Transcranial Magnetic Stimulation of Wernicke's and Right Homologous Sites to Curtail "Voices:" A Randomized Trial

Supplemental Information

Deriving Talairach Coordinates for Sites, W and rW

Talairach coordinates for W and rW were computed from various sites in Wernicke's area and the right homologue in different hallucinating subjects showing improvements per Hallucination Change Scale (HCS) when targeted with repetitive magnetic stimulation (rTMS) in our 2007 *Cerebral Cortex* study (1). In that study, target sites were delineated according to patient-specific maps of functional magnetic resonance imaging (fMRI) activation and functional connectivity during hallucination events. Sites in Wernicke's area, and to a lesser extent, the right Wernicke's homologue, appeared most likely to produce improved auditory/verbal hallucinations (AVHs) following rTMS. For the current study, weighted averages of Talairach coordinates were computed across all patient-sites in the 2007 study within these two regions that yielded improvement in AVHs when targeted with rTMS. Weights were based on observed rate of improvement computed for each patient-specific site delineated in the earlier study. For instance, the Y-coordinate for W was computed as follows:

$$\left(\sum_{i=1}^{N} HCSR(i) * YCorr(i)\right) \Big/ \sum_{i=1}^{N} HCSR(i) \qquad (1)$$

where $N = 6$ was the number of patient-sites showing improvement in Wernicke's area in the 2007 study, $HCSR(i)$ is the rate of improvement detected in that study for patient-site, i, measured according to HCS, i.e., $HCSR(i) = (HCS_{baseline} - HCS_{final})/(\#\text{ stim sessions} * HCS_{baseline})$, and Ycorr(i) is the Y-coordinate of the i-th site. Other coordinates for W and rW were computed similarly.

Subject-specific variation in gyral folding was considered in pinpointing W and rW in order to retain location relative to individualized cortical surface topology per the two Talairach atlas locations. The derived W-coordinate was (-65, -41, 9). The Talairach atlas shows this site as located along the posterior aspect of an inferior-directed gyral fold located in the posterior superior temporal gyrus (STG; Brodmann Area 22, see [2], p. 57). Equivalent sites in this gyral fold were located for each individual patient. The right-sided site (rW) per the Talairach atlas

(+65,-38, 11) was located on a patient-specific basis in a slightly more anterior location in the right homologue of this fold. We estimate that W was within 1.0 cm radius of previous fMRI sites in Wernicke's area showing improvement, whereas rW was within 1.6 cm radius of previous fMRI sites in the right homologue of Wernicke's area showing improvement (see Figures 4 and 5 in [1]).

Adjusting Stimulation Strength Delivered to W/rW Based on Distance to Cortical Surface

Stimulation strength was adjusted upward for rTMS/sham if a given site (either W or rW) exhibited distance from skin-surface-to-cortical-surface that exceeded the equivalent distance measurement for the ipsilateral motor cortex at site of motor threshold determination. This adjustment, which reflected the fact that magnetic field strength falls off exponentially relative to distance to the center of the coil, used the following formula (personal communication, D. Bohning, Department of Radiology, Medical University of South Carolina):

$$\textit{Adjusted Stimulation Strength} = 0.90 * MT * e^{0.036*(dc-dm)} \qquad (2)$$

where MT = motor threshold, dc = distance between scalp-to-cortical-surface measured (in mm) for the target region, and dm = scalp-to-cortical-surface distance (in mm) for the ipsilateral motor cortex surface. Figure S1 shows this adjustment graphically:

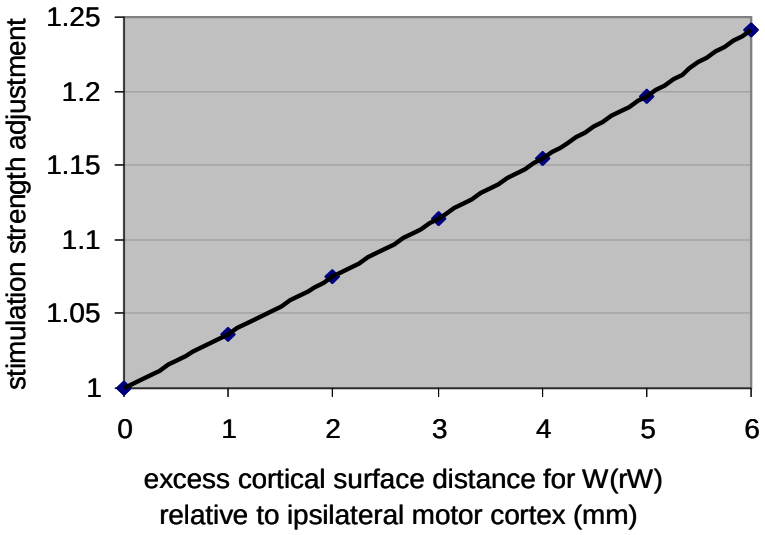

Figure S1. X-axis reflects relatively greater scalp-to-cortical-surface distance for the target site (W or rW) compared to motor cortex corresponding to (dc-dm) in equation 2 above. One unit on the Y-axis corresponds to unadjusted stimulation strength = 0.9 x motor threshold. If scalp-to-cortical-surface distance for W/rW compared to motor cortex was less than zero, the minimum stimulation strength of 0.9 x motor threshold was still delivered.

Statistics

Power Analysis for Primary Outcome Variable (HCS)

Estimated effect size for improved HCS scores following active rTMS using a stereotactically determined site in Wernicke's area ("W") relative to sham stimulation was based in part on our single-site rTMS clinical trial targeting a left temporoparietal site (TP3, approximately 2 cm distant from W) derived from the International 10/20 electrode positioning method (3). In that study of 50 subjects, an effect size estimate for HCS was calculated to be 0.83 using last observation carried forward analysis, and 0.81 using slope estimates per a mixed model. Our second estimate was based on our fMRI-guided rTMS study (1) yielding an effect size of 0.79 for rTMS delivered to W and .67 for rTMS delivered to our right Wernicke's homologous site ("rW"). These estimates were based on % improvement in HCS scores following rTMS delivered to all sites for left and right temporoparietal rTMS averaged over total

days of stimulation for each. A comparable % improvement statistic was calculated for the sham condition that was used to calculate effect size.

Randomized blocks of stimulation in the current trial were limited to 5 sessions. However, our fMRI-guided trial (1) averaging only 3.7 sessions per Wernicke's site, yielded an estimated effect size of 0.79. Thus our estimated power at the end of the 5 sessions for W, assuming 30 subjects receiving active and 30 subjects receiving sham stimulation, was then 0.82 based on Cohen (4). We were less confident regarding rTMS response delivered to rW given the lower estimated effect size in our fMRI-guided rTMS study (d = 0.67). However, this site was included given that some of the subjects in the fMRI-guided rTMS study (1) demonstrated robust clinical improvement associated with rTMS to right Wernicke's regions; moreover, our neuroimaging data (5) strongly implicated this region in the genesis of AVHs. In terms of endpoint HCS, our primary outcome variable - given that our protocol contrasted responses to W versus rW within the same subjects and added 5 more sessions delivered to the best site - we projected that effect sizes for the full 15-session protocol including all subjects (60 active versus 30 sham) would at least match that estimated for W alone described above and would demonstrate effect sizes equal to or exceeding endpoint effect sizes for our earlier TP3 controlled trial (3), which estimated power to be in the 0.90 for the entire sample.

Secondary variables used to assess rTMS effects included Clinical Global Improvement (CGI) score and change in hallucination frequency. A comparison of CGI scores and slope change in hallucination frequency for active vs. sham patients in our earlier left temporoparietal rTMS study of 50 patients (3) generated effect sizes of 1.09 and .94 respectively, suggesting that we would have power well in excess of 0.90 to detect group differences for these variables assuming 60 patients in the active and 30 patients in the sham group. These variables were not assessed in our fMRI-guided trial (1). The effect size for total Auditory Hallucination Rating Scale score based on the TP3 trial (3) was smaller and yielded only borderline power. However, this variable was included as a secondary outcome measure to provide a contrasting index of

hallucination improvement anchored to specific experiential dimensions that have been studied extensively.

Multiple Imputation for Endpoint Data

Multiple imputation was employed (6) where 10 data sets were imputed (PROC MI in SAS) using a Markov chain Monte Carlo method (7). Results from the 10 sets of analyses were combined using PROC MIANALYZE in SAS (Cary, NC). The required assumption of data missing at random in the above analyses seems plausible as the reason for missingness is unlikely related to the unobserved values themselves. Sensitivity analyses, albeit based on a small sample of subject with missing data, uncovered no distinctions from completers.

Table S1. Reasons for Not Re-Enrolling Patients

Patient no longer interested in participating or unable to be re-contacted	64
Insufficient frequency of AVHs	19
Worsening psychosis, suicidality, behavioral problems	17
History of seizures	14
Unlikely to fit in MRI scanner	10
Active drug or alcohol abuse	7
Evidence of reduced capacity to understand study and give informed consent	6
Subnormal IQ, history of serious head trauma, coma	4
Inability to differentiate AVHs from verbal thoughts	4
Metal risk	2
Poor reporter of symptoms	2
Too old	2
Other (pregnancy, pituitary tumor, prior rTMS)	3

AVH, auditory/visual hallucinations; MRI, magnetic resonance imaging; rTMS, repetitive transcranial magnetic stimulation.

Success of Mask

Patients were debriefed to assess success of masking. Results after the first session are especially informative since biases due to induced change in symptoms are relatively minimal compared to debriefing after the full 15 sessions. After the first session, 4/28 in the sham group guessed correctly, compared to 18/55 in the active group.

Table S2. Distribution of types of guessing after one session of repetitive transcranial magnetic stimulation

	Sham		Active	
	Frequency	**Percent**	**Frequency**	**Percent**
Not sure, don't know, or no basis for guessing	9	32.1	26	47.3
Correctly guessed active due to somatic sensation present	–	–	14	25.5
Incorrectly guessed active due to somatic sensation detected	6	21.4	–	–
Incorrectly guessed active due to change in voice experience	2	7.1	–	–
Correctly guessed active due to change in voice experience	–	–	2	3.6
Incorrectly guessed sham due to absent or reduced somatic sensation	–	–	5	9.1
Correctly guessed sham due to absent or reduced somatic sensation	2	7.2	–	–
Incorrectly guessed sham due to no change in conscious experience	–	–	3	5.5
Incorrectly guessed active due to altered consciousness	1	3.6	–	–
Correctly guess sham due to absent changes in voice experience	1	3.6	–	–
Guessed sham or active correctly – other reasons	2	7.1	2	3.6
Guessed sham or active incorrectly – other reasons	5	17.9	3	5.5
Totals	28	100	55	100

Table S2 shows that after session 1, effects of somatic sensation caused 25.5% accurate guessing in the active group, with a near equivalent rate of inaccurate guessing of active stimulation due to somatic experience for sham subjects (21.4%). Similarly, 9.1% of active patients incorrectly guessed sham due to reduced or absent somatic sensation, while 7.1% of patients in the sham group correctly guessed sham due to reduced or absent somatic sensation during the first session. Thus, rates of guessing based on somatic sensation in both directions (sensation causing guess of "active" and reduced sensation causing guess of "sham") were similar for the two groups:

Debriefing from drop-outs were folded into endpoint debriefing data. Correct guessing at masked trial endpoint (session 15) for the active group was increased (32/55) - almost entirely based on improvement in AVHs (16/55) rather than somatic sensation. There was a modest increase in correct guessing for the sham group at trial endpoint (10/27) credited to lack of improvement in AVHs (5/27). Drop-outs during the course of trial were asked to guess their allocation with one exception, a drop-out patient in the sham group who could not be debriefed due to agitation/paranoia; these data were folded into endpoint data.

Crossing Over Sham Patients to Unmasked rTMS

Table S3. Within-subject improvements in the four outcome variables for unmasked rTMS crossing over from sham stimulation (n = 18)[1]

	Baseline	After 15 sessions (endpoint)	$p_{\text{Time effect}}$
HCS	10[2]	7.64 (0.73)	0.01
Hallucination frequency	5.39 (0.74)	4.56 (0.75)	0.12
AHRS total	27.0 (1.97)	25.4 (2.00)	0.36
CGI	4[2]	3.21 (0.25)	0.01

[1]This trial component offered to patients allocated to the sham group. Models included time (baseline, endpoint) as within-subjects factor; block 3 data derived from multiple imputations; Data shown as model-based least-square means (standard error). HCS, Hallucination Change Scale, which was reset to 10 at initiation of this trial arm; AHRS, Auditory Hallucination Rating Scale; CGI, Clinical Global Improvement also reset to four at the commencement of this trial arm. Hallucination frequency corresponds to the frequency variable reflected in the AHRS.
[2]Standardized values at baseline.

Table S4. Within-subject improvements in outcome variables for unmasked rTMS extending from 15 to 20 sessions (n = 38)[1]

	After 15 sessions	After 20 sessions	$t_{(37)}$ =, (p)
HCS	6.7 (0.50)	5.5 (0.47)	3.0 (0.005)
Hallucination frequency	4.6 (0.47)	4.2 (0.49)	2.9 (0.006)
AHRS total	24.3 (1.13)	21.8 (1.47)	3.6 (0.001)
CGI	2.8 (0.18)	2.3 (0.18)	4.3 (<0.0001)

[1]Data from patients who received both masked and unmasked rTMS; data shown as mean (standard error), paired *t*-test (2-tailed *p* value). HCS, Hallucination Change Scale; AHRS, Auditory Hallucination Rating Scale; CGI, Clinical Global Improvement. Hallucination frequency corresponds to the frequency variable reflected in the AHRS.

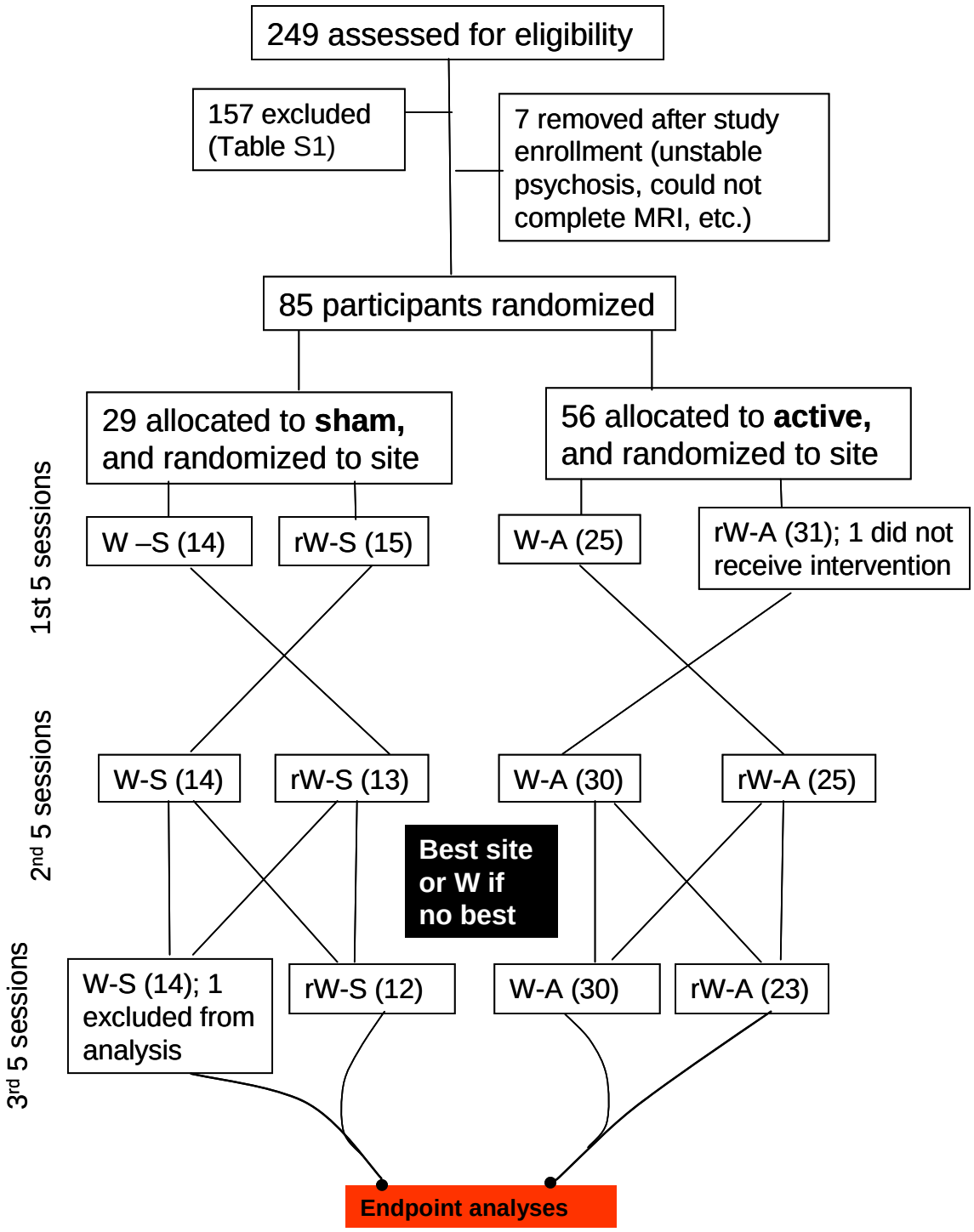

Figure S2. Consort flow diagram. W, Wernicke's site; rW, site in right homologue to Wernicke's area; A, active; S, sham.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	Title page
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	2-3
	2b	Specific objectives or hypotheses	3
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	4
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	none
Participants	4a	Eligibility criteria for participants	3
	4b	Settings and locations where the data were collected	Single site
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	4-5
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	5-6
	6b	Any changes to trial outcomes after the trial commenced, with reasons	none
Sample size	7a	How sample size was determined	6
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	4
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	4
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	4
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	4
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	4

	11b	If relevant, description of the similarity of interventions	4-5
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	6-8
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	7-8
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Supplement Consort flow diagram
	13b	For each group, losses and exclusions after randomisation, together with reasons	Table S1, 8-9
Recruitment	14a	Dates defining the periods of recruitment and follow-up	4
	14b	Why the trial ended or was stopped	4
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Consort flow diagram, supplement, Tables 2-5
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Tables 2-3
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	NA
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	9-10, Tables 4,5
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	8-9
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	13-14
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	14
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	11-14
Other information			
Registration	23	Registration number and name of trial registry	abstract
Protocol	24	Where the full trial protocol can be accessed, if available	Correspond-ing author
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	14

Supplemental References

1. Hoffman RE, Hampson M, Wu K, Anderson A, Gore J, Buchanan RJ, *et al.* (2007): Probing the pathophysiology of auditory hallucinations by combining functional magnetic resonance imaging and transcranial magnetic stimulation. *Cereb Cortex* 17:2733-2743.
2. Talairach J, Tournoux P (1988): *Co-planar stereotaxic atlas of the human brain*. Thieme Medical Publishers, New York.
3. Hoffman RE, Gueorguieva R, Hawkins KA, Varanko M, Boutros NN, Wu Y-T, *et al.* (2005): Temporoparietal transcranial magnetic stimulation for auditory hallucinations: safety, efficacy and predictors in a fifty patient sample. *Biol Psychiatry* 58:97-104.
4. Cohen J (1988): *Statistical power analysis for the behavioral sciences, second edition.* Lawrence Erlbaum Associates, Publishers. Hillsdale, NJ.
5. Hoffman RE, Pittman B, Constable RT, Bhagwagar Z, Hampson M (2011): Time-course of regional brain activity accompanying auditory/verbal hallucinations in schizophrenia. *Br J Psychiatry* 198:277-283.
6. Rubin DB (1987): *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.
7. Schafer JL (1997): *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.