The results of Level 3 and Level 4 were very recently publicly reported (Fava, et al, 2006, Nierenberg, et al, 2006, and McGrath, et al, 2006), and therefore not previously included in 510(k) K061053. These data are included in the STAR*D information given in Figures 19 and 20 below.

In Level 3, two potential options were offered. In the first option, those patients (N=235) who had not achieved remission on their chosen option at Level 2, were offered 14 weeks of monotherapy with either mirtazapine (up to 60 mg per day) or nortriptyline (up to 200 mg per day). The second Level 3 option randomly assigned patients (N=142) to augmentation with either lithium (up to 900 mg per day) or tri-iodothyronine (T3, up to 50 mcg/day). The remission rates using the HAMD17 (total score at endpoint < 8) are reported for Level 3 as 12.3% for mirtazapine and 19.8% for nortriptyline, with an overall pooled remission rate of

16.2% across both treatment options. For the second strategy, the pooled remission rate for augmentation was 20.4%. The final Level in the STAR*D algorithm for patients who had not achieved remission at any of the preceding Levels was random assignment in Level 4 to either tranylcypromine (N=58; up to 60 mg/day) or the combination of venlafaxine XR (N=51; up to 300 mg/day) plus mirtazapine (up to 45 mg/day). The remission rates at this Level were reported as 6.9% for the tranylcypromine group and 13.7% for the venlafaxine plus mirtazapine group.

Figure 19 compares the open-label data from Study 44-01102 (Group B patients who were treated with sham in Study 44-01101 before receiving open-label TMS in Study 44-0102) using remission outcomes at the Week 6 and Week 9 time points (i.e., after 3 weeks of taper from TMS to pharmacologic antidepressant treatment) to the STAR*D trial data using similar clinical outcomes (i.e., remission as defined by the 17-item HAMD rating scale). As shown, the Neuronetics Study 44-01102 data compare favorably against the STAR*D data, an important benchmark study of treatment resistant depression.

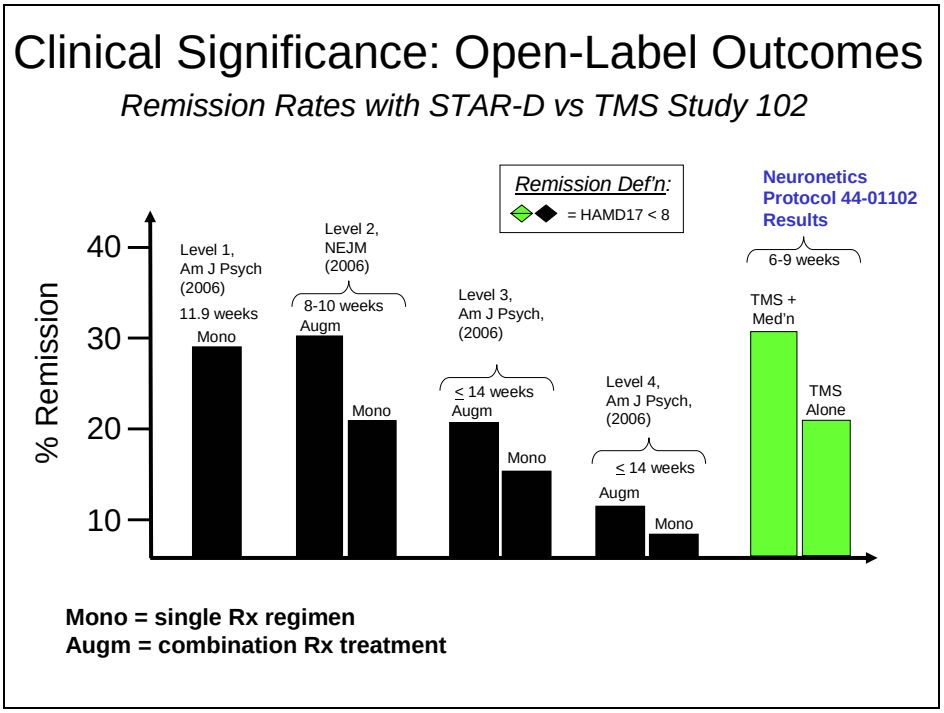


Figure 19. STAR*D Clinical Outcomes (Level 1 through 4) Compared to Clinical Outcomes Observed in Study 44-01102 for Group B Patients (Remission defined as HAMD17 < 8)

Because the study sample of Study 44-01102 is composed of a composite of patients whose treatment resistance history spans across STAR*D Levels 2 through 4, remission outcomes from Study 44-01102 are shown in the following 3

figures, stratified by the ATHF treatment resistance level of the patients, in order to more precisely compare the outcomes observed relative to the relevant treatment-resistance-matched cohort observed in STAR*D. Figures 20, 21, and 22 show the results for ATHF treatment resistance Levels 1, 2 and 3, respectively, for Group B patients (sham treated in Study 44-01101 before open-label TMS in Study 44-01102). As shown, the results obtained in Study 44-01102 compare favorably to the outcomes reported in the STAR*D project, when stratified for the degree of historical treatment resistance [for the ATHF category of 4 treatment failures in current episode, there were no patients who achieved remission (N=3)].

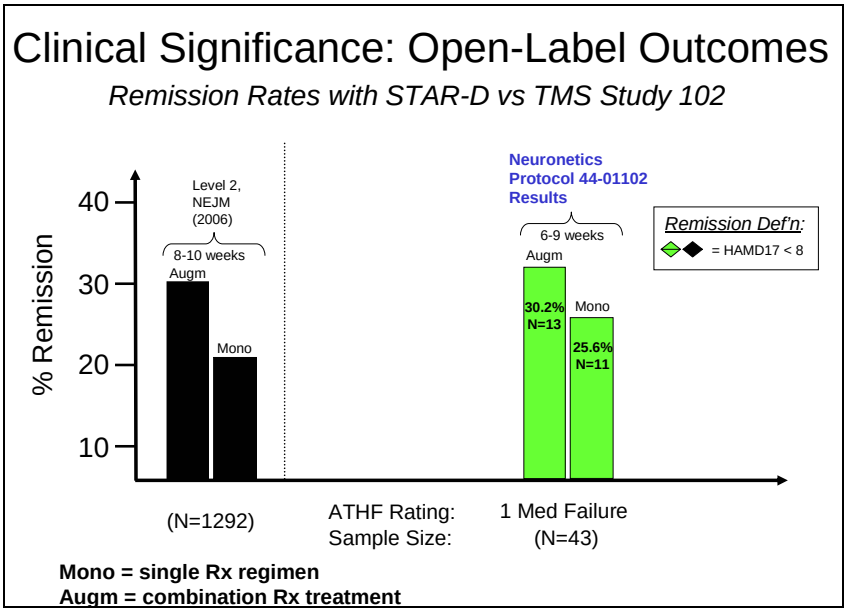


Figure 20. STAR*D Clinical Outcomes (Level 2 - 1 medication failure) Compared to Clinical Outcomes Observed in Study 44-01102 for Group B Patients (ATHF Level 1 – 1 medication failure)

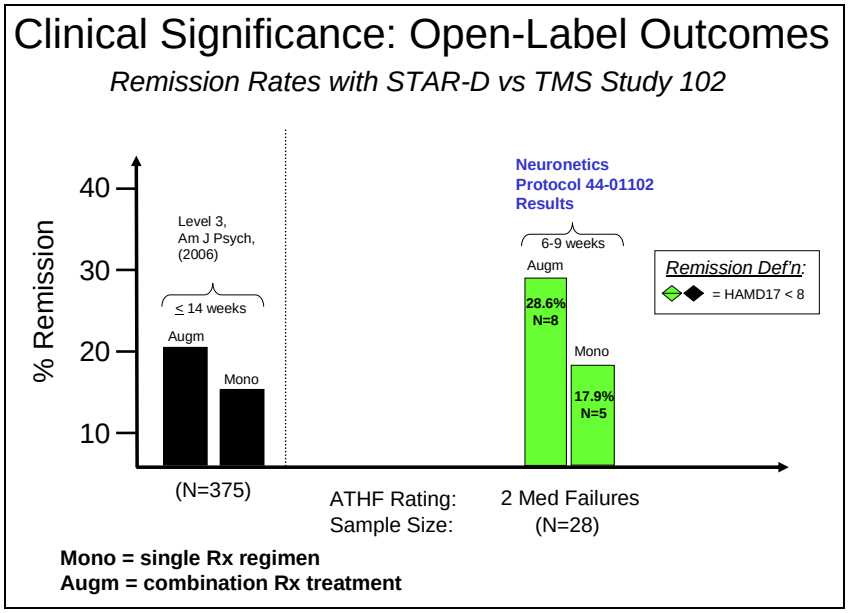


Figure 21. STAR*D Clinical Outcomes (Level 3 – 2 medication failures) Compared to Clinical Outcomes Observed in Study 44-01102 for Group B Patients (ATHF Level 2 – 2 medication failures)

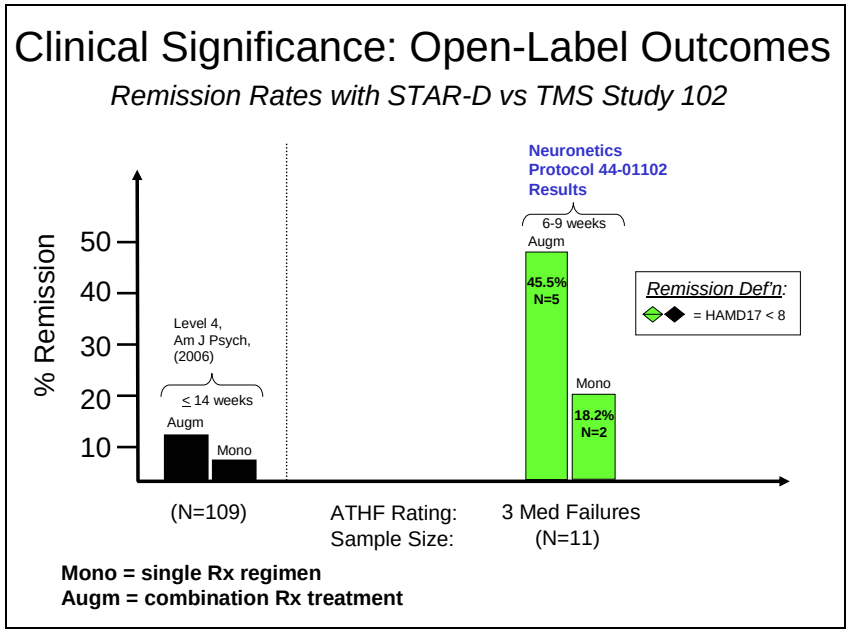


Figure 22. STAR*D Clinical Outcomes (Level 4 – 3 medication failures) Compared to Clinical Outcomes Observed in Study 44-01102 for Group B Patients (ATHF Level 3 – 3 medication failures)

Comparison of TMS Therapy Efficacy to Clinical Outcomes with ECT

The published literature on the use of ECT in the treatment of depression is extensive. In 510(k) K061053, Section 12: Substantial Equivalence, we provided a justification for focusing on several authoritative, peer-reviewed references in order to provide a comparative summary of the efficacy and safety of ECT relative to the outcomes observed with TMS Therapy in the Neuronetics studies. In particular, two important sources provide comprehensive, organized, and authoritative summaries of the ECT literature. These are the 2nd edition of the American Psychiatric Association's Committee on Electroconvulsive Therapy Report on The Practice of ECT: Recommendations for Treatment, Training, and Privileging, and the recently published findings of the United Kingdom ECT Review Group (The UK ECT Review Group, 2001), that employed a systematic review and meta-analysis of the ECT literature modeled after the search strategies of the Cochrane Library.

More recent studies of ECT are also available from the peer-reviewed, published reports from several important US federally-funded studies of ECT practice. The first report is the publication from Sackeim and colleagues from Columbia University, referred to here as the Optimization of ECT (OPT/ECT) Study (Sackeim, HA,, et al., 2001). This study was a randomized, double-blind, placebo-controlled trial of continuation pharmacotherapy following ECT response. The durability of the acute clinical response to ECT was compared among three different treatment conditions, nortriptyline, the combination treatment of nortriptyline and lithium, and placebo. The second study was also conducted by Sackeim and colleagues and is referred to as the Community ECT Study (Prudic, J., et al., 2004). This study was conducted as a prospective, naturalistic study of a large sample of depressed patients treated at clinically diverse community treatment settings. The overall effectiveness of ECT was measured, and patients were followed for up to 24 weeks to assess their outcome and its relation to the variations in the clinical practice of ECT in the community.

The third study reports results from the Consortium for Research in ECT Study (CORE Study), conducted by Kellner and colleagues (Kellner, CH, et al., 2005; 2006; and Petrides, G, et al., 2001). This study was a collaborative, randomized controlled study conducted in two phases. The first phase examined the acute efficacy of ECT in achieving remission in patients with major depression, and the second phase was a continuation phase comparing continuation ECT or continuation pharmacotherapy (nortriptyline and lithium) in sustaining the acute remission over 6 months of treatment. In aggregate, these reports provide supportive, uncontrolled evidence of acute efficacy and randomized, controlled evidence of the durability of response to ECT.

There are several important observations that can be made regarding ECT. First, similar to the observations discussed above for pharmaceutical antidepressants, the same relationship between prior treatment failure and subsequent treatment

response has also been reported in the ECT literature. For instance, Prudic and colleagues (Prudic, J, et al., 1996) reported the acute antidepressant outcome to open-label ECT treatment in a cohort of 100 patients who met Research Diagnostic Criteria for unipolar major depression. In this study, patients were rigorously staged in terms of their antidepressant resistance using the Antidepressant Treatment History Form (ATHF), the same methodology used in the Neuronetics clinical studies. The HAMD (24-item version) was used to assess outcome to acute treatment immediately after the last ECT session and then in follow up one week later. Treatment was *to effect* and was an average of 8.9 (SD=2.8) ECT sessions in patients who had ATHF-confirmed medication resistance, and 9.9 (SD=4.0) ECT sessions in patients who did not have such confirmed resistance. The overall remission rate was reported and was operationally defined as having at least a 60% reduction of HAMD score from baseline, and a maximum total score of no more than 10. Immediately after the last treatment, 73.0% of the overall sample met these remission criteria, while only 63.1% of those patients who had confirmed medication resistance met such criteria. After one week, 57.0% of the overall sample continued to meet remission criteria, while 47.7% of the confirmed medication resistant sample continued to meet these criteria.

A second major observation is that ECT is clearly an effective short term antidepressant. Indeed, the summary analysis provided in the UK ECT Review Group Report reports a standardized effect size outcome using the HAMD17 for ECT of 0.91 among the group of randomized, simulated ECT-controlled studies that withstood rigorous methodologic inspection for trial validity used by that study group. When compared to the observed effect size on the same instrument (HAMD17 = 0.55) in the Neuronetics controlled trial Study 44-01101, the comparison suggests that TMS may be estimated as roughly two-thirds as effective as ECT in acute outcome.

This observation regarding the acute efficacy of ECT from the historical controlled literature has also been borne out in results from the more contemporary open-label studies noted above (Sackeim, 2001; Prudic, 2004; Kellner, 2005; 2006; and Petrides, 2001). It is notable, however, that the outcomes reported in the open-label acute efficacy literature for ECT span a wide range. For instance, research samples such as the OPT-ECT and CORE studies generally report remission rates using the HAMD24 in excess of 70% at the end of acute treatment. In contrast, the remission rates observed in the large, hospital-based Community ECT study are more modest and generally less than 50% (36.4% - 57.1%). When compared to the open-label remission rates with the HAMD24 observed in the Neuronetics studies (27.1% - 36.5%), it can similarly be noted that the acute effect of TMS approaches the lower end of the estimates reported for ECT.

Finally, a review of the recently published continuation of effect literature for ECT has highlighted that the acute effect of ECT is difficult to sustain, despite

aggressive attempts with either combination pharmacotherapy (Sackeim, 2001), or even in the face of continuation ECT, as recently discussed in the outcome from the CORE Study (Kellner, 2006). In the continuation of effect data summarized for the NeuroStar TMS system in Study 44-01103, it can be seen that the persistence of benefit achieved with TMS Therapy (4 week relapse rate = 9.1%) is at least comparable, if not favorable, to that observed in the ECT literature (4 week relapse rate = 4.5% - 36%).

Overall Comparison of Efficacy of TMS Therapy with Pharmacotherapy or ECT

Taken together, the data summarized in this section indicates that NeuroStar TMS Therapy is at least as effective as currently available pharmaceutical interventions for the treatment of major depression, and achieves an acute outcome that appears roughly two-thirds as effective as ECT devices. This is graphically summarized in Figure 23 which shows the reported standardized effect size observed with the HAMD17 outcome measure in the Neuronetics Study 44-01101, compared to the effect sizes observed in both the large FDA SBA database for pharmaceutical antidepressants analyzed by Khan (2000), and the controlled trial data summarized for ECT in the UK ECT Review Group Report (2003).

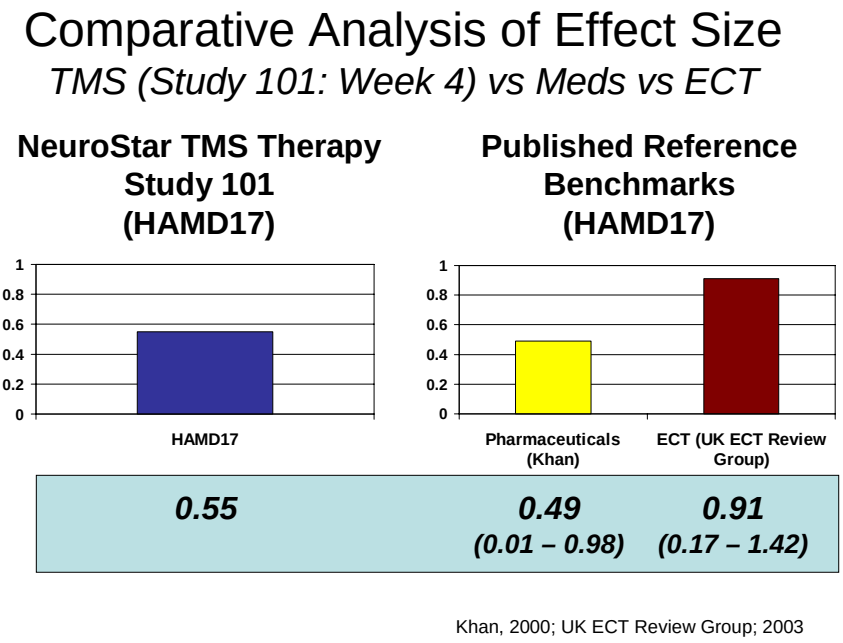


Figure 23. Comparative Analysis of Effect Sizes: NeuroStar TMS Therapy versus Pharmacotherapy and ECT

5.3. Clinical Significance: Risk-to-Benefit

The clinical significance of any treatment can only be answered by an assessment of the *clinical benefit* of the treatment, i.e., the effect size, in relationship to its *associated risks*, i.e., an appraisal of the risk-benefit trade-off afforded by the treatment option. The safety and tolerability profile of antidepressants currently used in the clinical management of major depression are known to be increasingly intolerable to the patient with each incremental increase in treatment resistance, and bring with them an associated increasing risk of medical toxicity. This can be seen in the observed discontinuation rate due to adverse events reported in the large STAR*D dataset. As patients proceed from Level 1 through Level 3 alone, the discontinuation due to adverse events increases in an almost monotonic manner (8.6% at Level 1, ~20.5% (range: 12.5% - 27.2%) at Level 2, and 35.2% (range: 34.2% - 36.2%) at Level 3. This discontinuation rate is due to the known adverse effects associated with the pharmaceutical options offered to patients in that study as the Levels advanced (e.g., tricyclic antidepressants, lithium and thyroid hormone augmentation). In clinical practice, the range of combination therapeutics and augmentation strategies, including the use of such agents as the newer atypical antipsychotic agents, only aggravates this issue for patients. With regard to electroconvulsive therapy, the limiting toxicity in clinical practice is even more dramatic, with short-term confusion and cognitive dysfunction seen in all patients, and sustained deficits in long-term memory seen in a clinically significant subset of patients. Therefore, to fully consider the clinical significance of an antidepressant treatment, it is the *combination of effect size and risk* that must be considered.

As discussed below, the overall risk-to-benefit ratio, using an analysis of effect size versus risk, is approximately 2-3 fold more favorable for TMS than for antidepressant medications.

In this section, we use currently accepted metrics in evidence-based medicine research to present a detailed analysis of the overall risk-to-benefit ratio of TMS compared to pharmaceutical antidepressants. We believe this is a helpful reference guidepost to gauge clinical significance for several reasons. First, as noted above, the datasets available for pharmaceuticals are large and current. Second, the outcome measures and study methodologies used in those studies closely resemble the methods used in the Neuronetics studies, lending credibility to the analyses. Finally, since pharmaceuticals are probably the most commonly used approach to the treatment of major depression, what magnitude of benefit represents a clinically significant effect in the eyes of clinicians is well understood. Using this analysis of effect size versus risk, it is demonstrated that TMS Therapy has approximately a 2-3 fold more favorable risk-benefit profile than pharmaceutical antidepressant medications. Though study design limitations make a similar comparison to ECT less easy to perform, comment is provided on the similar conclusions that can be drawn for the comparative risk-benefit profile for TMS and ECT from the available literature.

Various methods to quantify effect size have been proposed in the large field of evidence-based medicine. For example, the Cochrane Collaboration provides an extensive discussion of these concepts on their website (http://www.cochrane-net.org) and Kraemer and Kupfer (2006) have also recently provided a useful summary of analytical techniques for the assessment of clinical significance. During their review of 510(k) K060153, the FDA suggested that *relative effect size* should be used to compare the effect in the active TMS group with the sham TMS group, and therefore this was used above to present the Neuronetics data with several comparison datasets.

Many authors in the field of evidence-based medicine advocate the use of *absolute effect size*, using the number needed to treat (NNT) as an easy-to-calculate metric for this estimate. NNT is expressed as the inverse of the risk difference between two treatment options, and can be calculated from dichotomous outcome data, such as the differences in response rates. Stated another way, the NNT represents the number of patients one would expect to treat with T (active treatment) to have one more success (or one less failure) than if the same number were treated with C (control treatment) (Altman and Andersen 1999; Cook and Sackett, 1995). A related metric, the number needed to harm (NNH), represents the absolute risk difference between two treatment options with regard to a specific, clinically meaningful outcome that can be similarly dichotomized, such as the discontinuation rate due to adverse events. The ratio of these two estimates of effect size can then be compared in order to appraise the clinical benefit obtained (i.e., NNT) for a specific intervention, compared to the clinical risk of that option (i.e., NNH).

The NNT (Number Needed to Treat) and NNH (Number Needed to Harm) used in the analysis of RCTs are especially useful because they emphasize the effort that must be expended to accomplish a single, tangible outcome. NNT conveys the effort required to achieve a positive outcome without distinguishing between the presence or absence of treatment-related adverse events. Similarly, NNH conveys harm without accounting for the achievement or lack of achievement of the benefit of therapy. More recently, analyses of RCTs include the combination of NNT and NNH to represent the effort required to achieve trial success or failure in the context of treatment-induced side effects. This can be achieved by the determining the ratio of NNH/NNT using the discontinuation rate data at a given study time point as an estimate of the NNH. Using this metric, the *larger the ratio* of NNH to NNT, *the more favorable the risk to benefit profile* of the treatment option.

The discontinuation rates for currently used pharmaceutical antidepressants as described in the relevant product labels, and similar data observed in Neuronetics Study 44-01101 at 4 weeks and their calculated values for Number Needed to Harm (NNH) are shown in Table 23.

Table 23. Discontinuation Rates and Number Needed to Harm for Currently Marketed Pharmaceutical Antidepressants and in Neuronetics Study 44-01101

Antidepressant	D/C rate due to Adverse Events		NNH
	Active Treatment	Placebo Treatment	
Duloxetine	10%	4%	16.7
Fluoxetine	12%	9%	33.3
Effexor XR	11%	6%	20
Remeron	16%	7%	11.1
Symbyax	10%	4.6%	18.5
Paroxetine	20%	Not reported	-
Bupropion	11%	4%	14.3
Sertraline	Not reported	Not reported	-
Average			**19.0**
Neuronetics Study 44-01101 @ 4 weeks	4.5%	3.4%	**91.0**
Neuronetics Study 44-01101 @ 6 weeks	5.8%	3.4%	**41.7**

Table 24 provides the calculated NNT for the categorical outcomes on active and placebo treatment groups in Neuronetics' Study 44-01101 as compared to the same calculations for the categorical outcomes described in the reference pharmaceutical datasets discussed above, i.e., from Walsh (2000) and Thase (2005). The table also provides the ratio of NNH/NNT for the three datasets.

Table 24. NNH and NNT for Pharmacologic Antidepressants and Neuronetics TMS

Effect Size Metric	Neuronetics Study 44-01101		Antidepressants	
	Week 4	Week 6	(Walsh, 2002)	(Thase, 2002)
HAMD17 Response Rate - Active	20.6%	24.5%	50.1%	62.8%
HAMD17 Response Rate - Sham	11.6%	13.7%	29.7%	50.8%
NNT - Number Needed to Treat	11.1	9.3	4.9	8.3
NNH - Number Need to Harm	91	41.7	19	19
NNH/NNT (Risk/Benefit Ratio)	**8.2**	**4.5**	**3.9**	**2.3**

As shown in Table 24, Neuronetics' TMS Therapy shows a more favorable risk-to-benefit ratio as compared to FDA-approved pharmaceutical antidepressants. NNH/NNT for antidepressant medications ranged from 2.3 to 3.9, while for Neuronetics' TMS Therapy, it ranged from 4.5 to 8.2.

This analysis demonstrates that the overall risk/benefit ratio for TMS is about 2-3 fold more favorable than for antidepressant medications.

The estimate provided here indicates that, for pharmaceutical antidepressants which are the most commonly used treatments for depression, for every ~3 patients, one will have discontinued treatment because of intolerance to the adverse effects of the treatment compared to one success. For TMS Therapy, this equation is much more favorable, namely for every ~6 patients, one will have discontinued treatment because of intolerance to the adverse effects of the treatment compared to each treatment success. Neuronetics' TMS Therapy clearly compares favorably as a treatment option for major depression.

A similarly detailed quantitative estimate of NNT and NNH for ECT is not as easily described. There are several methodologic reasons for this. For example, in the older controlled trial literature, response rates were often not reported.

Similarly, the clinical study context for the administration of ECT raises a question as to what the appropriate comparative endpoint for estimate of NNH should be. If one uses the nearly universal consequence of post-treatment confusion and cognitive dysfunction, the estimates of the NNH/NNT ratio may be excessively biased against ECT. On the other hand, focusing on only clinically more problematic outcomes such as persistent long-term amnesia or death may be underestimating the risk profile for this treatment.

Nevertheless, a reasonable clinical conclusion can be drawn regarding the comparative risk-benefit profile of ECT as compared to TMS using the evidence previously summarized in 510(k) K061053, (see CD-ROM Attachment 8, Section 12: Substantial Equivalence). This summary indicates that there are a number of well-known and important medical risks associated with ECT treatment. We assert that these risks, and the ubiquity of the short-term cognitive deficits associated with this treatment, lead to the self-evident conclusion that the risk-benefit ratio offered by TMS Therapy represents a substantially more favorable option for the patient with major depression who is considering ECT.

In summary, TMS Therapy provides a favorable option for patients with MDD and, in particular, for the patient who may be considering, among the next available options, the use of complex pharmaceutical intervention or the more invasive intervention afforded by ECT.

5.4. Potential Role of TMS Therapy in Treatment Planning for Major Depressive Disorder

The data presented in Section 4 of this Executive Summary indicates that TMS Therapy as delivered by the NeuroStar System is efficacious in the treatment of depression and therefore offers a logical treatment option for patients with major depressive disorder, including those patients who have failed to receive clinical benefit from previous antidepressant therapy.

MDD patients in Neuronetics studies had to have failed to receive benefit from at least one adequate dose and duration of antidepressant treatment and represented a moderate-to-severe disease, difficult-to-treat patient population, consistent with the patient population considered eligible for ECT. Treatment of this MDD population with the NeuroStar System in the multicenter Study 44-01101 resulted in a *clinical response* rate (i.e., 50% reduction in depression symptoms) after 4 to 6 weeks of treatment that ranged from ~20-25%. In the open-label Study 44-01102, a clinical response rate after 4 to 6 weeks of treatment ranged from ~20-40%.

The response rates observed in Neuronetics Studies 44-01101 and 44-01102 are equivalent to or greater than the response rates reported in the available literature for a patient population (i.e., STAR*D) that has shown limited clinical benefit using available pharmacotherapeutic interventions.

The data presented here demonstrate that the efficacy achieved with the NeuroStar System lies within the range of effectiveness expected for currently available treatment options for this MDD population, including the difficult-to-treat patient. Therefore, the NeuroStar System provides the depressed patient an opportunity for an effective clinical outcome that compares favorably to the likelihood of response to further courses of pharmacotherapy for treatment of their illness. In addition, for those patients who experience intolerable adverse effects with pharmacotherapy, the virtual absence of systemic adverse effects with TMS Therapy provides an additional rationale for clinical consideration in these patients.

In general clinical practice, the use of ECT as a treatment option requires the ***weighing of the potential benefit of ECT against the risks*** of ECT treatment for a particular patient. The substantially unfavorable tolerability profile of ECT relative to other therapeutic options, as well as the societal stigma associated with its use, places ECT in the later portion of the continuum of treatment choices for the patient with major depression and may not be accessed by many patients.

The safety and efficacy data presented support the opportunity for the Neuronetics NeuroStar System to expand the range of potentially effective treatment options for patients with major depressive disorder. As shown in Figure 24, patients whose only treatment options are combination pharmacotherapy, ECT, or VNS are faced with the choice of potential wellness versus the troubling invasiveness and side effects of these treatments. As a treatment option for MDD, TMS Therapy as delivered by Neuronetics' NeuroStar System should appropriately occupy a position in the armamentarium of available antidepressant treatments, and may demonstrate highest use and benefit when placed intermediate between one or more medication trials on the one hand and ECT on the other.

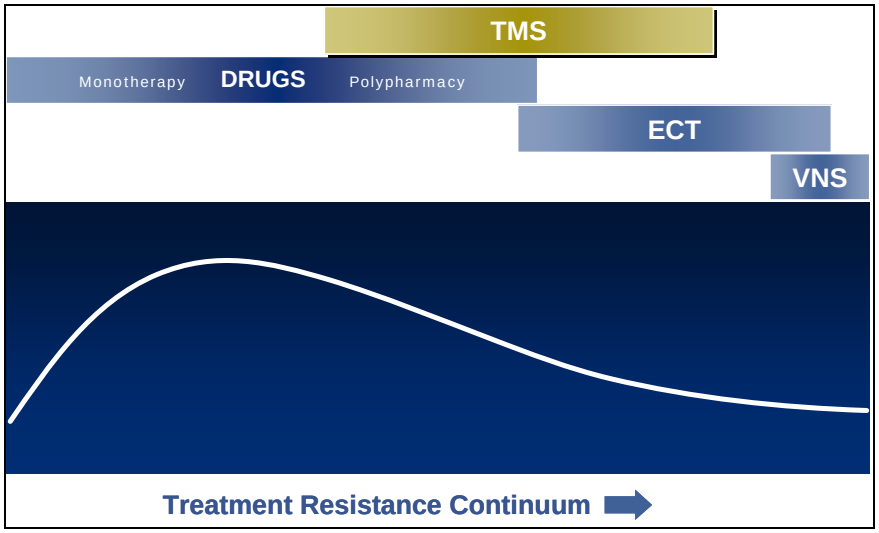


Figure 24. Placement of Neuronetics' TMS Therapy as a Therapeutic Option for Major Depressive Disorder

SECTION 6. SUBSTANTIAL EQUIVALENCE OF THE NEUROSTAR SYSTEM TO ECT DEVICES

6.1. Basis for Substantial Equivalence

In Neuronetics' premarket 510(k) notification, the NeuroStar System, which delivers transcranial magnetic stimulation (TMS), is shown to be substantially equivalent to Electroconvulsive Therapy (ECT) devices, the predicate devices, for the intended use of treatment of Major Depressive Disorder.

The NeuroStar System is substantially equivalent to ECT devices because:

- the NeuroStar System and ECT devices have the ***same intended use***;
- technological differences such as differences in design, materials, and in energy source (magnetic vs. electrical stimulation, respectively) between TMS and ECT devices ***do not adversely affect safety or effectiveness*** in a way that is consequential under the conditions of intended use;
- the NeuroStar System does not raise any new types of questions about safety or effectiveness in treatment of MDD as compared to ECT devices;
- *an* ***accepted scientific method*** (i.e., a randomized, sham-controlled clinical trial, Study No. 44-01101, see Final Study Report for Study 44-01101) has demonstrated that the NeuroStar System *is safe and effective* in the treatment of Major Depressive Disorder; and
- the safety and effectiveness of the NeuroStar System provides a ***clinical advancement*** as compared to ECT devices with a *risk-to-**benefit ratio that is favorable*** to the predicate ECT devices .

There is substantial regulatory precedence for FDA's clearance of non-implantable *magnetic* stimulators for the same purposes (indications) as non-implantable *electrical* stimulators. This includes several indications for use including evoked response or nerve conduction velocity studies, e.g. (K002889, K992911), stimulation of the pelvic floor as a treatment of incontinence (K973096) and for muscle stimulation (K973929). The NeuroStar System uses magnetic stimulation for the same intended use as electrical stimulators, ECT devices, for the treatment of major depressive disorder.

6.2. Comparison of NeuroStar TMS Therapy and ECT in Treatment of Major Depressive Disorder

A head-to-head study of NeuroStar TMS Therapy with ECT was not practically possible and therefore, as discussed with the FDA, comparison of TMS and ECT effects was obtained from a comparison of the results of Neuronetics' studies with available ECT data from the medical literature.

The clinical response to ECT has been shown in single-site clinical studies to be high but variable (UK ECT Review Group, 2003). ECT is clearly an effective antidepressant, however, it is difficult to compare ECT treatment with other antidepressant therapies given that there are no well-designed and methodologically sound head-to-head comparative studies and given that the design of ECT studies differs considerably from typical antidepressant studies. For example, ECT treatment is given "to effect", while other antidepressant studies select a standard treatment time for effect (i.e., 1-3 months). Given the difficulty in recruiting patients for ECT studies, it is also likely that ECT studies select a particular patient population (i.e., ECT responders) that may contribute to the high efficacy values reported in some studies. The risk of a large placebo contribution to the clinical effect is considerable with ECT, given the elaborate clinical ritual, inpatient care, and substantial personal attention provided during the course of an ECT treatment sequence (APA Committee on ECT, 2001). The rapid loss of effect in the early days and weeks after a course of ECT has also been suggested to be due to the placebo contribution to the overall clinical effect of this treatment. Indeed, all of these variables raise questions regarding the interpretation of the true effect size for ECT.

In controlled studies of ECT where effect was determined, simulated ECT (i.e., general anesthesia without seizure induction) was employed as the control condition (UK ECT Review Group, 2003). Since those early studies, the acute efficacy of ECT has been presumptively established, and for ethical considerations, no additional randomized, controlled trials of acute efficacy have been recently conducted.

With the limitations of the existing datasets for ECT in mind, Neuronetics' premarket 510(k) notification, Section 12: Substantial Equivalence (CD-ROM Attachment 8), provides an extensive discussion of the relevant ECT literature that can be used to compare effect size and safety to similar data obtained from Neuronetics' clinical studies of the NeuroStar System. This is summarized below.

6.3. Demographics and Efficacy of TMS Therapy versus Electroconvulsive Therapy (ECT)

Demographic Comparison of ECT and TMS Patient Populations

Evidence for the efficacy of ECT in the treatment of major depressive disorder was drawn from 6 randomized, controlled clinical trials cited in the UK ECT Review Group Report (2003). In those studies, simulated ECT (i.e., general anesthesia without seizure induction) was employed as the control condition.

The patient populations in these ECT studies and in the Neuronetics studies were compared for patient demographic variables and for treatment resistance as described in 510(k) Section 12: Substantial Equivalence (CD-ROM Attachment

8). Overall, the ECT and Neuronetics studies included patients that were similar in general demographics and in the level of treatment resistance (ECT studies = range (0.7 to 1.07 adequate medication failures; Neuronetics studies = 1.6 adequate medication failures).

Two of the more recent ECT studies, the Optimization of ECT Study (Sackeim, et al., JAMA [2001]), and the Community ECT Study (Prudic, et al., Biological Psychiatry [2004]) were of particular use in comparing ECT patient characteristics with the Neuronetics sample because they used the same instrument for assessment of treatment resistance, the ATHF. Comparison of these datasets also indicated the two populations were comparable in demographic features such as baseline illness history, treatment resistance profile, symptom severity, and clinical outcomes (see CD-ROM Attachment 8).

One question raised by the FDA regarding these analyses was that a portion of the reference ECT populations also contained patients diagnosed with either psychotic depression, bipolar illness or both, which were both exclusions in the Neuronetics studies. These patients had been excluded, at the specific request of the FDA, in order to focus Neuronetics clinical studies on the population with the single indication of DSM-IV-defined major depressive disorder.

Neuronetics was able to address the FDA's question with an analysis that was conducted of the two recent ECT studies by Dr. H. Sackeim, the principal investigator of the two reference datasets. This analysis provided a summary of the subset of patients in each study that were most comparable to the inclusion criteria used for the Neuronetics study, by *excluding* those patients who met one of the following 4 descriptive criteria; diagnosis of psychosis, diagnosis of bipolar disorder, age > 70, or current illness duration > 3 years. The methodology of this analysis and further details of the results are provided in Tab 4.

In brief, a total of 47.9% percent (139/290) of the eligible entry population in the OPT-ECT Study, and 37.2% (129/347) of the eligible entry population for the Community ECT Study overlap with the Neuronetics study population on the major inclusion/exclusion criteria used in these studies. Data from these subgroups were then compared to the active TMS and sham TMS treatment arms for the Neuronetics sample for age, gender, recurrent illness course, duration of current episode, current episode duration > 2 years, # adequate antidepressant treatments in current episode (ATHF criteria), and baseline symptom severity by HAMD24.

The results of this analysis confirmed the conclusions of the original comparison of the reference ECT datasets with the Neuronetics population as described in Section 12: Substantial Equivalence (CD-ROM Attachment 8). They indicate that the patient population included in the Neuronetics study overlaps substantially with the patients recruited for ECT treatment in both research and community settings, and share a significant degree of clinical similarity to those

patient samples. Areas of difference are subtle and suggest that the Neuronetics patient sample shows a ***slightly greater degree of illness recurrence and chronicity*** than in the ECT reference populations.

Efficacy

Evidence for the efficacy of ECT in the treatment of major depressive disorder was drawn from the 6 randomized, controlled clinical trials cited in the UK ECT Review Group Report (2003). The data review focused on the primary outcome measure cited in most reports which was the 17-Item HAMD. The UK ECT Review Group also computed the standardized effect size for the individual studies and pooled for all studies.

For the 6 ECT studies cited in the UK ECT Group report, the ECT pooled outcome measure across studies resulted in an estimated effect size of 0.91 (range of 0.17 to 1.42) using HAMD17 response rates. For Neuronetics' studies, using the same metric for HAMD17, the standardized effect size is 0.55. For the most responsive subgroup in Study 44-01101, the ATHF Level 1 (see Section 4.3.1.5.2.1), the standardized effect size using HAMD17 response rate is 0.83. These data indicate that response to NeuroStar TMS Therapy is consistent with the range of effect sizes observed in controlled clinical trials of ECT.

A more contemporary dataset for ECT efficacy is available from the "open-label" community ECT study (Prudic, 2004). In this study, which evaluated remission (criterion: >60% decrease from baseline in HAMD24 score and endpoint score <10), remission rates ranged from 36.4% to 57.1% across community sites. In the comparable Neuronetics' open-label Study 44-01102, remission rates at 6 weeks (end of acute phase) and 9 weeks (end of taper phase) (criterion: endpoint HAMD24 score <10) were 27.1% to 36.5%, respectively, and therefore are consistent with the remission rates observed with ECT.

In summary,

- The Neuronetics study population shows clinically meaningful and substantive overlap with the population treated with ECT with regard to clinical diagnosis, demographics, symptom severity and prior treatment failure.
- Results of the randomized, controlled efficacy trial of the NeuroStar TMS System, Study 44-01101, shows a standardized effect size using the HAMD17 of 0.55, that is within the range of HAMD17 values observed in randomized controlled trials of ECT in the UK ECT Review Group report (0.17 to 1.42)
- Similarly, comparison of the categorical clinical remission rates (HAMD24 < 11) observed after 6 weeks (end of acute phase) and 9 weeks (end of taper phase) in the Neuronetics open-label extension Study 44-01102 (27.1% to 36.5%) met the lower range of HAMD24 remission rates observed in the open-label Community ECT Study (36.4% to 57.1%).

6.4. Durability of Effect: TMS Therapy versus ECT

Persistence of clinical effect after acute response to treatment with the NeuroStar System is comparable to that seen with ECT.

Continuation course and relapse rates for ECT from two peer-reviewed, published studies are described in detail in 510(k) Section 12: Substantial Equivalence (CD-ROM Attachment 8). In these studies, after acute ECT, maintenance with pharmacotherapy resulted in a 4.5% relapse rate at 4 weeks for one study (Prudic, 2004) and 26-36% at 4 weeks for the second study (Sackeim, 2001). ***Using the same definition for relapse as used in these studies, Study 44-01103 showed a relapse rate of 9.1% at 4 weeks***.

At 24 weeks, in the absence of formal continuation treatment, relapse of illness after acute response to ECT is inevitable in nearly all patients. With pharmacotherapy, relapse after ECT occurred in 50-63% of patients (Sackeim, 2001). ***Interim results from Study 44-01103, using the same definition of relapse, show that relapse occurred in 20% of patients in this study after 24 weeks.***

These data indicate that the persistence of effect at 1 month with pharmacotherapy after acute treatment with the NeuroStar System is at least comparable to that seen with ECT with pharmacotherapy (Neuronetics Study 44-01101 = 9.1% relapse, ECT = 4.5%-36% relapse rates). Furthermore, persistence of effect with NeuroStar TMS Therapy at 6 months after treatment was sustained and also at least comparable to that observed with ECT (Neuronetics Study 44-01103 = 20% relapse, ECT = 50-63%).

6.5. Safety: TMS Therapy versus ECT

ECT is associated with significant and medically consequential adverse events, largely related either to the morbidity of the intentionally-induced seizure or the effects of the anesthesia procedures normally required for each ECT procedure (APA Committee on ECT, 2001). ECT is associated with measurable and clinically significant effects on cognitive function including anterograde amnesia, retrograde amnesia, concentration difficulties and postictal delirium. Cardiovascular complications (hypertension, hypotension), cardiac arrhythmias and pulmonary complications (prolonged apnea) are infrequent but serious complications of the ECT procedure. Rarely, deaths have been reported with the use of ECT. The adverse effects of ECT are summarized in Section 12: Substantial Equivalence and are based on Neuronetics review of the existing literature regarding ECT safety, including review of 407 reported adverse events available in FDA's medical device reporting databases. This detailed safety information was provided to the FDA as part of the 510(k) K061053 Class III Summary of Safety and Effectiveness (not included in the Panel Package).

Treatment with the NeuroStar System is associated with predictable, and clinically mild adverse effects (see Section 4.3.1.8.3). No deaths or seizures were reported in Neuronetics clinical studies. The most frequently reported events were headache (58.2% of active TMS treated patients vs. 55.1% of sham treatment) and application site pain (35.8% of active TMS treated patients vs. 3.8% of sham treatment). Treatment with the NeuroStar System is not associated with any measurable change in cognitive function. There was no effect of treatment with the NeuroStar System on formally assessed auditory threshold, as implemented with ear protection during the treatment.

These data indicate that the safety of the NeuroStar TMS System in the treatment of major depressive disorder exceeds the safety of ECT.

6.6. Summary of Substantial Equivalence

Determination of substantial equivalence requires consideration and balance of risk and benefit in the treatment population as shown by ***efficacy*** of treatment, its ***durability of effect***, and overall ***safety***.

- Efficacy, durability and safety of NeuroStar TMS Therapy has been shown in a patient population whose illness type and severity is comparable to that of patients normally referred for ECT

- NeuroStar TMS Therapy is effective in the treatment of MDD and is roughly 2/3 as effective as ECT, as determined from the extant ECT literature.

- NeuroStar TMS Therapy is durable in its effect over 1 month and appears at least comparable, if not favorable to, ECT in persistence of effect over 1-6 months with maintenance pharmacotherapy.

- NeuroStar TMS Therapy has a superior safety profile as compared to ECT.

- TMS Therapy as delivered by Neuronetics' NeuroStar System presents a highly attractive therapeutic option for patients with major depressive disorder.

Given these considerations of efficacy, durability, and safety, the NeuroStar TMS System is substantially equivalent to the predicate device, ECT.

SECTION 7. PROPOSED INDICATIONS FOR USE

The proposed indication for use for the NeuroStar TMS System is ***"for the treatment of major depressive disorder".***

The data provided in 510(k) K061053 support the proposed indication as follows:

- TMS Therapy as delivered by the NeuroStar System has proven efficacy and safety in the treatment of patients meeting the DSM-IV criteria for major depressive disorder.
    - The NeuroStar System labeling indicates that "major depressive disorder" refers to DSM-IV criteria and provides this definition.

- The labeling for the NeuroStar System should be the same as for ECT devices and not be accompanied by specification of a minimum level of treatment resistance.
    - The level of treatment resistance in the Neuronetics studies is comparable to that of patients treated with ECT as described in Section 6.3. ECT labeling is not restricted to a particular level of treatment resistance.
    - Patients intolerant to pharmacologic antidepressant therapy were not excluded from the Neuronetics studies and could benefit from its use.
    - The strong safety profile of TMS Therapy supports its use in a patient population as equally broad as indicated for ECT.

- ECT devices are also indicated for the treatment of patients with Major Depressive Disorder with psychosis or with bipolar disorder. The MDD patient populations with psychosis or bipolar disorder were not studied in Neuronetics studies at the request of the FDA. Therefore, Neuronetics is not seeking these additional label indications.

- ECT devices are also indicated for patients with MDD with emergent suicidal symptoms. Patients with emergent suicidal symptoms in imminent risk of life-threatening danger were excluded from Neuronetics studies due to safety concerns relative to the conduct of a randomized, sham-controlled outpatient trial. Therefore, Neuronetics is not seeking this additional label indication.

- The TMS Therapy delivered by the NeuroStar System is ***by prescription only***. Conditions regarding safe clinical use of the NeuroStar System are provided as part of the product labeling. Psychiatrist training on NeuroStar System procedures and conditions for use are included as part of the device instructions for use.

The intended use, contraindications, warnings, precautions, a summary of clinical efficacy and safety data obtained from Neuronetics' clinical studies for the NeuroStar System, and a description of the device with detailed step-by-step, illustrated, conditions for its clinical use are provided in the NeuroStar User Manual (CD-ROM Attachment 15).

SECTION 8. BIBLIOGRAPHY

The following articles are referenced in this panel package. All articles are available upon request.

1. American Psychiatric Association Committee on Electroconvulsive Therapy. (2001) The practice of electroconvulsive therapy: recommendations for treatment, training, and privileging. 2nd edition, (Text, not included) Washington, DC, American Psychiatric Association.
2. Altman DG, Andersen K (1999) Calculating the number needed to treat for trials where the outcome is time to an event. Br Med J. 1999; 319:1492–1495.
3. Avery DH, et al. (2006) A Controlled Study of Repetitive Transcranial Magnetic Stimulation in Medication-Resistant Major Depression. BIOL PSYCHIATRY 2006; 59:187–194
4. Buschke, H, et al. (1974): Evaluating storage, retentions and retrieval in disordered memory and learning. Neurology. November 1974; 24:1019-1025,
5. Burt T, et al. (2002) Neuropsychiatric Applications of transcranial magnetic stimulation: a meta analysis. Int J Neuropsych . 2002; 5:73-103.
6. Cochrane Collaboration: website (http://www.cochrane-net.org)
7. Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences, 2nd Edition, Hillsdale N.J.
8. Cook, JA, et al. (2002): Effects of depressive symptoms and mental health quality of life on use of highly active antiretroviral therapy among HIV-seropositive women. J Acquir Immune Defic Syndr. 2002; 30:401-409.
9. Cook, RJ, Sackett, DL (1995): The number needed to treat: A clinically useful measure of treatment effect. Br Med J. 1995; 310:452– 454.
10. Eaton, W, et al. (1996): Depression and Risk for Onset of Type II Diabetes: A Prospective Population-Based Study. Diabetes Care. 1996; 19:1097-1102.
11. Endicott, J, et al. (1993) Quality of Life Enjoyment and Satisfaction Questionnaire: A New Measure. Psychopharm Bull. 1993; 29(2):321-326.
12. Maurizio Fava, M.D., et al. (2006): A Comparison of Mirtazapine and Nortriptyline Following Two Consecutive Failed Medication Treatments for Depressed Outpatients: A STAR*D Report. Am J Psychiatry. 2006; 163:1161–1172
13. I. Fawzy, M.D., et al. (2003): Malignant Melanoma: Effects of a Brief, Structured Psychiatric Intervention on Survival and Recurrence at 10-Year Follow-up. Arch Gen Psychiatry. 2003; 60:100-103
14. Fitzgerald P.B., et al. (2003): Transcranial magnetic stimulation in the treatment of depression: A double-blind, placebo-controlled trial. Arch Gen Psychiatry 2003; 60: 1002–1008.

15. Fitzgerald, P.B. , et al. (2006) A Randomized, Controlled Trial of Sequential Bilateral Repetitive Transcranial Magnetic Stimulation for Treatment-Resistant Depression. Am J Psychiatry January 2006; 163:1,

16. First, Michael B., et al. (2002): Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P) New York: Biometrics Research, New York State Psychiatric Institute. November 2002

17. Folstein MF, et al. (1975) "Mini-Mental State" A Practical Method for Grading the Cognitive State of patients for the Clinician. J Psychiat. Res. 1975; 12:189-198.

18. Guy, (1976): ECDEU Assessment Manual for Psychopharmacology Revised. National Institute of Mental Health. 1976

19. Hamilton, MA. (1960) A Rating Scale for Depression. J. Neurol. Neurosurg. Psychiatry. 1960; 23: 56-62

20. Katon, WJ. (2003) Clinical and health services relationships between major expression, depressive symptoms, and general medical illness. Biol Psychiatry. 2003; 54:216-226.

21. Kessler, RC, et al. (2003) The epidemiology of major depressive disorder: Results from the national co-morbidity survey replication (NCS-R). JAMA. 2003; 289:3095-3105.

22. Khan, A, Warner, HA, Brown, WA. (2000) Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials. Arch Gen Psychiatry. 2000; 57:311-317.

23. Kraemer and Kupfer (2006) Size of Treatment Effects and Their Importance to Clinical Research and Practice. BIOL PSYCHIATRY. 2006; 59:990–996

24. Kraemer, HC, et al. (2006) Caution Regarding the Use of Pilot Studies to Guide Power Calculations for Study Proposals. Archives of General Psychiatry 2006; 63:484-489

25. Lisanby, SH, et al. (2000) The effects of electroconvulsive therapy on memory of autobiographical and public events. Arch Gen Psychiatry. 2000; 578:581-590

26. Loo, C, et al. (2001) Effects of a 2 to 4 week course or repetitive transcranial magnetic stimulation on neuropsychological functioning, electroencephalogram, and auditory threshold in depressed patients. Biol. Psychiatry. 2001; 49:615-623.

27. McGrath, PJ, M.D., et al. (2006) Tranylcypromine Versus Venlafaxine Plus Mirtazapine Following Three Failed Antidepressant Medication Trials for Depression: A STAR*D Report. Am J Psychiatry. 2006; 163:1531-1541

28. Montgomery, SA, Asberg, MA. (1979) A New Depression Scale Designed to be Sensitive to Change. Br. J. Psychiatry. 1979; 134:382-389.

29. Murray, CJL, Lopez, AD, eds. (1996) The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk

factors in 1990 and projected to 2020. Cambridge, Mass: Harvard University Press. 1996

30. Nierenberg, A, M.D., et al. (2006) A Comparison of Lithium and T3 Augmentation Following Two Failed Medication Treatments for Depression: A STAR*D Report. Am J Psychiatry. 2006; 163:1519-1530
31. Pascual-Leone, A., Cohen, L.G., et al (1992) No evidence of hearing loss in humans due to transcranial magnetic stimulation. Neurology 1992; 42: 647-651.
32. Pascual-Leone, A., Houser, C.M., et al. (1993) Safety of rapid-rate transcranial magnetic stimulation in normal volunteers. Electroencephalography and Clinical Neurophysiology. 1993; 89(2): 120-30.
33. Prudic, J, et al. (1996) Resistance to antidepressant medications and short-term clinical response to ECT. Am J Psychiatry. 1996; 153:985-992.
34. Prudic, J, et al. (2004) Effectiveness of electroconvulsive therapy in community settings. Biol Psychiatry. 2004; 55:301-312.
35. Rugulies, R. (2002) Depression as a predictor for coronary heart disease: A review and meta-analysis. Am J Prev Med. 2002; 23:51-61.
36. Rush, AJ, et al. (1996) The Inventory of Depressive Symptomatology (IDS): Psychometric Properties. Psychological Medicine 1996; 26:477-486.
37. Rush, AJ, et al. (2005) Vagus Nerve Stimulation for Treatment-Resistant Depression: A Randomized, Controlled AcutePhase Trial. BIOL PSYCHIATRY. 2005; 58:347–354
38. Rush, AJ, et al. (2006) Neuropsychopharmacology. 2006; 31:1841–1853
39. Rush, AJ, et al. (2006) Bupropion-SR, Sertraline, or Venlafaxine-XR after Failure of SSRIs for Depression N Engl J Med. 2006; 354:1231-42
40. Sackeim, et al. (1993) Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. New England Journal of Medicine. 1993; 328(12):839-846.
41. Sackeim, et al. (2000) A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. Arch Gen Psychiatry. 2000; 57:425-434.
42. Sackeim, et al. (2001) Vagus nerve stimulation (VNS) for treatment-resistant depression: Efficacy, side effects, and predictors of outcome. Neuropsychopharmacology. 2001; 25(5):713-728.
43. Sackeim, et al. (2001) Continuation pharmacotherapy in the prevention of relapse following electroconvulsive therapy: A randomized controlled trial. JAMA. 2001; 285:1299-1307.
44. Thase, ME, Rush, AJ. (1997) When at first you don't succeed: sequential strategies for antidepressant non-responders. J Clinical Psychiatry. 1997; 58[suppl.13]:23-29.

45. The UK ECT Review Group. (2003) Efficacy and safety of electroconvulsive therapy in depressive disorders: A systematic review and meta-analysis. Lancet. 2003; 361:799-808.

46. Trivedi, MH, M.D., et al. (2006) Medication Augmentation after the Failure of SSRIs for Depression. N Engl J Med. 2006; 354:1243-52.

47. Walsh, BT, et al. (2002) Placebo response in studies of major depression: variable, substantial and growing. JAMA. 2002; 287(14):1840-1847.

48. Ware, J.E., and Kosinski, M. (2001) SF-36 Physical and Mental Health Summary Scales: A Manual for Users of Version 1. Second Edition, Lincoln, RI: Quality Metric Incorporated, 2001