Supplementary material

SAINT protocol development:

Precise targeting to L-DLPFC-sgACC: Both the L-DLPFC and the sgACC have been shown to be dysfunctional in MDD; L-DLPFC has been reliably found to be hypoactive [96–98] and the sgACC is consistently found to be hyperactive in MDD patients [99–101]. In addition to the distinct dysfunction in both the L-DLPFC and the sgACC, resting-state fMRI data suggest that MDD patients exhibit reduced anti-correlation between these two regions, thought to reflect reduced indirect inhibitory control of the sgACC from the L-DLPFC [15]. rTMS delivered to the L-DLPFC has been shown to normalize hypoactivity in L-DLPFC [102,103], reduce sgACC hyperactivity [104,105] and strengthen indirect inhibitory connections between the L-DLPFC and sgACC [15,16]. Neuroimaging data suggest the higher the anti-correlation between the stimulated region of the L-DLPFC and the sgACC, the better the clinical outcome [15]. This finding has since been confirmed prospectively in other cohorts [25,106]. We selected BA46 as this area has been shown to have the highest average anti-correlation to the sgACC [15]. A recent TMS-fMRI study used resting-state fMRI to target the region of the L-DLPFC which showed greatest functional connectivity with the sgACC and showed stimulation propagated to the sgACC in all participants [69]. In comparison, another study which targeted L-DLPFC using anatomical MRI coordinates (the border of BA9 and BA46), stimulation propagated to the sgACC in only 44% of participants [70]. These data suggest that using resting-state fMRI to identify and stimulate the region of the L-DLPFC which is most anti-correlated with the sgACC in each individual could increase the efficacy of TMS protocols (rTMS/iTBS).

Accelerated delivery: Multiple spaced stimulation sessions appear to produce non-linear additive effects. Basic neuroscience research conducted using hippocampal brain slices has shown that multiple iTBS sessions have a cumulative effect on dendritic spine enlargement [20–22]. Studies which have applied theta-burst stimulation protocols (iTBS/cTBS) to the motor cortex in humans have also shown that two spaced stimulation sessions produce longer lasting changes in cortical excitability than single stimulation sessions [61,63]. Improvements in visual perception after one session of continuous theta-burst delivered to the left posterior parietal cortex in individuals with hemi-spatial neglect last 30-40 minutes, two spaced applications last 3-8 hours, 4 sessions last 32 hours, and 8 sessions last 3 week [107,108]

1800 pulses per session: 1800 pulses per session was chosen as it is the only pulse dose that has been explored in a blinded iTBS trial [71]. Additionally, 1800 pulses has been shown to produce long-lasting changes in cortical excitability [109] and optimally produce intended cellular changes [110].

Overall Dose: 18,000 pulses were applied each day of the SAINT protocol to match the number of pulses of a six week FDA-approved iTBS protocol [18]. In total, across the five consecutive days of stimulation, 90,000 pulses were used to match the total number of pulses in a 6-week standard rTMS course [111,112]. Given that 90,000 iTBS pulses equates to 5X the standard rTMS dose, this approach allowed for a dose-response curve to be developed.

Inter-session interval: Stimulation sessions were delivered hourly (50-minute inter-session interval) based upon the evidence from basic neuroscience research showing that iTBS sessions delivered to hippocampal slices with inter-session intervals of 50-90 minutes have a cumulative effect on dendritic spine enlargement, a process involved in synaptic strengthening. In comparison, iTBS sessions delivered with inter-session intervals of 40 minutes or less do not have a cumulative effect on dendritic spines [20–22]. Two iTBS sessions delivered to the prefrontal

cortex [64] or the motor cortex [62] with a 15-minute inter-session interval have been shown not to increase cortical excitability further than a single iTBS session.

Stimulation intensity: Stimulation was delivered at 90% resting motor threshold as it has been demonstrated that theta-burst stimulation applied <100% rMT produces the optimal change in prefrontal cortical excitability [30,31]. A depth correction [32] was applied to the resting motor threshold to adjust for difference in the cortical depth of the individual's functional target compared to the primary motor cortex in order to consistently achieve 90% rMT in the intended target.

Detailed information about the neuropsychological test battery:

Verbal learning and memory
The Hopkins Verbal Learning Test – Revised (HVLT-R) was given to assess learning and recall of verbal information. The HVLT-R [39] is a list-learning task with three learning trials, a 20-minute delayed recall, and a recognition paradigm following the delayed recall. There are six alternate forms that allow for serial evaluation.

Visuospatial learning and memory
The Brief Visuospatial Memory Test – Revised (BVMT-R) was administered to measure learning and memory of visuospatial stimuli. The BVMT-R [38] is a task that requires the participant to learn an array of simple geometric figures over three learning trials. There is a delayed recall after 25 minutes and a recognition task following the delay. There is also a copy task following the memory recall and recognition portions of the test. There are six alternate forms that allow for serial evaluation.

Executive Functioning
Subtests of the Delis Kaplan Executive Function System (D-KEFS)[40] were used to assess executive functioning. The tests used from the D-KEFS were the Trail Making Test, Color-Word Interference, Verbal Fluency, and the Tower Test. The Trail Making Test (five conditions) was used to measure combined visuomotor and executive functioning including sequencing and cognitive switching. The test also provides measures of visual scanning and motor speed. The Color-Word Interference test (four conditions) provides a measure of cognitive inhibition and cognitive switching. There are also word reading and color naming trials that provide a measure of reading and color naming speed. The Verbal Fluency test (three conditions) provides measures of phonemic and semantic fluency, as well as a trial that involves cognitive switching combined with semantic fluency. The Tower Test provides a measure of learning, planning, and problem solving. This test involves building towers matching pictured models, according to a set of rules that must be followed. The test taker must try to build the tower with the fewest number of moves possible.

Attention and Working Memory
The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Fourth Edition[40])[41] Digit Span subtest provides a measure of simple attention and working memory. The test requires the test taker to recall strings of numbers and then manipulate them into backward order and sequenced orders.

Processing Speed
The Wechsler Adult Intelligence Scale- Fourth Edition ((WAIS-IV; Fourth Edition[40])[41] Coding and Symbol Search subtests combine to provide the Processing Speed Index score. These timed

subtests provide a measure of processing speed and visual perception. The Coding subtest requires the participant to fill in individual symbols assigned to represent each number. There is a component of learning that would improve the participant's speed and efficiency of test performance. Symbol Search requires the participant to determine if one of two figures is present within a group of symbols.

Resting-state fMRI scan details

Participants were instructed to keep their eyes open and their attention focused on a central fixation point, which consisted of a black screen with a white fixation cross. Participants were also instructed to let their minds wander freely and to avoid any repetitive thoughts such as counting. Scans were collected with a 3X simultaneous multi-slice (i.e. multiband) acquisition echo planar imaging (EPI) sequence, TR=2000 ms, TE=30 ms, flip angle=77°, FOV=230x230, 128×128 voxel matrix, 1.8×1.8 mm2 in-plane resolution, 87 horizontal slices, yielding >1.4M voxels every 2 seconds. A structural anatomical 3D T1-weighted scan was also collected with a 0.9 mm3 voxel size with 256x256x176 voxel dimensions. Head motion of participants was effectively minimized with the use of memory foam and inflatable padding. Participant alertness during the resting state task was monitored using in-scanner video cameras.

Targeting algorithms

The first algorithm used the Spearman correlation coefficient between voxel time series as the linkage measurement. Functional sub-units were defined as all voxel pairs being correlated with each other with a Spearman's correlation coefficient of rho=0.5 or above. Each participant's DLPFC was subdivided into a number of functional subunits. This same process was repeated for the sgACC, with the hierarchical agglomerative clustering algorithm defining the size and shape of each functional subunit based on the correlation coefficient patterns within the sgACC. For each functional subunit in the DLPFC and each subunit in the sgACC, a single time-series value was created by finding the single voxel time series that was most correlated with the median time series of the subunit. Once a single time series was identified for each subunit, we then calculated all of the Spearman correlation coefficients were calculated between the DLPFC functional subunits and sgACC subunits to form a correlation matrix. A second algorithm was then deployed to choose the optimal L-DLPFC subunit. The decision-making algorithm considers the net correlation/anti-correlation amount for each L-DLPFC subunit with the sgACC. This is value is calculated using the sum of all the correlation coefficients multiplied by all the sizes of the sgACC subunits. The decision-making algorithm also considers the size of the L-DLPFC subunit (larger clusters are easier to target) and the spatial concentration of voxels that make up the subunit. The spatial concentration value was generated by calculating the average of all the 3D Euclidean distances between each of the voxels that make up a subunit and dividing the subunit voxels size by the Euclidean distance measure. The net correlation/ anti-correlation amount, the L-DLPFC subunit size, and the spatial concentration of the subunits were the 3 parameters contributing to the decision-making algorithm.

Supplementary Tables

Supplementary Table 1: Current medications

Participant	Current Medications
1	Fluoxetine (PROZAC) 20mg, amphetamine salt 20mg, quetiapine (SEROQUEL) 25mg, zolpidem (AMBIEN) 6.25mg
2	None
3	Dextroamphetamine- amphetamine (ADDERALL XR) 30mg, alprazolam 0.25-0.5mg/day, vortioxetine (TRINTELLIX) 20mg, zolpidem 10mg, propranolol 30mg
4	Quetiapine (SEROQUEL) 100 mg/night
5	Fluoxetine (PROZAC) 40mg
6	Clonazepam (KLONOPIN) 1mg, zolpidem (AMBIEN) 10mg, tazodone 10mg, levothyroxine (SYNTHROID) 50mcg, lurasidone (LATUDA) 60mg, pregabalin (LYRICA) 50mg, modafinil (PROVIGIL) 400mg, lisdexamfetamine (VYVANSE) 60mg, amitriptyline 50mg
7	Lamotrigine (LAMICTAL) 200 mg, carbidopa-levodopa 50 mg
8	Amitriptyline 100mg, lamotrigine 150mg, dextroamphetamine 45mg, levothyroxine 100mcg, liothyronine sodium 10mg
9	Clonazepam (KLONOPIN) 0.5 mg, dextroamphetamine (DEXEDRINE SPANSULE) 10 mg, levomilnacipran (FETZIMA) 40 mg, mirtazapine (REMERON) 30 mg, olanzapine (ZYPREXA) 5 mg, levothyroxine (LEVOXYL) 100 mcg, montelukast (SINGULAIR) 10
10	Duloxetine 120mg, amlodipine 5mg, liothyronine 25mcg, cyanocobalamin 1000mcg, folic acid 400 mcg, zolpidem 10 mg (every other night)
11	Lamotrigine (LAMICTAL) 100 mg, pramipexole 200 mg, dextroamphetamine- amphetamine (ADDERALL XR) 20 mg, clonazepam (KLONOPIN) 1 mg
12	Fluoxetine (PROZAC) 20mg, bupropion (WELLBUTRIN) 300mg
13	Lamotrigine (LAMICTAL) 150mg, aripiprazole (ABILIFY) 2mg
14	Lamotrigine 200mg, levothyroxine 100mcg, liothyronine 10mcg
15	Bupropion (WELLBUTRIN) 522mg, dextroamphetamine- amphetamine (ADDERALL), Lisinopril
16	Duloxetine (CYMBALTA) 120mg, alprazolam (XANAX) 2mg
17	Lamotrigine (LAMICTAL) 175mg
18	Divalproex Sodium (DEPAKOTE) 500mg, lamotrigine (LAMICTAL) 125mg, gabapentin 300mg, quetiapine (SEROQUEL) 25mg, Armour Thyroid 60 mg, diazepam 10mg, lorazapam 5mg, pregabalin 1050mg, valacyclovir (VALTREX) 500mg, tramadol (ULTRAM) 100mg
19	None
20	Venlafaxine 225mg, risperidone .5mg, pantoprazole 20mg
21	None
22	Bupropion (WELLBUTRIN) 400mg, hydroxyzine (ATARAX) 50mg, flexeril 10mg, doxylamine succinate 25mg, estradiol .0375, tirosint 50mcg, cytomel 5mcg,

	moxifloxacin HCL 400mg, ondansetron 8mg, oxycodone HCL 10-650mg, propranolol HCL 10mg
23	Fluoxetine 20mg, atorvastatin 20mg, diazepam 5mg-15mg
24	Carbidopa-Levodopa 25- 100mg, carbidopa 25mg, hydrochlorothiazide 25 mg, primidone 50mg, rosuvastatin 5mg, gabapentin 100mg, clonazepam PRN
25	Quetiapine (SEROQUEL) 300mg, escitalopram (LEXAPRO) 20mg
26	Bupropion (WELLBUTRIN) 450mg, dextroamphetamine- amphetamine (ADDERALL) 10mg, vortioxetine (TRINTELLIX) 20mg, ramelteon (ROZEREM) 8mg, zolpidem (AMBIEN) 6.25mg, atorvastatin (LIPITOR) 10mg, hydrochlorothiazide (HYDRODIURIL) 0.5mg, lisinopril (PRINIVIL) pantoprazole (PROTONIX) 40mg, acetaminophen (TYLENOL) 500-1000mg, cetirizine (ZYRTEC fluticasone propionate (FLONASE) 100 mcg, melatonin 5mg
27	Lamotrigine (LAMICTAL) 400mg, gabapentin 2000mg, clonazepam 0.5mg, hydroxozine 25mg, Carbidopa-Levodopa 250mg, rasagiline 1mg, dexlansoprazole (DEXILANT) 30mg
28	Vortioxetine (TRINTELLIX) 20mg, alprazolam (XANAX .25mg, simvastatin 10mg, oestrodiol .1mg
29	Duloxetine (CYMBALTA) 60mg, aripiprazole (ABILIFY) 5mg, clonazepam 2mg
30	None
31	Sertraline (ZOLOFT) 200 mg, Bupropion (WELLBUTRIN) 100 mg, clonazepam (KLONOPIN) 1.5 atorvastatin 10 mg, celecoxib 200 mg

Supplementary Table 2: Demographic information and treatment history for the MDD subsample, n=22

Participant info	Mean (SD)
Gender (male:female)	10:12
Age	48.36 (15.98)
Age of MDD onset	21.95 (8.90)
Duration of MDD	26.41 (15.48)
Number of antidepressant failures (lifetime)[1]	8.77 (5.51)
Number of adjunctive medications (lifetime)[2]	1.36 (1.47)
Number of participants attempted rTMS	11*
Number of participants attempted ECT	4**
Maudsley Staging Method Score	11.32 (2.12)

[1]Adequate antidepressant trials defined as a minimum of 6 weeks at an appropriate dose according to the Anti-depressant Treatment History Form (ATHF).
[2]Medications defined as adequate augmentation strategies according to the ATHF.
[*]1 remitter, all other participants did not respond.
[**]All participants were non-responders.

Supplementary Table 3: Clinical assessment scores for MDD participants, n=22; mean (SD)

	Pre-SAINT	Post-SAINT	Responders (%)[1]	Remission (%)[2]
MADRS	36.36 (6.82)	4.82 (5.45)	90.91	90.91
HAMD-17	27.14 (5.61)	4.36 (4.22)	90.91	81.82
HAMD-6	14.77 (2.96)	2.50 (3.04)	86.36	81.82
BDI[3]	32.59 (11.21)	6.90 (7.08)	88.24	85.00
C-SSRS[4]	1.632 (1.34)	.00 (.00)	100.00	100.00
HAMD-Q3	1.50 (.96)	.00 (.00)	100.00	100.00
MADRS-Q10	2.32 (.99)	.00 (.00)	100.00	100.00

[1]Response defined as >50% reduction in score.
[2]Remission defined <8 on HAMD-17 [48]<5 on the HAMD-6 [49], ≤10 on MADRS [43], BDI<13 [50] and C-SSRS=0 [51].
[3]n=20, 3 participants only had a post-SAINT BDI score so for remission % calculation n=20, for responder % n=17.
[4]Current suicidal ideation subscale n=18, one participant only had a post-SAINT score so for % remission calculation n=18, for % responder n=17

Supplementary Table 4: Clinical assessment scores one-month post-SAINT for all participants n=31; mean (SD)

	Pre-SAINT	Month post-SAINT	Responders (%)[1]	Remission (%)[2]
MADRS	37.71 (7.24)	14.23 (14.68)	60.00	53.33
HAMD-17	27.87 (5.23)	10.93 (10.35)	63.33	53.33
HAMD-6	15.23 (2.80)	6.07 (5.74)	60.00	53.33
BDI[3]	32.61 (12.37)	16.24 (15.58)	50.00	56.00
C-SSRS[4]	1.52 (1.29)	0.09 (.42)	95.24	95.65
HAMD-Q3	1.52 (.85)	0.2 (.48)	93.33	83.33
MADRS-Q10	2.39 (.99)	.57 (.94)	86.67	66.67

[1]Response defined as >50% reduction in score.
[2]Remission defined <8 on HAMD-17 [48]<5 on the HAMD-6 [49], ≤10 on MADRS [43], BDI<13 [50] and C-SSRS=0 [51].
[3]n=25, three participants only had post-SAINT scores so for % responders n=22 and for % remission n=25.
[4]Current suicidal ideation subscale, n=23 and two participants only had a post-SAINT score (n=21 for % responders whereas n=23 and for % remission).

Supplementary Table 5: Clinical assessment scores at one-month post-SAINT for MDD participants n=22; mean (SD)

	Pre-SAINT	Month post-SAINT	Responders (%)[1]	Remission (%)[2]
MADRS	37.71 (7.24)	16.00 (15.13)	54.55	45.45
HAMD-17	27.87 (5.23)	11.55 (10.27)	59.09	50.00
HAMD-6	15.23 (2.80)	6.50 (5.78)	59.09	50.00
BDI[3]	32.61 (12.37)	19.05 (16.46)	41.18	47.37
C-SSRS[4]	1.52 (1.29)	0.12 (.49)	93.75	94.12
HAMD-Q3	1.52 (.85)	.27 (.55)	90.91	77.27
MADRS-Q10	2.39 (.99)	.73 (1.05)	81.82	59.09

[1]Response defined as >50% reduction in score.
[2]Remission defined <8 on HAMD-17 [48]<5 on the HAMD-6 [49], ≤10 on MADRS [43], BDI<13 [50] and C-SSRS=0 [51].
[3]n=19, 2 participants only had post-SAINT scores so for % responders n=17 and for % remission n=19.
[4]Current suicidal ideation subscale, n=17, one participant only had post-SAINT scores so for % responders n=17 and for % remission n=18).

Supplementary Table 6: Clinical assessment scores at one-month post-SAINT for all participants excluding rTMS non-responders, n=18; mean (SD)

	Pre-SAINT	Month post-SAINT	Responders (%)[1]	Remission (%)[2]
MADRS	37.39 (7.91)	7.33 (10.23)	83.33	72.22
HAMD-17	27.44 (5.56)	6.55 (8.54)	83.33	72.22
HAMD-6	15.17 (2.96)	3.44 (4.54)	83.33	72.22
BDI[3]	30.27 (11.42)	8.31 (9.67)	80.00	76.92
C-SSRS[4]	1.54 (1.27)	.00 (.00)	100.00	100.00
HAMD-Q3	1.78 (.81)	.06 (.24)	100.00	94.44
MADRS-Q10	2.56 (1.10)	.33 (.69)	94.44	77.78

[1]Response defined as >50% reduction in score.
[2]Remission defined <8 on HAMD-17 [48], <5 on the HAMD-6 [49], ≤10 on MADRS [43], BDI<13 [50] and C-SSRS=0 [51].
[3]n=13, three participants only had post-SAINT scores so for % responders n=10 and for % remission n=13.
[4]Current suicidal ideation subscale, n=14, two participants only had post-SAINT scores (n=12 for % responders whereas n=14 for % remission).

Supplementary Table 7: Clinical assessment scores at one-month post-SAINT for MDD participants excluding rTMS non-responders, n=12; mean (SD)

	Pre-SAINT	Month post-SAINT	Responders (%)[1]	Remission (%)[2]
MADRS	34.92 (6.21)	7.17 (8.27)	83.33	66.67
HAMD-17	26.17 (5.65)	5.92 (6.75)	83.33	76.92
HAMD-6	14.33 (2.96)	2.92 (3.53)	91.67	75.00
BDI[3]	31.57 (8.02)	10.33 (11.05)	71.43	66.67
C-SSRS[4]	1.67 (1.41)	.00 (.00)	100.00	100.00
HAMD-Q3	1.75 (.97)	.08 (.29)	100.00	91.67
MADRS-Q10	2.41 (1.08)	.50 (.80)	91.67	66.67

[1]Response defined as >50% reduction in score.
[2]Remission defined <8 on HAMD-17 [48] <5 on the HAMD-6 [49], ≤10 on MADRS [43], BDI<13 [50] and C-SSRS=0 [51].
[3]n=9, two participants only had post-SAINT scores so for % responders n=7 and for % remission n=9.
[4]Current suicidal ideation subscale, n=9, one participant only had post-SAINT scores (n=8 for % responders whereas n=9 for % remission).