Daily Left Prefrontal Transcranial Magnetic Stimulation Therapy for Major Depressive Disorder

A Sham-Controlled Randomized Trial

Mark S. George, MD; Sarah H. Lisanby, MD; David Avery, MD; William M. McDonald, MD; Valerie Durkalski, PhD; Martina Pavlicova, PhD; Berry Anderson, PhD, RN; Ziad Nahas, MD; Peter Bulow, MD; Paul Zarkowski, MD; Paul E. Holtzheimer III, MD; Theresa Schwartz, MS; Harold A. Sackeim, PhD

Arch Gen Psychiatry. 2010;67(5):507-516

APPENDIX. RATER CERTIFICATION PROCEDURES

Ten clinical raters met the certification criteria on first attempt. Raters were clinically experienced personnel at the 4 sites, all with college degrees. Although many were MDs, RNs, and social workers, we did not require this. These criteria required high levels of agreement with the external rater in evaluating 5 videotapes to derive Hamilton Scale for Depression (HAM-D) and Montgomery-Åsberg Depression Rating Scale ratings. Eight other raters failed on their first attempt (3 at the Medical University of South Carolina, 1 at Emory University, and 4 at the New York State Psychiatric Institute). Six of these 8 raters made a second attempt on another set of videotapes, and 3 passed. The 3 who failed again were permanently disqualified. Once certified, all rating interviews were videotaped, and a subset at key time points were rated by a masked external rater. A computer program then compared the HAM-D scores of the clinical raters and the expert rater, and the study chairs were notified if there was significant drift. One rater was removed because of poor performance.

eFigure. Transcranial magnetic stimulation (TMS) method. A and B, Repetitive TMS (rTMS) was delivered using the Neuronetics model 2100 therapy system investigational device and figure-eight, solid-core coil (Neuronetics Inc, Malvern, Pennsylvania) demonstrated on one of us (B.A.). The novel active sham condition consisted of the sham coil, noise-dampening earphones for the patient and the teacher, and electrical pads inserted under the coil on the patient’s head (left panel). (Each site used 3 magnetic coils with a metal insert blocking the field, identical in weight and external appearance and with similar acoustic properties when actively pulsed. One coil, coil A, was unblinded and was used to determine motor thresholds. The remaining 2 coils were distinguishable only by external labels as coil B or coil C, with one being active and the other sham. Sham coils contained a magnetic shield limiting the magnetic energy reaching the cortex to 10% or less of that of the active coil. The active and sham coils had identical weight and appearance and similar acoustic properties. Triplets of coils were periodically rotated from the coordinating center at the Medical University of South Carolina across the 4 clinical sites to reduce the possibility that inadvertent unmasking would result in knowledge of the nature of the B and C coils. On 4 occasions [3 at Emory University and 1 at the New York State Psychiatric Institute], a treater encountered equipment problems with the TMS coils; the sham and active coils were replaced at that site, and the TMS administrator was changed until that patient completed treatment. This procedure ensured that the TMS administrator was not accidentally unblinded. There were no instances in which treaters indicated that they were unblinded or were highly confident of the randomized assignment of any patient.) C, Coincident with each discharge of the sham coil, an electrical pulse (unidirectional, square wave, 0.5 mA) was administered from a modified Mecta Spectrum 5000Q machine (Mecta Corp, Lake Oswego, Oregon) connected to the TMS machine through a signaling interface created by the James Long Co (Caroga Lake, New York) that mimicked the active rTMS sensation and also caused focal scalp twitching. Patients, treaters, clinical raters, the offsite expert rater, and all other study personnel were masked to coil functionality. The integrity of the mask was assessed immediately at phase 1 exit. Patients, treaters, and local raters made “best guesses” as to assignment to active or sham rTMS and indicated their confidence in this guess (not at all, slightly, moderately, considerably, or extremely).
### eTable. Integrity of the Blind

<table>
<thead>
<tr>
<th>Person Asked, by Treatment Arm</th>
<th>Active</th>
<th>Sham</th>
<th>Active</th>
<th>Sham</th>
<th>Active</th>
<th>Sham</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active</strong></td>
<td>72</td>
<td>83</td>
<td>72</td>
<td>84</td>
<td>74</td>
<td>87</td>
</tr>
<tr>
<td><strong>Sham</strong></td>
<td>83</td>
<td></td>
<td>84</td>
<td></td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

- **End of phase 1 accurate guess, No. (%)**
  - Active Sham Active Sham Active Sham Active Sham
  - 35 (49) 55 (66) 48 (67) 54 (64) 26 (35) 51 (59)
- **Extremely confident in guess, No.**
  - 9 13 0 0 0 0
- **Correct guess when extremely confident, No.**
  - 5 10 0 0 0 0

#### Confidence in guess, No. (%)

- **Extremely**
  - 9 (13) 13 (16) 0 0 0 0
- **Considerably**
  - 19 (26) 26 (31) 4 (6) 3 (4) 6 (8) 2 (2)
- **Moderately**
  - 25 (35) 20 (24) 14 (19) 7 (8) 6 (8) 11 (13)
- **Slightly**
  - 8 (11) 15 (18) 23 (32) 30 (36) 16 (22) 21 (24)
- **Not at all**
  - 11 (15) 9 (11) 31 (43) 44 (36) 46 (62) 53 (61)

---

**Twenty-two patients reported that they were “extremely confident” in their guess, but only 15 of these guessed correctly. Treaters guessed correctly 65% of the time (67% correct for active repetitive transcranial magnetic stimulation [rTMS] and 64% for sham). No treaters were extremely confident in their guesses, with 82% of guesses rated as slightly or not at all confident. Clinical raters guessed correctly 48% of the time (35% correct for active rTMS and 59% for sham). No clinical rater was extremely confident, with 84% of clinical rater guesses reported to be slightly or not at all confident. We conducted logistic regression on the primary outcome of the actual stated guess (active vs sham). Independent variables included treatment (active vs sham), percentage change from baseline in the last recorded Hamilton Scale for Depression score of the fixed-treatment phase and site (categorical). Interactions between model covariates were examined but were not significant at the level of \( P < .15 \). For patient guesses, the effect was significant for the percentage change in Hamilton Scale for Depression score \( (P < .001) \) and not for treatment \( (P = .11) \) or site \( (P = .24) \). That is, patients’ guesses were affected by their degree of clinical change in the direction of guessing that active treatment was administered when greater improvement was achieved. However, there was no effect of actual randomized assignment. Similar results were obtained for the guesses of the clinical raters. For treaters, the effect was significant for treatment \( (P < .001) \) and percentage change \( (P = .04) \) but not for site \( (P = .12) \). This analysis, plus the lack of any “confident” or “extremely confident” ratings by treaters and clinical raters (and random accuracy in patients with confident or extremely confident marks), suggest that blinding was effectively maintained.**