Data supplement for Cole et al., Stanford Neuromodulation Therapy (SNT): A Double-Blind Randomized Controlled Trial. Am J Psychiatry (doi: 10.1176/appi.ajp.2021.20101429)

CONTENTS

Supplementary Methods

Recruitment details

Recruitment strategies included local clinic referrals, advertisement through social media, flyers and local events. Study participants were pre-screened online before being invited for an in-person screening visit to determine eligibility. Participants were excluded if they had a diagnosis of autism spectrum disorder, severe cognitive impairment, or epilepsy. Participants were also excluded if they presented with mania, psychosis, active suicidal behavior, a structural neurological lesion that could increase seizure risk or affect brain connectivity, or contraindications to either rTMS or MRI such as history of seizures, metallic implants, or severe insomnia (<4 hours sleep per night with hypnotic). Participants underwent a urine drug screen, and if female a urine pregnancy test, both of which were required to be negative.

Power analysis

The study was powered for an estimated effect size of Cohen's d=0.8 for change in Montgomery-Åsberg Depression Rating Scale (MADRS) between sham and active SNT groups. We set the probability of rejecting a true null hypothesis (Type 1 error, alpha) at 0.05 (two-sided). Thus, assuming independent groups and ~10% attrition, we aimed to recruit 30 participants in each treatment group, 60 total. Considering the vulnerable nature of this population, we planned to conduct an interim analysis after 30 participants completed the trial to assess for superiority, inferiority, or futility of the active compared to sham treatment. The primary outcome for the study and this planned interim analysis was proportional change from baseline MADRS 1-month following the end of treatment. The MADRS was selected as the primary outcome measure prior to the end of data collection when investigators were blinded to participants' treatment assignments, to allow comparison to emerging data demonstrating TMS site-specific reductions in depressive symptoms as measured by the MADRS(1). Proportional change was used rather than raw scores to be consistent with the previous blinded iTBS trial(2).

Randomization

The Cool-B65 A/P Butterfly coil (MagVenture MagPro; Farum, Denmark) device holds a blinded key code delivered from a computerized random number generator with a permuted block design ratio (1:1), to ensure balancing between arms. The randomization codes were sent directly from MagVenture to a study collaborator who did not take part in any study procedures beyond this role. The printed randomization codes for treatment group assignment were stored in a secure, locked cabinet and retrieved by the study coordinator at each baseline visit. The

study coordinator did not take part in data acquisition or analysis procedures. Participants were assigned to their randomization code during a baseline inclusion visit, and this code was available to the TMS operators in each participant's protected, online RedCap database chart. During the rTMS setup, the operators entered each participant's randomization code, and operators were instructed to rotate the coil to correspond with the key code.

MRI acquisition and processing information

All MRI scans were acquired using a 3-tesla GE Discovery MR750 scanner with a 32-channel imaging coil at the Center for Cognitive Neurobiological Imaging at Stanford, using a simultaneous multi-slice (SMS) imaging sequence with an acceleration factor of 3 (SMS=3), which collects 3 equally spaced axial slices simultaneously. A total of 29 sets of 3 slices, for a total of 87 slices, were collected within each repetition time of 2 seconds. During the 8-minute resting state scans, participants were instructed to let their minds wander, avoid repetitive thought, keep their eyes open and focus their attention on a central fixation point.

Statistical Parametric Mapping segmentations based on tissue probability maps were used to calculate estimation parameters to warp the T1-weighted structural images into Montreal Neurological Institute (MNI) space. These normalization parameters were inverted and applied to MNI space regions of interest (ROIs) for the left DLPFC (Brodmann area 46) and the sgACC (Brodmann area 25) to map these ROIs onto the individual participant's brain. The participant-space ROIs were then resliced, smoothed, and binarized to match the dimensions of the resting-state scans.

All analyses were conducted in a participant's own brain space. Statistical Parametric Mapping (SPM12) software was used for all pre-processing. Two separate algorithms were used to determine the individualized target location within the participant space left DLPFC ROI. The first algorithm sorted each of the left DLPFC and bilateral sgACC voxels into functional sub-units using a hierarchical agglomerative clustering algorithm. The voxel time series that most accurately reflected the median time series was created for each functional subunit, and the correlation coefficients were calculated between all selected time series extracted from all functional subunits of the left DLPFC and sgACC. The second algorithm determined the optimal left DLPFC subunit to target based on 3 factors: the net correlation/anticorrelation of the left DLPFC subunit with sgACC subunits, the size of the subunit and the spatial concentration of the subunit.

<u>Additional details</u>

Resting motor threshold (rMT) was determined at baseline and checked by 2 experienced rTMS operators prior to the first stimulation session. rMT was checked again if there were changes in medication schedule or sleep duration (>2 hours the night prior). All participants were offered mouth guards (but allowed to refuse it) for roughly the first half of this RCT and were required to wear mouth guards in the second half of this RCT after a participant chipped their tooth in a different trial running concurrently in the lab.

Study data were collected and stored using paper forms and scales for all assessments and questionnaires and then entered into a secure RedCAP database. Data were verified by 2 study coordinators who were otherwise not involved in data acquisition or analysis.

<u>Analysis of neurocognitive side effects</u>

Changes in standardized scores for Hopkins Verbal Learning Test - Revised (verbal immediate recall and delayed recall) and Delis-Kaplan Executive Function System Trail Making Test and Color Word subtests (processing speed and executive functioning) immediately following SNT were analyzed. List-wise deletion was used for cases of missing data. Linear mixed models were used for outcome measures for which the model produced normally-distributed residuals. Outcome measures for which a linear mixed model produced non-normal residuals were assessed with generalized linear mixed models (GLMM) that used a compound symmetry covariance structure, Satterthwaite approximation of degrees of freedom, and robust estimation of coefficients to handle violations of model assumptions. All mixed models assessed fixed effects of time (baseline vs. immediate post-treatment) and group (active vs. sham), as well as the time by group interaction. All post-hoc pairwise comparisons were Bonferroni-corrected.

Supplementary Results

GLMM results for secondary outcomes and last observation carried forward method

Compared to the primary outcome of mean %baseline MADRS scores, equivalent results were found for mean %baseline HDRS-17 scores (group: $F_{1,6}$=35.4, p=0.001, time: $F_{4,11}$=30.8, p<0.001, group*time: $F_{4,12}$=29.6, p<0.001, autoregressive covariance) and mean %baseline HDRS-6 scores (group: $F_{1,9}$=18.9, p=0.002, time: $F_{5,12}$=15.4, p<0.001, group*time: $F_{5,12}$=6.9, p=0.003, CS). Participants in the active group showed significantly greater %baseline post-SNT reductions in all scores at all follow-up time points (Bonferroni, p<0.05).

Results were also equivalent when missing data were imputed using the last observation carried forward method for the mean %baseline MADRS scores (group: $F_{1,20}$=27.2, p<0.001, time: $F_{5,30}$=15.3, p<0.001, group*time: $F_{5,30}$=6.4, p<0.001, CS), mean %baseline HDRS-17 scores (group: $F_{1,17}$=26.1, p<0.001, time: $F_{4,37}$=21.5, p<.001, group*time: $F_{4,37}$=8.9, p<0.001, AR), and mean %baseline HDRS-6 scores (group: $F_{1,15}$=21.9, p<0.001, time: $F_{5,23}$=15.7, p<0.001, group*time: $F_{5,23}$=8.3, p<0.001, CS). Participants in the active group showed significantly greater %baseline post-SNT reductions in all scores at all follow-up time points (Bonferroni, p<0.05).

There were 2 participants without pretreatment scores for the QIDS, 1 from each group. Without imputing missing data for the remaining participants, under an autoregressive covariance structure, there was a main effect of time ($F_{5,37}$=10.2, p<0.001) and treatment group ($F_{1,12}$=6.0, p=0.03), but the group*time interaction ($F_{5,37}$=1.7, p=0.17) did not reach significance for changes in mean %baseline QIDS scores. Participants in the active group displayed significantly greater %baseline reductions in QIDS scores at the immediate post-SNT as well as the 1-week post-SNT time points (Bonferroni, p<0.05). When the last observation carried forward method was used to account for missing data, there was a significant effect of treatment group ($F_{1,16}$=7.6, p=0.01) and time ($F_{5,37}$=9.8, p<0.001) but not a significant group*time interaction ($F_{5,37}$=2.3, p=0.06) on QIDS score change. Participants in the active group displayed significantly greater %baseline reductions in QIDS scores at the immediate post-SNT as well as the 1-, 2- and 4-week post-SNT time points (Bonferroni, p<0.05).

Generalized linear mixed model analysis of raw assessment scores revealed a significant effect of treatment on mean MADRS score change (group: $F_{1,23}$=32.8, p<0.001, time: $F_{5,53}$=13.3, p<0.001, group*time interaction: $F_{5,53}$=4.8, p=0.001, AR). Equivalent results were found for

mean HDRS-17 score change (group: $F_{1,14}$=33.8 p<0.001, time: $F_{4,30}$=21.8, p<0.001, group*time: $F_{4,30}$=16.1, p<0.001, AR), and mean HDRS-6 score change (group: $F_{1,22}$=22.0 p<0.001, time: $F_{5,55}$=16.4, p<0.001, group*time: $F_{5,55}$=7.7, p<0.001, AR). For mean QIDS score change under a compound symmetry covariance structure there was a significant effect of time $(F_{5,26}$=10.4, p<0.001) and treatment group ($F_{1,20}$=7.5, p=0.01) but no significant group*time interaction ($F_{5,26}$=1.5, p=0.22). Participants in the active group showed significantly greater post-SNT reductions in all scores at all follow-up time points for mean raw scores (Bonferroni, p<0.05).

Using the last observation carried forward method, generalized linear mixed model analysis of raw assessment scores revealed a significant effect of treatment on mean MADRS score change (group: $F_{1,22}$=34.2, p<0.001, time: $F_{5,54}$=12.8, p<0.001, group*time: $F_{5,54}$<5.3, p<0.001, AR). Equivalent results were found for mean HDRS-17 score change (group: $F_{1,14}$=32.7 p<0.001, time: $F_{4,22}$=19.0, p<0.001, group*time: $F_{4,22}$=7.7, p=0.001, AR) and mean HDRS-6 score change (group: $F_{1,23}$=24.0, p<0.001, time: $F_{5,58}$=17.2, p<0.001, time*group: $F_{5,58}$=9.0, p<0.001, AR). For mean QIDS score change under a compound symmetry covariance structure there was a significant effect of group ($F_{1,16}$=9.1, p=0.01) and time ($F_{5,30}$=8.7, p<0.001) but the group*time interaction ($F_{5,30}$=2.2, p=0.09) did not meet significance. Participants in the active group showed significantly greater post-SNT reductions in all scores at all follow-up time points for mean raw scores (Bonferroni, p<0.05).

Supplementary Tables

TABLE S1. Side effects

Type	SNT Group	
	Active (n=14)	Sham (n=15)
Fatigue	57% (8)	53% (8)
Neck/Back discomfort	50% (7)	33% (5)
Discomfort at treatment site	36% (5)	27% (2)
Post-SNT headache	57% (8)	13% (2)
Nausea	0% (0)	0% (0)
Anxiety	29% (4)	20% (3)
Dental Issues	7% (1)	0% (0)
Jaw discomfort	14% (2)	0% (0)
Other	7% (1)	0% (0)

Any other spontaneous effects reported during SNT were included under "Other." One participant in the active SNT group experienced polydipsia during SNT, with no prior history.

TABLE S2. Numbers and percentages of medications taken by participants during SNT

Medication category	Patients taking specified medication by group, n(%)		
	Active (n=14)	Sham (n=15)	Overall (n=29)
SSRI	5 (36)	2 (13)	7 (24)
SNRI	4 (29)	2 (13)	6 (21)
MAOI	0 (0)	0 (0)	0 (0)
TCA	1 (7)	0 (0)	1 (7)
Atypical Antidepressant	6 (43)	7 (47)	13 (45)
Atypical Antipsychotic	4 (29)	0 (0)	4 (14)
Anti-Epileptic	3 (21)	3 (20)	6 (21)
Benzodiazepine	7 (50)	6 (40)	13 (45)
Psychostimulant	2 (14)	1 (7)	3 (10)
Lithium	3 (21)	1 (7)	4 (14)
Other	2 (14)	0 (0)	2 (7)
No Concurrent Psychotropic Medication	1 (7)	4 (27)	5 (17)

Examples of each medication category are: SSRI: Fluoxetine (Prozac) and Escitalopram (Lexapro); SNRI: Desvenlafaxine (Pristiq); MAOI, N/A; TCA, Amitriptyline (Elavil); Atypical Antidepressant, Bupropion (Wellbutrin/Wellbutrin SR) and Mirtazapine (Remeron); Atypical antipsychotic: Quetiapine (Seroquel) and Aripiprazole (Abilify); Anti-Epileptic: Gabapentin (Neurontin) and Lamotrigine (Lamictal); Benzodiazepine, Lorazepam (Ativan) and Clonazepam (Klonopin); Psychostimulant, Modafinil (Provigil), Methylphenidate (Concerta, Ritalin, Methylin); Lithium, Lithium; and Other: Testosterone, Estradiol, Progesterone and Melatonin

TABLE S3. Psychiatric comorbidities

Comorbid psychiatric condition	Active (n=14)		Sham (n=15)		
	n	%	n	%	p
Anxiety	3	21	6	40	0.29
ADHD	1	7	1	7	0.96
PTSD	1	7	1	7	0.96
Fibromyalgia	1	7	1	7	0.96
Substance use disorder (in remission)	0	0	3	20	0.08
Eating disorder	1	7	0	0	0.34
No. participants with comorbidities	5	36	8	53	0.20

TABLE S4. Neurocognitive testing results

Test	N Active: Sham	Time	Treatment Group	Interaction	Post-hoc testing Bonferroni-corrected
Color-Word Interference Condition 2: Word Reading Standardized Scores	10:13	$F_{1,21}=1.83$ p=0.19	$F_{1,21}=1.15$ p=0.30	$F_{1,21}=4.70$ p=0.04*	Pre v. post Active p=0.03 Sham p=0.54 Sham v. Active Pre p=0.11 Post p=0.72
Color-Word Interference Condition 3: Inhibition Standardized Scores[#]	10:13	$F_{1,8}=5.60$ p=0.31	$F_{1,1}=2.07$ p=0.19	$F_{1,1}=1.65$ p=0.47	NA
Color-Word Interference Condition 4: Inhibition Switching Standardized Scores[#]	10:13	$F_{1,10}=14.74$ p<0.01*	$F_{1,15}=1.58$ p=0.23	$F_{1,10}=2.61$ p=0.14	Pre v. post Active p<0.01 Sham p=0.10 Sham v. Active Pre p=0.14 Post p=0.49
Hopkins Verbal Learning Delayed Recall T scores[#]	10:13	$F_{1,9}=2.89$ p=0.125	$F_{1,19}=0.58$ p=0.45	$F_{1,9}=0.42$ p=0.535	NA
Hopkins Verbal Learning Immediate Recall T scores	10:13	$F_{1,21}=0.60$ p=0.45	$F_{1,21}<0.01$ p=0.97	$F_{1,21}=1.2$ p=0.29	NA
Trail Making Test Condition 2: Number Sequence Standardized Scores	10:13	$F_{1,21}=11.60$ p<0.01*	$F_{1,21}=0.47$ p=0.50	$F_{1,21}=0.57$ p=0.46	Pre v. post Active p=0.01 Sham p=0.06 Sham v. Active Pre p=0.32 Post p=0.98
Trail Making Test Condition 4: Switching Standardized Scores[#]	10:13	$F_{1,4}=1.74$ p=0.25	$F_{1,4}=1.97$ p=0.23	$F_{1,4}<0.01$ p=0.99	NA
Trail Making Test Condition 5: Motor Speed Standardized Scores	10:13	$F_{1,21}=5.32$ p=0.03*	$F_{1,21}=0.10$ p=0.76	$F_{1,21}=0.09$ p=0.77	Pre v. post Active p=0.10 Sham p=0.14 Sham v. Active Pre p=0.87 Post p=0.64

#Generalized Model

TABLE S5. Response and remission rates at each time point

	Immediate-post	1 week	2 weeks	3 weeks	4 weeks	Any week of follow-up
Active group						
Response rate	71.4%	71.4%	78.6%	64.3%	64.3%	85.7%
Remission rate	57.1%	64.3%	50.0%	57.1%	42.9%	78.6%
Sham group						
Response rate	13.3%	13.3%	6.7%	6.7%	6.7%	26.7%
Remission rate	0%	6.7%	6.7%	6.7%	0%	13.3%

Supplementary Figures
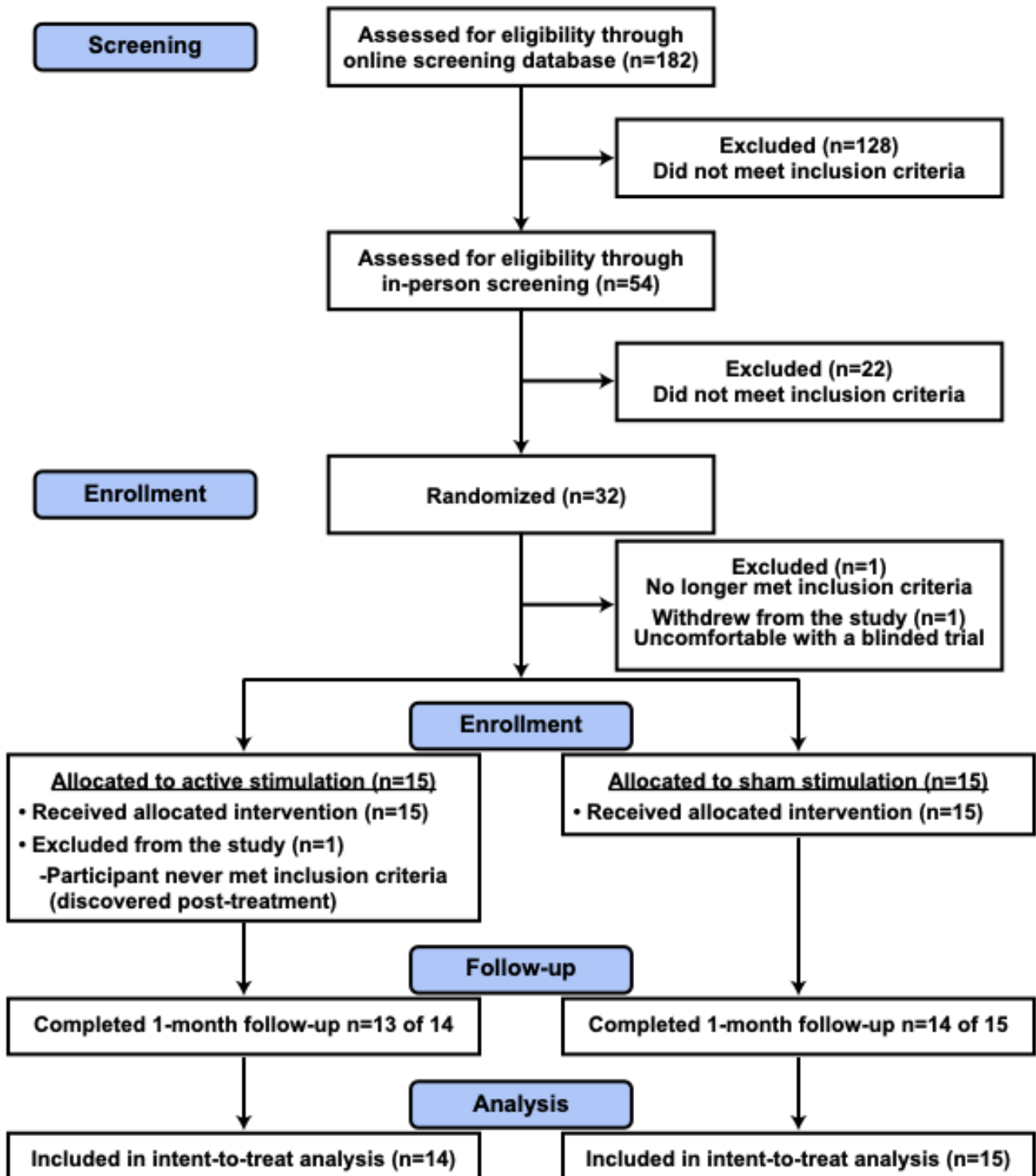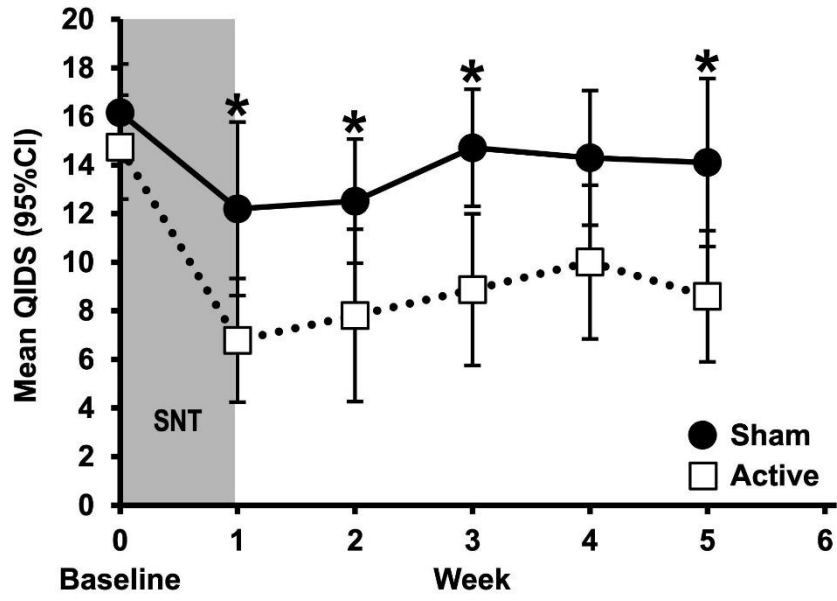
FIGURE S1. CONSORT diagram

FIGURE S2. Mean Quick Inventory of Depressive Symptoms (QIDS) score

References

1. Siddiqi S, Taylor S, Cooke D, et al.: Distinct symptom-specific treatment targets for antidepressant neuromodulation. Brain Stimul Basic Transl Clin Res Neuromodulation 2019; 12:537

2. Blumberger DM, Vila-Rodriguez F, Thorpe KE, et al.: Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. The Lancet 2018; 391:1683–1692