Journal of Affective Disorders 277 (2020) 65–74

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Clinical outcomes in a large registry of patients with major depressive disorder treated with Transcranial Magnetic Stimulation

Harold A. Sackeim[a,*], Scott T. Aaronson[b], Linda L. Carpenter[c], Todd M. Hutton[d], Miriam Mina[e], Kenneth Pages[f], Sarah Verdoliva[g], W. Scott West[h]

[a] Departments of Psychiatry and Radiology, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, USA
[b] Sheppard Pratt Health System and the Department of Psychiatry, University of Maryland, Baltimore, MD, USA
[c] Butler Hospital and Department of Psychiatry, Brown University, Providence, RI, USA
[d] Southern California TMS Center, Pasadena, CA, USA
[e] Neuronetics Inc., Malvern, PA, USA
[f] TMS of South Tampa, Tampa, FL, USA
[g] NAMSA - St. Louis Park, Minneapolis, MN, USA
[h] Nashville NeuroCare Therapy, Franklin, TN, USA

ARTICLE INFO

Keywords:
Major depressive disorder
Transcranial magnetic stimulation
Antidepressant
Efficacy
Gender
Aging

ABSTRACT

Background: Randomized clinical trials have demonstrated that Transcranial Magnetic Stimulation (TMS) is an effective treatment for episodes of major depressive disorder (MDD). However, characterization of outcomes in routine clinical practice is needed, as well as identification of patient- and treatment-related outcome predictors. This study documented patient-rated (PHQ-9) and clinician-rated (CGI-S) clinical outcomes in the NeuroStar® Advanced Therapy System Clinical Outcomes Registry.

Methods: Registry data were collected at 103 practice sites. Of 7759 participants, 5010 patients were included in an intent-to-treat (ITT) sample, defined as a primary MDD diagnosis, age ≥ 18, and completion of the PHQ-9 before TMS and with at least one PHQ-9 assessment after baseline. Completers (N = 3,814) were responders or had received ≥ 20 sessions and had an end of acute treatment PHQ-9 assessment. CGI-S ratings were obtained in smaller samples.

Results: In the total ITT and Completer samples, response (58–83%) and remission (28–62%) rates were notably high across self-report and clinician-administered assessments. Female patients and those treated with a larger number of pulses per session had superior clinical outcomes.

Limitations: Site participation in the registry was voluntary and treatment was open label.

Conclusions: The extent of clinical benefit reported by patients and clinicians following TMS in routine practice compares favorably with alternative interventions for treatment-resistant depression. Strong efficacy and the low side effect and medical risk profile suggest that TMS be evaluated as a first-line treatment for MDD. The findings derive from the largest registry of clinical outcomes in MDD for any treatment.

1. Introduction

Transcranial Magnetic Stimulation (TMS) is widely regarded as an effective treatment for episodes of major depressive disorder (MDD) (Perera et al., 2016; McClintock et al., 2018; Mutz et al., 2019). It is cleared by the US Federal Drug Administration (FDA) specifically for adult patients who have failed to receive satisfactory improvement from prior antidepressant medication in the current episode.

Since its first use in MDD, TMS has evolved in terms of the neural targeting methods, stimulation parameters, and treatment schedule. It also appears that the antidepressant effects of TMS have changed. The first meta-analyses of TMS treatment studies in MDD concluded that, while TMS had greater antidepressant properties than sham stimulation, its therapeutic effects were modest in magnitude and of doubtful clinical importance (Sackeim, 2000; Burt et al., 2002; Loo and Mitchell, 2005). The multi-site, sham-controlled trials that

Abbreviations: CGI-S, Clinical Global Impression – Severity scale; DLPFC, dorsolateral prefrontal cortex; ITT, intention-to-treat; ITI, inter-train interval; MDD, major depressive disorder; MT, motor threshold; PHQ-9, Patient Health Questionnaire-9; TMS, transcranial magnetic stimulation

* Corresponding author.
E-mail address: has1@columbia.edu (H.A. Sackeim).

https://doi.org/10.1016/j.jad.2020.08.005
Received 26 May 2020; Received in revised form 30 July 2020; Accepted 3 August 2020
Available online 07 August 2020

subsequently led to regulatory clearance further demonstrated that TMS had superior antidepressant effects than sham stimulation. However, in the studies by O'Reardon et al. (2007) and George et al. (2010), response and remission rates following the blinded, sham-controlled phase were only in the range of 15–20% and 7–14%, respectively.

Once TMS became available in routine clinical practice, new information emerged regarding its efficacy. Carpenter et al. (2012) collected a sample of 307 treatment-resistant MDD patients treated at 42, mostly private practice, sites, and reported response and remission rates of approximately 50% and 30% across self-report and clinician-rated measures. An average of 28 TMS sessions was administered in this open label, naturalistic study, a longer treatment course than in many of the earlier controlled trials, and patients were also treated with concomitant antidepressant medications. In a lead-in to a randomized study of continuation/maintenance TMS treatment conducted at 6 academic sites, Philip et al. (2016) administered 30 TMS sessions to 67 treatment-resistant MDD patients, who were free of psychotropics other than limited use of sleep and anxiolytic medications. They also reported robust response and remission rates, underscoring the possibility that longer courses of TMS may be responsible for the improved efficacy (Sackeim, 2016). In a large randomized non-inferiority trial conducted at 3 academic centers in treatment-resistant MDD, Blumberger et al. (2018) found that traditional fast-frequency (10 Hz) TMS and intermittent Theta Burst Stimulation (iTBS) delivered to the left dorsolateral prefrontal cortex (DLPFC) for 20 sessions (4 weeks) in medicated patients did not differ in antidepressant effects, with response and remission rates ranging from 39–49% and 20–32%, respectively across self-report and clinician-rated measures. Thus, it appears that in recent years the antidepressant effects of TMS have increased in magnitude.

In 2016, Neuronetics Inc. initiated the NeuroStar® Advanced Therapy System Clinical Outcomes Registry. This registry documents demographic features, treatment parameters, and clinical outcomes of MDD patients treated with the NeuroStar® Advanced Therapy System. To date, this registry has collected information on more than 8000 patients treated with TMS, and represents the largest registry documenting clinical outcomes for any treatment of MDD (Aaronson et al., 2017). Over 100 sites, almost all private practitioner or private practice TMS centers, have contributed data.

Treatments for medical disorders often show superior clinical outcomes in research trials and at academic centers compared to the same treatments delivered in community practice settings (Hewitt et al., 1999; Institute of Medicine (U.S.), Committee on Quality of Health Care in America., 2001; Hoekstra et al., 2002). This pattern also holds for the treatment of psychiatric disorders, including MDD (Dixon et al., 1995; National Advisory Mental Health Council, 1999; Unutzer et al., 1999; Weersing and Weisz, 2002; Prudic et al., 2004). Thus, the first objective of this study was to examine clinical outcomes in this large population of MDD patients treated in routine practice, using both self-report and clinician-rated measures. This information should be useful in benchmarking the effectiveness of TMS against alternative interventions, and for informing patients of likely treatment outcomes.

The second objective was to examine potential demographic factors and treatment-related correlates of TMS clinical outcomes. Some studies have suggested that geriatric depressed patients have less favorable clinical outcomes with TMS than younger adults (Figiel et al., 1998; Fregni et al., 2006; Jorge et al., 2008; Carpenter et al., 2012; Pallanti et al., 2012), perhaps due to cortical atrophy in older patients resulting in a greater coil-to-cortex distance and lower induced current density (Nahas et al., 2004). TMS is cleared by the FDA only for MDD patients between the ages of 22-70 years even though several recent investigations and meta-analyses have not observed an inverse relationship between age and TMS clinical outcome (Ciobanu et al., 2013; Sabesan et al., 2015). Gender has rarely been examined as a predictor of TMS outcome, although some small studies have suggested that premenopausal women are especially likely to benefit (Huang et al., 2008; Malik et al., 2016), and a meta-analysis of 54 sham-controlled trials suggested that females have superior clinical benefit (Kedzior et al., 2014).

FDA clearance of TMS devices for treatment-resistant MDD was based on evidence from sham-controlled, multisite RCTs of “fast” or high frequency (10 Hz) TMS delivered to the left DLPFC (O'Reardon et al., 2007; George et al., 2010). All patients included here received left DLPFC, fast frequency TMS, with the recommended protocols specifying 10 Hz stimulation, at an intensity 120% of motor threshold, and delivery of 3000 pulses. Meta-analyses have also supported the efficacy of slow frequency (1 Hz), right DLPFC TMS in the treatment of MDD (Burt et al., 2002; Berlim et al., 2013; Chen et al., 2013), and there has been considerable interest in whether sequential bilateral treatment (e.g., left fast frequency TMS followed by right-sided slow frequency TMS within the same session) enhances efficacy (Fitzgerald et al., 2013; Blumberger et al., 2016). A substantial proportion of the registry sample received sequential bilateral treatment. We conducted analyses of the total sample of MDD patients treated with TMS, as well as analyses restricted to patients treated only with left DLPFC, fast frequency TMS. In addition to demographic correlates, we examined the extent to which the number of TMS sessions, number of pulses delivered per session, motor threshold level, and stimulation intensity (% power relative to MT) were related to clinical outcomes.

2. Methods

2.1. Clinical outcomes registry

Site selection for participation in the NeuroStar® Advanced Therapy System Clinical Outcomes Registry required that clinical facilities treated a minimum of 24 patients the year before joining the registry, used TrakStar® Cloud software for recording patient characteristics and treatment parameters, and had a secure link for electronic data transfer. In addition, sites used the Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001) and/or the Clinical Global Impression – Severity scale (CGI-S) (Guy, 1976) to assess the severity of depressive symptoms by self-report and clinician rating, respectively.

Patient inclusion in the registry started on May 5, 2016 and this report concerns all data collected until October 4, 2019. Patients treated prior to the start of the registry were included, if the required clinical data were collected at the appropriate intervals. The first treatment included in the registry was administered on May 5, 2010. However, 95.3% of the patients in the intent-to-treat (ITT) sample were treated between the dates noted above. Prior to the first acute phase TMS treatment, site personnel entered patient demographic information (date of birth, gender), site identifier, patient primary diagnosis and in some cases diagnoses of co-morbid psychiatric conditions (using ICD-9, ICD-10, or DSM-IV), and PHQ-9 and CGI-S scores. Treatment parameters were captured passively at each treatment session, and included session date, stimulation target location (i.e., left DLPFC, right DLPFC, or both), motor threshold (MT) level, number of pulses delivered per treatment location or session, stimulation intensity (i.e., treatment level, % device output relative to MT), pulse frequency (e.g., 10 Hz vs. 1 Hz), and the number of acute phase treatment sessions that comprised the course of therapy. The acute phase treatment period was defined as starting with the patient's first recorded TMS treatment and continuing until there was a period of at least seven days without any treatment. It was expected that the PHQ-9 and CGI-S assessments would be completed at baseline and at acute phase treatment termination. PHQ-9 and CGI-S assessments completed at weekly intervals during the acute treatment course were also recorded.

The Clinical Outcomes Registry was maintained in compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA). All patient data were de-identified prior to electronic transfer. Collection and analysis of clinical care data in this way does not require local Institutional Review Board approval or informed consent.

2.2. Sample definitions

The NeuroStar Clinical Outcomes Registry collected data on 7,759 patients treated at 103 U.S. sites (mean per site = 51.1 patients, SD = 73.1). These registry participants were all unique individuals who received at least 1 treatment with the NeuroStar TMS Therapy System. The 103 sites were a substantial proportion of the more than 800 sites using this system in the U.S. The sites were primarily private practice practitioners (N = 51) or private practice TMS Centers (N = 49), with few hospital-based practices (N = 2) or academic institutions (N = 1). Registry entries included patients treated off-label for indications other than MDD, outside the indicated age range, and patients who received maintenance TMS. Once a site joined the registry, treatment data were captured from all patients treated at the site.

An ITT sample was defined, identifying all evaluable patients treated for MDD. Exclusions included age less than 18 years at the time of the first treatment (N = 85), invalid age entry (N = 1), and no MDD diagnosis (n = 825) or a primary diagnosis other than MDD (N = 12). In addition, to ensure that the treatment objective was management of an acute episode of MDD, patients with co-morbid diagnoses of psychiatric disorders other than generalized anxiety disorder, panic disorder, and unspecified anxiety disorder were also excluded (e.g., schizophrenia, bipolar disorder, autism, attention-deficit hyperactivity disorder) (N = 193). In addition, patients were excluded who did not have a PHQ-9 assessment within 14 days prior to the first TMS session (N = 1010), or who did not have at least one PHQ-9 assessment following the start of TMS (N = 394). Finally, individuals were excluded whose baseline PHQ-9 was less than 10, indicating insufficient severity of baseline depressive symptoms (N = 229). The ITT sample comprised 5,010 patients. In the ITT sample, the last clinical assessment was used to evaluate efficacy, regardless of when it occurred during the acute treatment course.

The Completer sample met the same eligibility criteria as the ITT sample. In addition, to ensure a minimally adequate course of TMS (Sackeim et al., 2019), individuals were excluded if classified as PHQ-9 nonresponders and had ended TMS after fewer than 20 sessions (N = 301), or if a PHQ-9 assessment was not conducted near the end of acute phase treatment (N = 895), i.e., within ± 4 days relative to the final session in the acute phase. The Completer sample comprised 3,814 patients. Of this sample, 110 patients (2.9%) were classified as responders and received fewer than 20 sessions. Their inclusion had minor impact on the findings

The ITT and Completer samples were further subgrouped, identifying patients treated only with left DLPFC stimulation, by excluding patients who had received any right-sided treatment during the acute course, either as the only treatment or in combination with left-sided treatment in the same session. Of the 5010 patients in the ITT sample, 2158 patients (43.1%) received some form of right-sided treatment (all low frequency, 1 Hz), with 30 of these patients treated with only right-sided TMS over the course of treatment. Of the 3814 patients in the Completer Total sample, 1694 patients (44.4%) received right-sided TMS at one or more sessions. In addition, patients were also excluded from the “Left Only” subgroup if they received more than one left-sided treatment per session (N = 83). The ITT and Completers samples were reduced to 2764 and 2050 patients, respectively, when limited to “Left Only” patients, i.e., those treated only with left-sided, fast frequency TMS with a single treatment per session.

Clinician-rated CGI-S scores were contributed principally by psychiatrists supervising TMS and occasionally by other clinical staff. These ratings were completed in a considerably smaller number of patients than the self-report PHQ-9 scale. The samples examined in the CGI-S analyses were thus subsets of those examined in PHQ-9 analyses. In addition, patients with CGI-S scores ≤ 3 prior to starting TMS (n = 9) were excluded from the CGI-S analyses due to insufficient baseline illness severity.

2.3. TMS procedures

MT was determined at the first treatment, using single pulse stimulation over motor cortex representation of the left abductor pollicis brevis muscle and visual observation of thumb twitch. MT level was defined as the minimum device power that induced an observable motor response in 50% of the applied pulses, using an iterative algorithm (MT Assist). MT was expressed in Standardized MT (SMT) units. SMT are calibrated so that a SMT of 1.0 corresponds to the average MT level observed in a large patient sample (NeuroStar® System Instructions for Use, Rev. F, Apr. 2019) and reflects an estimated electric field of 135 V/m at a point located 2.0 cm along the central axis of the treatment coil from the surface of the scalp into the patient's cortex.

External coordinates for coil placement over the DLPFC target were usually calculated by the device, identifying a site 5.5 cm anterior to the MT location, along a left superior oblique plane. However, practitioners could use other target localizing procedures (Beam et al., 2009). The recommended and FDA cleared protocols consisted of left DLPFC stimulation at 120% of the MT level, using a pulse frequency of 10 Hz, 4 s train duration, and 75 pulse trains, thereby delivering 3000 pulses per session. The inter-train interval (ITI) was 26 s in the original Standard protocol, resulting in a session duration of 37.5 min. In 2016, the FDA cleared a new protocol termed “Dash” for the NeuroStar device. The Dash protocol uses the same parameters as the Standard protocol but with a variable (shorter) ITI, ranging from 11 s to 25 s (Neuronetics Inc., 2016). This change reduced the minimum duration of a 3000 pulse session to 18.75 min. However, practitioners did not necessarily apply the parameter configurations specified by the Standard or Dash protocols, resulting in variability in stimulation parameters and number of sessions. For example, at the discretion of practitioners, some patients were administered a considerably larger number of pulses per session, up to a maximum of 5000.

Comparison of outcomes with the Standard and Dash protocols which varied in ITI, and comparison of bilateral versus left-sided only treatment, will be the subject of future reports. The NeuroStar TrakStar software captured the age, gender, and MT level of each patient, as well as the treatment parameters at each session, including stimulation intensity, pulse frequency, stimulation time, ITI, number of pulse trains, and total number of prescribed and delivered pulses.

2.4. Statistical analyses

The primary analyses were conducted in the ITT Total sample and were repeated for confirmation in the Completer Total sample and the ITT and Completer Left Only samples. Descriptive statistics on demographic features, treatment parameters, and clinical outcomes are reported for the ITT and Completer Total and Left Only samples.

Response was defined as ≥ 50% reduction in PHQ-9 scores at final assessment relative to pre-TMS baseline and remission was defined as a final PHQ-9 score less than 5. On the CGI-S, response corresponded to a score of 3 (“mildly ill”) or less, while remission corresponded to a score of 2 (“borderline mentally ill”) or less. The PHQ-9 and CGI-S have established reliability and validity and have been commonly used in large population-based studies of antidepressant effects (Kroenke et al., 2001; Busner and Targum, 2007). To examine concordance between the self-report and clinician ratings, the correlation between these scores was computed using a mixed effects linear model based on the method of Lam et. al. (Lam et al., 1999; Hamlett et al., 2003) to account for multiple measurements per subject. Additionally, Spearman's rho was calculated which assumes independence of multiple measurements from the same subject but considers the ordinal nature of the scales. These calculations were performed on the 3583 paired scores contributed by 1474 patients from baseline until the end of the acute course, using all measures obtained on the same day. Cohen's kappa coefficient (Cohen, 1960) was used to test the association between

PHQ-9 and CGI-S categorization of response and remission.

Categorical outcomes on the PHQ-9 were compared in patients who did and did not have CGI-S ratings using z-tests of independent proportions. McNemar's test was used to compare categorical outcomes in the patients rated with both the PHQ-9 and CGI-S.

To examine the relations of clinical outcomes to demographic variables and treatment parameters, parallel analyses were performed on PHQ-9 and CGI-S scores. A simultaneous multiple regression analysis was conducted on the final PHQ-9 or CGI-S score with baseline score, age, gender, MT, stimulation intensity, number of pulses delivered per session, and number of treatment sessions in the course as predictors. These same predictors were used in logistic regression analyses on response and remission classifications.

All analyses were conducted using SAS v9.4 (SAS Institute Inc., Cary, NC, USA). Results are reported as mean ± SD. Treatment parameters were averaged over all treatment sessions in the acute course. Significance values are two-sided with an alpha of 0.05. All *p*-values reported are without multiplicity adjustment.

3. Results

3.1. Demographic characteristics and treatment parameters

Table 1 presents demographic characteristics and treatment parameters of the ITT and Completer Total and Left Only samples. Approximately two-thirds of patients were female, and the average age was nearly 50 years. The average baseline PHQ-9 score indicated moderate-to-severe symptom severity, which was also reflected in the CGI-S ratings. On average, patients received over 30 TMS sessions during the acute course, over a period of 7–8 weeks.

3.2. PHQ-9 and CGI-S concordance

There was strong agreement between the PHQ-9 and CGI-S ratings, with a correlation coefficient across all paired ratings of 0.81, and a Spearman's rho of 0.80. Cohen's kappa coefficients for response and remission categorizations were 0.57 (95% CI: 0.52–0.63) and 0.52 (95% CI: 0.47–0.57), respectively.

3.3. Antidepressant Effects

Table 2 presents the PHQ-9 and CGI-S clinical outcome measures in the ITT and Completer samples. For the PHQ-9, response and remission rates varied from 58 to 69% and 28 to 36%, respectively. For the CGI-S, these rates were 69 to 83% and 47 to 63%, respectively.

Table 2
PHQ-9 treatment outcomes in the intent-to-treat and completer samples.

	PHQ-9 outcomes			
	Intent-to-treat sample		Completer sample	
	Total (*N* = 5010)	Left only (*N* = 2764)	Total (*N* = 3814)	Left only (*N* = 2053)
Baseline PHQ-9	19.8 ± 4.1	19.4 ± 4.2	19.8 ± 4.1	19.3 ± 4.1
LOCF PHQ-9	9.6 ± 6.8	9.0 ± 6.7	8.6 ± 6.4	7.9 ± 6.2
Difference (Pre-Post)	10.2 ± 6.8	10.4 ± 6.8	11.1 ± 6.6	11.4 ± 6.6
Response rate	57.7%	60.6%	65.0%	68.9%
Remission rate	27.9%	31.2%	31.7%	35.8%
	CGI-S Outcomes			
	Intent-to-Treat Sample		*Completer Sample*	
	Total (*N* = 1489)	Left Only (*N* = 735)	Total (*N* = 1170)	Left Only (*N* = 615)
Baseline CGI-S	5.5 ± 0.8	5.3 ± 0.7	5.4 ± 0.8	5.3 ± 0.7
LOCF CGI-S	2.8 ± 1.5	2.5 ± 1.3	2.6 ± 1.4	2.3 ± 1.3
Difference (Pre-Post)	2.7 ± 1.5	2.8 ± 1.4	2.8 ± 1.4	2.9 ± 1.3
Response rate	69.4%	79.0%	75.0%	83.1%
Remission rate	46.5%	57.8%	52.5%	62.3%

Response and remission rates were higher in the CGI-S than PHQ-9 ratings. Table 3 presents the response and remission rates on the PHQ-9 in patients who did and did not have CGI-S ratings, and the PHQ-9 and CGI-S categorical outcomes in the patients who had both ratings. PHQ-9 response and remission rates were higher in patients assessed with both measures than in patients without CGI-S ratings (all *p's* < 0.001). In addition, when restricted to patients with both ratings, response and remission rates were significantly higher in the CGI-S ratings (*p's* < 0.045 to 0.001). Thus, clinicians completed CGI-S scores especially for patients who reported substantial clinical improvement. Nonetheless, clinicians also reported superior outcomes compared to the self-ratings for patients in whom both sets of ratings were completed. This pattern was also observed in the ITT and Completer Left Only samples (see Table 1 of the online supplement).

3.4. Demographic and treatment-related predictors of outcome

The regression analyses for the ITT and Completer Total samples are presented in Table 4 for the PHQ-9 outcomes and Table 5 for the CGI-S outcomes. The major effects were confirmed in the Left Only samples

Table 1
Demographics and treatment parameters of the intent-to-treat and completer samples.

	Intent-to-treat sample		Completer sample	
	Total (*N* = 5010)	Left Only (*N* = 2764)	Total (*N* = 3814)	Left Only (*N* = 2053)
Age	47.8 ± 16.0	49.4 ± 16.0	47.7 ± 15.9	49.5 ± 15.8
Gender (% female)	64.0%	65.1%	63.6%	64.4%
Baseline PHQ-9	19.8 ± 4.1	19.4 ± 4.2	19.8 ± 4.1	19.3 ± 4.1
Baseline CGI-S[a]	5.5 ± 0.8	5.3 ± 0.8	5.5 ± 0.8	5.3 ± 0.7
Baseline CGI-S (N)	(2371)	(1098)	(1936)	(892)
TMS treatment location				
Left only (%)	56.9%	100.0%	55.6%	100.0%
Right only (%)	0.6%	0.0%	0.5%	0.0%
Left and right (%)	42.5%	0.0%	43.9%	0.0%
Number of sessions in acute course	32.0 ± 8.8	31.5 ± 8.9	34.2 ± 6.3	33.9 ± 6.0
Acute course duration (days)	52.3 ± 17.2	50.9 ± 16.7	55.8 ± 13.9	54.6 ± 12.7
Number of treatments per session	1.3 ± 0.5	1.0 ± 0.0	1.3 ± 0.5	1.0 ± 0.0
Motor threshold (MT, % device output)	1.04 ± 0.23	1.04 ± 0.23	1.03 ± 0.23	1.04 ± 0.23
Stimulation intensity (% MT across sessions)	115.3 ± 8.9	115.3 ± 8.0	116.3 ± 7.8	116.4 ± 6.8
Interval between pulse trains (ITI, s)	13.5 ± 7.3	16.6 ± 7.3	13.5 ± 7.3	16.7 ± 7.3
Pulse frequency per session	8.7 ± 2.4	10.2 ± 1.5	8.6 ± 2.3	10.2 ± 1.2
Number of pulses delivered per session	2925.1 ± 726.9	3280.9 ± 685.0	2923.1 ± 719.9	3289.2 ± 673.4

[a] Includes all patients with CGI-S scores at baseline, including patients with baseline CGI-S score ≤ 3.

Table 3
Comparison of response and remission rates based on PHQ-9 and CGI-S scores for the same patients or PHQ-9 scores only.

Outcome (Sample)	PHQ-9 No CGI-S	PHQ-9 both rated	CGI-S both rated
Response (ITT Total sample)	53.8% (1893/3521)	67.2%[a] (1000/1489)	69.4%[b] (1034/1489)
Response (Completer total sample)	62.7 % (1657/2644)	70.4%[a] (824/1170)	75.0%[c] (878/1170)
Remission (ITT total sample)	25.0% (879/3521)	34.8%[a] (518/1489)	46.5%[c] (692/1489)
Remission (Completer total sample)	29.7% (785/2644)	36.2%[a] (423/1170)	52.5%[c] (614/1170)

[a] $p < 0.001$ comparing PHQ-9 outcomes with and without concomitant CGI-S ratings based on z-test for independent proportions.
[b] $p < 0.05$ comparing outcomes on PHQ-9 and the CGI-S using McNemar's test for dependent proportions.
[c] $p < 0.001$ comparing outcomes on PHQ-9 and the CGI-S using McNemar's test for dependent proportions.

(see Tables 2 and 3 of the online supplement). Three variables had consistent associations with the three clinical outcome measures (final score, response and remission) across the ITT and Completer samples and the PHQ-9 and CGI-S assessments. Patients with more severe baseline symptomatology had higher PHQ-9 and CGI-S scores following TMS and were less likely to attain response or remission. Females had superior clinical outcomes compared to males. Patients who received more pulses per session also had superior clinical outcomes. In contrast, the number of TMS sessions administered was positively associated with clinical outcomes only in the ITT samples, which included patients who dropped out early in treatment. This association did not hold in the Completer samples, which required a minimum of 20 sessions for nonresponse classification. Stimulation intensity was negatively associated with clinical outcomes, but only in the CGI-S ratings and this negative association was not observed in the unadjusted data. Age and MT level did not show consistent association with clinical outcomes.

Females had significantly higher response and remission rates than males in all samples ($0.03 > P's < 0.001$), with approximately 5–10% more female patients achieving these outcomes than males. Fig. 1 presents unadjusted response and remission rates for women and men across the age range in the ITT and Completer Total samples (see Fig. 1 of the online supplement for the Left Only samples). Superior clinical outcomes in female patients were evident starting at 50 years of age and above. In all samples except Left Only CGI-S ratings, responders and remitters received a larger number of pulses per session than non-responders or non-remitters ($0.02 > p's < 0.001$). As seen in Fig. 2, patients who averaged 4000 or more pulses per session had superior outcomes (see Fig. 2 in the online line supplement for the Left Only samples).

4. Discussion

The NeuroStar® Advanced Therapy System Clinical Outcomes Registry collected clinical outcome data on a large sample of patients

Table 4
Regression analyses in the ITT and completer total samples on PHQ-9 outcomes with demographic and treatment-related predictors.

	ITT total sample $N = 4{,}999$[a]			Completer total sample $N = 3{,}809$[b]		
	Post-TMS PHQ-9 Score					
	Coefficient	SE	*p*	Coefficient	SE	*p*
Baseline PHQ- 9 score	**0.48**	**0.02**	**<0.0001**	**0.43**	**0.02**	**<0.0001**
Age	−0.01	0.01	0.29	−0.01	0.01	0.17
Gender (Female = 1)	**−1.08**	**0.19**	**<0.0001**	**−1.06**	**0.21**	**<0.0001**
Motor threshold (MT) Level	0.01	0.40	0.98	0.53	0.45	0.24
Stimulation intensity[c]	−0.08	0.06	0.15	0.02	0.07	0.75
Total pulses per session[d]	**−0.20**	**0.06**	**0.001**	**−0.28**	**0.07**	**<0.0001**
Number of TMS sessions	**−0.18**	**0.01**	**<0.0001**	−0.01	0.02	0.74
	PHQ-9 response					
	Odds ratio [95% CI]		*p*	Odds ratio [95% CI]		*p*
Baseline PHQ- 9 Score	**0.98 [0.97, 0.996]**		**0.01**	0.99 [0.97, 1.00]		0.11
Age	1.00 [1.00, 1.01]		0.26	1.00 [1.00, 1.01]		0.16
Gender (Female = 1)	**1.34 [1.19, 1.52]**		**<0.0001**	**1.35 [1.17, 1.55]**		**<0.0001**
Motor Threshold (MT) Level	1.06 [0.81, 1.38]		0.67	0.96 [0.71, 1.31]		0.81
Stimulation Intensity[c]	1.03 [0.99, 1.07]		0.15	0.99 [0.94, 1.04]		0.66
Total pulses per session[d]	**1.09 [1.04, 1.13]**		**0.0001**	**1.13 [1.07, 1.18]**		**<0.0001**
Number of TMS sessions	**1.06 [1.05, 1.06]**		**<0.0001**	**0.99 [0.98, 1.00]**		**0.03**
	PHQ-9 remission					
	Odds ratio [95% CI]		*p*	Odds ratio [95% CI]		*p*
Baseline PHQ- 9 score	**0.92 [0.90, 0.93]**		**<0.0001**	**0.92 [0.90, 0.93]**		**<0.0001**
Age	**1.00 [1.00, 1.01]**		**0.04**	1.00 [1.00, 1.01]		0.06
Gender (Female = 1)	**1.37 [1.19, 1.57]**		**<0.0001**	**1.31 [1.13, 1.52]**		**0.0003**
Motor Threshold (MT) Level	0.98 [0.73, 1.32]		0.91	0.89 [0.64, 1.22]		0.45
Stimulation Intensity[c]	**1.06 [1.01, 1.10]**		**0.02**	1.02 [0.97, 1.08]		0.36
Total pulses per session[d]	**1.07 [1.02, 1.11]**		**0.004**	**1.08 [1.03, 1.13]**		**0.003**
Number of TMS sessions	**1.04 [1.03, 1.05]**		**<0.0001**	1.01 [0.99, 1.02]		0.35

[a] The ITT Total sample for the PHQ-9 was reduced from 5010 to 4999 patients due to missing predictor data for 11 individuals: 6 patients had gender reported as "Other", 4 patients had missing data for Stimulation Intensity, and 1 patient was missing both Stimulation Intensity and Total Pulses per Session.
[b] The Completer Total sample for the PHQ-9 was reduced from 3814 to 3809 patients due to missing predictor data for 5 individuals: 3 patients had gender reported as "Other", 1 patients had missing data for Stimulation Intensity, and 1 patient was missing both Stimulation Intensity and Total Pulses per Session.
[c] Stimulation Intensity values were binned into units of 5% increments.
[d] Total Pulses per Session was binned into units of 500 pulses.

Table 5
Regression analyses in the ITT and completer total samples on CGI-S outcomes with demographic and treatment-related predictors.

	ITT total sample N = 1,486[a]			Completer total sample N = 1168[b]		
	Post-TMS CGI-S Score					
	Coefficient	SE	*p*	Coefficient	SE	*p*
Baseline CGI-S score	**0.44**	**0.05**	**<0.0001**	**0.36**	**0.05**	**<0.0001**
Age	0.00	0.00	0.72	0.00	0.00	0.68
Gender (Female = 1)	**−0.18**	**0.08**	**0.02**	**−0.24**	**0.08**	**0.004**
Motor threshold (MT) level	0.12	0.18	0.51	0.18	0.19	0.35
Stimulation intensity [c]	**0.09**	**0.03**	**0.003**	**0.10**	**0.03**	**0.001**
Total pulses per session[d]	**−0.14**	**0.02**	**<0.0001**	**−0.11**	**0.03**	**<0.0001**
Number of TMS sessions	**−0.03**	**0.01**	**<0.0001**	0.01	0.01	0.29
	CGI-S response					
	Odds ratio [95% CI]		*p*	Odds ratio [95% CI]		*p*
Baseline CGI-S score	**0.59 [0.50, 0.68]**		**<0.0001**	**0.58 [0.49, 0.70]**		**<0.0001**
Age	1.00 [0.99, 1.00]		0.20	1.00 [0.99, 1.01]		0.36
Gender (Female = 1)	1.17 [0.92, 1.49]		0.21	1.26 [0.95, 1.68]		0.11
Motor threshold (MT) level	0.99 [0.56, 1.75]		0.97	0.96 [0.50, 1.86]		0.90
Stimulation Intensity [c]	**0.86 [0.78, 0.95]**		**0.002**	**0.83 [0.74, 0.93]**		**0.002**
Total pulses per session[d]	**1.21 [1.12, 1.31]**		**<0.0001**	**1.19 [1.09, 1.31]**		**0.0003**
Number of TMS sessions	**1.06 [1.04, 1.09]**		**<0.0001**	0.99 [0.95, 1.03]		0.55
	CGI-S remission					
	Odds ratio [95% CI]		*p*	Odds ratio [95% CI]		*p*
Baseline CGI-S score	**0.65 [0.56, 0.74]**		**<0.0001**	**0.66 [0.57, 0.77]**		**<0.0001**
Age	1.00 [0.99, 1.01]		0.87	1.00 [0.99, 1.01]		0.99
Gender (Female = 1)	**1.42 [1.13, 1.77]**		**0.002**	**1.47 [1.15, 1.89]**		**0.003**
Motor Threshold (MT) level	0.89 [0.53, 1.49]		0.65	0.71 [0.40, 1.27]		0.25
Stimulation intensity [c]	**0.91 [0.83, 0.98]**		**0.02**	**0.87 [0.79, 0.96]**		**0.004**
Total pulses per session[d]	**1.20 [1.12, 1.28]**		**<0.0001**	**1.17 [1.08, 1.27]**		**<0.0001**
Number of TMS sessions	**1.02 [1.00, 1.04]**		**0.03**	0.98 [0.95, 1.01]		0.19

[a] The ITT Total sample for the CGI-S was reduced from 1489 to 1486 patients due to missing predictor data for 3 individuals: 2 patients had missing data for Stimulation Intensity, and 1 patient was missing both Stimulation Intensity and Total Pulses per Session.
[b] The Completer Total sample for the CGI-S was reduced from 1170 to 1168 patients due to missing predictor data for 5 individuals: 1 patient had missing data for Stimulation Intensity, and 1 patient was missing both Stimulation Intensity and Total Pulses per Session.
[c] Stimulation Intensity values were binned into units of 5% increments.
[d] Total Pulses per Session was binned into units of 500 pulses.

treated with TMS in private practice settings in the United States. Site participation was voluntary, and sites were required to demonstrate significant experience using the NeuroStar® Advanced Therapy System in the year before joining the registry. Clinical outcomes were examined for patients who received treatment for a major depressive episode, with minimal exclusions.

By both patient self-report and clinician-rated measures, the efficacy of TMS in this population was striking. In the patient self-ratings, PHQ-9 response and remission rates in the ITT samples were approximately 60% and 30%, respectively, and approximately 5% higher in the Completer samples. The clinician ratings, albeit in a substantially smaller sample, yielded higher estimates, with CGI-S response and remission rates of approximately 70–80% and 50–60%, respectively. The fact that the self-report estimates were derived from the treatment of several thousand patients at over 100 facilities increases confidence in their accuracy.

The antidepressant effects in this study are similar in magnitude to those reported by Carpenter et al. (Carpenter et al., 2012) in an independent sample of 307 patients also treated with the NeuroStar® Advanced Therapy System, and who had documented treatment resistance. Other recent trials have also reported strong therapeutic results (Perera et al., 2016; Philip et al., 2016; Blumberger et al., 2018; Mutz et al., 2019). Thus, the efficacy of TMS for episodes of major depression appears considerably more substantial than in its earlier history (Sackeim, 2000; Burt et al., 2002; Gross et al., 2007), and the findings of this study indicate superior clinical outcomes with routine community-based treatment relative to the outcomes documented earlier with research-based treatment in academic settings. In this study, registry patients averaged about 30 TMS sessions, and longer courses of TMS have been suggested as contributing to enhanced efficacy (Gershon et al., 2003; Gross et al., 2007). In the original multi-site sham-controlled trials, extending the course of TMS also resulted in increased rates of response and remission (O'Reardon et al., 2007; George et al., 2010). However, in addition to longer courses of treatment, changes have also occurred in other potentially relevant factors, such as the accuracy of target engagement, dosing parameters, and the medication status of patients.

The fact that these registry patients received open-label treatment may have elevated the response and remission rates relative to sham-controlled trials. In general, antidepressant medications show stronger efficacy when randomized treatment conditions do not involve a placebo control (Sneed et al., 2008; Rutherford et al., 2013). Furthermore, the fact that patients and clinician raters were aware of the treatment and the protocols being administered may have impacted on ratings. However, open-label registry data of this type may better reflect the outcomes occurring in routine practice.

All registry studies have limitations (O'Brien and Keil, 2020), and one of the most significant is the limited information on registry participants. In this registry, only patient age and gender were recorded, so clinical characteristics of the sample are unknown, other than psychiatric diagnosis and PHQ-9 and CGI-S scores. It is presumed that the patient sample was largely composed of individuals with treatment-resistant MDD, as insurance reimbursement for TMS often requires an extensive recent history of failed antidepressant treatment (Optum United Behavior Health. Behavioral Clinical Policy. Transcranial Magnetic Stimulation), and previous studies of similar TMS samples in private care have documented extensive treatment resistance (Carpenter et al., 2012). The PHQ-9 response and remission rates obtained here compare favorably to the clinical outcomes with the 7 antidepressant interventions used in Level 2 of the STAR*D trial, after failure to benefit sufficiently from standardized treatment with citalopram (Rush et al., 2006). In other words, the efficacy findings of this study are especially striking given the presumption that the sample was largely composed of patients with treatment-resistant MDD.

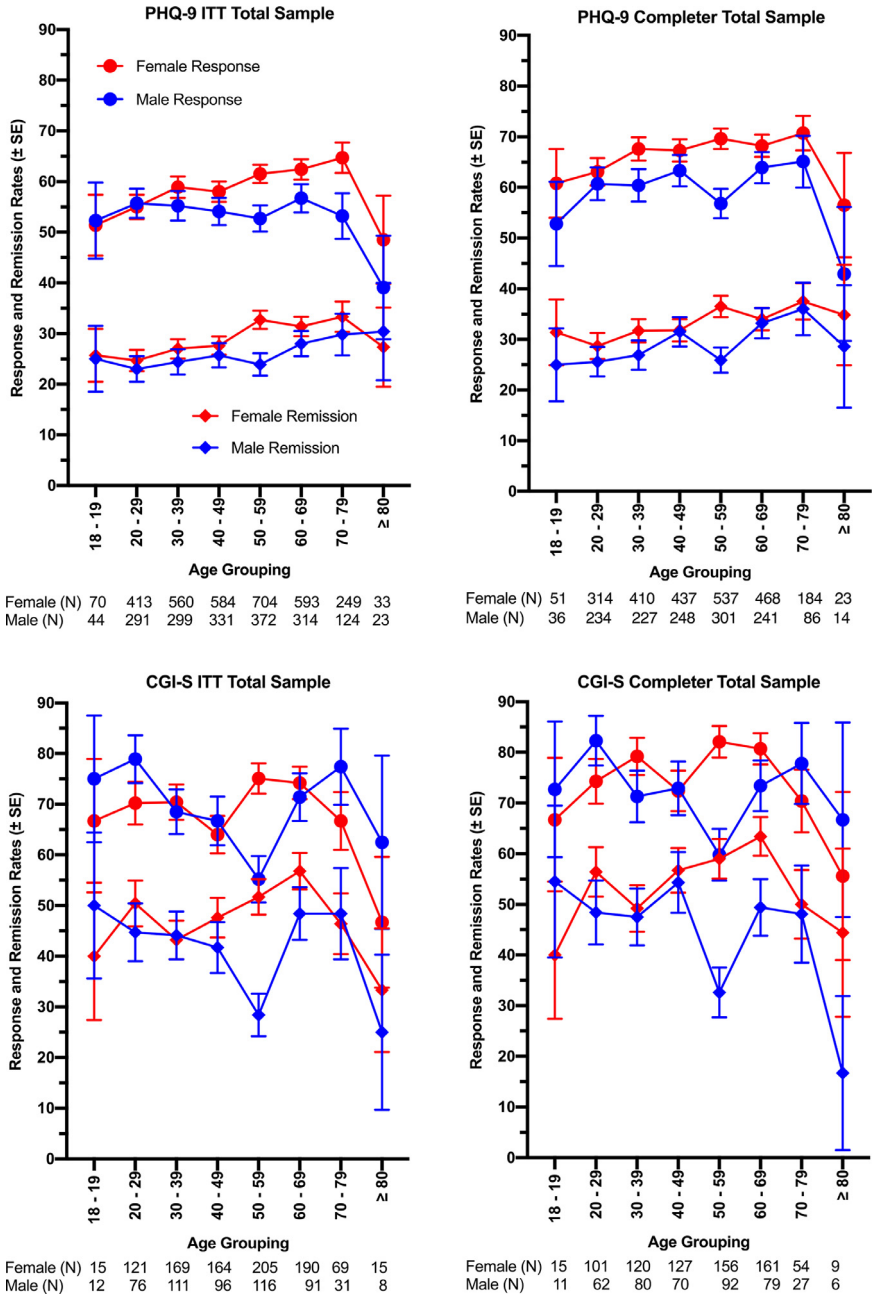

Fig. 1. Response and remisison rates for female and male patients as a function of age grouping for the ntent-to-treat (ITT) and Completer Total samples, separately for self-report (PHQ-9) and clinican-rated (CGI-S) outcomes.

The strong efficacy findings also raise the question of whether the use of TMS in treatment-resistant depression is unnecessarily restrictive. Interventions that are effective in the context of treatment-resistant depression are not expected to lose efficacy when administered to patients who are treatment naïve or who have been treated inadequately (Rush et al., 2006; Sackeim et al., 2019). As in the case of ECT, effective treatments may be reserved for treatment-resistant patients because increased side effect burden, medical risk, cost or convenience limits their use as a treatment of first choice. TMS was originally developed and tested in treatment-resistant depressed samples (George et al., 1999; Burt et al., 2002). However, its side effect and risk profile is now well characterized (Perera et al., 2016; McClintock et al., 2018) and in many respects compares favorably with antidepressant medications (Wang et al., 2018). The strong efficacy shown in this study, and the high safety margin of TMS, suggest consideration be given to broadening its indication to include MDD patients without established treatment resistance. Trials comparing outcomes of TMS and pharmacotherapy as first-line treatments may be needed.

The clinical outcomes were superior in the clinician ratings compared to the self-report ratings. This comparison is tenuous, however, since disparate scales were used to derive response and remission rates. For example, the PHQ-9 evaluates severity of discrete symptoms over a 2-week timeframe, while the CGI-S provides a cross-sectional, global evaluation of illness severity. Nonetheless, it is often the case that effect sizes for depressive symptom reduction are greater for clinician-ratings than self-ratings across studies of psychotherapy, pharmacology, and ECT (Sayer et al., 1993). The largest discrepancy in this study was the higher rate of remission in the clinician ratings. In a study of ECT in community settings, it was noted that treatment was not infrequently terminated by providers with the judgment that patients were in remission, while structured interviews by observers documented significant residual symptomology (Prudic et al., 2004). The clinician

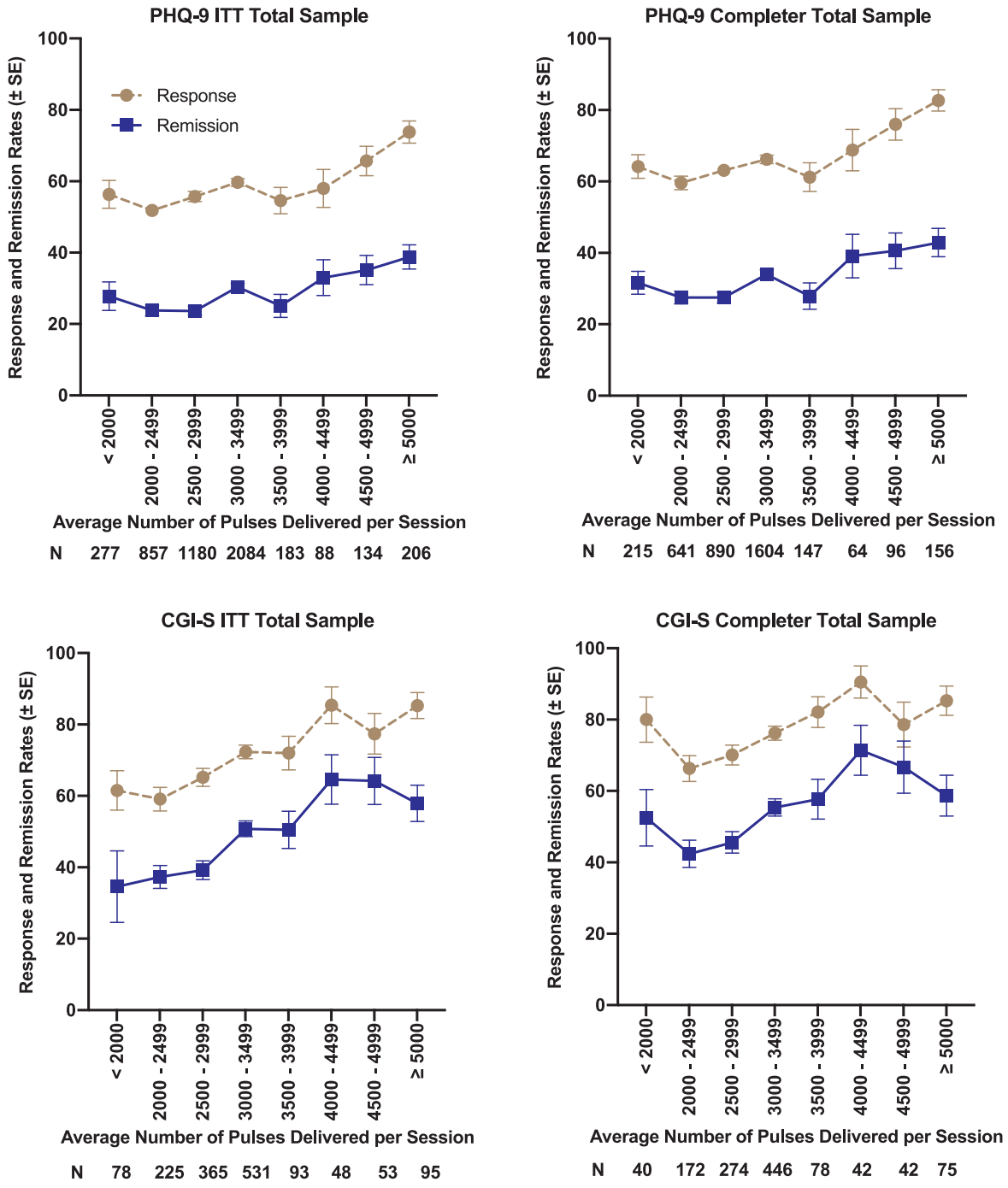

Fig. 2. Response and remission rates as a function of the average number of pulses delviered per treatment session for the intent-to-treat (ITT) and Completer Total samples, separately for self-report (PHQ-9) and clinican-rated (CGI-S) outcomes.

ratings were also obtained in a much small number of registry participants, and there was selection bias in terms of the patients in whom these ratings were completed. Thus, the clinician-rating data should be taken principally as supportive of the PHQ-9 results.

The PHQ-9 and CGI-S are commonly used and validated measures of depression symptom severity; they demand less time and training than instruments frequently used in clinical trials, such as the clinician-rated Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1967), or the Inventory of Depressive Symptoms-Self-Report (IDS-SR) scale (Trivedi et al., 2004). Regardless of the instrument used, clinical outcomes appear to be improved when standardized rating scales are obtained in routine practice (Wood and Gupta, 2017). It is unknown how outcomes would have differed had other instruments been used. As noted, one limitation of the PHQ-9 is the 2-week time period for assessing symptom severity. For treatments like TMS, where symptomatic improvement change can be rapid, using a 2-week reporting period may result in underestimation of therapeutic effects.

The clinical outcomes following TMS were consistently associated with two patient-related and one treatment-related predictors. As in most prospective studies of antidepressant treatment, severity of depressive symptoms at baseline and post-treatment was correlated. Female patients had greater therapeutic effects, supporting earlier suggestions of a gender difference (Huang et al., 2008; Kedzior et al., 2014; Malik et al., 2016). In this study, the gender difference was most manifest after 50 years of age. Contrary to earlier concerns (Figiel et al., 1998; Fregni et al., 2006; Jorge et al., 2008; Pallanti et al., 2012) and in

line with recent meta-analyses (Ciobanu et al., 2013; Sabesan et al., 2015), there was no evidence in this study that therapeutic effects were reduced with advancing age.

Practitioners varied in the TMS protocols they implemented. While all patients received some form of fast frequency TMS delivered to the left DLPFC (with or without additional slow frequency right-sided treatment), there was variability in patients' MT, stimulation intensity relative to MT, pulse frequency, number of pulses delivered per session, and the number of treatment sessions. Only the number of pulses per session showed consistent associations, as patients who received more pulses per session had superior results. Previous research has reported inconsistent findings regarding the relations of pulse number to efficacy (Schulze et al., 2018). As this was a registry-based, naturalistic, observational study, it is possible that the patients treated at sites using high number of pulses differed from those treated at the remaining sites. Nonetheless, the findings of gender differences in clinical outcomes and the relations of efficacy to the number of pulses per session may inform efforts to enhance further the efficacy of TMS in MDD. Indeed, this large registry with prospective outcome data provides opportunity to examine how variation in clinical practice is related to clinical outcomes. Here we observed that a larger number of pulses per session was associated with superior outcome. We did not examine the relations of ITI or laterality of treatment with outcome. Forthcoming reports will compare outcomes with the Standard (26 second ITI) versus Dash protocols (11–25 second ITI), and sequential bilateral treatment versus left-sided only treatment.

5. Conclusions

The NeuroStar Clinical Outcomes Registry documented clinical outcomes in the largest sample of individuals prospectively treated for an episode of major depressive disorder. By both self-report and clinician ratings, there were robust antidepressant effects, which compared favorably to those of alternative interventions for treatment-resistant depression. Given the low side effect and medical risk profile of TMS, the findings support of the use of TMS earlier in the treatment of MDD. Females and patients who received a larger number of pulses per session had stronger antidepressant effects. In contrast, there was no evidence that TMS efficacy declined with advancing age. These findings regarding outcome predictors may inform future efforts to further enhance the antidepressant efficacy of TMS.

CRediT Authorship Contribution Statement

All authors approved the final version of the manuscript, and participated in its conceptualization. Statistical analyses were carried out by SV, and supervised by HAS and MM. HAS drafted the manuscript with input from STA and LLC, and critical review by all authors. TMH, SP, and WSW participated in data collection and MM was responsible for establishing and coordinating the patient registry.

Declaration of Interest

Dr. Sackeim serves as a scientific adviser to LivaNova PLC, MECTA Corporation, and Neuronetics Inc. He receives honoraria and royalties from Elsevier, Inc. and Oxford University Press. He is the inventor on non-remunerative US patents for Focal Electrically-Administered Seizure Therapy (FEAST), titration in the current domain in ECT, and the adjustment of current in ECT devices, each held by the MECTA Corporation. He is also the originator of magnetic seizure therapy (MST).

Dr. Aaronson serves as a scientific adviser to Genomind Inc, LivaNova PLC, Neuronetics Inc, Janssen Pharmaceuticals Inc, and Sage Therapeutics and has received research support from Compass Pathways Inc and Neuronetics Inc. He is a member of the Speaker Bureau for Sunovion Pharmaceuticals Inc. and Janssen Pharmaceuticals Inc. Dr. Carpenter serves as a scientific advisor to Neuronetics Inc, Nexstim PLC, Affect Neuro Inc, Neurolief LTD, and Janssen Pharmaceuticals Inc.

Dr. Carpenter serves as a scientific advisor to Neuronetics Inc, Nexstim PLC, Affect Neuro Inc, Neurolief LTD, Sage Therapeutics, and Janssen Pharmaceuticals Inc. Dr. Carpenter has received research support (to Butler Hospital) from Neuronetics Inc, Neosync Inc, Nexstim PLC, Affect Neuro Inc, and Janssen Pharmaceuticals Inc.

Drs. Hutton, Pages, and **West** serve as consultants to Neuronetics, Inc. and **Ms. Mina** is an employee of Neuronetics, Inc.

Ms. Verdoliva reports no financial relationship with commercial interests.

Funding

The NeuroStar® Advanced Therapy System Clinical Outcomes Registry, analysis of the registry data, and the drafting of this manuscript were supported by Neuronetics Inc.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2020.08.005.

References

Aaronson, S.T., Sears, P., Ruvuna, F., Bunker, M., Conway, C.R., Dougherty, D.D., Reimherr, F.W., Schwartz, T.L., Zajecka, J.M., 2017. A 5-year observational study of patients with treatment-resistant depression treated with vagus nerve stimulation or treatment as usual: comparison of response, remission, and suicidality. Am. J. Psychiatry 174, 640–648.

Beam, W., Borckardt, J.J., Reeves, S.T., George, M.S., 2009. An efficient and accurate new method for locating the F3 position for prefrontal TMS applications. Brain Stimul. 2, 50–54.

Berlim, M.T., Van den Eynde, F., Jeff Daskalakis, Z., 2013. Clinically meaningful efficacy and acceptability of low-frequency repetitive transcranial magnetic stimulation (rTMS) for treating primary major depression: a meta-analysis of randomized, double-blind and sham-controlled trials. Neuropsychopharmacology 38, 543–551.

Blumberger, D.M., Maller, J.J., Thomson, L., Mulsant, B.H., Rajji, T.K., Maher, M., Brown, P.E., Downar, J., Vila-Rodriguez, F., Fitzgerald, P.B., Daskalakis, Z.J., 2016. Unilateral and bilateral MRI-targeted repetitive transcranial magnetic stimulation for treatment-resistant depression: a randomized controlled study. J. Psychiatry Neurosci. 41, E58–E66.

Blumberger, D.M., Vila-Rodriguez, F., Thorpe, K.E., Feffer, K., Noda, Y., Giacobbe, P., Knyahnytska, Y., Kennedy, S.H., Lam, R.W., Daskalakis, Z.J., Downar, J., 2018. Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. Lancet 391, 1683–1692.

Burt, T., Lisanby, S.H., Sackeim, H.A., 2002. Neuropsychiatric applications of transcranial magnetic stimulation: a meta analysis. Int. J. Neuropsychopharmacol. 5, 73–103.

Busner, J., Targum, S.D., 2007. The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry 4, 28–37.

Carpenter, L.L., Janicak, P.G., Aaronson, S.T., Boyadjis, T., Brock, D.G., Cook, I.A., Dunner, D.L., Lanocha, K., Solvason, H.B., Demitrack, M.A., 2012. Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice. Depress. Anxiety 29, 587–596.

Chen, J., Zhou, C., Wu, B., Wang, Y., Li, Q., Wei, Y., Yang, D., Mu, J., Zhu, D., Zou, D., Xie, P., 2013. Left versus right repetitive transcranial magnetic stimulation in treating major depression: a meta-analysis of randomised controlled trials. Psychiatry Res. 210, 1260–1264.

Ciobanu, C., Girard, M., Marin, B., Labrunie, A., Malauzat, D., 2013. rTMS for pharmacoresistant major depression in the clinical setting of a psychiatric hospital: effectiveness and effects of age. J. Affect. Disord. 150, 677–681.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.

Dixon, L.B., Lehman, A.F., Levine, J., 1995. Conventional antipsychotic medications for schizophrenia. Schizophr. Bull. 21, 567–577.

Figiel, G.S., Epstein, C., McDonald, W.M., Amazon-Leece, J., Figiel, L., Saldivia, A., Glover, S., 1998. The use of rapid-rate transcranial magnetic stimulation (rTMS) in refractory depressed patients. J. Neuropsychiatry Clin. Neurosci. 10, 20–25.

Fitzgerald, P.B., Hoy, K.E., Singh, A., Gunewardene, R., Slack, C., Ibrahim, S., Hall, P.J., Daskalakis, Z.J., 2013. Equivalent beneficial effects of unilateral and bilateral prefrontal cortex transcranial magnetic stimulation in a large randomized trial in treatment-resistant major depression. Int. J. Neuropsychopharmacol. 16, 1975–1984.

Fregni, F., Marcolin, M.A., Myczkowski, M., Amiaz, R., Hasey, G., Rumi, D.O., Rosa, M.,

Rigonatti, S.P., Camprodon, J., Walpoth, M., Heaslip, J., Grunhaus, L., Hausmann, A., Pascual-Leone, A., 2006. Predictors of antidepressant response in clinical trials of transcranial magnetic stimulation. Int. J. Neuropsychopharmacol. 9, 641–654.
George, M.S., Lisanby, S.H., Avery, D., McDonald, W.M., Durkalski, V., Pavlicova, M., Anderson, B., Nahas, Z., Bulow, P., Zarkowski, P., Holtzheimer 3rd, P.E, Schwartz, T., Sackeim, H.A., 2010. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. Arch. Gen. Psychiatry 67, 507–516.
George, M.S., Lisanby, S.H., Sackeim, H.A., 1999. Transcranial magnetic stimulation: applications in psychiatry. Arch. Gen. Psychiatry 56, 300–311.
Gershon, A.A., Dannon, P.N., Grunhaus, L., 2003. Transcranial magnetic stimulation in the treatment of depression. Am. J. Psychiatry 160, 835–845.
Gross, M., Nakamura, L., Pascual-Leone, A., Fregni, F., 2007. Has repetitive transcranial magnetic stimulation (rTMS) treatment for depression improved? A systematic review and meta-analysis comparing the recent vs. the earlier rTMS studies. Acta Psychiatr. Scand. 116, 165–173.
Guy, W., 1976. ECDEU Assessment Manual for Psychopharmacology. Superintendent of Documents, U.S. Government Printing Office, U.S. Department of Health, Education, and Welfare Publication, No. 76-338, Washington, D.C.
Hamilton, M., 1967. Development of a rating scale for primary depressive illness. Br. J. Soc. Clin. Psychol. 6, 278–296.
Hamlett, A., Ryan, L., Serrano-Trespalacios, P., Wolfinger, R., 2003. Mixed models for assessing correlation in the presence of replication. J. Air Waste Manag. Assoc. 53, 442–450.
Hewitt, M.E., Simone, J.V., 1999. Ensuring Quality Cancer Care. National Cancer Policy Board (U. S.)National Academy Press, Washington, D.C.
Hoekstra, J.W., Pollack Jr., C.V., Roe, M.T., Peterson, E.D., Brindis, R., Harrington, R.A., Christenson, R.H., Smith, S.C., Ohman, E.M., Gibler, W.B., 2002. Improving the care of patients with non-ST-elevation acute coronary syndromes in the emergency department: the CRUSADE initiative. Acad. Emerg. Med. 9, 1146–1155.
Huang, C.-C., Wei, I.H., Chou, Y.-H., Su, T.-P., 2008. Effect of age, gender, menopausal status, and ovarian hormonal level on rTMS in treatment-resistant depression. Psychoneuroendocrinology 33, 821–831.
Institute of Medicine (U.S.). Committee on Quality of Health Care in America, 2001. Crossing the Quality Chasm: A New Health System for the 21st Century. National Academy Press, Washington, D.C.
Jorge, R.E., Moser, D.J., Acion, L., Robinson, R.G., 2008. Treatment of vascular depression using repetitive transcranial magnetic stimulation. Arch. Gen. Psychiatry 65, 268–276.
Kedzior, K.K., Azorina, V., Reitz, S.K., 2014. More female patients and fewer stimuli per session are associated with the short-term antidepressant properties of repetitive transcranial magnetic stimulation (rTMS): a meta-analysis of 54 sham-controlled studies published between 1997–2013. Neuropsychiatr. Dis. Treat. 10, 727–756.
Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16, 606–613.
Lam, M., Webb, K.A., O Donnell, D.E., 1999. Correlation between two variables in repeated measures. In: Proceedings of the American Statistical Association Biometrics SectionAmerican Statistical Association, pp. 213–218.
Loo, C.K., Mitchell, P.B., 2005. A review of the efficacy of transcranial magnetic stimulation (TMS) treatment for depression, and current and future strategies to optimize efficacy. J. Affect. Disord. 88, 255–267.
Malik, A.M., Haque, Z., Ide, G., Farley, A., 2016. Gender and age as factors in response and remission of depression treated with transcranial magnetic stimulation. Brain Stimul. 9, e7.
McClintock, S.M., Reti, I.M., Carpenter, L.L., McDonald, W.M., Dubin, M., Taylor, S.F., Cook, I.A., O'Reardon, J., Husain, M.M., Wall, C., Krystal, A.D., Sampson, S.M., Morales, O., Nelson, B.G., Latoussakis, V., George, M.S., Lisanby, S.H., 2018. Consensus recommendations for the clinical application of repetitive Transcranial Magnetic Stimulation (rTMS) in the treatment of depression. J. Clin. Psychiatry 79 (1), 35–48.
Mutz, J., Vipulananthan, V., Carter, B., Hurlemann, R., Fu, C.H.Y., Young, A.H., 2019. Comparative efficacy and acceptability of non-surgical brain stimulation for the acute treatment of major depressive episodes in adults: systematic review and network meta-analysis. BMJ 364, l1079.
Nahas, Z., Li, X., Kozel, F.A., Mirzki, D., Memon, M., Miller, K., Yamanaka, K., Anderson, B., Chae, J.-H., Bohning, D.E., Mintzer, J., George, M.S., 2004. Safety and benefits of distance-adjusted prefrontal transcranial magnetic stimulation in depressed patients 55–75 years of age: a pilot study. Depress. Anxiety 19, 249–256.
National Advisory Mental Health Council, 1999. Bridging Science and Service: A Report by the National Advisory Mental Health Council's Clinical Treatment and Services Research Workgroup. National Institute of Health, Bethesda, MD.
Neuronetics Inc., 2016. Neurostar® TMS Therapy System: 510(k) Premarket Notification. Malvern, PA.
O'Brien, K.M., Keil, A.P., 2020. Design and interpretation considerations in registry-based studies. JAMA Psychiatry 77, 15–16.
Optum United Behavior Health. Behavioral Clinical Policy. Transcranial Magnetic Stimulation, BH803TMS062019. 2019, (Accessed 10 Aug 2020).
O'Reardon, J.P., Solvason, H.B., Janicak, P.G., Sampson, S., Isenberg, K.E., Nahas, Z., McDonald, W.M., Avery, D., Fitzgerald, P.B., Loo, C., Demitrack, M.A., George, M.S., Sackeim, H.A., 2007. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. Biol. Psychiatry 62, 1208–1216.
Pallanti, S., Cantisani, A., Grassi, G., Antonini, S., Cecchelli, C., Burian, J., Cauli, G., Quercioli, L., 2012. rTMS age-dependent response in treatment-resistant depressed subjects: a mini-review. CNS Spectr. 17, 24–30.
Perera, T., George, M.S., Grammer, G., Janicak, P.G., Pascual-Leone, A., Wirecki, T.S., 2016. The Clinical TMS society consensus review and treatment recommendations for TMS therapy for major depressive disorder. Brain Stimul. 9, 336–346.
Philip, N.S., Dunner, D.L., Dowd, S.M., Aaronson, S.T., Brock, D.G., Carpenter, L.L., Demitrack, M.A., Hovav, S., Janicak, P.G., George, M.S., 2016. Can medication free, treatment-resistant, depressed patients who initially respond to TMS be maintained off medications? A prospective, 12-month multisite randomized pilot study. Brain Stimul. 9, 251–257.
Prudic, J., Olfson, M., Marcus, S.C., Fuller, R.B., Sackeim, H.A., 2004. Effectiveness of electroconvulsive therapy in community settings. Biol. Psychiatry 55, 301–312.
Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163, 1905–1917.
Rutherford, B.R., Cooper, T.M., Persaud, A., Brown, P.J., Sneed, J.R., Roose, S.P., 2013. Less is more in antidepressant clinical trials: a meta-analysis of the effect of visit frequency on treatment response and dropout. J. Clin. Psychiatry 74, 703–715.
Sabesan, P., Lankappa, S., Khalifa, N., Krishnan, V., Gandhi, R.T., Palaniyappan, L., 2015. Transcranial magnetic stimulation for geriatric depression: promises and pitfalls. World J. Psychiatry 5, 170–181.
Sackeim, H.A., 2000. Repetitive transcranial magnetic stimulation: what are the next steps? Biol. Psychiatry 48, 959–961.
Sackeim, H.A., 2016. Acute, continuation, and maintenance treatment of major depressive episodes with transcranial magnetic stimulation. Brain Stimul. 9, 313–319.
Sackeim, H.A., Aaronson, S.T., Bunker, M.T., Conway, C.R., Demitrack, M.A., George, M.S., Prudic, J., Thase, M.E., Rushi, A.J., 2019. The assessment of resistance to antidepressant treatment: Rationale for the Antidepressant Treatment History Form: Short Form (ATHF-SF). J. Psychiatr. Res. 113, 125–136.
Sayer, N.A., Sackeim, H.A., Moeller, J.R., Prudic, J., Devanand, D.P., Coleman, E.A., Kiersky, J.E., 1993. The relations between observer-rating and self-report of depressive symptomatology. Psychol. Assess. 5, 350–360.
Schulze, L., Feffer, K., Lozano, C., Giacobbe, P., Daskalakis, Z.J., Blumberger, D.M., Downar, J., 2018. Number of pulses or number of sessions? An open-label study of trajectories of improvement for once-vs. twice-daily dorsomedial prefrontal rTMS in major depression. Brain Stimul. 11, 327–336.
Sneed, J.R., Rutherford, B.R., Rindskopf, D., Lane, D.T., Sackeim, H.A., Roose, S.P., 2008. Design makes a difference: a meta-analysis of antidepressant response rates in placebo-controlled versus comparator trials in late-life depression. Am. J. Geriatr. Psychiatry 16, 65–73.
Trivedi, M.H., Rush, A.J., Ibrahim, H.M., Carmody, T.J., Biggs, M.M., Suppes, T., Crismon, M.L., Shores-Wilson, K., Toprac, M.G., Dennehy, E.B., Witte, B., Kashner, T.M., 2004. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol. Med. 34, 73–82.
Unutzer, J., Katon, W., Sullivan, M., Miranda, J., 1999. Treating depressed older adults in primary care: narrowing the gap between efficacy and effectiveness. Milbank Q. 77, 225–256.
Wang, S.-M., Han, C., Bahk, W.-M., Lee, S.-J., Patkar, A.A., Masand, P.S., Pae, C.-U., 2018. Addressing the side effects of contemporary antidepressant drugs: a comprehensive review. Chonnam Med. J. 54, 101–112.
Weersing, V.R., Weisz, J.R., 2002. Community clinic treatment of depressed youth: benchmarking usual care against CBT clinical trials. J. Consult. Clin. Psychol. 70, 299–310.
Wood, J.M., Gupta, S., 2017. Using rating scales in a clinical setting. Curr. Psychiatry 16, 21–25.