

Treatment of Bipolar Depression with Deep TMS: Results from a Double-Blind, Randomized, Parallel Group, Sham-Controlled Clinical Trial

Diego F Tavares^{1,2}, Martin L Myczkowski¹, Rodrigo L Alberto¹, Leandro Valiengo^{1,3}, Rosa M Rios¹, Pedro Gordon¹, Bernardo de Sampaio-Junior^{1,4}, Izio Klein^{1,4}, Carlos G Mansur¹, Marco Antonio Marcolin⁵, Beny Lafer¹, Ricardo A Moreno², Wagner Gattaz³, Zafiris J Daskalakis⁶ and André R Brunoni^{*,1,3,4}

¹Service of Interdisciplinary Neuromodulation (SIN-EMT), Department and Institute of Psychiatry, Faculty of Medicine, University of São Paulo, São Paulo, Brazil; ²Mood Disorders Unit (GRUDA), Department and Institute of Psychiatry, Faculty of Medicine, University of São Paulo, São Paulo, Brazil; ³Laboratory of Neuroscience (LIM27), Department and Institute of Psychiatry, University of São Paulo, São Paulo, Brazil; ⁴University Hospital, University of São Paulo, São Paulo, Brazil; ⁵Department and Institute of Neurology, Faculty of Medicine, University of São Paulo, São Paulo, Brazil; ⁶Centre for Addiction and Mental Health, Clarke Division, Toronto, Ontario, Canada

Bipolar depression (BD) is a highly prevalent condition with limited therapeutic options. Deep (H1-coil) transcranial magnetic stimulation (dTMS) is a novel TMS modality with established efficacy for unipolar depression. We conducted a randomized sham-controlled trial to evaluate the efficacy and safety of dTMS in treatment-resistant BD patients. Patients received 20 sessions of active or sham dTMS over the left dorsolateral prefrontal cortex (H1-coil, 55 18 Hz 2 s 120% MT trains). The primary outcome was changes in the 17-item Hamilton Depression Rating Scale (HDRS-17) from baseline to endpoint (week 4). Secondary outcomes were changes from baseline to the end of the follow-up phase (week 8), and response and remission rates. Safety was assessed using a dTMS adverse effects questionnaire and the Young Mania Rating Scale to assess treatment-emergent mania switch (TEMS). Out of 50 patients, 43 finished the trial. There were 2 and 5 dropouts in the sham and active groups, respectively. Active dTMS was superior to sham at end point (difference favoring dTMS = 4.88; 95% CI 0.43 to 9.32, $p = 0.03$) but not at follow-up. There was also a trend for greater response rates in the active (48%) vs sham (24%) groups (OR = 2.92; 95% CI = 0.87 to 9.78, $p = 0.08$). Remission rates were not statistically different. No TEMS episodes were observed. Deep TMS is a potentially effective and well-tolerated add-on therapy in resistant bipolar depressed patients receiving adequate pharmacotherapy.

Neuropsychopharmacology (2017) **42**, 2593–2601; doi:10.1038/npp.2017.26; published online 31 May 2017

INTRODUCTION

Bipolar disorder is a prevalent and disabling condition, with a worldwide prevalence of around 2–3%, considering both bipolar I and II subtypes. Although hypomania and mania episodes characterize the disorder, depressive episodes in fact exceed them in duration and frequency (Holtzman *et al*, 2015; Perich *et al*, 2016; Zimmerman, 2016). Moreover, there are limited first-line therapies for treating bipolar depression (BD), treatment-resistance being two times higher compared to unipolar depression (Li *et al*, 2012; Tondo *et al*, 2014). Furthermore, antidepressant drugs (ADs), for example,

selective serotonin reuptake inhibitors (SSRIs), tricyclic antidepressants and others, which are the main treatment for unipolar depression, have limited efficacy and adverse effects in BD (Sachs *et al*, 2007). Moreover, pharmacotherapy for BD also presents side effects that can hinder optimal treatment adherence and produce long-term clinical comorbidities. These issues highlight the need for developing novel therapeutic strategies combining efficacy, acceptability, and tolerability (Sienaert *et al*, 2013).

Repetitive transcranial magnetic stimulation (rTMS) is a non-invasive brain stimulation therapy with established efficacy and acceptability for unipolar depression (George *et al*, 2010). Notwithstanding, the technique has been under constant development to increase its efficacy. One novel approach includes the ‘deep’ (H1 coil) TMS (dTMS) (Levkovitz *et al*, 2015). This coil generates electrical fields that, although less focal, have greater penetration depth compared to the standard, figure-of-eight coil (3 cm vs < 1 cm, respectively, according to simulations using the same TMS intensity (Roth *et al*, 2007). Therefore, this coil could

*Correspondence: Dr AR Brunoni, Service of Interdisciplinary Neuromodulation (SIN-EMT), Department and Institute of Psychiatry, University of Sao Paulo (USP), Laboratory of Neurosciences (LIM-27), R. Dr Ovidio Pires de Campos 785, 2o andar, Ala Sul, São Paulo 05403-010, Brazil, Tel: +55 11 2661 8159, Fax: +55 11 2661 8159, E-mail: brunoni@usp.br

Received 19 October 2016; revised 19 January 2017; accepted 27 January 2017; accepted article preview online 1 February 2017

stimulate deeper dorsolateral and ventrolateral prefrontal areas that also projects into other brain areas that present impaired functioning in depression (Luborzewski *et al*, 2007). In fact, a large randomized controlled trial showed that dTMS was effective and well-tolerated in depression treatment, with response and remission rates of, respectively, 38.4 and 32.6% (Levkovitz *et al*, 2015).

However, only preliminary studies have addressed rTMS efficacy in BD. A recent meta-analysis suggested that rTMS is effective in BD (McGirr *et al*, 2016). Nonetheless, negative results were also found in recent trials (Fitzgerald *et al*, 2016). Regarding dTMS, although case series suggest it might be effective in BD treatment (Harel *et al*, 2011; Rapinesi *et al*, 2015), no results from randomized clinical trials have been reported so far.

Therefore, we conducted a randomized, sham-controlled clinical trial to evaluate the effectiveness and tolerability of deep-TMS as an add-on therapy to pharmacological treatment of resistant bipolar depressed patients. Our primary hypothesis was that BD patients in the active group would present a statistically greater improvement in their depressive symptoms compared to those in the sham group at the end of the acute intervention phase (four weeks of treatment). Our secondary hypothesis was that this improvement would be maintained after a 4-week follow-up.

MATERIALS AND METHODS

Study Design

We conducted a single-center, double-blind, randomized, parallel-group, sham-controlled clinical trial (Clinicaltrials.gov identifier: NCT01962350) that lasted 8 weeks, comprising 4 weeks of the acute intervention phase, in which patients received 20 dTMS or sham sessions once daily excluding weekends; and 4 weeks of the follow-up phase, in which patients received no intervention. The protocol was reviewed and approved by the Ethics Committee of the Institute of Psychiatry—Clinics Hospital of the University of São Paulo and was conducted in accordance with the principles of the Helsinki Declaration (Williams, 2008) and the American Document of Good Clinical Practice (Castelein *et al*, 2014). This study report conforms to CONSORT guidelines (Stevely *et al*, 2015).

Participants were randomized using a computer-generated list in a 1:1 ratio. Allocation concealment consisted of sequentially numbered cards, which determined whether the TMS machine would produce real or sham stimulation. A secretary not directly participating in the research was responsible for handling the numbered cards to the staff before each session. Participants and personnel were therefore fully blinded to allocation group status.

Participants

All participants signed informed consent form. We enrolled 50 adults aging from 18 to 65 years old diagnosed with bipolar disorder types I or II in an acute depressive episode. All patients presented treatment-resistant depression (TRD). Although there is not a universally accepted definition for TRD, it has been proposed that failure to achieve remission with ≥ 2 adequate AD trials defines TRD in unipolar

disorder. Here TRD was conceptualized as the failure to achieve remission with ≥ 2 interventions (Parker and Graham, 2016) approved as first (lithium, lithium+divalproex, quetiapine, or lamotrigine), second (divalproex, lithium+lamotrigine, or divalproex+lamotrigine), or third line (carbamazepine, olanzapine, lithium+carbamazepine, and quetiapine+lamotrigine) therapies for BD according to CANMAT guidelines ((Yatham *et al*, 2013).

Recruitment strategies included referrals from physicians, patients from academic mood disorders clinics and advertisement through social media and local newspapers.

All clinical assessments were performed by a certified psychiatrist (DFT) and a certified psychologist (MLM) who are trained in the application of the structured questionnaires and interviews used in the present study. The diagnoses were performed per DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) criteria and confirmed using the Mini-International Neuropsychiatric Interview (M.I.N.I.; Sheehan *et al*, 1998). The main eligibility criterion was the presence of a depressive episode of at least moderate intensity, corresponding to a Hamilton Depression Rating Scale (17-items; HDRS-17) > 17 (Hamilton, 1960). We only enrolled treatment-resistant bipolar depressed patients who were AD-free and using a stable drug regimen for at least 2 weeks. Benzodiazepine drugs were allowed, although only at low doses (< 3 mg per day of lorazepam or equivalent). Patients who were on ADs had this class of medication washed out and were reassessed after 4 weeks.

Exclusion criteria included other neuropsychiatric conditions per DSM-IV criteria (such as unipolar depression, schizophrenia, substance dependence, dementias, traumatic brain injury, epilepsy, and others—although anxiety disorders as comorbidities were included, provided the primary diagnosis was bipolar disorder); severe personality disorders; presence of (hypo)manic symptoms at baseline and/or a Young Manic Rating Scale (YMRS) > 12 points; rapid-cycling bipolar disorder; acute suicidal ideation; pregnancy; specific contraindications to rTMS and motor threshold (MT) $> 70\%$ of maximum stimulator output assessed at the screening visit. Patients presenting psychotic depression at the time of assessment (but not those with a prior history of mood episodes with psychotic features) were also excluded, since ECT was consistently found to be better than rTMS in this subgroup of patients (Milev *et al*, 2016). Moreover, due to the severity of this condition, psychotic patients were not considered eligible to participate in an 8-week placebo-controlled trial for ethical reasons.

Complementary diagnostic exams consisted of, at physicians' discretion, brain MRI, general laboratory tests (including pregnancy testing, thyroid stimulating hormone levels, lithium levels and others) and a 12-lead ECG. They were used as a clinical safety parameter, to exclude possible decompensated clinical conditions that could cause or worsen secondary depressive symptoms and to perform differential diagnoses.

Interventions

The TMS sessions were delivered using the Brainsway dTMS system with the H1-coil investigational device (Brainsway Ltd, Jerusalem, Israel). The coil is situated inside a helmet to achieve effective cooling during stimulation. A sham coil is

also included in the same helmet. The sham coil mimics scalp sensations and the acoustic artifact of the active stimulation without inducing neuronal activation.

Before the screening interview, potential participants had their MT (the lowest stimulation intensity necessary to evoke a motor potential with at least 50 μ V amplitude in 50% of attempts) assessed to determine eligibility. MTs were reassessed at the first day of treatment and then every week. The coil was positioned over the left dorsolateral prefrontal cortex, which was found 6 cm anteriorly to the 'hot spot' (ie, the optimal location on the scalp to evoke a maximum TMS response with minimum stimulator intensity), per a ruler attached to the subject's cap. The subjects were stimulated every day for 4 weeks (except weekends). Patients who presented two consecutive missing visits were considered washouts. Missing visits were replaced at the end of the acute stimulation phase; therefore, all patients received 20 dTMS sessions.

The active stimulation consisted of 55 18 Hz, 2 s trains at 120% MT intensity, with a between-train interval of 20 s (1980 pulses per day or 39 600 pulses per treatment). The sham stimulation was performed using the same procedures, with the sham coil.

Assessments

Demographic and clinical data were collected, including age, age at onset of the first episode, marital status, occupational status, diagnosis subtype, duration of illness, medication use, and others.

The HDRS-17 (Hamilton, 1960) was the scale used for our primary efficacy outcome and also for defining response ($\geq 50\%$ improvement from baseline), and remission status ($\text{HDRS-17} \leq 7$).

Other instruments included the Clinical Global Impression Scale of severity (CGI-S) (Guy, 1976), the Global Assessment Functioning (GAF) scale, the Hamilton Anxiety Scale (HAM-A; Hamilton, 1959) and the YMRS (Young *et al*, 1978). These clinical assessments were performed every week until week 4, then every other week until week 8. Adverse events were assessed using a TMS side effects questionnaire, in which participants were actively asked regarding the presence of an adverse event and its relationship with the stimulation (Bersani *et al*, 2013). They were assessed every day during the first week and then every week during the acute treatment phase.

Outcomes

The primary efficacy outcome was defined as the change in HDRS-17 from baseline to week 4. Secondary efficacy outcomes included response and remission status at week 4, depression improvement from baseline to week 8, and response and remission status at week 8. Other outcomes included HAM-A and CGI-S improvement.

The presence of treatment-emergent mania switch (TEMS) was assessed according to the ISBD recommendations that consider TEMS as likely when there are 2 or more manic symptoms (eg, irritability or euphoria (racing thoughts, grandiosity, decreased need for sleep), and $\text{YMRS} > 12$; Tohen *et al*, 2009).

Statistical Analyses

Analyses were performed in Stata 14 (Statacorp, College Station, TX, USA). Clinical and demographic variables were compared between groups using *t*-tests, Mann-Whitney test, χ^2 tests, or the Fisher's exact test and described using mean (standard deviation), median (interquartile range), or number of events (frequency) according to the type of the variable and its normality (assessed using the Shapiro-Wilk test).

We performed an intention-to-treat (ITT) analysis using the last observation carried forward (LOCF) approach. Missing data were considered to be at random. We also performed per protocol (PP) analyses. The sample size was calculated based on a preliminary study evaluating the efficacy of dTMS in unipolar depression (Levkovitz *et al*, 2009). We estimated that the effects of active *vs* sham dTMS would be similar than in findings from that preliminary study that compared the efficacy of the H1 *vs* H1L coil groups. For a power of 90% and a two-tailed δ of 5%, we obtained a sample size of 40 patients, which was enriched to a final number of 50 participants (25 per group), considering attrition.

The primary analysis was a multilevel mixed-effects linear regression (mixed command in Stata) with group (2 levels: active and sham) and time (6 levels; baseline, weeks 1, 2, 4, 6, and 8) as independent variables and subject as a random-effects variable. HDRS was the dependent variable. Our primary hypothesis was that the interaction of time with group would be significant, with active dTMS being superior to sham at week 4. After that, pairwise comparisons were performed at each time point (contrast command in Stata). Similar analyses were performed for the other outcome scales and for the week 8 follow-up end point.

Logistic regressions were performed to assess response and remission rates between groups.

Frequency of TEMS and adverse events were compared among groups using Fisher's exact test or the χ^2 test. We considered an adverse event to be present when it was of at least mild intensity, at least subjectively remotely associated with the intervention and reported in ≥ 3 occasions (out of 8) and absent if otherwise.

To verify blinding integrity, we asked, at week 8, for patients and raters to guess whether the allocation group was active on a 0–100 scale; guessing scores were compared using a *t*-test.

Regarding predictors of response, we performed general linear models using the difference between baseline and week 4 or week 8 HDRS scores as the dependent variable. For the independent variables, we compared 1 predictor variable at a time and group. The number of failed treatments in the current episode was used to assess the degree of treatment-resistance. These analyses were only conducted in the ITT sample.

RESULTS

Participants

Of ~280 volunteers, 268 were screened and 216 were excluded for several reasons (Figure 1). Out of 50 patients included, 43 finished the trial. There were 2 dropouts in the

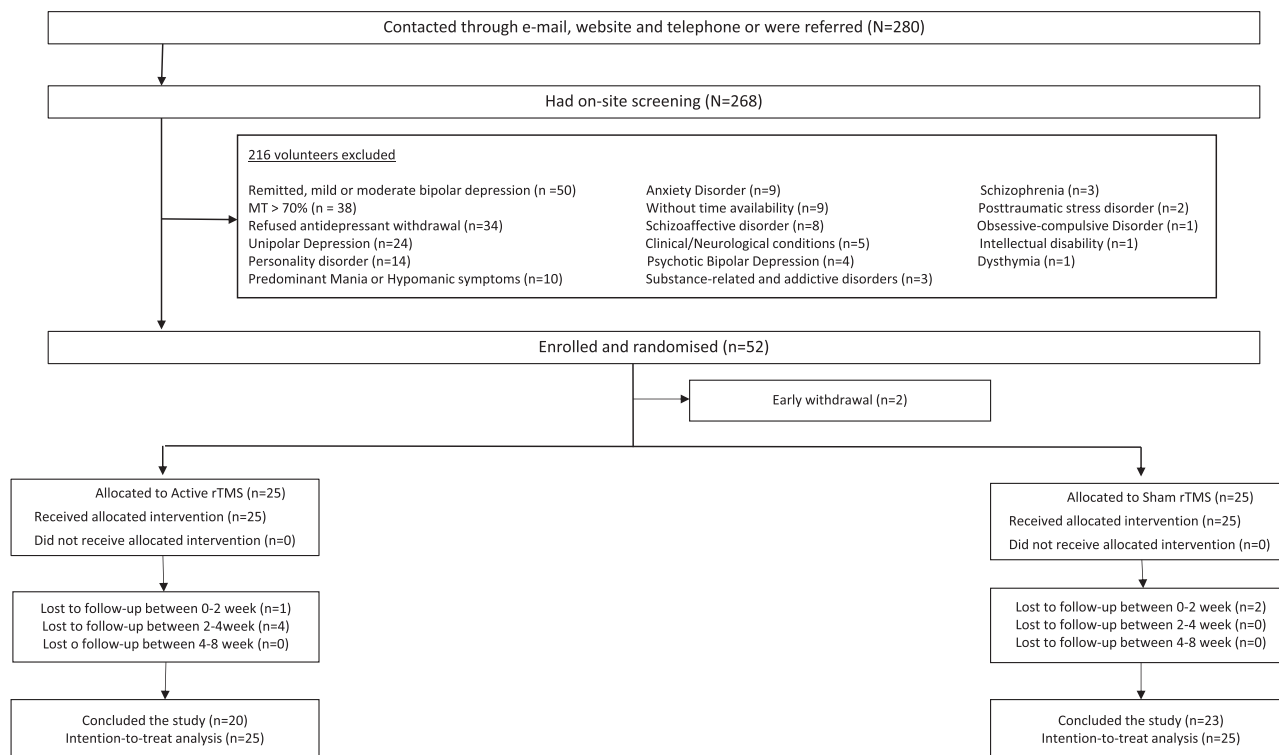


Figure 1 Flow chart.

sham group (both because of consecutive missing visits) and 5 dropouts in the active group (two were drop-outs for consecutive missing visits, two because of the severity of depressive symptoms and one because of side effects such as headache and burning sensation over the scalp), which was not statistically different ($p = 0.21$; Figure 1).

The groups were similar in all main clinical and demographic characteristics at baseline. The frequency of bipolar disorder types I and type II was 50% for each group. Only 10 (20%) patients, 4 in sham and 6 in active, were on ADs at baseline and needed a drug washout. The sample was composed mainly of women (70%), with a mean age of 42.34 (SD = 10.54) years. Patients presented a median of 2 (IQR 2–4) previous depressive episodes. The median duration of the current depressive episode was 6 months (IQR 3–12). Half of the sample was on lithium therapy, 20% on valproate, 30% on lamotrigine, and 36% on quetiapine. There were also a few patients (<10%) on aripiprazole, topiramate, olanzapine, risperidone, asenapine, carbamazepine, or ziprasidone. Forty-three patients (86%) were using at least one treatment considered a first-line therapy per CANMAT guidelines (Table 1).

The mean percentage of missing visits was 4.5%, being 7.1% in the sham group and 1.7% in the ($p = 0.044$). All patients received 20 dTMS sessions, as missing visits were replaced at the end of the acute treatment phase.

Main Findings

In the ITT analysis, results from our mixed model revealed a significant main effect of time ($F_{5,240} = 25.38$, $p < 0.001$) and a significant time \times group interaction ($F_{5,240} = 2.26$, $p = 0.046$)

(Figure 2). Further contrast comparisons revealed that active dTMS was superior to sham at weeks 4 (difference favoring dTMS = 4.88; 95% CI 0.43 to 9.32, $p = 0.03$) and 6 (5.2; 95% CI 0.75 to 9.64, $p = 0.02$) but not at other time points. Results were similar in the PP analyses (Table 2).

Response and Remission

There was a trend for greater response rates in the active (48%) vs sham (24%) groups (OR = 2.92, 95% CI 0.87–9.78, $p = 0.08$) at week 4. Comparisons regarding response and remission at week 8 were not statistically significant (Table 3).

Other Scales

At week 4, patients in the active group presented significant greater improvement compared to sham in the GAF (percentage of improvement 65.37% (53.46) vs 34.07% (48.62), $p = 0.03$, respectively) and CGI scores (36.47% (22.87) vs 19.2% (30.96), $p = 0.03$, respectively). Comparisons at week 8 and regarding HAM-A were not statistically significant (Supplementary Table 1).

Adverse Events and Treatment-Emergent (hypo)Mania

Scalp pain rates were higher in the active (20%) vs sham (0%) groups ($p = 0.05$). Other adverse events such as headache, neck pain, burning sensation, hearing complaints and concentration difficulties presented non-significantly different rates between groups (Supplementary Table 2).

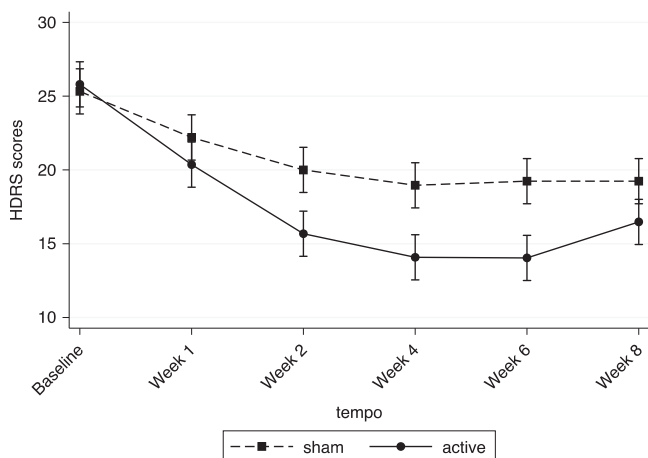
No clinical episodes of TEMS were observed.

Table 1 Baseline Clinical and Demographic Characteristics of the Study Sample

	Active (n = 25)	Sham (n = 25)	p	Total (n = 50)
<i>Clinical and demographic characteristics</i>				
Gender (% Fem)	17 (68)	18 (72)	0.76	35 (70)
Age, mean (SD)	43.5 (12)	41.2 (8.9)	0.43	42.3 (10.5)
Years at school, mean (SD)	14.34 (4.06)	15.08(4.33)	0.53	14.71 (4.17)
Ethnicity (% Caucasian)	22 (88)	19 (82.61)	0.69	41 (85.42)
Marital status (% not married)	13 (52)	9 (37.5)	0.31	22 (44.9)
Income (R\$)	5521(3703)	4313(3394)	0.23	4917(3568)
Age onset (SD)	22.16(7.78)	19.84(8.97)	0.33	21 (8.39)
<i>Depression characteristics</i>				
Recurrent depression, mean (SD)	9 (36)	10 (40)	0.77	19 (38)
Depression episodes, median (IQR)	2 (2–4)	2 (2–4)	0.71	2 (2–4)
Chronic depression, mean (SD)	10 (40)	13 (52)	0.39	23 (46)
Depressive episode duration, months, median (IQR)	6 (2–11)	8 (3–12)	0.35	6 (3–12)
Severe depression, n(%)	14 (56)	15 (60)	0.77	29 (58)
Depression Type I, n(%)	11 (44)	14 (56)	0.4	25 (50)
Depression Type II, n(%)	14 (56)	11 (44)	0.4	25 (50)
<i>Treatment history, n (%)</i>				
First-line therapy	22 (88)	21 (87.5)	0.96	43 (87.76)
Lithium	15 (60)	15 (60)	1	30 (60)
Quetiapine	11 (44)	7 (28)	0.24	18 (36)
Valproate	5 (20)	5 (20)	1	10 (20)
Lamotrigine	7 (28)	8 (32)	0.76	15 (30)
Benzodiazepine	11 (44)	7 (29.17)	0.28	18 (36.73)
<i>Baseline scores, mean (SD)</i>				
HDRS17	25.8 (5.25)	25.32(3.76)	0.71	25.56 (4.53)
HAMA	27.76 (7.67)	26.92(6.16)	0.67	27.34 (6.9)
YMRS	5.32 (2.19)	5.36 (2.39)	0.95	5.34 (2.27)
CGI-S	3.8 (0.91)	3.68 (0.75)	0.61	3.74 (0.83)
GAF	40.4 (11.38)	41.16(8.18)	0.78	40.78 (9.81)

Abbreviations: GAF, Global assessment of functioning; CGI-S, Clinical global impression-severity of illness; HAMA, Hamilton anxiety rating scale; HDRS17, Hamilton depression rating scale, 17 items; R\$, reais (Brazilian currency); YMRS, Young mania rating scale.

P values represent the significance of one-way ANOVAs or χ^2 -tests

**Figure 2** Primary outcome.

Predictors of Response

No interactions between groups with any predictor variable, including type of bipolar disorder and number of failed effective treatments in the present episode, were found (Supplementary Table 3).

Integrity of Blinding

The degree of confidence of active group allocation was, for raters, 52.17 (29.53) and 60 (31.46) in patients that received sham and active stimulation, respectively; while for patients it was 45.65 (30.46) and 53.57 (35.57), respectively. Both differences were not statistically significant ($t = 0.85$, $p = 0.4$; $t = 0.8$, $p = 0.43$ for raters and patients, respectively). In other words, both raters and patients were unable to identify the allocation group beyond chance.

Table 2 Main Outcomes of the Study at Different Time Points

	Baseline	HDRS				
		Week 1	Week 2	Week 4	Week 6	Week 8
<i>Sham</i>						
PP	25.32 (3.76)	21.83 (6.67)	19.39 (7.42)	18.26 (9.88)	18.56 (10.11)	18.56 (9.64)
ITT	25.32 (3.76)	22.2 (5.84)	20 (7.5)	18.96 (9.83)	19.24 (10.02)	19.24 (9.59)
<i>dTMS</i>						
PP	25.8 (5.25)	20.36 (5.72)	15.56 (8.22)	13.5 (9.41)	13.43 (10.91)	15.95 (11.20)
ITT	25.8 (5.25)	20.36 (5.72)	15.68 (7.9)	14.08 (8.99)	14.04 (10.13)	16.48 (10.17)
<i>Contrast (difference (95% CI); p</i>						
PP	0.48 (-3.97-4.93); <i>p</i> = 0.83	-1.61 (-6.09-2.87); <i>p</i> = 0.48	-4.18 (-8.74-0.38); <i>p</i> = 0.07	-5.03 (-9.62--0.45); <i>p</i> = 0.03	-5.54 (-10.16--0.92); <i>p</i> = 0.01	-2.79 (-7.44-1.85); <i>p</i> = 0.23
ITT	0.48 (-3.96-4.92); <i>p</i> = 0.83	-1.84 (-6.28-2.60); <i>p</i> = 0.41	-4.32 (-8.76-0.12); <i>p</i> = 0.056	-4.88 (-9.32--0.43); <i>p</i> = 0.03	-5.2 (-9.64--0.75); <i>p</i> = 0.02	-2.76 (-7.20-1.68); <i>p</i> = 0.22

Abbreviations: dTMS, deep transcranial magnetic stimulation; PP, per protocol analyses; ITT, intention to treat analyses.

p values represent results for the mixed model ANOVA time × group interaction (for the main analysis) or time × dTMS (for the factorial analysis) interaction at each week; values in bold represent significant interactions at a *p* ≤ 0.05 level.

Table 3 Response and Remission Rates According to HDRS Scores

	<i>n</i> (%)	OR (CI)	<i>p</i>
<i>Response (active × sham)</i>			
Week 4			
PP	12 (54.55) vs 6 (26.09)	3.4 (0.97-11.90)	0.06
ITT	12 (48) vs 6 (24)	2.92 (0.87-9.78)	0.08
Week 8			
PP	8 (40) vs 6 (26.09)	1.89 (0.52-6.87)	0.33
ITT	8 (32) vs 6 (24)	1.49 (0.43-5.17)	0.63
<i>Remission (active × sham)</i>			
Week 4			
PP	7 (31.82) vs 4 (17.39)	2.21 (0.54-9.01)	0.27
ITT	7 (28) vs 4 (16)	2.04 (0.51-8.12)	0.31
Week 8			
PP	6 (30) vs 6 (26.09)	1.21 (0.32-4.61)	0.78
ITT	6 (24) vs 6 (24)	1 (0.27-3.66)	1

Abbreviations: ITT, intention to treat analyses; PP, per protocol analyses.

Remission was defined as HDRS ≤ 7 and response as a score change > 50% from baseline. All remitters were responders, but not all responders were remitters. *P* values represent results for the logistic regression of time × group interaction.

DISCUSSION

We performed the first randomized, sham-controlled clinical assessing the efficacy, safety and tolerability of the H1-coil TMS for the treatment of resistant bipolar depression. Our primary hypothesis was confirmed as active dTMS was superior to sham at the end of the acute treatment phase, with a mean difference in means of 4.88 in the HDRS. For comparative purposes, the National Institute of Clinical

Excellence (NICE) states that a 3-point between group difference in HDRS scores translates into a clinically meaningful difference (Middleton *et al*, 2005). There was also a trend for greater response rates in the active (48%) vs sham (24%) groups. Moreover, patients in the active group presented significantly greater improvement in the GAF and CGI scores. Furthermore, no clinical episodes of TEMS were observed during the treatment. Finally, dTMS was similarly effective for both bipolar I and bipolar II patients.

This clinical trial was designed to evaluate the efficacy of deep-TMS as an add-on therapy to resistant BD patients, due to the paucity of studies assessing the effectiveness of treatment options in this group of patients. In fact, 86% of our sample was composed of patients using at least one treatment considered as a first-line therapy according to CANMAT guidelines (Yatham *et al*, 2013) at trial onset, with 50% of BD patients using lithium in clinically effective doses. Moreover, the frequency of type I and type II BD patients was equally distributed. Therefore, our findings point out that dTMS is a valid therapeutic option in such patients, whose treatment is particularly challenging, with only a few options currently available (Goodwin *et al*, 2016; Grunze *et al*, 2010; Malhi *et al*, 2015; Pacchiarotti *et al*, 2013; Yatham *et al*, 2013). Importantly, dTMS was not only effective in treating depressive symptoms, but also in improving global functioning.

The OR for response and remission rates observed were lower than the ORs for rTMS in unipolar depression (Berlim *et al*, 2014), which can be explained given the lower response rates usually observed in BD (Tondo *et al*, 2014). In fact, our dTMS response rate (48%) was similar than the rTMS response rate for BD (44.3%) according to a recent meta-analysis (McGirr *et al*, 2016)—possibly, the lack of a significant finding for response in our study occurred due to an underpowered analysis owing to a low sample size. Also, this meta-analysis suggested that low-frequency rTMS

over the right DLPFC might be more effective than high-frequency over the left DLPFC. However, we opted for stimulating the left DLPFC as low-frequency dTMS over the right DLPFC was not investigated for unipolar depression yet. Moreover, the results of this meta-analysis were not available when our study was designed. Our findings showing promising results for dTMS in bipolar depression stimulate further non-inferiority trials comparing the efficacy of this technique with standard rTMS approaches, which, although are costly and require large sample sizes, are necessary to assess which non-pharmacological therapy produces greater depression improvement.

Scalp pain was the only adverse event more prevalent in active compared to sham dTMS. Even though, this adverse effect did not increase attrition and was considered mild by the participants who experimented it. This effect was also not important enough to harm blinding, as patients and raters did not guess the allocation group beyond chance.

Our study lasted 8 weeks, with a 4-week follow-up period after the acute treatment phase, when patients received no extra stimulation sessions. During this period, dTMS efficacy progressively decreased over time, with superiority at week 6, but not at week 8. Possibly, TMS protocols used for BD should be different than for unipolar depression. In fact, in a recent meta-analysis (Kedzior *et al*, 2015) assessing the antidepressant effects of rTMS after the acute treatment phase, rTMS effects in unipolar depression were stable; although no long-lasting antidepressant effects were observed in trials that also included BD. Moreover, recent naturalistic rTMS studies in unipolar depression found that maximizing the number of sessions to up to 30 could also bring significant clinical gains (Carpenter *et al*, 2012; McDonald *et al*, 2011), particularly for treatment-resistant subjects. Another study using rTMS in BD also observed that more rTMS sessions in the acute treatment phase were associated with lower relapse rates (Cohen *et al*, 2010). Furthermore, in the dTMS unipolar depression trial (Levkovitz *et al*, 2015), 20 stimulation sessions were followed by two sessions per week for 12 weeks. In our study, we employed no maintenance schedule from weeks 4 to 8 devising that the mood stabilizers the patients were in use would sustain clinical improvement after the acute treatment phase. Conceivably, bipolar depressed patients might also profit from a longer treatment regimen and/or a maintenance treatment after the acute treatment phase when receiving dTMS.

Although there was no interaction between allocation group and benzodiazepine use, a main effect of benzodiazepine use was found. In fact, large, pragmatic clinical trials in bipolar disorder showed that benzodiazepine use in patients with bipolar depression seems to be a marker for a more severe course of illness, presents a higher risk of recurrence and is associated with greater illness complexity and higher burden of disease (Bobo *et al*, 2015; Perlis *et al*, 2010). These effects were independent of anxiety levels or comorbidity. These observations might explain why benzodiazepine users in our study presented a worse clinical outcome regardless of allocation group.

A recent study (Iovieno *et al*, 2016) discussed that placebo responses higher than 30% in bipolar disorder placebo-controlled trials might harm trial performance. In the

present study, the placebo response was lower than 30%, which adds evidence on its internal validity.

Limitations

The small sample size was the main study limitation, which might not have been adequately powered for the secondary outcomes, such as response and remission status and long term follow-up. Therefore, our results should be interpreted as preliminary and hypothesis-driven for future, pivotal trials. In addition, our findings might not be generalizable to BD patients on concurrent AD therapy, as such patients were not included in the study. Even though the use of this drug class in BD remains controversial (Pacchiarotti *et al*, 2013; Yatham *et al*, 2013), in most real-life clinical setting these drugs are often used.

Patients who presented >70% of MT of maximum stimulator output at baseline were not included. This criterion was employed because higher applied intensities produce more adverse events such as head and local pain. To reduce local side effects and increase tolerability, rTMS pivotal studies (eg, (George *et al*, 2010)) allow to progressively up-titrate stimulation intensity during or over several sessions. However, this approach was associated with lower dTMS efficacy in the pivotal dTMS depression trial (Levkovitz *et al*, 2015), possibly because sessions at intensities <120%MT are less effective (Levkovitz *et al*, 2009). As our study used a low sample size, we adopted the maximum MT eligibility criterion aiming to increase patients' adherence and the trial's internal validity. However, our findings are not generalizable to patients presenting MT > 70% at baseline, an issue that should be investigated in future studies.

Finally, we employed no neuronavigated methods for target localization. In fact, as the H1-coil produces electrical fields that are relatively non-focal and deep (Deng *et al*, 2013), the stimulated brain area in our study was probably widespread beneath the coil.

CONCLUSION

The present randomized, sham-controlled trial showed that deep TMS is a potentially effective and well-tolerated add-on therapy in resistant bipolar depressed patients receiving adequate pharmacotherapy. The effects were most evident immediately after the end of the acute treatment phase (20 dTMS sessions), and progressively faded away during the 4-week follow-up, which suggests that extended dTMS treatment might be necessary after the acute phase for an enduring response.

FUNDING AND DISCLOSURE

This work was primarily sponsored by Brainsway, which provided the dTMS devices and financial support. The sponsor had no role in the collection, management, analysis, and interpretation of the data; and also no role in the preparation, review, or approval of the manuscript. ARB receives grants from the 2012 FAPESP Young Research Award from the São Paulo State Foundation (Grant no. 20911-5), the Brain and Behavior research foundation (Grant

Number 20493) and a National Council for Scientific and Technological Development Grant (CNPq, Grant no. 470904). The Laboratory of Neuroscience receives financial support from the Beneficent Association Alzira Denise Hertzog da Silva. In the last 3 years, ZJD received research and equipment in-kind support for an investigator-initiated study through Brainsway Inc and Magventure Inc. ZJD has also served on the advisory board for Sunovion, Hoffmann-La Roche Limited and Merck and received speaker support from Eli Lilly. The authors declare no conflict of interest.

REFERENCES

- Berlim MT, van den Eynde F, Tovar-Perdomo S, Daskalakis ZJ (2014). Response, remission and drop-out rates following high-frequency repetitive transcranial magnetic stimulation (rTMS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. *Psychol Med* **44**: 225–239.
- Bersani FS, Minichino A, Enticott PG, Mazzarini L, Khan N, Antonacci G et al (2013). Deep transcranial magnetic stimulation as a treatment for psychiatric disorders: a comprehensive review. *Eur Psychiatry* **28**: 30–39.
- Bobo WV, Reilly-Harrington NA, Ketter TA, Brody BD, Kinrys G, Kemp DE et al (2015). Complexity of illness and adjunctive benzodiazepine use in outpatients with bipolar I or II disorder: results from the Bipolar CHOICE study. *J Clin Psychopharmacol* **35**: 68–74.
- Carpenter LL, Janicak PG, Aaronson ST, Boyadjis T, Brock DG, Cook IA et al (2012). Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice. *Depress Anxiety* **29**: 587–596.
- Castelein S, de Kort SJ, van der Moolen AE, Houtjes W, Roodbol PF, van Meijel B et al (2014). Good clinical practice and the maintenance of ethical standards in medical research: advice for junior researchers working in mental health care. *Tijdschr Psychiatr* **56**: 533–538.
- Cohen RB, Brunoni AR, Boggio PS, Fregni F (2010). Clinical predictors associated with duration of repetitive transcranial magnetic stimulation treatment for remission in bipolar depression: a naturalistic study. *J Nerv Ment Dis* **198**: 679–681.
- Deng ZD, Lisanby SH, Peterchev AV (2013). Electric field depth-focality tradeoff in transcranial magnetic stimulation: simulation comparison of 50 coil designs. *Brain Stimul* **6**: 1–13.
- Fitzgerald PB, Hoy KE, Elliot D, McQueen S, Wambeek LE, Daskalakis ZJ (2016). A negative double-blind controlled trial of sequential bilateral rTMS in the treatment of bipolar depression. *J Affect Disord* **198**: 158–162.
- George MS, Lisanby SH, Avery D, McDonald WM, Durkalski V, Pavlicova M et al (2010). Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch Gen Psychiatry* **67**: 507–516.
- Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes T, Cipriani A et al (2016). Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* **30**: 495–553.
- Grunze H, Vieta E, Goodwin GM, Bowden C, Licht RW, Moller HJ et al (2010). The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the biological treatment of bipolar disorders: update 2010 on the treatment of acute bipolar depression. *World J Biol Psychiatry* **11**: 81–109.
- Guy W. *ECDEU Assessment Manual for Psychopharmacology*. US Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs. Volume DHEW Publ No ADM 76.338. 1976; 218–222.
- Hamilton M (1959). The assessment of anxiety states by rating. *Br J Med Psychol* **32**: 50–55.
- Hamilton M (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry* **23**: 56–62.
- Harel EV, Zangen A, Roth Y, Reti I, Braw Y, Levkovitz Y (2011). H-coil repetitive transcranial magnetic stimulation for the treatment of bipolar depression: an add-on, safety and feasibility study. *World J Biol Psychiatry* **12**: 119–126.
- Holtzman JN, Lolic M, Ketter TA, Vazquez GH (2015). Clinical characteristics of bipolar disorder: a comparative study between Argentina and the United States. *Int J Bipol Disor* **3**: 8.
- Iovieno N, Nierenberg AA, Parkin SR, Hyung Kim DJ, Walker RS, Fava M et al (2016). Relationship between placebo response rate and clinical trial outcome in bipolar depression. *J Psychiatr Res* **74**: 38–44.
- Kedzior KK, Reitz SK, Azorina V, Loo C (2015). Durability of the antidepressant effect of the high-frequency repetitive transcranial magnetic stimulation (rTMS) in the absence of maintenance treatment in major depression: a systematic review and meta-analysis of 16 double-blind, randomized, sham-controlled trials. *Depress Anxiety* **32**: 193–203.
- Levkovitz Y, Harel EV, Roth Y, Braw Y, Most D, Katz LN et al (2009). Deep transcranial magnetic stimulation over the prefrontal cortex: evaluation of antidepressant and cognitive effects in depressive patients. *Brain Stimul* **2**: 188–200.
- Levkovitz Y, Isserles M, Padberg F, Lisanby SH, Bystritsky A, Xia G et al (2015). Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial. *World Psychiatry* **14**: 64–73.
- Li CT, Bai YM, Huang YL, Chen YS, Chen TJ, Cheng JY et al (2012). Association between antidepressant resistance in unipolar depression and subsequent bipolar disorder: cohort study. *Br J Psychiatry* **200**: 45–51.
- Luborzewski A, Schubert F, Seifert F, Danker-Hopfe H, Brakemeier EL, Schlattmann P et al (2007). Metabolic alterations in the dorsolateral prefrontal cortex after treatment with high-frequency repetitive transcranial magnetic stimulation in patients with unipolar major depression. *J Psychiatr Res* **41**: 606–615.
- Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K et al (2015). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Austr N Z J Psychiatry* **49**: 1087–1206.
- McDonald WM, Durkalski V, Ball ER, Holtzheimer PE, Pavlicova M, Lisanby SH et al (2011). Improving the antidepressant efficacy of transcranial magnetic stimulation: maximizing the number of stimulations and treatment location in treatment-resistant depression. *Depress Anxiety* **28**: 973–980.
- McGirr A, Karmani S, Arsappa R, Berlim MT, Thirthalli J, Muralidharan K et al (2016). Clinical efficacy and safety of repetitive transcranial magnetic stimulation in acute bipolar depression. *World Psychiatry* **15**: 85–86.
- Middleton H, Shaw I, Hull S, Feder G (2005). NICE guidelines for the management of depression. *BMJ* **330**: 267–268.
- Milev RV, Giacobbe P, Kennedy SH, Blumberger DM, Daskalakis ZJ, Downar J et al (2016). Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 4. Neurostimulation treatments. *Can J Psychiatry* **61**: 561–575.
- Pacchiarotti I, Bond DJ, Baldessarini RJ, Nolen WA, Grunze H, Licht RW et al (2013). The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. *Am J Psychiatry* **170**: 1249–1262.

- Parker GB, Graham RK (2017). Clinical characteristics associated with treatment-resistant bipolar disorder. *J Ner Ment Dis* **205**: 188–191.
- Perich T, Hadzi-Pavlovic D, Frankland A, Breakspear M, Loo C, Roberts G et al (2016). Are there subtypes of bipolar depression? *Acta Psychiatr Scand* **134**: 260–267.
- Perlis RH, Ostacher MJ, Miklowitz DJ, Smoller JW, Dennehy EB, Cowperthwait C et al (2010). Benzodiazepine use and risk of recurrence in bipolar disorder: a STEP-BD report. *J Clin Psychiatry* **71**: 194–200.
- Rapinesi C, Bersani FS, Kotzalidis GD, Imperatori C, Del Casale A, Di Pietro S et al (2015). Maintenance deep transcranial magnetic stimulation sessions are associated with reduced depressive relapses in patients with unipolar or bipolar depression. *Front Neurol* **6**: 16.
- Roth Y, Amir A, Levkovitz Y, Zangen A (2007). Three-dimensional distribution of the electric field induced in the brain by transcranial magnetic stimulation using figure-8 and deep H-coils. *J Clin Neurophysiol* **24**: 31–38.
- Sachs GS, Nierenberg AA, Calabrese JR, Marangell LB, Wisniewski SR, Gyulai L et al (2007). Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med* **356**: 1711–1722.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E et al (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* **59** (Suppl 20): 22–33.
- Sienaert P, Lambrichts L, Dols A, De Fruyt J (2013). Evidence-based treatment strategies for treatment-resistant bipolar depression: a systematic review. *Bipol Disord* **15**: 61–69.
- Stevley A, Dimairo M, Todd S, Julious SA, Nicholl J, Hind D et al (2015). An investigation of the shortcomings of the CONSORT 2010 Statement for the Reporting of Group Sequential Randomised Controlled Trials: a methodological systematic review. *PLoS One* **10**: e0141104.
- Tohen M, Frank E, Bowden CL, Colom F, Ghaemi SN, Yatham LN et al (2009). The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders. *Bipol Disord* **11**: 453–473.
- Tondo L, Vazquez GH, Baldessarini RJ (2014). Options for pharmacological treatment of refractory bipolar depression. *Curr Psychiatry Rep* **16**: 431.
- Williams JR (2008). The Declaration of Helsinki and public health. *Bull World Health Organ* **86**: 650–652.
- Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Beaulieu S, Alda M et al (2013). Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipol Disord* **15**: 1–44.
- Young RC, Biggs JT, Ziegler VE, Meyer DA (1978). A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* **133**: 429–435.
- Zimmerman M (2016). Screening for bipolar disorder: lessons not yet learned. *Evid Based Ment Health* **19**: e16.