



Dosing transcranial magnetic stimulation in major depressive disorder: Relations between number of treatment sessions and effectiveness in a large patient registry

Todd M. Hutton^a, Scott T. Aaronson^{b,c}, Linda L. Carpenter^{d,e}, Kenneth Pages^f, David Krantz^g, Lindsay Lucas^g, Bing Chen^g, Harold A. Sackeim^{h,i,*}

^a Southern California TMS Center, Pasadena, CA, USA

^b Sheppard Pratt Health System, Baltimore, MD, USA

^c Department of Psychiatry, University of Maryland, Baltimore, MD, USA

^d Butler Hospital, Providence, RI, USA

^e Brown University Department of Psychiatry and Human Behavior, Providence, RI, USA

^f TMS of South Tampa, Tampa, FL, USA

^g NAMS, St. Louis Park, MN, USA

^h Department of Psychiatry, Columbia University, NY, USA

ⁱ Department of Radiology, Columbia University, NY, USA

ARTICLE INFO

Keywords:

Transcranial magnetic stimulation (TMS)
Efficacy
Treatment duration
Number of sessions
Registry

ABSTRACT

Background: The number of sessions in an acute TMS course for major depressive disorder (MDD) is greater than in the earlier randomized controlled trials.

Objective: To compare clinical outcomes in groups that received differing numbers of TMS sessions.

Methods: From a registry sample (N = 13,732), data were extracted for 7215 patients treated for MDD with PHQ-9 assessments before and after their TMS course. Groups were defined by number of acute course treatment sessions: 1–19 (N = 658), 20–29 (N = 616), 30–35 (N = 1375), 36 (N = 3591), 37–41 (N = 626), or >41 (N = 349) and compared in clinical outcomes at endpoint and at fixed intervals (after 10, 20, 30, and 36 sessions). The impact of additional treatments beyond 36 sessions was also examined.

Results: Groups that received fewer than 30 sessions had inferior endpoint outcomes than all other groups. PHQ-9 symptom reduction was greatest in the group that ended treatment at 36 sessions. The extended treatment groups (>36 sessions) differed from all other groups by manifesting less antidepressant response early in the course and had a slower but steady rate of improvement over time. Extending treatment beyond 36 sessions was associated with further improvement without evidence of a plateau.

Conclusions: In real-world practice, there are strong relations between the number of TMS sessions in a course and the magnitude of symptom reduction. Courses with less than 30 sessions are associated with diminished benefit. Patients with longer than standard courses typically show less initial improvement and a more gradual trajectory, but meaningful benefit accrues with treatment beyond 36 sessions.

1. Introduction

Compared to the original sham-controlled trials that supported regulatory approval of transcranial magnetic stimulation (TMS) in major depressive disorder (MDD) [1,2], response and remission rates are substantially greater in recent reports documenting “real-world” outcomes across diverse clinical settings [3–5]. For example, in both the

industry-sponsored O’Reardon et al. trial [1] and the NIMH-sponsored Optimization of TMS study [2] remission rates at the primary endpoints of 4 weeks and 3 weeks, respectively, were approximately 15% with active TMS and 5% with sham in MDD patients withdrawn from antidepressant medication. While this difference supported the efficacy of TMS, the remission rates with active treatment were modest. In contrast, more recent naturalistic and registry studies of TMS treatment of MDD in community settings document response and remission rates of

* Corresponding author. Vagelos College of Physicians and Surgeons of Columbia University, New York, NY, USA.

E-mail address: has1@cumc.columbia.edu (H.A. Sackeim).

<https://doi.org/10.1016/j.brs.2023.10.001>

Received 7 August 2023; Received in revised form 19 September 2023; Accepted 1 October 2023

Available online 11 October 2023

1935-861X/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Abbreviations

CGI-S	Clinical Global Impression – Severity scale
DLPFC	Dorsolateral Prefrontal Cortex
HF	High Frequency
MDD	Major Depressive Disorder
MT	Motor Threshold
PHQ-9	Patient Health Questionnaire-9
TMS	Transcranial Magnetic Stimulation

approximately 60% and 30%, respectively [3–5], substantially greater than in the blinded phases of the original sham-controlled trials.

This putative increase in the effectiveness of TMS could be due to multiple factors, including selection or reporting biases in the observational studies or differences in outcome measures, patient characteristics, concomitant treatments, targeting methods, and TMS dosing. There is some evidence that patients who have previously failed to benefit from multiple adequate antidepressant trials are less likely to benefit from TMS than patients with lesser degrees of treatment resistance [6–8]. Samples enrolled in randomized controlled trials (RCTs) that require that pharmacoresistant depression be verified by detailed source documentation likely differ substantially from those receiving routine TMS in community settings [9,10]. Additionally, the randomized sham-controlled TMS trials were mainly conducted in patients who were withdrawn from antidepressant medications prior to starting TMS [1,2,11]. There is evidence that the combination of ECT and an antidepressant medication is more effective than ECT alone [12,13]. When treating MDD, TMS is now routinely combined with ongoing pharmacotherapy and/or psychotherapy, and it is possible that additive or synergistic effects with these treatments enhance effectiveness [14,15].

The protocols used to administer TMS have also evolved over the past two decades. Much of the early controlled research administered TMS at lower stimulus intensities and with fewer pulses per session than is now standard, and there is some evidence that both factors may be related to clinical outcome [3,16,17]. Perhaps most critically, the number of treatment sessions in a course of TMS has changed dramatically. The earliest sham-controlled trials often administered TMS for a period of only one or two weeks [18]. The primary endpoints in the pivotal O'Reardon et al. [1] and Optimization of TMS [2] trials were assessed after 3–4 weeks of TMS with 5 sessions per week (i.e., after 15–20 sessions). In both trials, after the acute phase endpoint, patients could receive up to 30 additional TMS sessions plus 6 spaced taper sessions. In both studies, the additional treatment after the acute endpoint resulted in improved efficacy [19,20], resulting in the FDA-cleared protocol for 36 total sessions. Similarly, Blumberger et al. [21] randomized MDD participants to intermittent Theta Burst Stimulation (iTBS) or standard 10 Hz TMS for 20 sessions over 4 weeks, with an additional 10 sessions delivered only for those who had 30% or greater reduction in scores by week 4. In contrast, in a sample of over 5000 participants treated for MDD in the NeuroStar® Advanced Therapy System Clinical Outcomes Registry, the average acute treatment course was composed of 32.0 TMS sessions, delivered over 7.5 weeks [3]. Patients now routinely receive longer courses of TMS than in the original RCTs.

There has been surprisingly little documentation of the relations between number of sessions in a TMS course and clinical outcomes in MDD [22,23] and few studies have attempted to characterize the point at which maximal symptom reduction is typically achieved and maximal response and remission rates typically observed [24,25]. This study addressed this knowledge gap with three main objectives. First, we sought to determine the number of treatment sessions at which symptom reduction peaks for the modal patient. Intolerance, perceived lack of efficacy, worsening depression symptoms, logistical and access

challenges, and other factors leading to early termination would be expected to produce relative short treatment courses with diminished antidepressant effects, while early termination due to rapid remission would have the opposite effect. With progressively longer treatment courses, TMS antidepressant effect would be expected to reach a peak. However, whether maximal antidepressant effect in group data is seen after 20, 30, 40 or more treatment sessions has not been established, and yet is critical in defining what constitutes an adequate TMS course [26,27].

The second objective of this study was to compare patient subgroups, defined by total number of treatment sessions in their acute course, in their trajectories of symptom improvement during the acute course. To inform measurement-based care, many TMS clinical practices have incorporated serial assessment with self-report depression severity scales. In addition to the acute course endpoint, clinical outcome data were available at fixed assessment intervals, i.e., after 10, 20, 30, and 36 TMS sessions. These data provide novel information in a very large real-world sample on the time course of improvement with TMS, both overall and in relation to groups receiving different lengths of treatment.

Our third objective focused on the patients that received especially long courses of TMS (>36 sessions and termed “extended treatment”) and examined whether the additional sessions resulted in further meaningful clinical improvement. In some cases, a commercial insurance company will approve a clinician’s request to cover an extension of the acute course beyond 36 sessions. This is typically done when the patient has had delayed onset of response or has steadily improved but not yet achieved remission. In other cases, patients self-pay for additional sessions. Specifically, we sought to determine whether extended courses of treatment resulted in meaningful increases in response or remission rates after what is typically considered a full course of treatment. In addressing this third objective, we also determined whether at any time point there was evidence that the effectiveness of TMS diminished or plateaued despite additional treatment or, to the contrary, evidence that TMS exerted additional antidepressant effects regardless of how many treatments were already administered.

This naturalistic, open-label study was conducted with data from the NeuroStar® Advanced Therapy System Clinical Outcomes Registry, involving more than 100 private practice sites [3,28–30]. In community practice, multiple factors determine the number of sessions administered in an acute TMS course besides the rate and extent of clinical improvement (e.g., tolerability, side effects, cost, convenience). Indeed, in the United States third-party insurance policies often pre-authorize a fixed number of TMS treatments, and in recent years 36 sessions has been commonly specified [31–33]. This study can be considered an observational, natural experiment comparing “real-world” outcomes in groups defined by receipt of different lengths of treatment. While identification of dose-response relationships in a naturalistic study of this type raises the possibility that differences in clinical outcomes are determined, at least in part, by dosage effects, it also possible that such effects reflect reverse causality, since clinical outcomes often determine the duration of exposure to treatments.

2. Methods

2.1. Clinical outcomes registry

As described previously [3,28–30], site selection for the NeuroStar® Advanced Therapy System Clinical Outcomes Registry required that the clinical facilities treated at least 24 patients the year before joining the registry, used TrakStar® Cloud software for recording de-identified patient characteristics and treatment parameters, and had a secure link for electronic data transfer. Sites used the Patient Health Questionnaire-9 (PHQ-9) [34] and/or the Clinical Global Impression – Severity scale (CGI-S) [35] to assess the severity of depressive symptoms by self-report and clinician rating, respectively. Once a site joined the registry, all patients treated at the site were included in the database.

Data entry in the registry started on May 5, 2016, and this report concerns all data collected until July 7, 2022. Site personnel entered patient demographic information (date of birth, gender), site identifier, primary diagnosis and diagnoses of co-morbid psychiatric conditions, and PHQ-9 and CGI-S scores. Treatment parameters were captured automatically, and included session date, treatment location of stimulation [i.e., left dorsolateral prefrontal cortex (DLPFC), right DLPFC, or both], motor threshold (MT), number of pulses per treatment location or session, treatment level (% device output relative to MT), pulse frequency, duration of pulse trains, intertrain interval (ITI), and the number of treatment sessions during the acute phase treatment course. The acute phase treatment period was retrospectively defined as starting with the patient's first recorded TMS treatment and continuing until there was a period of at least 8 or more days without any treatment or two consecutive weeks each with a gap of 7 days.

The 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. The study was documented at [ClinicalTrials.gov](https://clinicaltrials.gov) (Identifier: NCT05541302).

2.2. Sample definitions

Data were collected on 13,732 patients treated at 110 U.S. sites (see [Table 1](#)). These registry participants were all unique individuals who received at least one treatment with the NeuroStar TMS Therapy System. The registry sites were a substantial proportion of the approximately 900 sites using this system in the U.S. and were primarily administered by private practice practitioners (N = 51), larger group practices (N = 12) or private practice TMS Centers (N = 51), with few hospital-based practices (N = 2) or academic institutions (N = 1).

The total study sample excluded patients with age less than 18 years,

Table 1

Inclusion/exclusion criteria for the total sample and the left only and CGI-S subsamples.

Registry Sample (N = 13,732; 110 US sites)
Unique individuals with 1 TMS treatment before July 7, 2022
Total Study Sample (N = 7215; 105 sites)
Exclusions from Registry Sample (N = 13,732):
Age <18 years or invalid (N = 132)
No MDD diagnosis (N = 783)
Primary diagnosis other than MDD (N = 13)
Comorbid diagnosis (any non-MDD and non-anxiety diagnosis) (N = 651)
No baseline PHQ-9 within 14 days prior to first treatment (N = 1590)
No PHQ-9 within ± 4 days of last treatment session (N = 2349)
Baseline PHQ-9 score <10 (N = 407)
Received 3 or more treatments on at least 1 Day (N = 453)
In acute treatment at close of database (N = 139)
High Frequency, Left DLPFC Stimulation (Left Only) Subsample (N = 3805; 95 sites)
Exclusions from Total Study Sample (N = 7215):
Received Sequential Bilateral TMS (Left and Right) or Right Only TMS (N = 2677)
Received 2 or more left-sided treatments in a day (N = 325)
Any treatment at frequency other than 10 Hz or a target other than left DLPFC (N = 127)
Average treatment level <100% of motor threshold or average number of pulses/session <2000 (N = 281)
Total Study Sample CGI-S Subsample (N = 2303; 39 sites)
Exclusions from Total Study Sample (N = 7215):
No available baseline CGI-S within 14 days prior to first treatment (N = 3619)
No CGI-S within ± 4 days of last treatment session (N = 1281)
Baseline CGI-S Score ≤ 3 (N = 12)
CGI-S and Left Only Subsample (N = 1169; 34 sites)
Exclusions from Total Study Sample CGI-S Subgroup (N = 2303):
Not represented in the Left Only subgroup (N = 1134)

CGI-S = Clinical Global Impression-Severity Scale; DLPFC = dorsolateral prefrontal cortex; MDD = major depressive disorder; PHQ-9 = Patient Health Questionnaire-9.

no MDD diagnosis, or a primary diagnosis other than MDD ([Table 1](#)). To ensure that the treatment objective was management of an acute episode of MDD, patients with comorbid psychiatric diagnoses other than generalized anxiety disorder (GAD), panic disorder, and unspecified anxiety disorder were also excluded (e.g., post-traumatic stress disorder, obsessive compulsive disorder, schizophrenia, bipolar disorder, autism, attention-deficit hyperactivity disorder). Patients were excluded who did not have a PHQ-9 assessment within 14 days prior to the first TMS session or a PHQ-9 assessment within 4 days of the final session of the acute course. Individuals were also excluded whose baseline PHQ-9 was less than 10, indicating insufficient severity of baseline depressive symptoms, or who received 3 or more treatments in a day. After these exclusions, the total study sample with PHQ-9 ratings comprised 7215 patients.

Within the total sample with PHQ-9 ratings, a subsample was defined of patients treated with protocol-defined high frequency (10 Hz), left dorsolateral prefrontal cortex (DLPFC) TMS throughout the acute treatment course (see [Table 1](#)). Patients in this "Left Only" subsample (N = 3805) did not receive more than one treatment protocol in any session, all treatments delivered 10 Hz stimulation to the left DLPFC, the intensity of stimulation (treatment level) did not average less than 100% of MT, and the average number of pulses delivered per session was not less than 2000. Another subsample was comprised of patients who had CGI-S ratings both before and following the acute TMS treatment course (N = 2303), while otherwise applying the same exclusions as in the total sample with PHQ-9 ratings.

[Fig. 1](#) presents the distribution in the total sample of the number of TMS sessions in the acute course. Patients were grouped as having received 1–19 (N = 658), 20–29 (N = 616), 30–35 (N = 1375), 36 (N = 3591), 37–41 (N = 626) or 42–102 (N = 349) TMS sessions. Twenty sessions are often the threshold used to define a minimally adequate TMS course [[26,27](#)], and the 1–19 session group was considered as especially likely to have received insufficient treatment. Nearly 50% of the sample received exactly 36 sessions as this corresponded to the number of sessions pre-authorized by most insurance carriers [[31–33](#)]. These patients constituted a separate group. The two groups that received either 37–41 sessions or 42 or more sessions were considered as having received extended treatment, longer than typical TMS courses.

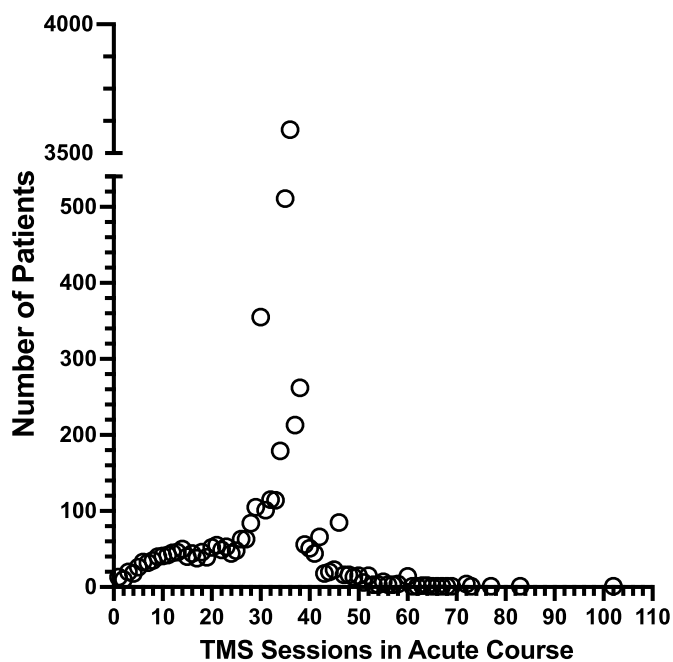


Fig. 1. Distribution of the number of sessions in the acute TMS course in 7215 patients treated with TMS for major depressive disorder (range = 1 to 102 sessions).

2.3. TMS procedures

MT was determined at the first treatment, using single pulse stimulation of the motor cortex area corresponding to the abductor pollicis brevis (or other muscle in a finger of the contralateral hand) and visual observation of the elicited twitch. MT level was defined as the minimum device intensity that induced an observable motor response in 50% of applied pulses, using an iterative algorithm for sequential testing. MT was expressed in Standardized MT units (SMTs) wherein 1.0 SMT corresponds to the average MT level observed in a large patient population (NeuroStar® System Instructions for Use, Rev. F, Apr. 2019).

External coordinates for coil placement over the DLPFC target were calculated by the device (coordinate for a site 5.5 cm anterior to the MT location, along a left superior oblique plane), but practitioners could use other methods to localize the stimulation target. The TrakStar software captured the age, gender, and MT level of each patient, as well as the treatment parameters at each session, including TMS intensity (% relative to MT), pulse frequency, duration of pulse trains, inter-train interval (ITI) duration, and total number of delivered pulses.

2.4. Statistical analyses

The six groups defined by number of TMS sessions in the acute course were compared in baseline demographics and treatment parameters, endpoint clinical outcomes, and clinical outcomes at fixed intervals of 10, 20, 30 and 36 sessions. For continuous measures, omnibus one-way analyses of variance (ANOVAs) were conducted testing whether groups differed. Significant effects of group membership were followed by Tukey-Kramer post hoc comparisons identifying the significant pairwise differences among the groups [36,37]. The same strategy was applied with the categorical measures; first conducting an omnibus Pearson chi-square test followed by post hoc comparisons using an adaptation of the Tukey-Kramer method.

Clinical outcomes at all time points included the raw PHQ-9 score, the difference from baseline (*Pre – Post*), and the percent change in score (*Pre – Post*)/*Pre*. Response was defined as $\geq 50\%$ reduction in PHQ-9 scores at follow-up assessment relative to pre-TMS baseline, and remission was defined as a follow-up PHQ-9 score of less than 5. Percentage change in PHQ-9 score was the primary outcome in each set of comparisons, while response and remission rates were the secondary outcomes. In addition, a metric reflecting the degree of clinical improvement per TMS session was calculated for each group by determining the change in mean percentage improvement in PHQ-9 score over an assessment interval (e.g., 1–10 sessions, 30–35 sessions) divided by the number of sessions in that interval (e.g., 10, 6). This metric also served as a secondary outcome measure to compare the groups in rate of improvement.

The first objective was to determine whether the groups differed in endpoint clinical outcomes, and if so, what was the length of treatment associated with peak effectiveness. A between-group, one-way ANOVA was conducted on the primary outcome measure and chi-square tests and one-way ANOVAs on the secondary outcome measures.

The second objective was to determine whether the groups differed in the rate or trajectory of improvement over the course of TMS. The six groups were compared in the clinical outcomes calculated following 10, 20, 30, 36 and endpoint TMS sessions, with separate between-group analyses conducted at each assessment time point. PHQ-9 assessments occurred within ± 4 days of each assessment time point, and the maximum available sample was identified for each group at each time point. There was little data loss as the groups varied from 89 to 99% in representation of the total assessable group at the interim time points. To confirm the findings regarding the group differences in degree and rate of improvement over time, a repeated measures mixed model was conducted specifically comparing the three groups that received 36 sessions, 37–41 sessions, and >41 sessions in percentage change in PHQ-9 scores at the assessments following 10, 20, 30, and 36 sessions. This

analysis required that participants contributed evaluations at all 4 time points.

To examine the possibility that there was a higher rate of extreme clinical outcomes in the groups with early termination, the four groups that received ≤ 36 TMS sessions were compared in the variance of the primary outcome measure after 10 and 20 sessions using Levene's F-test for variance equality [38]. These groups were also compared at these assessment occasions in the shape of their relative frequency distributions for percentage change in PHQ-9 scores binned into 10% change intervals.

The third objective was to determine whether the groups who received extended treatment (37–41 sessions and >41 sessions) showed meaningful clinical improvement beyond what they obtained after 30 or 36 sessions. To test for a main effect of time point, a repeated measures mixed model was conducted contrasting the two extended treatment groups in the extent of improvement in PHQ-9 scores at assessments following 30, 36 and endpoint sessions requiring that participants contributed evaluations at all 3 time points. Paired t-tests were used for comparison of adjacent time points. A repeated measures mixed model was also used to contrast the time points in rates of response and remission. The changes in response and remission rates at the assessments following session 30 and endpoint were also examined in the extended treatment groups.

The analyses of the Left Only and CGI-S subgroups were restricted to retesting the findings regarding the first objective: determining whether the groups differed in outcomes at endpoint and the number of sessions associated with peak effectiveness. Examination of the trajectory of improvement in these subgroups during the acute course was not conducted due to limited sample size at specific assessment time points. On the CGI-S, response corresponded to an endpoint score of 3 (“mildly ill”) or less, while remission corresponded to a score of 2 (“borderline mentally ill”) or less.

The analyses were conducted using SAS v9.4 (SAS Institute Inc., Cary, NC, USA). Results are reported in tables as mean \pm SD and in figures as mean \pm SEM. Treatment parameters were averaged over all treatment sessions in the acute course. Significance values are two-sided with an alpha of 0.05.

3. Results

3.1. Group differences in demographics and treatment parameters

Table 2 presents the demographic and TMS parameters for the total sample and the six groups defined by number of TMS sessions in the acute course. The groups did not differ in the distributions of gender or MT. The two extended treatment groups (“37–41 Sessions” and “ >41 Sessions”) were more likely to be treated with sequential bilateral TMS and consequently received more treatment protocols per session than each of the remaining groups. The extended treatment groups were also exceptional in receiving a larger number of pulses per session, and, on average, they were treated at a higher treatment level relative to MT. Due to the greater use of sequential bilateral TMS in the extended treatment groups, they also differed from the other groups in intertrain interval and pulse frequency.

Each of the groups were treated at an average rate greater than 4 sessions per week. The “1–19 Sessions” group was treated at a higher rate (4.56 sessions/wk) than all other groups, while the “ >41 Sessions” group had the lowest rate of treatment (4.06 session/wk). The “1–19 Sessions” group was also exceptional in being treated at a lower magnetic intensity relative to MT than all other groups. The low level of treatment intensity in this group likely reflected issues with tolerability.

3.2. PHQ-9 endpoint clinical outcomes

Endpoint PHQ-9 clinical outcomes for the groups differing in length of the TMS acute course are documented in Table 3 and Fig. 2a and b.

Table 2
Demographics, treatment schedule, and TMS parameters in groups that received differing durations of treatment with TMS.

Parameter	1–19 Sessions N = 658	20–29 Sessions N = 616	30–35 Sessions N = 1375	36 Sessions N = 3591	37–41 Sessions N = 626	>41 Sessions N = 349	All N = 7215	Test Statistic	p-value
Demographics									
Female	66.1% A	65.7% A	64.3% A	64.1% A	59.6% A	61.9% A	64.0%	$\chi^2(5) = 8.13$	0.15
Age	45.1 ± 15.7 ABC	44.9 ± 15.6 BC	46.8 ± 16.1 AB	47.5 ± 16.4 A	43.6 ± 15.1 C	47.0 ± 15.5 AB	46.6 ± 16.1	$F(5, 7209) = 9.18$	<.001
Treatment Schedule									
Number of Sessions	11.5 ± 4.8 F	25.1 ± 3.0 E	32.9 ± 2.1 D	36.0 ± 0.0 C	38.1 ± 1.2 B	48.2 ± 7.3 A	33.0 ± 8.5	$F(5, 7209) = 14657.41$	<.001
Days in Acute Course	18.9 ± 9.5 F	41.5 ± 8.3 E	54.1 ± 8.5 D	59.5 ± 6.5 C	63.0 ± 8.8 B	86.0 ± 24.2 A	54.8 ± 16.8	$F(5, 7209) = 3373.75$	<.001
Sessions per Week	4.56 ± 1.01 A	4.33 ± 0.60 B	4.32 ± 0.51 B	4.28 ± 0.43 B	4.30 ± 0.50 B	4.06 ± 0.62 C	4.31 ± 0.56	$F(5, 7209) = 44.22$	<.001
Treatments per Session	1.22 ± 0.38 BC	1.27 ± 0.37 B	1.25 ± 0.36 B	1.19 ± 0.34 C	1.44 ± 0.43 A	1.41 ± 0.41 A	1.24 ± 0.37	$F(5, 7208) = 67.43$	<.001
TMS Parameters									
Treatment Location									
Left	66.1% AB	58.0% C	61.8% BC	70.1% A	39.8% D	37.5% D	62.9%	$\chi^2(10) = 360.61$	<.001
Left and Right	30.7% CD	40.3% B	37.1% BC	29.1% D	58.6% A	61.6% A	35.9%		
Right	3.2% A	1.8% AB	1.1% AB	0.8% B	1.6% AB	0.9% AB	1.2%		
Motor Threshold (MT), % Device Output	1.04 ± 0.24 A	1.04 ± 0.22 A	1.04 ± 0.23 A	1.04 ± 0.22 A	1.03 ± 0.23 A	1.02 ± 0.22 A	1.04 ± 0.22	$F(5, 7208) = 0.80$	0.55
Treatment Level (% MT)	105.4 ± 13.8 E	112.6 ± 8.7 D	114.2 ± 7.6 C	115.5 ± 7.0 C	118.4 ± 9.5 B	119.9 ± 10.3 A	114.5 ± 9.1	$F(5, 7207) = 220.07$	<.001
Pulse Frequency, Hz	8.7 ± 2.7 AB	8.5 ± 2.4 B	8.8 ± 2.4 AB	9.0 ± 2.1 A	7.6 ± 2.5 C	7.9 ± 2.5 C	8.7 ± 2.4	$F(5, 7207) = 50.54$	<.001
Train Duration, s	4.0 ± 0.2 A	4.0 ± 0.1 AB	4.0 ± 0.2 AB	4.0 ± 0.2 AB	4.0 ± 0.2 B	4.0 ± 0.1 B	4.0 ± 0.2	$F(5, 7009) = 4.51$	<.001
Inter-Train Interval (ITI, s)	13.4 ± 5.3 B	13.7 ± 5.2 B	14.6 ± 5.8 A	13.3 ± 4.9 BC	12.5 ± 4.4 CD	12.3 ± 4.3 D	13.5 ± 5.1	$F(5, 7009) = 22.27$	<.001
Number of Pulses Delivered Per Session	3470.2 ± 922.6 B	3453.3 ± 813.9 B	3440.2 ± 727.9 B	3372.2 ± 763.9 B	4154.7 ± 978.7 A	4139.4 ± 927.8 A	3506.1 ± 846.1	$F(5, 7207) = 146.45$	<.001

Values are mean ± SD for continuous variables and percentage of group for categorical variables. One-way analyses of variance were conducted on the continuous variables and chi-square tests on the categorical variables. The letters following each group’s value describe the post-hoc comparisons. Groups that differ in all their following letters (e.g., AB vs DE) differed significantly from each other, while groups that shared any letter did not differ (e.g., B vs. BC).

Table 3
Endpoint PHQ-9 clinical outcomes in groups that received differing durations of treatment with TMS.

Parameter	1–19 Sessions N = 658	20–29 Sessions N = 616	30–35 Sessions N = 1375	36 Sessions N = 3591	37–41 Sessions N = 626	>41 Sessions N = 349	All N = 7215	Test Statistic	p-value
Baseline PHQ-9	20.1 ± 4.4 A	20.1 ± 4.2 A	19.6 ± 4.1 A	19.5 ± 4.1 A	19.5 ± 4.2 A	19.9 ± 4.3 A	19.7 ± 4.2	$F(5, 7209) = 3.36$	0.005
Endpoint PHQ-9	14.0 ± 6.6 A	11.1 ± 6.9 B	8.7 ± 6.5 DE	8.0 ± 6.3 E	9.0 ± 6.5 D	10.0 ± 6.4 C	9.1 ± 6.7	$F(5, 7209) = 111.47$	<.001
Difference (Pre-Post) Percentage Change	6.1 ± 6.2 D	9.0 ± 6.7 C	10.9 ± 6.8 AB	11.6 ± 6.7 A	10.6 ± 6.6 AB	9.9 ± 6.8 BC	10.5 ± 6.8	$F(5, 7209) = 84.81$	<.001
Response Rate (%)	29.8 ± 30.2 E	45.0 ± 31.5 D	55.1 ± 32.9 AB	58.9 ± 31.5 A	53.8 ± 32.4 BC	49.1 ± 30.9 CD	53.4 ± 32.8	$F(5, 7209) = 105.27$	<.001
Remission Rate (%)	26.4% E	47.6% D	63.9% AB	68.1% A	60.9% BC	55.0% CD	60.5%	$\chi^2(5) = 459.41$	<.001
	8.5% E	20.6% D	32.1% B	37.4% A	29.6% BC	22.1% CD	30.9%	$\chi^2(5) = 269.74$	<.001

Values are mean ± SD for continuous variables and percentage of group for categorical variables. One-way analyses of variance were conducted on the continuous variables and chi-square tests on the categorical variables. The letters following each group’s value describe the post-hoc comparisons. Groups that differ in all their following letters (e.g., AB vs DE) differed significantly from each other, while groups that shared any letter did not differ (e.g., B vs. BC).

None of the groups differed from each other in post hoc comparisons of baseline PHQ-9 scores. In contrast, there were marked differences in the post-treatment PHQ-9 score, change from baseline in this score (absolute and percentage change), and rates of response and remission. Effectiveness was poorest in the “1–19 Sessions” group which had significantly inferior outcomes in all measures relative to every other group. While the “20–29 Sessions” group had superior outcomes compared to the “1–19 Sessions” group, it had inferior outcomes on all measures when compared to each remaining group, except for a lack of difference in some measures with the “>41 Sessions” group. Antidepressant effect increased in the groups that received an increasing number of TMS sessions until it peaked in the “36 Sessions” group (Fig. 2a and b). Except

for the adjacent group that received 30–35 sessions, the “36 Sessions” group had significantly superior outcomes on all measures than each of the other groups. The “36 Sessions” group had a significantly higher rate of remission than the “30–35 Sessions” group, but the two groups did not differ in other PHQ-9 outcomes. The extended treatment groups (“37–41 Sessions” and “>41 Sessions”) showed intermediate efficacy outcomes, superior in most cases to those observed with short treatment courses (“1–19 Sessions” and “20–29 Sessions”), but generally inferior outcomes to the groups that received 30–35 or 36 sessions.

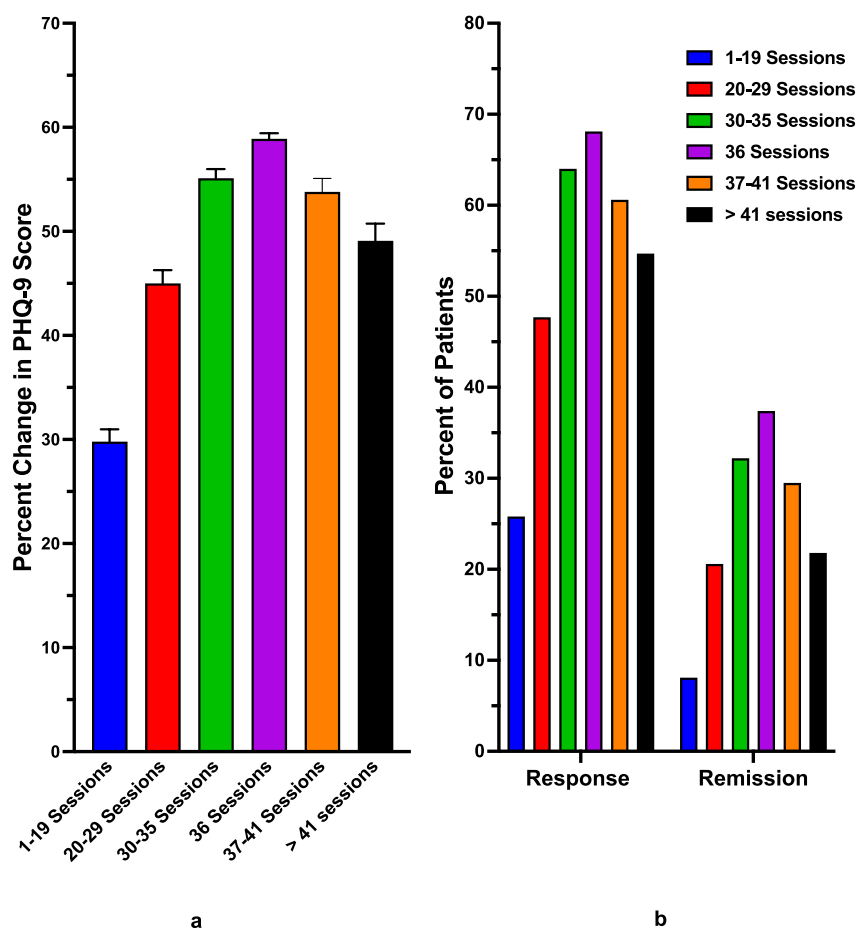


Fig. 2. Endpoint clinical outcomes for six groups defined by receipt of differing number of sessions in the acute TMS course. (a) Mean \pm SEM of percent change in PHQ-9 scores over the treatment course for the six groups. (b) Rates of response and remission in the six groups.

3.3. Clinical outcomes at fixed intervals during the acute course

Using the maximal data available at each time point, Table 4 and Fig. 3a–c present the clinical outcomes following assessments after 10, 20, 30, and 36 sessions, as well as at endpoint for the extended treatment groups. The analyses were consistent across time points and outcome measures. At the 10, 20, 30 and 36 sessions time points, the extended treatment groups had inferior antidepressant effects compared to all other groups. Furthermore, in most cases the group with the longest duration of treatment, “>41 Sessions”, had inferior outcomes compared to the other extended treatment group, “37–41 Sessions”. Thus, the groups that ultimately received especially long treatment courses had poorer antidepressant effects from the earliest assessment time point. Also of note, at the early time points, the groups that received the shortest treatment courses (“1–19 Sessions” and “20–29 Sessions”) improved to the same extent as groups that received longer courses (“30–35 Sessions” and “36 Sessions”). Thus, the groups that received the shortest courses of TMS were as responsive in symptomatic change at the early shared timepoints as the groups that later manifested peak effectiveness.

A repeated measure mixed model was conducted to verify that the extended treatment groups manifested less improvement from early in treatment and subsequently had a slower rate of improvement when restricted to participants with complete data sets. The “36 Sessions” group and the two extended treatment groups were compared in the percentage change from baseline in PHQ-9 score at the assessments following 10, 20, 30, and 36 sessions (Fig. 4). There was a significant main effect of time point, $F(3, 616) = 465.04$, $P < 0.0001$, indicating

that across the groups PHQ-9 scores improved over sessions. There was a significant main effect of group, $F(2, 609) = 67.35$, $P < 0.0001$, indicating that across the time points the groups differed in extent of improvement. There was also a significant interaction between group and time point, $F(6, 701) = 5.23$, $P < 0.0001$, indicating that the rate or slope of improvement differed among the groups. As seen in Fig. 4, after 10 sessions, the “36 Sessions” group ($N = 2848$) averaged a 30.0% improvement ($SD = 28.3$) in PHQ-9 scores, or 3.0% improvement in PHQ-9 score per session. The “37–41 Sessions” group ($N = 511$) averaged a 24.9% improvement ($SD = 27.1$), while the “>41 Sessions” group ($N = 285$), averaged a 17.1% improvement ($SD = 24.8$). One-way ANOVA indicated that the groups differed in extent of improvement at this Session 10 time point, $F(2, 618) = 37.84$, $P < 0.0001$, with post hoc comparisons indicating the three groups differed from each other. These three groups also differed in rate or slope of improvement between the 10th session assessment and the Session 36 assessment. During this interval the “36 Sessions”, “37–41 Sessions”, and “>41 Sessions” groups had an average additional improvement in PHQ-9 scores of 29.63% ($SD = 29.0$), 26.48% ($SD = 28.70$), and 20.93% ($SD = 27.47$), respectively. One-way ANOVA indicated that the groups also differed in the extent of improvement between the 10 and 36 Sessions time points, $F(2, 3641) = 13.24$, $P < 0.001$, with the post hoc comparisons indicating that the “>41 Sessions” group ($20.9\% \pm 27.5$) had significantly less improvement over this interval than either the “36 Sessions” ($29.6\% \pm 29.0$) or “37–41 Sessions” ($26.5\% \pm 28.7$) groups.

It is possible that the groups that received the shortest TMS courses had more extreme clinical outcomes, with higher rates of both marked benefit and ineffectiveness resulting in earlier termination. This could

Table 4

PHQ-9 clinical outcomes in groups that received differing durations of treatment after 10, 20, 30, 36, and endpoint TMS sessions.

Parameter	1–19 Sessions N = 658	20–29 Sessions N = 616	30–35 Sessions N = 1375	36 Sessions N = 3591	37–41 Sessions N = 626	>41 Sessions N = 349	Test Statistic	p-value
Assessment after 10 Sessions								
N	396	559	1234	3332	592	324		
Percentage Change	27.8 ± 29.5 A	29.8 ± 27.3 A	30.0 ± 28.7 A	29.9 ± 28.5 A	25.3 ± 26.8 A	17.9 ± 24.3 B	F(5, 6431) = 13.18	<.001
Interval % Change per Session	2.8 ± 3.0 A	3.0 ± 2.7 A	3.0 ± 2.9 A	3.0 ± 2.8 A	2.5 ± 2.7 A	1.8 ± 2.4 B	F(5, 6431) = 13.18	<.001
Response Rate (%)	24.5 AB	24.0 AB	26.8 A	25.8 A	20.1 B	10.8 C	X ² (5) = 45.73	<.001
Remission Rate (%)	7.1 A	6.4 A	6.3 A	7.1 A	4.7 AB	1.2 B	X ² (5) = 20.57	<.001
Assessment after 20 Sessions								
N	–	576	1227	3284	595	339		
Percentage Change	–	42.3 ± 30.8 A	41.5 ± 30.1 A	42.0 ± 30.5 A	36.4 ± 29.3 B	27.4 ± 28.8 C	F(4, 6016) = 21.48	<.001
Interval % Change per Session	–	1.2 ± 2.2 A	1.2 ± 2.4 A	1.2 ± 2.3 A	1.1 ± 2.3 A	1.0 ± 2.3 A	F(4, 5660) = 0.51	0.73
Response Rate (%)	–	43.9 A	42.5 A	44.4 A	34.8 B	23.3 C	X ² (4) = 70.32	<.001
Remission Rate (%)	–	16.3 A	14.8 A	15.3 A	9.7 B	5.9 B	X ² (4) = 34.46	<.001
Assessment after 30 Sessions								
N	–	–	1251	3188	585	327		
Percentage Change	–	–	51.9 ± 32.5 A	51.7 ± 31.0 A	44.8 ± 31.3 B	35.7 ± 29.3 C	F(3, 5347) = 32.97	<.001
Interval % Change per Session	–	–	1.0 ± 2.3 A	1.0 ± 2.1 A	0.9 ± 2.3 A	0.8 ± 2.6 A	F(3, 5001) = 1.16	0.32
Response Rate (%)	–	–	58.4 A	58.2 A	48.0 B	34.9 C	X ² (3) = 83.67	<.001
Remission Rate (%)	–	–	28.5 A	25.8 A	17.3 B	8.6 C	X ² (3) = 75.25	<.001
Assessment after 36 Sessions								
N	–	–	–	3591	583	321		
Percentage Change	–	–	–	58.9 ± 31.5 A	51.2 ± 32.6 B	38.4 ± 30.3C	F(2, 4492) = 71.58	<.001
Interval % Change per Session	–	–	–	1.3 ± 3.1 A	1.1 ± 3.2 A	0.6 ± 4.0 B	F(2, 4041) = 6.50	0.002
Response Rate (%)	–	–	–	68.1 A	56.3 B	39.6 C	X ² (2) = 124.31	<.001
Remission Rate (%)	–	–	–	37.4 A	26.8 B	11.8 C	X ² (2) = 101.77	<.001
Endpoint Assessment after > 36 Sessions								
N	–	–	–	–	626	349		
Percentage Change	–	–	–	–	53.8 ± 32.4 A	49.1 ± 30.9 B	F(1, 973) = 4.89	0.03
Interval % Change per Session	–	–	–	–	1.0 ± 6.2 A	1.0 ± 2.7 A	F(1, 902) = 0.02	0.89
Response Rate (%)	–	–	–	–	60.9 A	55.0 A	X ² (1) = 3.16	0.08
Remission Rate (%)	–	–	–	–	29.6 A	22.1 B	X ² (1) = 6.40	0.01

Values are mean ± SD for continuous variables and percentage of group for categorical variables response and remission rate). Rate of change per session was calculated as the change in the efficacy measure from the previous assessment interval divided by the number of sessions since the last assessment [e.g., (Assessment after 36 sessions – Assessment after 30 sessions)/6]. One-way analyses of variance were conducted on the continuous variables and chi-square tests on the categorical variables. The letters following each group's value describe the post-hoc comparisons. Groups that differ in all their following letters (e.g., AB vs DE) differed significantly from each other, while groups that shared any letter did not differ (e.g., B vs BC).

produce average symptom improvement scores equivalent to other groups, but would result in increased variance or altered peaks in the distribution of scores. Comparison of the variances in PHQ-9 percentage change scores for the groups that had ≤36 sessions did not demonstrate inequality after 10 sessions, $F(3, 5517) = 0.92$, $P = 0.43$, or 20 sessions, $F(2, 5084) = 0.29$, $P = 0.75$. [Supplemental Fig. 1](#) presents for each group the relative frequency distributions for the percentage change in PHQ-9 scores after 10 and 20 sessions, binned into 10% change intervals. There was no indication that the shape of these distributions differed in the groups receiving shorter or longer TMS courses. Patients who received shorter and longer TMS courses did not differ in the average extent of symptom improvement at early assessments nor in the variance or shape of the distributions of these scores.

3.4. Effectiveness associated with extended treatment

Using all available data, [Table 4](#) also presents the clinical outcomes for the extended treatment groups after 30, 36 and endpoint sessions. [Supplemental Table 1](#) presents the clinical outcome data restricted to participants who contributed evaluations at all three time points. The repeated measures mixed model contrasting the two extended groups in percentage change in PHQ-9 scores at these three time points yielded main effects of group $F(1, 571.2) = 17.29$, $P < 0.001$, time point, $F(2, 551.6) = 70.63$, $P < 0.001$, and a significant group by time point interaction, $F(2, 551.6) = 5.01$, $P = 0.007$. Both groups showed statistically significant and clinically meaningful improvement at each subsequent time point ([Supplemental Table 1](#)). Across time points, the “>41 Sessions” group had inferior outcomes compared to the “37–41 Sessions” group. However, the group by time point interaction reflected the fact that the “>41 Sessions” group had a significantly greater

improvement than the “37–41 Sessions” group in the interval between Session 36 and endpoint. During this interval (Session 36 to endpoint), the “37–41 Sessions” group averaged 2.7 (SD = 1.2) additional sessions, whereas the “>41 Sessions” group averaged an additional 12.2 (SD = 7.1) sessions, $t(583) = 21.95$, $P < 0.001$.

Around the 30th session, clinicians and patients often anticipate that a pre-authorized limit of 36 sessions will be insufficient and seek authorization for a treatment extension beyond 36 sessions. As documented in [Supplemental Table 1](#), in the “37–41 Sessions” group the response rate following the 30th session increased at the endpoint assessment from 45.9 to 61.6% and the remission rate increased from 14.7 to 29.7%. In the “>41 Sessions” group the response rate increased from 34.0 to 53.6% and the remission rate increased from 8.5 to 20.9%. These changes reflect clinically meaningful improvement in outcomes.

3.5. Time course of improvement with TMS

The foregoing analyses indicated that all patient groups showed their greatest rate of PHQ-9 improvement at the assessment after ten sessions and then the rate of improvement declined at the assessment following 20 sessions and remained relatively stable at subsequent assessments ([Table 4](#)). This pattern of symptom improvement was confirmed in within-group analyses requiring complete data at each assessment period ([Supplemental Table 2](#), [Fig. 5](#)). As seen in [Fig. 5](#), all but the extended treatment groups averaged approximately 3% improvement per session in percentage change in PHQ-9 score at the assessment following 10 sessions. In all groups, including extended treatment, this rate of improvement dropped to approximately 1.0% PHQ-9 improvement per session at the next assessment interval, 20 sessions, and remained relatively stable thereafter. There was no indication of a

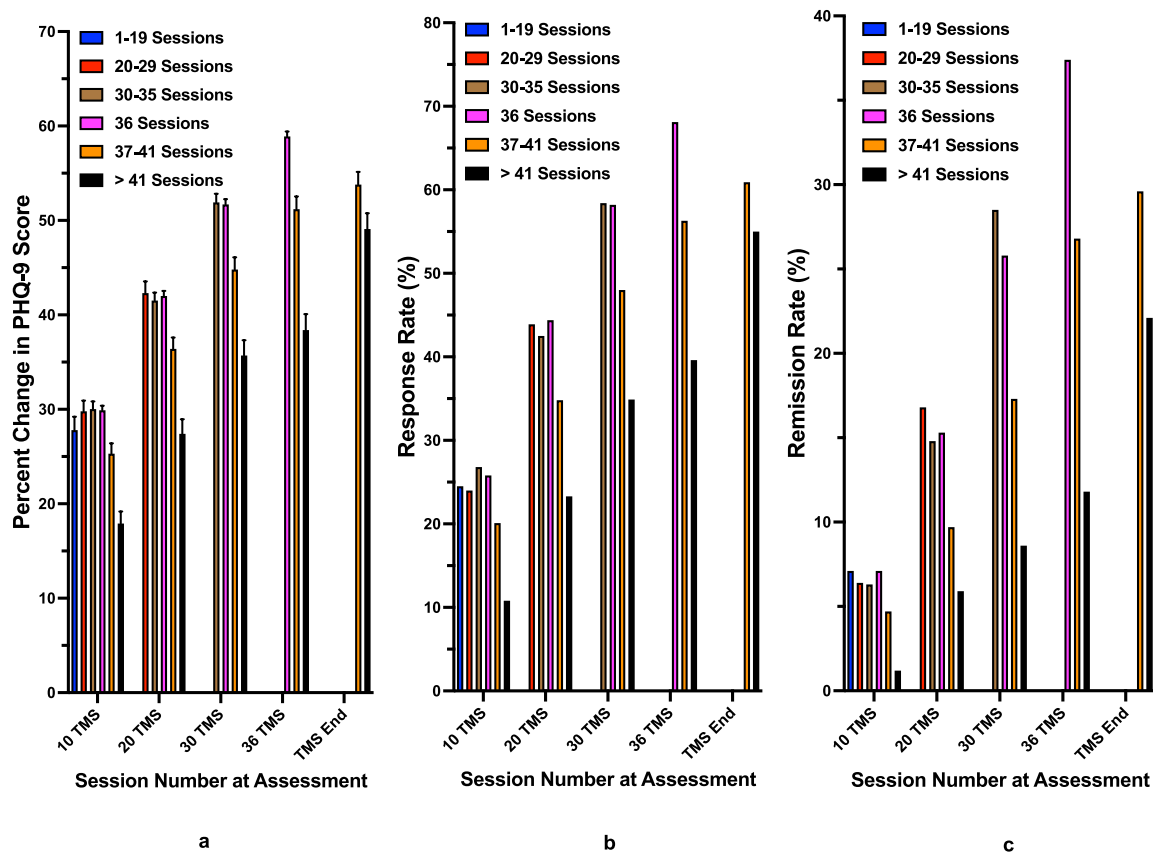


Fig. 3. Clinical outcomes for the six groups after 10, 20, 30 and 36 sessions and at endpoint in the groups with extended treatment (>36 sessions). (a) Percent change in PHQ-9 scores for the six groups at the fixed time points. (b) Response rates for the six groups at the fixed time points. (c) Remission rates for the six groups at the fixed time points.

plateau in efficacy with extended treatment as the rate of improvement per session appeared unchanged during the extended treatment periods.

3.6. Left only and CGI-S subsamples

The preceding analyses were conducted without constraint on the TMS treatment protocols administered. The findings also pertained to outcomes based on patient self-report. Therefore, generalizability was tested with respect to the subsample exclusively treated with the classic 10 Hz left DLPFC protocol and the subsample that had CGI-S ratings (see Table 1 for inclusion criteria in the Left Only and CGI-S subsamples). In both subsamples we retested the associations between total number of TMS sessions and endpoint clinical outcomes, as was reported in Table 3 and Fig. 2a and b for the total sample.

Supplemental Table 3 reports the findings for the Left Only subsample, while Supplemental Table 4 reports the findings for the CGI-S subsample. The CGI-S findings mirrored those in the total sample with PHQ-9 scores. The groups that received less than 30 sessions had the weakest antidepressant effects, effectiveness peaked around 36 sessions, and the extended treatment groups had intermediate outcomes. A similar pattern was also obtained in the PHQ-9 scores of the Left Only subsample, except that the effectiveness peak was broader as the 30–35 Sessions, 36 Sessions, and 37–41 Sessions groups each manifested strong antidepressant effects that did not differ.

4. Discussion

This naturalistic study, conducted in a large registry sample, found consistent associations between the number of TMS sessions patients were administered during an acute course of treatment for MDD and the

antidepressant effects of the intervention. Groups that received 1–19 sessions or 20–29 sessions had reduced benefit compared to groups that received longer treatment courses. In general, effectiveness peaked in the group that received 36 sessions and was still substantial, but reduced, in the groups that received extended treatment (36–41 sessions and >41 sessions). The extended treatment groups showed less clinical benefit and a slower rate of improvement from early in the treatment course. However, the additional treatments in the extended treatment courses were accompanied by additional and clinically meaningful improvement in symptom severity, with substantial increases in response and remission rates. There was no evidence in any group of a plateau over time in the average rate of improvement. Across the sample, the average trajectory reflected a reduction in symptom severity of about 3% per session over the first 10 sessions, followed by a rate of about 1% with each subsequent treatment. The association between number of sessions and effectiveness was confirmed in analyses restricted to the subsample treated exclusively with classic 10 Hz left DLPFC TMS and in the subsample with clinician-rated CGI-S scores.

This study documented consistent dose-response effects in real-world settings, as the number of TMS sessions strongly covaried with the effectiveness of the intervention. However, observation of robust dose-response effects does not entail demonstration of causal relationships [39,40]. Our findings derive from an observational study and not a prospective, randomized trial, and causality of number of sessions in impacting effectiveness was not tested. Across medicine, scores of important dose-response associations have been detected or confirmed in observational studies [41–44], with one of the most salient being the association between smoking exposure and cancer risk [45]. However, dose-response effects in observational treatment studies may be due to biases, especially when subgroups with different prognoses receive

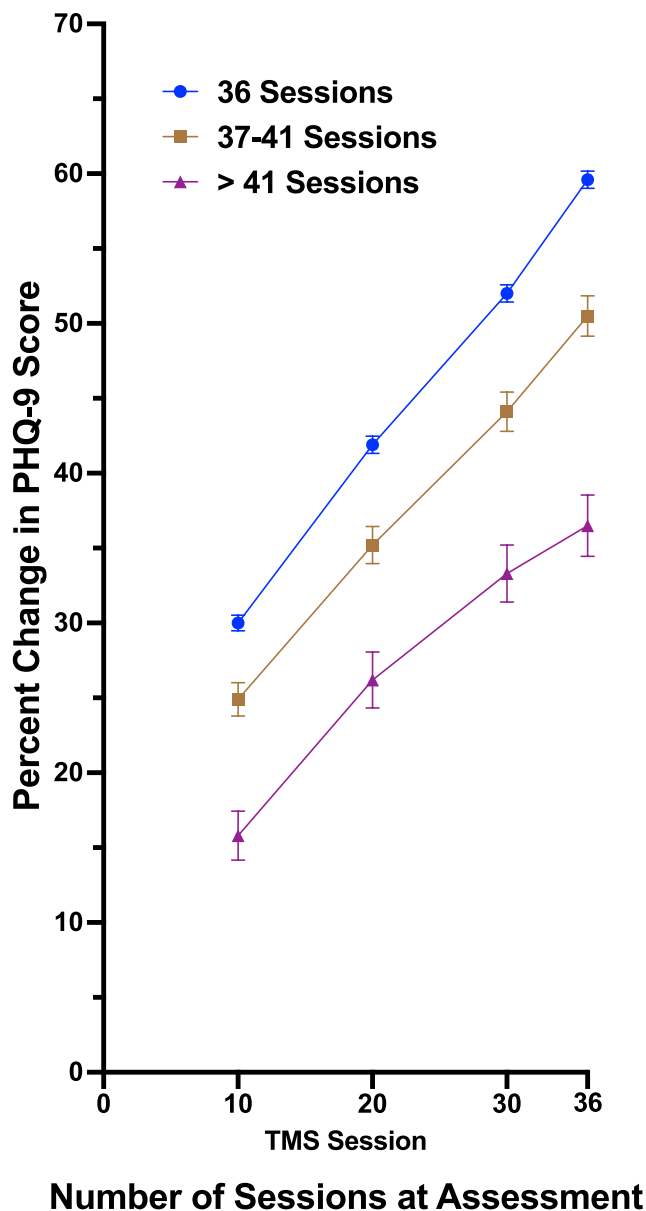


Fig. 4. Percent change in PHQ-9 scores for the “36 Sessions”, “37–41 Sessions”, and “>41 Sessions” groups after 10, 20, 30, and 36 sessions. To be included, patients contributed PHQ-9 scores at all 4 time points.

different dosing [39,40]. In our study, where clinicians were responding to their patients’ clinical state in real time, the decision on how many TMS sessions to administer was likely driven by the effectiveness of the intervention. Patients who showed little benefit, plateau in benefit, or symptomatic worsening would be expected to terminate sooner than patients showing progressive improvement. Not all patients had access to, or were offered, the same number of treatment sessions, and extensions beyond 36 sessions are might typically be requested from insurance companies for patients when the clinician believes they are likely to benefit further. Thus, our study describes dose-response relations in a large naturalistic sample where multiple factors determined treatment duration, and the extent to which the findings reflect a causal effect of treatment number or potential confounds is uncertain. Only randomized trials in which patients are assigned to different durations of treatment can definitively address the causal role of treatment number in impacting on effectiveness [46,47].

The findings regarding the trajectory of symptom change during the TMS course are compatible with the possibility that number of TMS sessions actively contributes to effectiveness, but since they are based on non-randomized, naturalistic data they are not determinative of causality. Patients who received fewer than 30 sessions had diminished clinical benefit at endpoint compared to all other groups, yet on a group level they appeared to be as responsive to TMS after 10 and 20 sessions as the groups that showed peak effects with longer treatment courses. The groups that received shorter TMS courses also did not differ from the groups with longer courses in the variance or shape of the distributions of symptom improvement scores at these assessments. Thus, the reduced endpoint effectiveness in the groups that received fewer than 30 sessions often may have been due to premature termination as opposed to treatment resistance. In contrast, the extended treatment groups, which showed reduced and slower improvement from early in the treatment course, experienced substantial clinical gains with longer courses of treatment. Their outcomes were superior to those that received short courses. This pattern is also compatible with an active effect of treatment number on effectiveness.

At the group level, there was no indication in this study that the antidepressant effects of TMS plateaued with prolonged treatment. This is consistent with the findings from the prospective extension trial that followed the O’Reardon et al. [1] sham-controlled trial; 73 participants without meaningful clinical benefit after 4–6 weeks of active TMS in the blinded RCT received open-label TMS for an additional 6 weeks [20]. The cumulative incidence of sustained response steadily rose over the 12 weeks of continuous (5 sessions/week) treatment. However, Zhang et al. [25] in a 12-week trial of patients randomized to 4 types of TMS reported no change in the magnitude of antidepressant effects after 5 weeks.

In this study, in all groups the average rate of symptomatic improvement was relatively uniform and stable after 10 sessions and until treatment termination, raising the possibility that additional treatment could result in further symptom reduction in patients who have shown incomplete improvement. In this and other TMS studies [3, 4,21,48], rates of response are often nearly double rates of remission, and providing longer TMS courses may be a useful strategy to maximize remission rates. It has been recently demonstrated that providing an additional second course of ECT results in marked clinical benefit in patients who were nonresponders to a standard course of ECT treatment [49]. In contrast, there is substantial evidence that the therapeutic effects of antidepressant medications generally plateau despite increasing oral dose [50–52] or extending the duration of treatment beyond 6–12 weeks [53–56].

In addition to the powerful associations between the number of sessions in the acute TMS course and antidepressant effectiveness, we characterized average trajectories of symptom improvement in groups defined by length of the TMS course. In patients who received 36 or fewer sessions, the average rate of improvement in symptom scores was about 3% per session after the first 10 sessions and then a steady 1.0–1.3% thereafter. Patients who received longer than typical courses of TMS showed less benefit after the first 10 sessions and a slower but steady rate of improvement thereafter. These patterns correspond well with the response trajectories empirically derived by Kaster et al. [24] when applying group-based trajectory modeling to serial depression scores of individual patients in a 6-week RCT. Similar to the average pattern of symptom change documented in this study, Kaster et al. [24] identified a group with Rapid Response where marked symptom reduction occurred over the first 2 weeks (10 sessions) followed by a stable and considerably slower rate of improvement. Kaster et al. [24] identified two groups that differed only in pre-TMS depression severity and who had slow and steady improvement from TMS outset. This pattern, labelled as Linear Response, resembles the trajectories we observed in the extended treatment groups, i.e., relatively diminished benefit after the first 10 sessions with a slow and steady rate of improvement thereafter. In addition, Kaster et al. [24] identified a

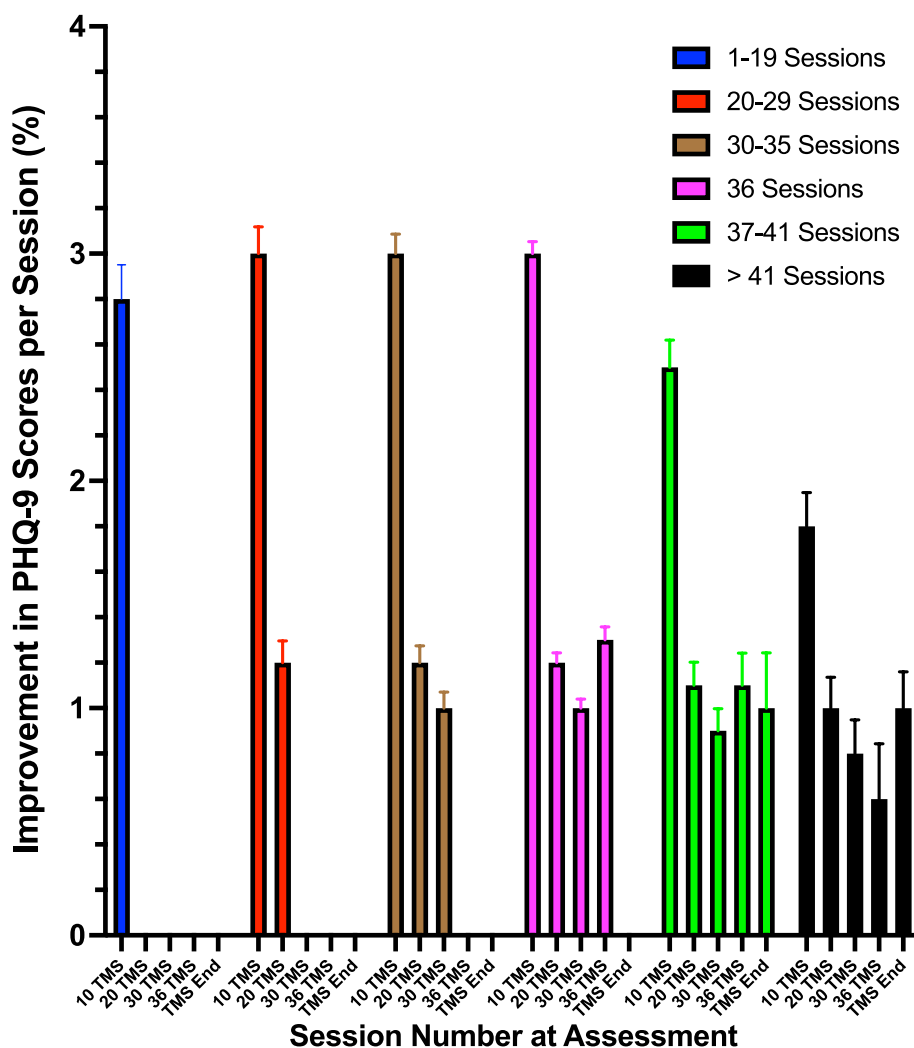


Fig. 5. Improvement in percentage change in PHQ-9 scores per session (Mean \pm SEM) for the six groups during the intervals of 1–10, 11–20, 21–30, 31–36 and 37-endpoint sessions. To be included, patients contributed PHQ-9 scores at all applicable time points.

fourth group of non-responsive individuals with minimal symptomatic change over time, which we did not attempt to isolate here. Despite differences in sample selection, outcome measures, TMS procedures, and study designs, our investigation and that of Kaster et al. [24] appear to describe similar rapid and slow patterns of symptom change during TMS for MDD.

This study found that exceptionally long treatment duration was associated with additional meaningful clinical benefit, and, in all groups, there was no drop-off in the rate of clinical improvement at the end of the treatment course. Presuming that lengthening the TMS course is a useful strategy to maximize effectiveness, there may be practical impediments with protocols involving a single session per day due to cost and time inefficiencies. Some accelerated protocols, in which multiple TMS sessions are administered in the same day, have shown rapid symptom improvement in MDD [57–59]. In studies implementing accelerated protocols, there appears to be a positive association between the total number of sessions and effectiveness in MDD [59], akin to the findings of this study. Another potential implication centers on the observation that after 10 TMS sessions the modal patient has a stable rate of improvement that is about one-third that displayed earlier. Adaptive neurophysiological changes contingent on treatment exposure are commonly observed with pharmacological and other neuromodulatory interventions that, in some cases, may interfere with therapeutic effects [60–62]. For example, ECT dynamically raises its own seizure threshold [63,64], and, yet, electrical dose relative to

threshold is a key determinant of both therapeutic and adverse effects [65–67]. Speculatively, identifying the adaptive mechanisms responsible for the marked reduction in the rate of improvement during the TMS course and potentially modifying TMS procedures to mitigate their impact could broadly enhance the efficiency and cost/effectiveness of this intervention.

The major strengths of this study, like its major limitations, derive from its observational nature utilizing a large patient registry in real-world settings. A strong association was demonstrated between the number of TMS treatments in an acute course and antidepressant effectiveness. Evidence was also provided that the rate of change in symptom severity during the TMS course has a characteristic pattern, typically declining sharply after 10 sessions and remaining stable thereafter until endpoint. In this study, patients who received exceptionally long TMS courses (>36 sessions) also had different trajectories of improvement, showing less benefit from early in the treatment course. This study also documented that the additional treatments administered in these extended courses were associated with substantial improvements in rates of response and remission.

Disclosures

The NeuroStar® Advanced Therapy System Clinical Outcomes Registry, analysis of the registry data, and the drafting of this manuscript were supported by Neuronetics Inc. (Malvern, PA, USA). While staff at

Neuronetics, Inc. provided comments on the manuscript, final approval of its content and the decision to submit were solely determined by the authors.

Previous presentation

Presented at the 5th International Brain Stimulation Conference, Lisbon, Portugal, February 20, 2023.

CRedit authorship contribution statement

Todd M. Hutton: Investigation, Conceptualization, Writing – review & editing. **Scott T. Aaronson:** Conceptualization, Writing – review & editing. **Linda L. Carpenter:** Conceptualization, Writing – review & editing. **Kenneth Pages:** Investigation, Writing – review & editing, Conceptualization. **David Krantz:** Formal analysis, Data curation. **Lindsay Lucas:** Formal analysis, Data curation. **Bing Chen:** Formal analysis, Data curation. **Harold A. Sackeim:** Conceptualization, Methodology, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Drs. Hutton and Pages serve as consultants to Neuronetics Inc. 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. Dr. Carpenter serves as a scientific advisor to, or received consulting income from, Neuronetics Inc, Nexstim PLC, Affect Neuro Inc, Neurolief LTD, Sage Therapeutics, Otsuka, Magnus Medical Inc, MAPS Public Benefit Corp, and Janssen Pharmaceuticals Inc. Dr. Carpenter has received research support (to Butler Hospital) from Neuronetics Inc, Neosync Inc, Nexstim PLC, Affect Neuro Inc, Neurolief, and Janssen Pharmaceuticals Inc. Mr. Krantz, Ms. Lucas and Dr. Chen report no financial relationship with commercial interests. Dr. Sackeim serves as a scientific adviser to and receives consulting fees from Cerebral Therapeutics Inc, Holmusk Inc, LivaNova PLC, MECTA Corporation, NeuroInsights Inc., Neurolief LTD, Neuronetics Inc, Parow Entheobiosciences LLC and SigmaStim LLC. 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 SigmaStim LLC. He is also the originator of magnetic seizure therapy (MST).

Acknowledgements

The authors thank Cory S. Anderson, MS and Steve Erickson, BS who facilitated the conduct of the statistical analyses and commented on the manuscript. Mr. Anderson and Mr. Erickson are employees of Neuronetics, Inc. and report no financial relationship with other commercial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.brs.2023.10.001>.

References

- O'Reardon JP, Solvason HB, Janicak PG, Sampson S, Isenberg KE, Nahas Z, et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatr* 2007;62(11):1208–16.
- George MS, Lisanby SH, Avery D, McDonald WM, Durkalski V, Pavlicova M, et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch Gen Psychiatr* 2010;67(5):507–16.
- Sackeim HA, Aaronson ST, Carpenter LL, Hutton TM, Mina M, Pages K, et al. Clinical outcomes in a large registry of patients with major depressive disorder treated with Transcranial Magnetic Stimulation. *J Affect Disord* 2020;277:65–74.
- Carpenter LL, Janicak PG, Aaronson ST, Boyadjis T, Brock DG, Cook IA, et al. Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice. *Depress Anxiety* 2012;29(7):587–96.
- Phillip NS, Dunner DL, Dowd SM, Aaronson ST, Brock DG, Carpenter LL, et al. 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* 2016;9(2):251–7.
- Lisanby SH, Husain MM, Rosenquist PB, Maixner D, Gutierrez R, Krystal A, et al. Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology* 2009;34(2):522–34.
- Fregni F, Marcolin MA, Myczkowski M, Amiaz R, Hasey G, Rumi DO, et al. Predictors of antidepressant response in clinical trials of transcranial magnetic stimulation. *Int J Neuropsychopharmacol* 2006;9(6):641–54.
- Gonterman F. A systematic review assessing patient-related predictors of response to Transcranial Magnetic Stimulation in major depressive disorder. *Neuropsychiatric Dis Treat* 2023;19:565–77.
- Thomas L, Kessler D, Campbell J, Morrison J, Peters TJ, Williams C, et al. Prevalence of treatment-resistant depression in primary care: cross-sectional data. *Br J Gen Pract* 2013;63(617):e852–8.
- Oquendo MA, Kamali M, Ellis SP, Grunebaum MF, Malone KM, Brodsky BS, et al. Adequacy of antidepressant treatment after discharge and the occurrence of suicidal acts in major depression: a prospective study. *Am J Psychiatr* 2002;159(10):1746–51.
- Levkovitz Y, Isserles M, Padberg F, Lisanby SH, Bystritsky A, Xia G, et al. Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial. *World Psychiatr* 2015;14(1):64–73.
- Sackeim HA, Dillingham EM, Prudic J, Cooper T, McCall WV, Rosenquist P, et al. Effect of concomitant pharmacotherapy on electroconvulsive therapy outcomes: short-term efficacy and adverse effects. *Arch Gen Psychiatr* 2009;66(7):729–37.
- Pluijms EM, Kamperman AM, Hoogendijk WJ, Birkenhager TK, van den Broek WW. Influence of an adjunct antidepressant on the efficacy of electroconvulsive therapy: a systematic review and meta-analysis. *Aust N Z J Psychiatr* 2021;55(4):366–80.
- Wei Y, Zhu J, Pan S, Su H, Li H, Wang J. Meta-analysis of the efficacy and safety of repetitive transcranial magnetic stimulation (rTMS) in the treatment of depression. *Shanghai Arch Psychiatr* 2017;29(6):328–42.
- Wilke SA, Johnson CL, Corlier J, Marder KG, Wilson AC, Pleman CM, et al. Psychostimulant use and clinical outcome of repetitive transcranial magnetic stimulation treatment of major depressive disorder. *Depress Anxiety* 2022;39(5):397–406.
- Chail A, Saini RK, Bhat PS, Srivastava K, Chauhan V. Transcranial magnetic stimulation: a review of its evolution and current applications. *Ind Psychiatr J* 2018;27(2):172–80.
- Caulfield KA, Brown JC. The problem and potential of TMS' infinite parameter space: a targeted review and road map forward. *Front Psychiatr* 2022;13.
- Burt T, Lisanby SH, Sackeim HA. Neuropsychiatric applications of transcranial magnetic stimulation: a meta analysis. *Int J Neuropsychopharmacol* 2002;5(1):73–103.
- McDonald WM, Durkalski V, Ball ER, Holtzheimer PE, Pavlicova M, Lisanby SH, et al. Improving the antidepressant efficacy of transcranial magnetic stimulation: maximizing the number of stimulations and treatment location in treatment-resistant depression. *Depress Anxiety* 2011;28(11):973–80.
- Avery DH, Isenberg KE, Sampson SM, Janicak PG, Lisanby SH, Maixner DF, et al. Transcranial magnetic stimulation in the acute treatment of major depressive disorder: clinical response in an open-label extension trial. *J Clin Psychiatry* 2008;69(3):441–51.
- Blumberger DM, Vila-Rodriguez F, Thorpe KE, Feffer K, Noda Y, Giacobbe P, et al. Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. *Lancet* 2018;391(10131):1683–92.
- Teng S, Guo Z, Peng H, Xing G, Chen H, He B, et al. High-frequency repetitive transcranial magnetic stimulation over the left DLPFC for major depression: session-dependent efficacy: a meta-analysis. *Eur Psychiatr* 2017;41:75–84.
- Valiengo L, Maia A, Cotovio G, Gordon PC, Brunoni AR, Forlenza OV, et al. Repetitive transcranial magnetic stimulation for major depressive disorder in older adults: systematic review and meta-analysis. *J Gerontol A Biol Sci Med Sci* 2022;77(4):851–60.
- Kaster TS, Downar J, Vila-Rodriguez F, Thorpe KE, Feffer K, Noda Y, et al. Trajectories of response to dorsolateral prefrontal rTMS in major depression: a THREE-D study. *Am J Psychiatr* 2019;176(5):367–75.
- Zhang T, Huang Y, Jin Y, Ma X, Liu Z. Treatment for major depressive disorder by repetitive transcranial magnetic stimulation in different parameters: a randomized double-blinded controlled trial. *Front Psychiatr* 2021;12:623765.
- Perera T, George MS, Grammer G, Janicak PG, Pascual-Leone A, Wirecki TS. The Clinical TMS Society consensus review and treatment recommendations for TMS therapy for major depressive disorder. *Brain Stimul* 2016;9(3):336–46.

- [27] Sackeim HA, Aaronson ST, Bunker MT, Conway CR, Demitrack MA, George MS, et al. The assessment of resistance to antidepressant treatment: rationale for the antidepressant treatment history form: short form (ATHF-SF). *J Psychiatr Res* 2019;113:125–36.
- [28] Aaronson ST, Carpenter LL, Hutton TM, Kraus S, Mina M, Pages K, et al. Comparison of clinical outcomes with left unilateral and sequential bilateral Transcranial Magnetic Stimulation (TMS) treatment of major depressive disorder in a large patient registry. *Brain Stimul* 2022;15(2):326–36.
- [29] Carpenter L, Aaronson S, Hutton TM, Mina M, Pages K, Verdoliva S, et al. Comparison of clinical outcomes with two Transcranial Magnetic Stimulation treatment protocols for major depressive disorder. *Brain Stimul* 2021;14(1):173–80.
- [30] Hutton TM, Aaronson S, Carpenter L, Pages K, West WS, Kraemer C, et al. The anxiolytic and antidepressant effects of transcranial magnetic stimulation in patients with anxious depression. *J Clin Psychiatry* 2023.
- [31] Optum Behavioral Health Solutions. Behavioral clinical policy: transcranial magnetic stimulation [updated 10/18/2022. Policy Number: BH803TMS102022: [Available from: <https://www.providerexpress.com/content/dam/ope-provexp/r/us/pdfs/clinResourcesMain/guidelines/bcp/tmsMCS.pdf>]; 2022.
- [32] Aetna. Transcranial Magnetic Stimulation and Cranial Electrical Stimulation. Aetna; 2022 [updated 7-12-2023. Available from: https://www.aetna.com/cpb/medical/data/400_499/0469.html#dummyLink2.
- [33] Cigna. Medical Coverage Policy: Transcranial Magnetic Stimulation. Cigna Companies; 2023 [updated 3/15/2023. Available from: https://static.cigna.com/assets/chcp/pdf/coveragePolicies/medical/mm_0383_coveragepositioncritria_transcranial_magnetic_stimulation.pdf.
- [34] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606–13.
- [35] Guy W. ECDEU Assessment Manual for Psychopharmacology. Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office, U.S. Department of Health, Education, and Welfare Publication; 1976. p. 76–338.
- [36] Tukey JW. *Exploratory Data Analysis*. Reading, MA: 1977.
- [37] Kramer CY. Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 1956.
- [38] Levene H. Robust tests for equality of variances. In: Olkin I, Hotelling H, editors. *Contributions to probability and statistics: essays in honor of Harold hotelling*. Palo Alto, CA: Stanford University Press; 1960. p. 278–92.
- [39] Rosenbaum PR. Does a dose-response relationship reduce sensitivity to hidden bias? *Biostatistics* 2003;4(1):1–10.
- [40] Gibbons RD. Bridging the gap between observational studies and randomized clinical trials. *Biol Psychiatr* 2023;93(12):1059–60.
- [41] Choi E, Park H, Lee K, Park J, Eisenhut M, van der Vliet H, et al. Body mass index and 20 specific cancers: re-analyses of dose-response meta-analyses of observational studies. *Ann Oncol: Off J Eur Soc Med Oncol/ESMO* 2018;29(3):749–57.
- [42] Liu M, Wu L, Yao S. Dose-response association of screen time-based sedentary behaviour in children and adolescents and depression: a meta-analysis of observational studies. *Br J Sports Med* 2016;50(20):1252–8.
- [43] Knott C, Bell S, Britton A. Alcohol consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of more than 1.9 million individuals from 38 observational studies. *Diabetes Care* 2015;38(9):1804–12.
- [44] Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993;4(3):218–28.
- [45] Lee PN, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 2012;12:385.
- [46] Lipkovich I, Adams DH, Mallinckrodt C, Faries D, Baron D, Houston JP. Evaluating dose response from flexible dose clinical trials. *BMC Psychiatr* 2008;8:3.
- [47] Berger VW, Bour LJ, Carter K, Chipman JJ, Everett CC, Heussen N, et al. A roadmap to using randomization in clinical trials. *BMC Med Res Methodol* 2021;21(1):168.
- [48] Berlin MT, van den Eynde F, Tovar-Perdomo S, Daskalakis ZJ. Response, remission and drop-out rates following high-frequency repetitive transcranial magnetic stimulation (rTMS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. *Psychol Med* 2014;44(2):225–39.
- [49] Sackeim HA, Prudic J, Devanand DP, Nobler MS, Haskett RF, Mulsant BH, et al. The benefits and costs of changing treatment technique in electroconvulsive therapy due to insufficient improvement of a major depressive episode. *Brain Stimul* 2020;13(5):1284–95.
- [50] Furukawa TA, Cipriani A, Cowen PJ, Leucht S, Egger M, Salanti G. Optimal dose of selective serotonin reuptake inhibitors, venlafaxine, and mirtazapine in major depression: a systematic review and dose-response meta-analysis. *Lancet Psychiatr* 2019;6(7):601–9.
- [51] Hieronymus F. Which antidepressant doses are optimal? *Lancet Psychiatr* 2019;6(7):552–4.
- [52] Furukawa TA, Salanti G, Cowen PJ, Leucht S, Cipriani A. No benefit from flexible titration above minimum licensed dose in prescribing antidepressants for major depression: systematic review. *Acta Psychiatr Scand* 2020;141(5):401–9.
- [53] Hirschfeld RM, Montgomery SA, Aguglia E, Amore M, Delgado PL, Gastpar M, et al. Partial response and nonresponse to antidepressant therapy: current approaches and treatment options. *J Clin Psychiatry* 2002;63(9):826–37.
- [54] Cheng Q, Huang J, Xu L, Li Y, Li H, Shen Y, et al. Analysis of time-course, dose-effect, and influencing factors of antidepressants in the treatment of acute adult patients with major depression. *Int J Neuropsychopharmacol* 2020;23(2):76–87.
- [55] American Psychiatric Association. Practice Guideline for the Treatment of Patients with Major Depressive Disorder. third ed. Washington, D.C.: American Psychiatric Publishing, Inc; 2010.
- [56] Department of Veterans Affairs, Department of Defense. VA/DoD Clinical Practice Guideline for the Management of Major Depressive Disorder. Washington DC: US VA/DoD; 2022.
- [57] Cole EJ, Phillips AL, Bentzley BS, Stimpson KH, Nejad R, Barmak F, et al. Stanford neuromodulation therapy (SNT): a double-blind randomized controlled trial. *Am J Psychiatr* 2022;179(2):132–41.
- [58] Cole EJ, Stimpson KH, Bentzley BS, Gulser M, Cherian K, Tischler C, et al. Stanford accelerated intelligent neuromodulation therapy for treatment-resistant depression. *Am J Psychiatr* 2020;177(8):716–26.
- [59] Caulfield KA, Fleischmann HH, George MS, McTeague LM. A transdiagnostic review of safety, efficacy, and parameter space in accelerated transcranial magnetic stimulation. *J Psychiatr Res* 2022;152:384–96.
- [60] Fang Y, Sun Y, Liu Y, Liu T, Hao W, Liao Y. Neurobiological mechanisms and related clinical treatment of addiction: a review. *Psychoradiology* 2022;2(4):180–9.
- [61] Targum SD. Identification and treatment of antidepressant tachyphylaxis. *Innov Clin Neurosci* 2014;11(3–4):24–8.
- [62] Fornaro M, Anastasia A, Novello S, Fusco A, Pariano R, De Berardis D, et al. The emergence of loss of efficacy during antidepressant drug treatment for major depressive disorder: an integrative review of evidence, mechanisms, and clinical implications. *Pharmacol Res: Off J Italian Pharmacol Soc* 2019;139:494–502.
- [63] Sackeim HA, Decina P, Prohovnik I, Malitz S. Seizure threshold in electroconvulsive therapy. Effects of sex, age, electrode placement, and number of treatments. *Arch Gen Psychiatr* 1987;44(4):355–60.
- [64] Sackeim HA. The convulsant and anticonvulsant properties of electroconvulsive therapy: towards a focal form of brain stimulation. *Clin Neurosci Rev* 2004;4:39–57.
- [65] Sackeim HA, Prudic J, Devanand DP, Kiersky JE, Fitzsimons L, Moody BJ, et al. Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *N Engl J Med* 1993;328(12):839–46.
- [66] Sackeim HA, Prudic J, Devanand DP, Nobler MS, Lisanby SH, Peyser S, et al. A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Arch Gen Psychiatr* 2000;57(5):425–34.
- [67] McCall WV, Reboussin DM, Weiner RD, Sackeim HA. Titrated moderately suprathreshold vs fixed high-dose right unilateral electroconvulsive therapy: acute antidepressant and cognitive effects. *Arch Gen Psychiatr* 2000;57(5):438–44.