Comparison of clinical outcomes with two Transcranial Magnetic Stimulation treatment protocols for major depressive disorder

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

a Butler Hospital, Providence, RI, USA
b Brown University, Department of Psychiatry and Human Behavior, Providence, RI, USA
c Sheppard Pratt Health System, Baltimore, MD, USA
d Department of Psychiatry, University of Maryland, Baltimore, MD, USA
e Southern California TMS Center, Los Angeles, CA, USA
f Neuronetics Inc., Malvern, PA, USA

g TMS of South Tampa, Tampa, USA
h FL USA University, Augusta, GA, USA
i Nashville NeuroCare Therapy, Nashville, TN, USA
j Department of Psychiatry, Columbia University, NY, USA
k Department of Radiology, Columbia University, NY, USA

Article history:
Received 21 August 2020
Received in revised form 7 December 2020
Accepted 9 December 2020
Available online 17 December 2020

Abstract

Background: Transcranial magnetic stimulation (TMS) is an effective treatment for major depressive disorder (MDD). The rest time between pulse trains is the inter-train interval (ITI). Since 2016, some TMS clinicians have adopted a stimulation protocol with shorter ITIs than were used in regulatory clinical trials.

Objective: To contrast treatment outcomes with the Standard TMS protocol (38.5 min per session) and the “Dash” protocol, which, at the shortest ITI, has a session duration of 18.75 min.

Methods: Registry data were collected at 103 practice sites. Of 7759 participants, 5010 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 = 3814) were responders or had received ≥ 20 sessions and had an end of acute treatment PHQ-9 assessment. Within the ITT sample, 613 patients were treated with the Standard NeuroStar 38-min protocol and 1493 patients with the new Dash protocol. CGI-S ratings were obtained in smaller samples. Treatment outcomes were also examined in subgroups considered Completers, as well as the subgroups who met criteria for Full Adherence to the Standard or Dash protocol parameters.

Results: In the ITT, Completer, and Fully Adherent samples, response (58–72%) and remission (28–53%) rates were notably high across PHQ-9 and CGI-S ratings. The Standard and Dash protocols did not differ in number of treatment sessions, and both manifested strong antidepressant effects.

Conclusions: The Standard and Dash protocols did not meaningfully differ in efficacy.

© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Introduction

Transcranial Magnetic Stimulation (TMS) is widely regarded as an effective treatment for episodes of major depressive disorder (MDD) [1–4]. In 2008, the first TMS device (NeuroStar® Advanced Therapy) was cleared by the US Food and Drug Administration (FDA) specifically for adult patients with treatment-resistant MDD. The original device “label” described a stimulation protocol targeting the left dorsolateral prefrontal cortex (DLPFC), and specified stimulation parameters that were used to establish effectiveness and safety in the pivotal regulatory clinical trials [5–7]. This “default” prescription stipulated 10 Hz stimulation at 120% intensity (relative to resting motor threshold), 4 s train duration, with each train separated by a 26 s inter-train “rest” interval (ITI), and 75 pulse trains, resulting in a total of 3000 pulses per session. Those same treatment parameters were used to confirm the efficacy and safety of TMS in a subsequent NHM-supported randomized sham-controlled trial by George et al. [8].

Stimulation protocols have evolved since the first use of TMS in MDD. Some of the earliest studies showing antidepressant efficacy (e.g., Ref. [9]) used conservative TMS dosing compared to modern standards, with relatively short trains and long ITIs (e.g., 2 s stimulation followed by 58 s rest), and far fewer pulses per session and fewer sessions per treatment course than commonly prescribed today. Over time, cumulative clinical trial experience revealed that a variety of stimulation protocols delivering higher net doses of stimulation could produce an antidepressant effect without jeopardizing safety [10].

Selection of specific parameters was largely informed by neurophysiology experiments in healthy volunteers that applied TMS to primary motor cortex with concurrent electromyographic (EMG) monitoring and manipulated magnetic pulse intensity, length of stimulation trains and ITI duration. Spread of excitation, facilitation of larger motor-evoked potentials, and risk for seizure induction were greatest when the ITI was short, i.e., 1 s or less [11]. Thus, ITI was recognized as an important safety parameter, particularly when targeting brain regions near motor cortex, and it was recommended that high-frequency protocols generally incorporate ITIs that are twice as long as the stimulation trains [12]. Those guidelines, together with technological considerations related to coil construction and cooling, influenced the selection of a fixed stimulation protocol that was applied for all TMS treatments delivered in the large regulatory clinical trials that lead to FDA clearance of the NeuroStar device.

In 2016, FDA cleared a new protocol termed “Dash” for the NeuroStar device. Dash uses the same parameters as the Standard protocol but allows a variable ITI, ranging from 11 to 25 s [13]. At the shortest ITI, this change reduces the duration of a 3000 pulse session to 18.75 min, i.e., half the time required for the Standard protocol. The Dash protocol was proposed after Neuronsenetics conducted an internal review of 44 studies with efficacy data in MDD and found that the shortest ITIs were associated with the highest response rates, while ITIs longer than 29 s were linked to lower response rates. In contrast, safety data from 79 studies suggested that adverse events, including seizure risk, did not vary with ITIs in the Dash range [13]. Consequently, in conjunction with an enhanced capacity NeuroStar coil, the Dash protocol was introduced to improve the acceptability and convenience of TMS therapy.

A relevant preclinical study evaluating the effect of ITI in intracortical short interval intracortical inhibition (SICI) and cortical excitability (motor evoked potentials) following a single session of four different protocols with varying ITIs [14]. In healthy subjects, the effects of TMS on excitability were found to be independent of ITI duration, and a greater level of cortical disinhibition was seen with shorter ITIs [14], raising the possibility that clinical protocols with shorter rest intervals between trains might also produce superior neuroplastic effects. However, we are aware of no clinical studies directly comparing MDD treatment outcomes associated with TMS protocols differing in ITI duration.

The NeuroStar® Advanced Therapy System Clinical Outcomes Registry has gathered demographic features, treatment parameters, and clinical outcomes of MDD patients treated on the NeuroStar device at over 100 clinical practice sites since 2016. To date, this registry has collected information on more than 9000 patients and constitutes the largest registry prospectively documenting treatment outcome in MDD patients [15]. Derived principally from private practitioner or private practice TMS centers, these data offer the large sample sizes necessary to meaningfully contrast the real-life clinical outcomes of patients treated specifically with the Standard and Dash TMS protocols.

Methods

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 de-identified 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) [16] and/or the Clinical Global Impression — Severity scale (CGI-S) [17] to assess the severity of depressive symptoms by self-report and clinician rating, respectively.

Data entry 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, as long as the required clinical data were collected at the appropriate intervals. Prior to the first acute phase TMS treatment, site personnel entered patient demographic information (date of birth, gender), site identifier, primary diagnosis and diagnoses of co-morbid psychiatric conditions (using ICD-9, ICD-10, or DSM-IV), and the PHQ-9 and CGI-S scores. Treatment parameters were captured at each treatment session passively with the NeuroStar TMS software which was synchronized to the TrakStar® Cloud software, 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, TMS intensity (treatment level, % device output relative to motor threshold), pulse frequency (e.g., 10 Hz vs. 1 Hz), and the number of treatment sessions during the acute phase treatment course. For this study, 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 following acute phase treatment termination. Additional PHQ-9 and CGI-S assessments were typically completed weekly during the course of treatment, per TMS standard of care.

Sample definitions

The NeuroStar Clinical Outcomes Registry collected data on 7759 patients treated at 103 U.S. sites (mean per site = 71.1 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 approximately 600 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).

The intent-to-treat (ITT) sample was defined by the following criteria. 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 comorbid diagnoses of psychiatric disorders other than generalized anxiety disorder (GAD), panic disorder, and unspecified anxiety disorder were also excluded (e.g., post-traumatic stress disorder, 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 5010 patients.

A subset of the ITT sample comprised the “Completer” sample (n = 3814). To ensure a minimally adequate course of TMS [18], individuals were excluded if classified as nonresponders and they ended TMS after fewer than 20 sessions (N = 301). Patients were also excluded if a PHQ-9 assessment was not conducted near the end of acute phase treatment (N = 895), i.e., within ±4 days of the final session.

**Dash and Standard protocol groups**

The initial analyses were designed to compare Dash and Standard protocols (both delivered over left DLPFC) with the largest possible sample sizes. Of the 5010 patients in the ITT sample, 2158 patients (43.1%) received some form of right-sided treatment and were thus excluded from this analysis. Of the 3814 patients in the Completer sample, 1697 patients (44.5%) received right-sided TMS and were excluded. Patients were also excluded if they received more than one left-sided treatment per session (N = 83). Thus, the ITT and Completers samples were reduced to 2764 and 2050 patients, respectively, when limited to the subgroups treated only with left-sided TMS with a single treatment per day.

Each left-sided, once-daily treatment session was classified as “Standard,” “Dash” or “Other”. Both Standard and Dash sessions required the left DLPFC target, 10 Hz pulse frequency, 4 s train duration, and >3000 pulses delivered in the session. The Standard protocol required an ITI of 26 s, while the Dash protocol required an ITI between 11 and 25 s. Any deviations from these parameters (e.g., 5 Hz stimulation, fewer pulses per session) resulted in the session being classified as “Other”. To be classified as having received a course of treatment with the Standard or Dash protocol, a patient must have received at least 75% of their sessions delivered with that protocol type. As seen in Table 1, when these criteria were applied, 22.2% of the ITT sample (N = 613) made up the Standard protocol group and 54% (N = 1493) made up the Dash protocol group. Excluded from analysis were 88 patients (3.2%) who received both the Standard and Dash protocols, with neither administered in 75% or more of sessions, and 570 patients (20.6%) who were excluded because 25% or more of their sessions were classified as “Other” and did not meet criteria for either the Standard or Dash protocol (e.g., stimulation frequency other than 10 Hz, too few pulses prescribed per session). The proportion of cases in each protocol group in the final Completer sample was similar to those for the ITT sample: 464 in the Standard group and 1111 in the Dash group.

### Table 1

Demographic and clinical characteristics of standard and dash protocol groups.

| Data Type | Intent-to-Treat Sample | P-Value* |
|-----------|------------------------|----------|
|           | Standard (N = 613)     | Dash (N = 1493) |
| Age       | 50.2 ± 15.1            | 50.1 ± 16.3 | 0.815 |
| Gender (% female) | 67.5%              | 65.6% | 0.389 |
| Baseline PHQ-9 | 19.6 ± 4.2           | 19.2 ± 4.2 | 0.046 |
| Baseline CGI-S | 5.3 ± 0.9            | 5.3 ± 0.7 | 0.484 |

| Data Type | Completer Sample | P-Value* |
|-----------|------------------|----------|
|           | Standard (N = 467) | Dash (N = 1112) |
| Age       | 50.9 ± 14.9      | 50.1 ± 16.1 | 0.329 |
| Gender (% female) | 66.2%              | 65.3% | 0.645 |
| Baseline PHQ-9 | 19.8 ± 4.2       | 19.1 ± 4.1 | 0.002 |
| Baseline CGI-S | 5.3 ± 0.9         | 5.3 ± 0.7 | 0.783 |

| Data Type | Fully Adherent Sample | P-Value* |
|-----------|----------------------|----------|
|           | Standard (N = 276)   | Dash (N = 354) |
| Age       | 50.4 ± 15.4          | 50.8 ± 16.9 | 0.756 |
| Gender (% female) | 64.1%              | 61.3% | 0.466 |
| Baseline PHQ-9 | 19.9 ± 4.3           | 19.1 ± 4.2 | 0.024 |
| Baseline CGI-S | 5.2 ± 0.9            | 5.2 ± 0.6 | 0.670 |
| Baseline CGI-S | 5.2 ± 0.9 (N = 118)  | 5.2 ± 0.6 (N = 109) | — |

* P-values compare Standard and Dash groups. 1 Includes all patients in the population with CGI-S score at baseline including patients with baseline CGI-S score ≤ 3.

### Fully Adherent sample

Secondary analyses compared subsets of the Dash and Standard Completer groups with strict protocol adherence. This comparison may better simulate the administration of TMS in a regulatory clinical trial with high protocol adherence. In addition to meeting the criteria that defined the Completer sample, the Fully Adherent subgroup received 100% of their sessions by the defined protocol (Standard or Dash), with an average of 2900–3500 pulses per session, and with a minimum mean treatment intensity of 115% (averaged across all sessions). Finally, to qualify for the Fully Adherent Dash subgroup, the ITT must have been shortened to 11 s within the first 5 treatment sessions. The Fully Adherent analyses were performed with 276 patients treated with the Standard Protocol and 354 treated with Dash.

### TMS procedures

MT was determined at the first treatment, using single pulse stimulation (of the left motor cortex area corresponding with the right abductor pollicis brevis muscle) and visual observation of thumb twitch. MT level was determined by the following System Instructions for Use, Rev. F, Apr. 2019) and is associated with an induced electric field of 135 V/m at a point located 2.0 cm from the surface of the scalp into the patient’s cortex along the central axis of the treatment coil.

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 Dash protocol differed from the Standard protocol only in the duration of the ITI. TMS treatments could be initiated at the first session with an ITI of 11 s; at this minimum, the duration of a Dash treatment period was 18.75 min. Practitioners also had the option of beginning TMS with a 26 s (default) ITI and progressively decreasing this parameter to produce shorter sessions over the treatment course. 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, pulse frequency, stimulation time, ITI duration, number of pulse trains, and total number of pulses.

Statistical analyses

The primary analyses were conducted in the ITT sample, then repeated in the Completer and Fully Adherent samples for confirmation. Descriptive statistics on demographic features, treatment parameters, and clinical outcomes are reported for each sample. The Standard and Dash groups were compared in demographic and treatment parameters using t-tests and chi-square analyses on continuous and categorical measures, respectively.

The primary contrast of the Standard and Dash groups in short-term efficacy was conducted on final PHQ-9 scores using analysis of covariance (ANCOVA). In the ITT sample, the final PHQ-9 was the last observation obtained after baseline and, in the Completer and Fully Adherent samples, the final PHQ-9 was obtained at the end of acute phase treatment. A simple “reduced” model was first tested, contrasting the Standard and Dash groups in final PHQ-9 scores, with only the baseline score as a covariate. The ANCOVA model was then expanded to include as covariates, age, gender, MT level, treatment level (% intensity relative to MT), number of pulses per session, and number of treatment sessions in the acute course. The expanded analysis was conducted to enhance sensitivity to potential differences in efficacy between the Standard and Dash protocols and to identify predictors or correlates of antidepressant effects.

The Standard and Dash groups were also compared in response and remission rates in the ITT, Completer, and Fully Adherent 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. In “reduced” models, logistic regression analyses were conducted on these rates with protocol (Standard vs. Dash) and baseline score as independent variables. These were followed with expanded analyses, adding age, gender, MT level, TMS intensity, number of pulses delivered, and number of treatment sessions as covariates.

Analyses with the CGI-S were conducted in a considerably smaller subset of the sample used for PHQ-9 analyses, due to missing data and application of exclusion criteria. Patients with pre-TMS CGI-S scores ≥ 3 were excluded from CGI-S analyses due to insufficient baseline illness severity (n = 9). On the CGI-S, categorical 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 analyses of the CGI-S scores followed the procedures used with the PHQ-9 scores.

Preliminary analyses indicated that site did not interact with treatment protocol group in influencing clinical outcomes. Within the Fully Adherent sample, outcomes for the Standard and Dash groups were compared specifically within 9 sites that treated at least 10 patients with each protocol. The findings of these analyses were consistent with those obtained in the larger samples.

Descriptive statistics are reported as mean ± SD for continuous variables and frequency counts and percentages for categorical variables. Treatment parameters were averaged over all treatment sessions in the acute course. Significance values are two-tailed with an alpha of 0.05. All p-values reported are without multiplicity adjustment. All analyses were conducted using SAS v9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Table 1 presents demographic and clinical characteristics of the ITT, Completer, and Fully Adherent samples for each treatment protocol. The Standard and Dash groups did not differ in age, baseline CGI-S scores, or gender representation. Baseline PHQ-9 scores were slightly higher in the Standard than the Dash group in the ITT sample, t(2,104) = 2.00, P = 0.046, Completer sample, t(1,577) = 3.07, P = 0.002, and the Fully Adherent sample, t(628) = 2.26, P = 0.024.

Table 2 presents the treatment parameters applied with each protocol. On average, patients received over 30 TMS sessions during the acute course, over a period of 7–8 weeks. The Standard and Dash groups did not differ in these measures, nor in pulse frequency. There was a substantial difference in MT between the groups. MT was higher in the Standard than Dash group in the ITT, t(2,104) = 6.41, P < 0.001, Completer, t(1,577) = 6.28, P < 0.001, and Fully Adherent, t(628) = 6.47, P < 0.001, samples. The Dash group had an average ITI of 13.4 s across sessions, indicating that many patients were treated from the outset with an ITI of 11 s (53%, ITT Dash sample). The Dash group also received slightly more pulses per session than the Standard group [ITT sample: t(2,104) = 6.75, P < 0.001; Completer sample: t(1,577) = 6.32, P < 0.001; Fully Adherent sample: t(628) = 2.30, P = 0.022]. Finally, the Dash group had a slightly lower average treatment level, i.e., the extent to which the magnetic intensity exceeded MT, only in the Fully Adherent sample, t(628) = 2.08, P = 0.038. Delivery of first-week sessions at a lower intensity to allow accommodation likely explains this finding, as there was no group difference in treatment intensity measured at the time of the final TMS session, t(628) = −0.06, P = 0.949.

Self-report depression severity

Table 3 presents the PHQ-9 clinical outcome measures for each sample. In the primary analysis in the ITT sample using the reduced ANCOVA model, there was a marked effect of baseline PHQ-9 score on final PHQ-9 score, F(1, 2103) = 161.2, P < 0.001. As expected, individuals with higher baseline scores had higher final PHQ-9 scores. In addition, there was a significant effect of treatment protocol, F(1, 2103) = 4.06, P = 0.04. Final PHQ-9 scores were higher in the Dash compared to Standard protocol group. Supplemental Table 1 presents the results of the reduced models applied for each sample, while Supplemental Tables 2–4 presents the results for the expanded models in which age, gender, MT level, treatment level, number of pulses per session, and number of treatment sessions were included as additional covariates. The difference between the Standard and Dash groups in final PHQ-9 scores remained significant in the expanded ANCOVA in the ITT sample, F(1, 2904) = 7.62, P = 0.006. In the expanded analysis, several covariates were also significantly associated with final PHQ-9 scores, including gender (P < 0.001), treatment level (P = 0.004), pulses per session (P = 0.006), and number of sessions in the acute course (P < 0.001). These findings were consistent with an earlier report on the total ITT and Completer samples [15], and indicated better clinical outcomes associated with female sex, higher treatment levels, more pulses per session, and a larger number of sessions in the acute course.

As seen in Supplemental Table 1, there was no effect of treatment group on final PHQ-9 scores in the reduced models in the
Table 2
TMS treatment parameters for standard and dash protocol groups.

|                         | Intent-to-Treat Sample |                     |                     |
|-------------------------|------------------------|---------------------|---------------------|
|                         | Standard (N = 613)     | Dash (N = 1493)     | P-Value*            |
| Number of Sessions in   |                        |                     |                     |
| Acute Course            | 31.4 ± 8.5             | 31.8 ± 8.6          | 0.368               |
| Acute Course Duration   | 50.0 ± 15.9            | 51.4 ± 16.1         | 0.076               |
| Number of Treatments per Session | 1.0 ± 0.0 | 1.0 ± 0.0 | –                   |
| Motor Threshold (MT, % device output) | 1.10 ± 0.24 | 1.03 ± 0.23 | –0.001              |
| Treatment Level (MT across sessions) | 114.2 ± 7.1 | 114.8 ± 7.6 | 0.103               |
| Interval Between Pulse Trains (ITI, sec) | 27.0 ± 2.9 | 13.4 ± 3.5 | N/A                 |
| Pulse Frequency per Session | 10.0 ± 0.1 | 10.0 ± 0.0 | 0.351               |
| Number of Pulses per Session | 3045.8 ± 184.2 | 3183.1 ± 490.2 | <0.001              |
|                         |                        |                     |                     |
| Completor Sample        |                        |                     |                     |
| Number of Sessions in   | 33.6 ± 5.5             | 34.0 ± 6.0          | 0.239               |
| Acute Course Duration   | 53.7 ± 12.3            | 54.7 ± 12.3         | 0.133               |
| Number of Treatments per Session | 1.0 ± 0.0 | 1.0 ± 0.0 | –                   |
| Motor Threshold (MT, % device output) | 1.10 ± 0.24 | 1.02 ± 0.23 | <0.001              |
| Treatment Level (MT across sessions) | 115.6 ± 5.1 | 115.7 ± 6.7 | 0.856               |
| Interval Between Pulse Trains (ITI, sec) | 27.0 ± 3.0 | 13.5 ± 1.5 | N/A                 |
| Pulse Frequency per Session | 10.0 ± 0.1 | 10.0 ± 0.0 | 0.124               |
| Number of Pulses per Session | 3049.9 ± 187.2 | 3204.3 ± 512.3 | <0.001              |
|                         |                        |                     |                     |
| Fully Adherent Sample   |                        |                     |                     |
| Number of Sessions in   | 34.2 ± 5.5             | 34.1 ± 5.7          | 0.688               |
| Acute Course Duration   | 54.5 ± 12.0            | 54.6 ± 11.5         | 0.953               |
| Number of Treatments per Session | 1.0 ± 0.0 | 1.0 ± 0.0 | –                   |
| Motor Threshold (MT, % device output) | 1.07 ± 0.24 | 0.95 ± 0.21 | <0.001              |
| Treatment Level (MT across sessions) | 118.1 ± 2.2 | 118.5 ± 2.2 | 0.038               |
| Interval Between Pulse Trains (ITI, sec) | 27.0 ± 2.9 | 11.3 ± 1.1 | <0.001              |
| Pulse Frequency per Session | 10.0 ± 0.0 | 10.0 ± 0.0 | N/A                 |
| Number of Pulses per Session | 3004.7 ± 39.4 | 3016.7 ± 78.9 | 0.022               |

* P-values for comparisons of Standard and Dash groups.

Table 3
PHQ-9 clinical outcomes for standard and dash protocol groups.

|                         | Intent-to-Treat Sample |                     |                     |
|-------------------------|------------------------|---------------------|---------------------|
|                         | Standard (N = 613)     | Dash (N = 1493)     | P-Value*            |
| Baseline PHQ-9          | 19.6 ± 4.2             | 19.2 ± 4.2          | 0.046               |
| LOCF PHQ-9              | 8.6 ± 6.4              | 9.0 ± 6.6           | 0.158               |
| Difference (Pre-Post)   | 11.0 ± 7.0             | 10.2 ± 6.7          | 0.009               |
| Response Rate           | 64.8%                  | 59.9%               | 0.037               |
| Remission Rate          | 22.7%                  | 30.8%               | 0.457               |
|                         |                        |                     |                     |
| Completor Sample        |                        |                     |                     |
| Baseline PHQ-9          | 19.8 ± 4.2             | 19.0 ± 4.1          | 0.002               |
| End of Acute Course PHQ-9 | 7.9 ± 6.1       | 7.9 ± 6.2           | 0.993               |
| Difference (Pre-Post)   | 11.8 ± 6.8             | 11.2 ± 6.5          | 0.057               |
| Response Rate           | 70.2%                  | 68.3%               | 0.438               |
| Remission Rate          | 35.3%                  | 35.9%               | 0.899               |
|                         |                        |                     |                     |
| Fully Adherent Sample   |                        |                     |                     |
| Baseline PHQ-9          | 19.9 ± 4.3             | 19.1 ± 4.2          | 0.024               |
| End of Acute Course PHQ-9 | 7.7 ± 5.8         | 7.9 ± 6.2           | 0.749               |
| Difference (Pre-Post)   | 12.2 ± 7.0             | 11.2 ± 6.6          | 0.089               |
| PHQ-9 Response Rate     | 71.7%                  | 66.9%               | 0.197               |
| PHQ-9 Remission Rate    | 37.3%                  | 37.9%               | 0.891               |

* P-values for comparisons of Standard and Dash groups.

Completer and Fully Adherent samples. Likewise, the Standard and Dash protocol groups did not differ in these scores in the expanded models in the Completer and Fully Adherent samples (Supplemental Tables 3-4). The same covariates were significantly related to final PHQ-9 scores in the ITT and Completer samples (Supplemental Tables 2-3).

In the ITT sample, the reduced simultaneous logistic regression analyses yielded a significant effect of Standard vs. Dash protocol on the LOCF response rate, \( F(1, 2103) = 4.50, P = 0.034 \), but not on the remission rate, \( F(1, 2103) = 1.15, P = 0.284 \) (Supplemental Table 1). These results were unchanged in the expanded logistic regressions, which also showed a significant pattern of covariate effects similar to the findings with final PHQ-9 scores. The simultaneous logistic regression analyses in the Completer and Fully Adherent samples produced highly similar results, except that the Standard and Dash protocol groups did not significantly differ in response or remission rates (Supplemental Tables 1, 3, and 4).

CGI-S scores

Table 4 presents the CGI-S clinical outcome measures for each sample, restricted to participants with a baseline CGI-S score \( \geq 4 \) and at least one post-baseline CGI-S score (ITT sample) or to patients with a baseline and end of acute treatment CGI-S score (Completer and Fully Adherent samples). Both the reduced and expanded covariate analyses indicated that final CGI-S scores were lower in the Dash than Standard group in the ITT, Completer, and Fully Adherent samples (Supplemental Tables 1-4).
Despite the large sample size, the two protocol groups did not differ in baseline CGI-S scores among patients treated with TMS principally in private practice settings in the United States. Data from this registry were used to examine treatment outcomes as function of inter-train interval (ITI), comparing outcomes with the 26-second ITI “default” protocol to “Dash” protocols that shorten the rest time between trains. Despite the large sample size, differences in efficacy between the protocols were relatively small in magnitude, differed in direction for final PHQ-9 and CGI-S scores, and are reported here without regard for multiple statistical comparisons. Across the three a priori samples (ITT, Completer, and Fully Adherent) and using both simple and expanded logistic regression models, there were no differences in remission rates and virtually no differences in response rates for PHQ-9 and CGI-S outcomes. Final PHQ-9 scores were higher in the Dash than Standard group, while the opposite finding was seen with final CGI-S scores. The larger antidepressant effect of Standard protocol treatment on final PHQ-9 scores was largely attributable to higher baseline PHQ-9 scores among patients in that group. Indeed, the Standard and Dash groups did not differ when directly compared in final PHQ-9 scores (Table 3). CGI-S ratings were obtained in considerably fewer patients; when analyzing CGI-S data, we found larger antidepressant effects in the Dash treatment group manifested as lower final CGI-S scores. However, this effect was small in magnitude and was not reflected in consistent differences between protocol groups for CGI-S response and remission rates.

Overall, marked antidepressant effects were documented in this study, in line with findings obtained in the companion report on the total registry sample [15]. The main finding of this study confirmed that protocols with shortened ITI, such as instantiated by the Dash protocol, are generally equivalent with regard to short-term antidepressant outcomes. Even when the data were constrained to subsets that had strictly adhered to either Standard (37.5 min session) or Dash (18.75 min session) protocols, our results did not detect a clinically meaningful difference in overall treatment outcomes.

Despite the large sample size, the two protocol groups did not differ meaningfully in outcomes defined by PHQ-9, a self-report scale. The Dash protocol was superior to the Standard protocol only in two of the clinician-rated outcomes (CGI-S), but these were likely attributable to the patient sample or the TMS practice characteristics at the relatively few sites contributing CGI-S data to the registry. The superior effects of Dash in the clinician ratings were also relatively small in magnitude and could have reflected a contemporaneous change in the treatment methods other than the change to Dash or a difference in unblinded rater expectancy. The two protocols were administered predominantly in different epochs of time, with the Dash patients more recently treated.

Taken together, the self-report and clinician-rating findings comprising the registry data indicate that the Standard and Dash protocols both exert substantial efficacy and, conservatively, that there is little evidence that either protocol is more or less effective than the other. The registry did not collect information on adverse effects or the durability of benefit in those who responded or remitted — important considerations when evaluating alternative interventions. Our finding of general equivalence between two 10 Hz protocols based on ITI duration adds to a growing number of reports describing fairly similar acute treatment outcomes across pairwise comparisons of TMS protocols for MDD that differed with regard to type of stimulation (e.g. conventional versus theta burst stimulation), pulse frequency, and laterality [19–21]. While findings from this large naturalistic dataset still indicate that certain parameters, such as greater number of total number of pulses per session, may confer superior treatment outcomes [15], the data suggest that patients who terminated the course of therapy before 20 sessions were simply less likely to do well than those who continued beyond that point. Results from the present analysis show that treatment with more than 4000 pulses per session was a strong predictor of positive response in the ITT sample, with less predictive value in the Completer sample, and no consistent significant predictive value in the Fully Adherent sample.

Several steps were taken to address the heterogeneity inherent in this naturalistic dataset. We excluded 88 patients (3.2%) patients who received both the Standard and Dash protocols during their course of treatment because neither was administered in 75% or

### Table 4

| Table 4 | CGI-S clinical outcomes for the standard and dash protocol groups. |
|---|---|
| | Intent-to-Treat Sample | Standard | Dash | P-Value* |
| | (N = 160) | (N = 403) | |
| Baseline CGI-S | 5.2 ± 0.9 | 5.2 ± 0.7 | 0.279 |
| LOCF CGI-S | 2.7 ± 1.4 | 2.4 ± 1.3 | 0.011 |
| Difference (Pre-Post) | 2.9 ± 1.3 | 2.9 ± 1.4 | 0.003 |
| Response Rate | 75.0% | 80.1% | 0.178 |
| Remission Rate | 52.5% | 60.3% | 0.091 |
| | Completer Sample | Standard | Dash | P-Value* |
| | (N = 137) | (N = 335) | |
| Baseline CGI-S | 5.2 ± 0.9 | 5.3 ± 0.7 | 0.286 |
| End of Acute Course CGI-S | 2.6 ± 1.4 | 2.2 ± 1.2 | 0.006 |
| Difference (Pre-Post) | 2.6 ± 1.3 | 3.0 ± 1.3 | 0.001 |
| Response Rate | 78.8% | 85.4% | 0.082 |
| Remission Rate | 57.7% | 66.0% | 0.089 |
| | Fully Adherent Sample | Standard | Dash | P-Value* |
| | (N = 96) | (N = 80) | |
| Baseline CGI-S | 5.2 ± 0.9 | 5.2 ± 0.6 | 0.782 |
| End of Acute Course CGI-S | 2.6 ± 1.4 | 2.2 ± 1.0 | 0.018 |
| Difference (Pre-Post) | 2.5 ± 1.3 | 3.0 ± 1.2 | 0.013 |
| Response Rate | 75.0% | 86.3% | 0.063 |
| Remission Rate | 55.2% | 65.0% | 0.187 |

*P-values derive from comparisons of Standard and Dash groups. P-values for post-baseline outcomes are based on a two-sample t-test without adjustment for base-line score; LOCF = last observation carried forward. Analyses include patients in the population with CGI-S scores at baseline and post-baseline. Patients with baseline CGI-S score <3 are excluded.
more of sessions. We also excluded 570 patients (20.6%) because 25% or more of their sessions were classified as “Other” and did not meet criteria for either the Standard or Dash protocol for various reasons (e.g., stimulation frequency other than 10 Hz, too few pulses prescribed per session). Finally, we defined and selected a “Fully Adherent” sample for additional supportive analysis to evaluate consistency of the results. The Fully Adherent sample received 100% of their sessions by the defined protocol (Standard or Dash) and their ITIs must have been shortened to 11 s within the first 5 treatment sessions. Even though we constrained the Fully Adherent groups to cases with a mean of 2900–3500 pulses per session, the Dash group had slightly more pulses per session. Evaluation of outcomes for protocols with less than the standard 3000 pulses per session was beyond the scope of this study, but the consistency of our efficacy findings across simple analyses comparing Dash and Standard, as well as results of our expanded analyses controlling for pulse number, suggest that the slight difference in mean total pulses per session did not influence the overall findings.

The major theoretical concern with the shorter ITI of Dash is that greater neuroexcitability could increase the risk of TMS-induced seizure [1–4]. To date, only 19 suspected seizures have been reported to Neuronetics since FDA clearance in 2008. There has been no trend for increased suspected seizure rate since clearance of Dash protocol (Neuronetics Inc, data on file, January 7, 2020).

This study was limited by its naturalistic, observational design. Indeed, the bulk of patients were treated at a later point in time with the Dash than Standard protocols, and secular changes other than the change in ITI may have influenced the results. For example, MT was substantially lower in the Dash than Standard group, which may have been due to changes in TMS coil design and other internal hardware that enabled implementation of the Dash protocol while maintaining system reliability. Nonetheless, MT was not associated with clinical outcome in the expanded models.

Similarly, variable methods used by clinicians for localizing the DLPFC may also limit interpretation of the outcomes with regard to unknown effects attributable to the targeting method used or the consistency of targeting across a course of treatment. A change over time from use of the “5.5 cm rule,” to the “Beam F3 method” or by placement of a simple EEG Brainnet to mark F3 could have been influenced by published consensus recommendations in 2018 [4] which were subsequently disseminated in national TMS training courses. There is no option in the registry database for sites to specify their targeting method, and while the Neurostar system provides default coordinates for a target 5.5 cm anterior to the MT location, often slight modifications (such as coil rotation) are made by the user to address patient discomfort. It is not known whether such user modifications were consistently entered into the system as updated coordinates at the time they were made, nor whether adjustments to the coordinates were consistently applied across subsequent treatment sessions. Future modifications to the registry software for collection of user data surrounding targeting method, adverse event management, and reasons for protocol selection or change would seem fruitful for parsing out these effects.

Conclusions

The NeuroStar® Advanced Therapy System Clinical Outcomes Registry collected prospective treatment parameter and treatment outcome information on a large sample of patients with MDD who received TMS therapy in private practice settings. Clinical outcomes were compared between patients treated with the Standard protocol (using an ITI of 26 s and a session duration of 37.5 min) and the Dash protocol (with a shortened ITI varying between 11 and 25 s, and a minimum session duration of 18.75 min). Both protocols yielded marked antidepressant effects without meaningful differences in efficacy. The strong efficacy obtained with the shorter ITI should enable clinician confidence in implementing the Dash protocol. Use of the more time efficient protocol allows treatment of a larger number of patients with a fixed number of devices and staff hours, and creates a more convenient treatment option for depressed patients.

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, MM, and LLC. LLC drafted the manuscript with input from STA and HAS, and with critical review and additional input by all authors. TMH, SP, and WSW participated in data collection and MM was responsible for establishing and coordinating the patient registry.

Declaration of competing interest

Dr. Carpenter serves as a scientific advisor to Neuronetics Inc, Nexstim PLC, Affect Neuro Inc, Neurelief 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.

Dr. Aaronson serves as a scientific advisor to Genomind Inc, LivaNova PLC, Neurotrends 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.

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

Ms. Verdoliva reports no financial relationship with commercial interests.

Dr. Sackeim serves as a scientific advisor to Cerebral Therapeutics Inc., 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).

Collection of registry data and statistical analyses for this study were funded by Neuronetics Inc.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2020.12.003.

References

[1] 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.
[2] Schlaepfer TE, George MS, Mayberg H, Stimulation WFtFoB. WFSBP guidelines on brain stimulation treatments in psychiatry. World J Biol Psychiatr 2010;11(1):2–18.
[3] Mutz J, Vipulananthan V, Carter B, Hurlemann R, Fu CHY, Young AH. 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 2019;364:j1079.
[4] McClintock SM, Reti IM, Carpenter LL, McDonald WM, Dubin M, Taylor SF, et al. Consensus recommendations for the clinical application of repetitive Transcranial Magnetic Stimulation (rTMS) in the treatment of depression. J Clin Psychiatr 2018;79(1).
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.

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.

Janicak PG, Nahas Z, Lisanby SH, Solvason HB, Sampson SM, McDonald WM, et al. Durability of clinical benefit with transcranial magnetic stimulation (TMS) in the treatment of pharmacoresistant major depression: assessment of relapse during a 6-month, multisite, open-label study. Brain Stimul 2010;3(4):187–99.

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 Psychiatry 2010;67(5):507–16.

George MS, Wassermann EM, Kimbrell TA, Little JT, Williams WE, Danielsen AL, et al. Mood improvement following daily left prefrontal repetitive transcranial magnetic stimulation in patients with depression: a placebo-controlled crossover trial. Am J Psychiatr 1997;154(12):1752–6.

Hadley D, Anderson BS, Borckardt JJ, Arana A, Li X, Nahas Z, et al. Safety, tolerability, and effectiveness of high doses of adjunctive daily left prefrontal repetitive transcranial magnetic stimulation for treatment-resistant depression in a clinical setting. J ECT 2011;27(1):18–25.

Chen R, Gerloff C, Classen J, Wassermann EM, Hallett M, Cohen LG. Safety of different inter-train intervals for repetitive transcranial magnetic stimulation and recommendations for safe ranges of stimulation parameters. Electroencephalogr Clin Neurophysiol 1997;105(6):415–21.

Rossi S, Hallett M, Rossini PM, Pascual-Leone A. Safety of TMS. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 2009;120(12):2008–39.

NeuroReonics Inc. Neurostar® TMS therapy system: 510(k) premarket notification. 2016.

Cash RFH, Dar A, Hui J, De Ruiter L, Baarbe J, Fettes P, et al. Influence of intertrain interval on the plastic effects of rTMS. Brain Stimul 2017;10(3):630–6.

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.

Kroeske K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606–13.

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.

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 2015;113:125–36.

Blumberger DM, Maller JJ, Thomson I, Mulsant BH, Rajji TK, Maher M, et al. Unilateral and bilateral MRI-targeted repetitive transcranial magnetic stimulation for treatment-resistant depression: a randomized controlled study. J Psychiatry Neurosci 2016;41(4):E38–66.

Zhang YQ, Zhu D, Zhou XY, Liu YY, Qin B, Ren GP, et al. Bilateral repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis of randomized controlled trials. Braz J Med Biol Res 2015;48(3):198–206.

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;391:1683–92. 10131.