Received: 5 April 2021 | Revised: 6 October 2021 | Accepted: 17 December 2021

DOI: 10.1002/da.23237

RESEARCH ARTICLE

ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA WILEY

Absence of early mood improvement as a robust predictor of rTMS nonresponse in major depressive disorder

Alex M. Mirman[1,2] | Juliana Corlier[1,2] | Andrew C. Wilson[1,2] | Reza Tadayonnejad[1,3] | Katharine G. Marder[1,2] | Christopher M. Pleman[1,2] | David E. Krantz[1,2] | Scott A. Wilke[1,2] | Jennifer G. Levitt[1,2] | Nathaniel D. Ginder[1,4] | Rashi Ojha[1,2] | Zafiris J. Daskalakis[5] | Andrew F. Leuchter[1,2] | Jonathan C. Lee[1,2]

[1]Neuromodulation Division, UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, California, USA

[2]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[3]Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California, USA

[4]VA Greater Los Angeles Healthcare System, Los Angeles, California, USA

[5]Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California, USA

Correspondence

Alex M. Mirman, Neuromodulation Division, Semel Institute for Neuroscience and Human Behavior, and the Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at UCLA, 760 Westwood Plaza, Los Angeles, CA 90024, USA.
Email: AMirman@mednet.ucla.edu

Abstract

Background: Symptoms of major depressive disorder (MDD) are reported to change early in treatment with repetitive transcranial magnetic stimulation (rTMS). We evaluated early changes in *sleep*, *anxiety*, and *mood* as predictors of nonresponse to rTMS treatment.

Methods: Three hundred twenty-nine subjects with nonpsychotic MDD completed a 6-week course of rTMS treatment. Subjects were stratified by the severity of their baseline depression, and had their overall depressive symptoms recorded every week of treatment. We evaluated lack of improvement in *sleep*, *anxiety*, and *mood* symptoms after 1 and 2 weeks as potential predictors of eventual nonresponse, defined as <50% improvement in compositive depressive symptoms after 6 weeks. This was measured as negative predictive value (NPV; the likelihood that lack of early symptom improvement accurately predicted eventual treatment nonresponse).

Results: Subjects with severe or very severe baseline depression achieving <20% improvement in *mood* at 1 week were correctly predicted as nonresponders with NPVs largely >90%. At 2 weeks, subjects with very severe baseline depression who failed to demonstrate any improvement in *mood* were all nonresponders. Lack of improvement in *sleep* at 2 weeks was also a significant predictor.

Conclusions: Identifying a lack of early *mood* improvement is a practical and robust method to predict rTMS nonresponse. This suggests a treatment protocol change may be indicated in patients with more severe baseline depression showing minimal early *mood* improvement.

KEYWORDS

clinical predictor, IDS-SR, major depressive disorder (MDD), nonresponse, PHQ-9, repetitive transcranial magnetic stimulation (rTMS), treatment outcome

1 | INTRODUCTION

Repetitive transcranial magnetic stimulation (rTMS) is increasingly used for patients suffering from treatment-resistant major depressive disorder (MDD). Research has demonstrated its safety, efficacy, and tolerability (Gaynes et al., 2014; George et al., 2010; O'Reardon et al., 2007; Perera et al., 2016). Therapeutic response, defined as a greater than 50% improvement from baseline depression, varies between 30% and 60% (Blumberger et al., 2018; Brakemeier et al., 2008; George et al., 2010).

Insurers typically approve 36 sessions for patients with MDD (Voigt et al., 2019). While this number may induce a therapeutic response in some, it may be insufficient for others (Yip et al., 2017). It is imperative to maximize benefit from the finite number of treatments which most patients can access (Demitrack, 2010). Early, accurate prediction of nonresponse could allow clinicians to change treatment approach, with potentially better outcomes. We recently reported that changing treatment parameters was associated with greater response in patients who did not achieve 20% improvement after 10 sessions (Lee et al., 2020).

The absence of improvement in composite scores of depression rating scales predicts nonresponse to several treatments (Bares et al., 2017; Lin et al., 2016). For SSRIs, <20% improvement after 2 weeks is strongly correlated with nonresponse, with negative predictive values (NPVs) >80% (Kudlow et al., 2014). A similar pattern exists for ECT (Lin et al., 2016). This "<20% by two weeks" is also a practical predictor of rTMS outcome. Patients who achieve <20% improvement in overall depressive symptoms by 2 weeks of rTMS were accurately predicted to be nonresponders with an NPV of 89.5% (Feffer et al., 2018).

Depressive symptoms are heterogeneous and may resolve at different times during treatment. Wardenaar and colleagues (Wardenaar et al., 2010) conducted a confirmatory factor analysis of the Inventory of Depressive Symptomatology Self Report (IDS-SR) (Rush et al., 1996) and identified three meaningfully discrete symptom subscales: *sleep*, *anxiety/arousal*, and *mood/cognition*. Few prior studies have examined early resolution of specific symptom clusters during rTMS or whether specific symptoms could predict outcome sooner than two weeks.

The current study was designed to expand our understanding of the predictive utility of early clinical response to rTMS. First, we examined clinical improvement at one week in addition to two weeks to determine if this might provide an earlier decision-point to modify treatment. Second, we examined not only baseline composite depression severity as measured by the IDS-SR, but also narrower subscales specifically measuring changes in *sleep*, *anxiety/arousal*, and *mood/cognition*. We aimed to determine whether this more granular approach to symptom assessment might identify specific symptom clusters relevant to predicting nonresponse. Third, we examined the clinically relevant measures of both NPV (the likelihood that a negative result correctly predicts nonresponse) and positive predictive value (PPV, the likelihood that a positive result correctly predicts response).

2 | METHODS

2.1 | Overview and patient population

This retrospective study included 329 patients treated with a full 6 weeks of rTMS in the UCLA TMS Clinical and Research Program between August 2015 and March 2020 for an episode of non-psychotic MDD. All subjects had failed to benefit adequately from at least three antidepressant trials with ATHF score of 3 or greater and the great majority had also failed to benefit from two augmentation medication trials (Sackeim et al., 2019). Subjects underwent pre- and posttreatment depressive symptom assessments. All subjects had a primary diagnosis of MDD confirmed on the MINI International Diagnostic Interview (MINI) (Sheehan et al., 1998). Subjects continued to receive previously prescribed psychotropic medications during rTMS. All subjects provided written consent to participate in this UCLA IRB-approved study and were treated in accordance with the 2013 Declaration of Helsinki.

2.2 | Clinical assessments

Symptoms were assessed at pretreatment baseline, after approximately every five treatment sessions, and at the end of treatment using the 30-item IDS-SR and Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001; Rush et al., 1986, 1996), self-rated instruments widely used in clinical settings (Rush et al., 2006). Total IDS-SR scores, ranging from 0 to 84, were used to measure depression severity with the primary outcome measure being percent improvement in IDS-SR score from baseline at the final treatment. Response was defined as a decline in pre- to posttreatment IDS-SR of ≥50%.

In addition to the overall IDS-SR, this study examined three validated symptom subscales: *sleep*, *mood/cognition*, and *anxiety/arousal* (Wardenaar et al., 2010). The *sleep* subscale includes questions 1–4; the *mood/cognition* subscale includes questions 5, 8, 10–12, 15–18, 20, 22, and 29; and the *anxiety/arousal* subscale includes questions 6, 23–28, and 30 of the IDS-SR. The sum of individual scores was calculated as the sum of their component questions and percent improvement in each subscale from baseline was calculated after five and ten treatments and evaluated as potential predictors of the final total IDS-SR score. Subjects with baseline *sleep* scores of 0 (n = 9, 2.1%) were excluded from the sleep subanalysis, as they could not demonstrate any improvement on the sleep subscale. No subjects had baseline *mood* scores of 0.

2.3 | rTMS procedures

All rTMS treatments were delivered with either a MagPro X100 (Magventure), Magstim Horizon (Magstim), Magstim Super Rapid2 (Magstim), or Neurostar (Neuronetics) device. Resting motor threshold (RMT) defined as the minimum stimulus intensity necessary to elicit an overt motor response in the right abductor pollicis brevis (APB) muscles for ≥50% of applied stimuli was determined for each participant before the first

TABLE 1 Demographic characteristics of subjects receiving rTMS to the left DLPFC[a]

Characteristic	Total	Nonresponders	Responders	Test statistic	*p*
N	423	235	94		
Female	223 (52.7%)	127 (54.0%)	54 (57.4%)	$\chi^2 = 2.94$	.23
Male	199 (47.0%)	108 (46.0%)	39 (41.5%)		
Transgender (F-to-M)	1 (0.2%)	0 (0.0%)	1 (1.1%)		
Age	46.0 ± 16.5	45.1 ± 16.0	46.2 ± 16.5	$t = 0.55$	.58
Pretreatment IDS-SR Score	43.2 ± 11.1	44.1 ± 10.8	41.0 ± 11.2	$t = 2.36$	.02
Pretreatment PHQ-9 Score	17.3 ± 5.3	17.6 ± 5.0	16.2 ± 5.4	$t = 2.17$	.03
Pretreatment *Sleep* Score	5.5 ± 2.4 (0–12)	5.5 ± 2.4 (0–12)	5.2 ± 2.5 (0–12)	$t = 0.90$	.37
Pretreatment *Anxiety* Score	10.1 ± 4.5 (0–24)	10.4 ± 4.6 (0–24)	9.2 ± 4.2 (1–19)	$t = 2.03$	.04
Pretreatment *Mood* Score	20.0 ± 5.0 (3–32)	20.4 ± 4.8 (3–31)	19.1 ± 5.0 (3–30)	$t = 2.21$	.03
% Improv in IDS-SR	33.7 ± 30.7%	19.7 ± 23.7%	68.6 ± 13.4%	$t = 18.82$	<.001*
% Improv in PHQ-9	34.5 ± 36.7%	21.1 ± 33.2%	70.2 ± 19.0%	$t = 12.70$	<.001*
% Improv in *Sleep*	21.6 ± 50.3%	10.3 ± 50.5%	51.0 ± 35.8%	$t = 6.95$	<.001*
% Improv in *Anxiety*	38.4 ± 35.4%	25.6 ± 31.8%	70.2 ± 20.4%	$t = 12.60$	<.001*
% Improv in *Mood*	32.0 ± 45.7%	16.4 ± 44.2%	71.1 ± 17.1%	$t = 11.66$	<.001*
# of Psychotropic Medications	2.6 ± 1.8	2.6 ± 1.8	2.5 ± 1.9	$t = 0.62$	.62

[a]Data are presented as mean ± standard deviation unless otherwise specified. The statistical significance of differences between responders and nonresponders was calculated using the χ^2 test for categorical gender and two-tailed *t* tests for continuous data. Given the number of comparisons, we used $^*p \leq .0038$ based on Bonferroni correction to avoid a Type I error. Medication data were missing for four subjects who completed treatment in full.

treatment. There was no difference in treatment outcome among the devices used (data not presented).

Subjects received 30 rTMS treatment sessions beginning with high frequency left (HFL) rTMS, consisting of 10 Hz stimulation (4-s trains, 26-s intertrain intervals, 75 trains, and 3000 pulses total) lasting 37.5 min daily to the left dorsolateral prefrontal cortex (L-DLPFC) using the Beam F3 localization method (Beam et al., 2009). We increased intensity to 120% RMT as tolerated over the first five treatments. Parameters could be adjusted anytime thereafter to optimize tolerability. Subjects completed IDS-SR ratings weekly. Those unable to tolerate 10 Hz stimulation after 10 treatments due to anxiety, agitation, pain, or worsening depressive symptoms could be transitioned to sequential bilateral treatment (Fitzgerald et al., 2006) or intermittent theta-burst stimulation priming (Lee et al., 2020).

2.4 | Data analysis

Data analyses were performed using SPSS, version 26. First, multivariate linear regression analyses were performed using percent improvement in *sleep*, *anxiety*, and *mood* subscales at 1 and 2 weeks (in separate analyses) as independent variables, baseline IDS-SR score as a covariate, and overall percent improvement in IDS-SR score as the dependent variable. Only the subscales significantly associated with the dependent variable were included in further analyses. Regression analyses were repeated using PHQ-9 percent improvement as the dependent variable to verify findings with an independent scale (Kroenke et al., 2001).

Second, we calculated the NPV of changes in *mood* and *sleep* subscales across different cut-off values ranging from 0% to 50% improvement at one and two weeks of treatment. To enhance the immediate clinical relevance of our predictive model and allow clinicians to make informed decisions based on baseline depression severity, we generated separate curves for subjects with moderate (26–38), severe (39–48), and very severe MDD (49–84) as measured by the IDS-SR (Rush et al., 1986). We did not generate curves representing subjects with mild depression severity (14–25) given the small number of subjects in this category. For the higher severity groups, the likelihood of nonresponse to treatment was examined across the range of percent improvement in each subscale after one and two weeks.

Lastly, we generated receiver operator characteristic (ROC) curves to assess how well early percent improvements in *sleep*, *mood*, and overall IDS-SR scores predicted the likelihood of response to rTMS. We assessed *mood*, *sleep*, and total IDS-SR scores independently at one and two weeks and calculated the area under the curve (AUC) for each model.

3 | RESULTS

3.1 | Demographics and baseline measurements

Four hundred twenty-three subjects (52.7% female, mean age 46.0 ± 16.5) were included in the study, 329 (78%) of whom completed their treatment course in full. Of these 329 subjects, 94 (28.6%) were responders. Nonresponders demonstrated significantly less improvement in *sleep*, *mood*, *anxiety*, and overall IDS-SR scores. There were no group differences between responders and nonresponders in gender, age, baseline depression severity on the IDS-SR, baseline subscale scores (*sleep*, *mood*, and *anxiety*), or number of concurrent psychotropic medications (Table 1).

3.2 | Linear regressions

Linear regression analyses revealed that after 1 week, baseline IDS-SR score and percent improvement in the *mood* subscale were the only significant predictors of percent IDS-SR total improvement at endpoint; percent improvement in *sleep* had a trend level for significance (Table 2). The overall regression model at week one significantly predicted outcome ($F_{(4,269)} = 17.24$, $p < .001$, $R^2 = .20$). After 2 weeks, percent improvement in both *sleep* and *mood* predicted percent IDS-SR improvement, while baseline IDS-SR score was no longer significant. The overall regression model at the two-week mark was significantly associated with percent IDS-SR improvement ($F_{(4,308)} = 28.13$, $p < .001$, $R^2 = .268$). We excluded the *anxiety/arousal* subscale from further analysis as it was unassociated with outcome in either regression analysis. We examined the variance inflation factor (VIF) for each variable examined in these models for evidence of multi-collinearity. VIFs greater than 2.5 are generally considered evidence of significant multi-collinearity (Johnston et al., 2018). All variables had VIFs less than 1.6 (data not shown).

We conducted additional linear regression analysis using the IDS-SR total score excluding *mood* and *sleep* subscale items to assess the predictive power of early changes in IDS-SR score independent of these factors. Early changes in this composite IDS-SR score excluding *mood* and *sleep* items were still predictive of outcome at both 1 week ($F_{(1,279)} = 87.57$, $p < .001$, $R^2 = .239$) and 2 weeks ($F_{(1,319)} = 75.11$, $p < .001$, $R^2 = .191$). This finding suggests the total of the remaining items on the IDS-SR are independently predictive of response.

Lastly, we conducted linear regression analysis utilizing baseline IDS-SR and percent improvement in *sleep*, *mood*, and *anxiety* to predict percent improvement in PHQ-9 scores (Table 2). Percent change in the *mood* subscale at one and two weeks predicted percent PHQ-9 improvement at endpoint but early change in *sleep* was not predictive.

3.3 | Negative predictive value analysis—sleep subscale

Subjects with severe or very severe baseline IDS-SR scores (NPVs > 75%) were more likely not to respond at endpoint than subjects with moderate baseline IDS-SR scores (NPVs < 70%) regardless of improvements in *sleep* quality (Figure 1). After 1 week, those with severe or very severe baseline IDS-SR scores whose *sleep* improved by <30% were predicted with 83%–88% certainty to be nonresponders; the NPV declined sharply beyond this 30% threshold (Figure 1a).

Analyses of *sleep* changes at Week 2 yielded similar results (Figure 1b). Subjects with severe or very severe baseline IDS-SR had a notably higher NPV (NPVs > 75%) than subjects with moderate

TABLE 2 Results of linear regression analysis examining subscale percent improvement after 1 and 2 weeks of treatment as predictors of percent improvement in IDS-SR and PHQ-9 score after completing treatment[a]

Predictor	Dependent variable	After 1 week		After 2 weeks	
		Coefficient	*p*	Coefficient	*p*
Sleep % Improv	IDS-SR	.06	.06	.09	.001*
	PHQ-9	.02	.70	.04	.32
Mood/Cognition % Improv	IDS-SR	.43	<.001*	.41	<.001*
	PHQ-9	.45	<.001*	.48	<.001*
Anxiety/Arousal % Improv	IDS-SR	.06	.23	.05	.24
	PHQ-9	.05	.44	−.01	.91
Pretreatment Total IDS-SR Score	IDS-SR	−.31	.03*	−.18	.16
	PHQ-9	−.26	.20	−.18	.35

[a]Baseline IDS-SR score is included as a covariate.

*$p \leq .05$.

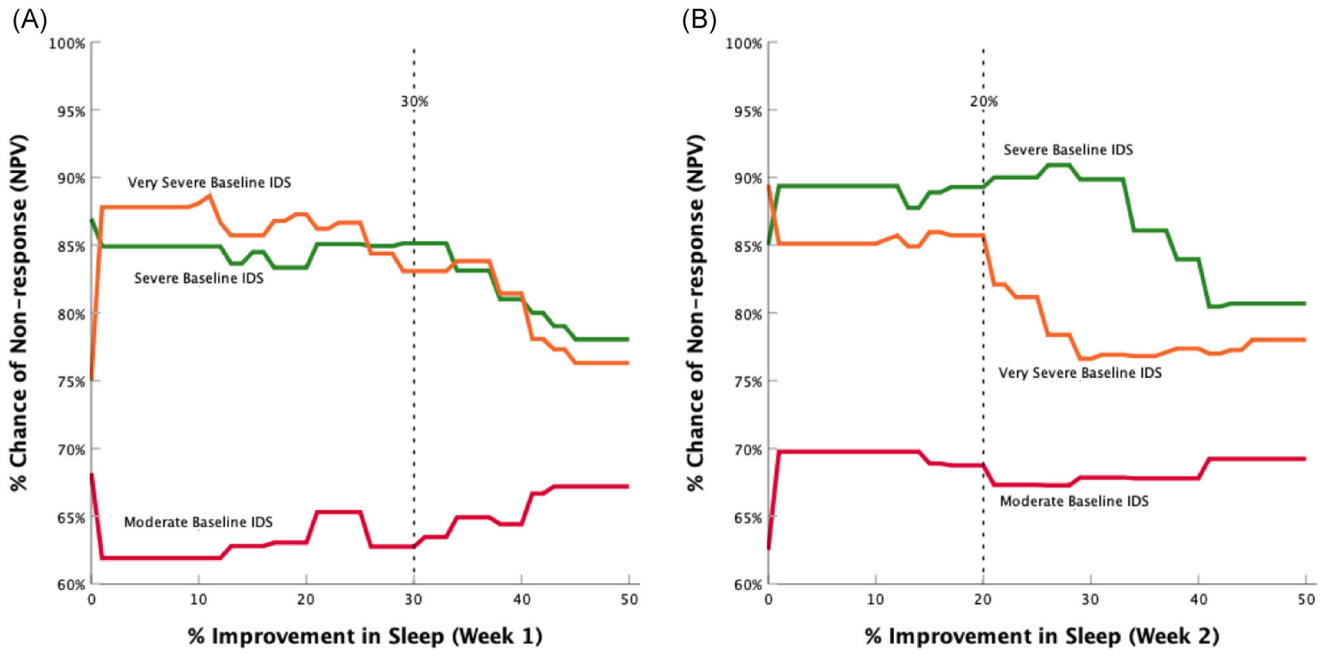

FIGURE 1 Plot illustrating the relationship between NPV and percent improvement in *sleep*, stratified by baseline IDS-SR severity, after 1 (a) or 2 weeks (b) of treatment

baseline IDS-SR (NPVs < 70%), regardless of percent improvement in their *sleep* subscale. Subjects with severe or very severe baseline IDS-SR whose sleep improved by <20% after 2 weeks could be categorized as nonresponders with 85%–90% certainty.

3.4 | Negative predictive value analysis—mood subscale

After 1 week, subjects with severe or very severe baseline IDS-SR scores (NPVs > 78%) more likely not to respond than subjects with moderate baseline IDS-SR scores (NPVs < 77%), regardless of improvements in *mood* subscale scores (Figure 2a). After 1 week, those with severe or very severe baseline IDS-SR scores whose *mood* improved by <20% could be predicted to be nonresponders to treatment with 90% certainty.

After 2 weeks, subjects with very severe baseline IDS-SR scores whose *mood* improved by <20% could be predicted to be nonresponders to treatment with an NPV > 92%, giving them the highest NPV of any subgroup (Figure 2b). Additionally, those with very severe baseline IDS-SR scores who failed to show any *mood* improvement by two weeks were all nonresponders (n = 19, 5.8%).

Similar results were seen for prediction of nonresponse as measured by the PHQ-9. The NPVs were lower than for IDS-SR outcomes but displayed a similar stepwise pattern of decreasing NPV with higher *mood* cutoff values (Figure 3). Subjects with severe or moderately severe baseline depression who experienced <20% improvement in *mood* by 1 week could be predicted to be nonresponders on the PHQ-9 with an NPV > 81%; a less than 20% improvement in *mood* in these groups by two weeks yielded NPVs >89%.

3.5 | ROC curve analysis

ROC curves for changes in *sleep*, *mood*, and total IDS-SR scores after 1 and 2 weeks are shown in Figure 4. After 1 week, the resulting AUCs were 63.1% for the *sleep* subscale, 70.5% for the *mood* subscale, and 71.6% for the IDS-SR total score. After 2 weeks, AUCs increased to 66.2%, 78.2%, and 78.4% for *sleep*, *mood*, and total IDS-SR, respectively.

3.6 | Prediction table for nonresponse

We generated a lookup table based on NPV graphs in Figures 1 and 2 (Table 3) to summarize clinically relevant information. The table provides the NPV based on percent improvement in *mood*, ranging from 0% to 50% improvement after one and two weeks, stratified by baseline IDS-SR severity. For example, a subject with a very severe baseline IDS-SR score that improved by 10% after 1 week would be predicted to be a nonresponder with 92.7% certainty.

4 | DISCUSSION

We evaluated the predictive power of early improvement in *sleep*, *anxiety*, and *mood* subscales of the IDS-SR during 6 weeks of rTMS. Subjects with severe or very severe baseline IDS-SR scores whose *mood* subscale score improved <20% by Week 1 were highly likely to be nonresponders at 6 weeks. Change in *mood* subscale was superior to change in *sleep* as a response predictor. Early change in *anxiety* was not predictive. Predictions based on *mood* closely approximated

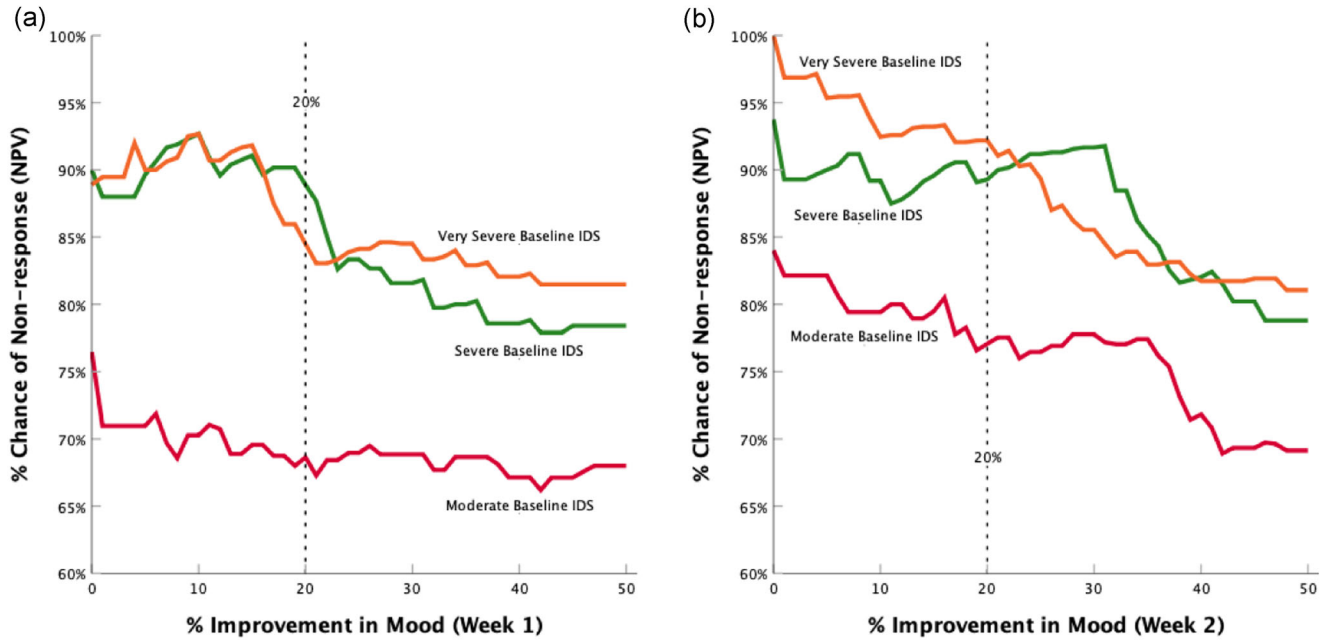

FIGURE 2 Plot illustrating the relationship between NPV and percent improvement in *mood*, stratified by baseline mood severity, after 1 week (a) or 2 weeks (b) of treatment

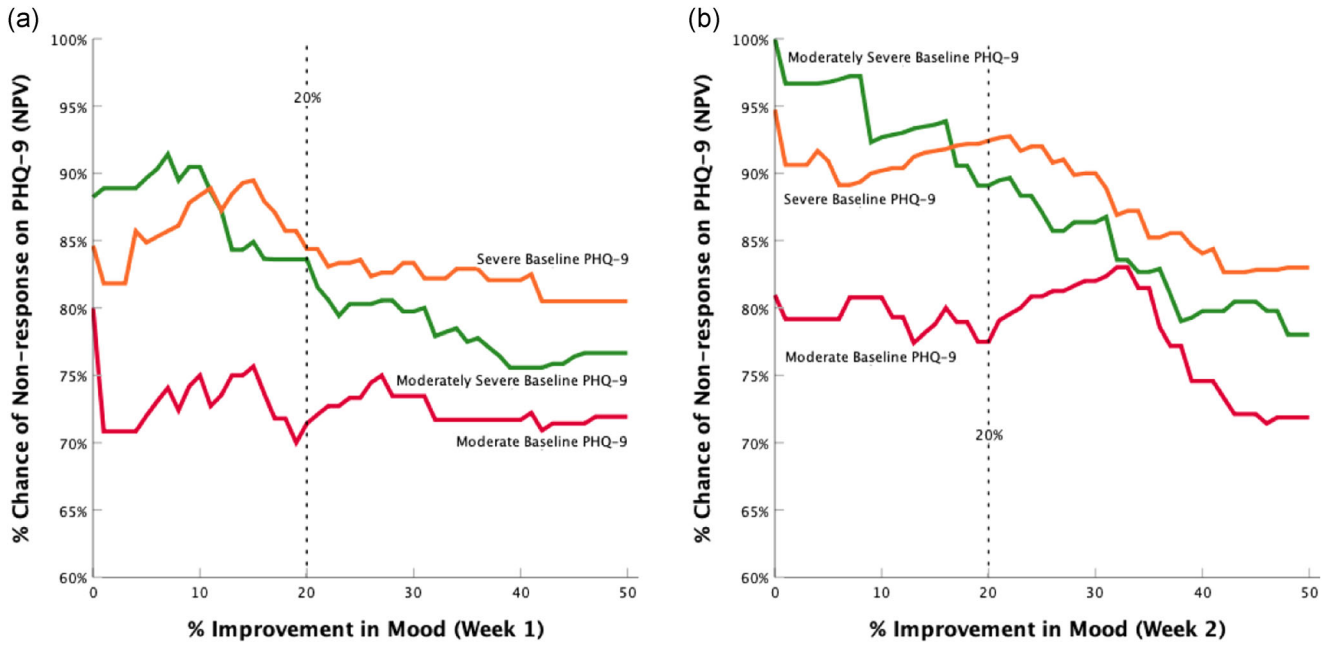

FIGURE 3 Plot illustrating the relationship between NPV based on PHQ-9 scores and percent improvement in *mood*, stratified by baseline PHQ-9 severity, after 1 week (a) or 2 weeks (b) of treatment

those based on total IDS-SR score, suggesting that simple *mood* items could be an easier and more rapidly assessed predictor of response.

Our study extends prior research that has examined early clinical changes in two ways. First, we considered the NPV of specific symptom clusters—*sleep*, *anxiety*, and *mood*—and stratified subjects by severity of baseline depression. Second, we examined the predictive power of symptom changes at one week instead of two, as in previous studies (Beck et al., 2020; Feffer et al., 2018). This delimited approach permitted accurate response prediction as early as one week into treatment.

To our knowledge, few studies have specifically explored early *sleep* changes on the IDS-SR as a predictor of rTMS non-response. Brakemeier et al. (2007) found that changes in sleep on the HAMD-24 predicted response but did not replicate this

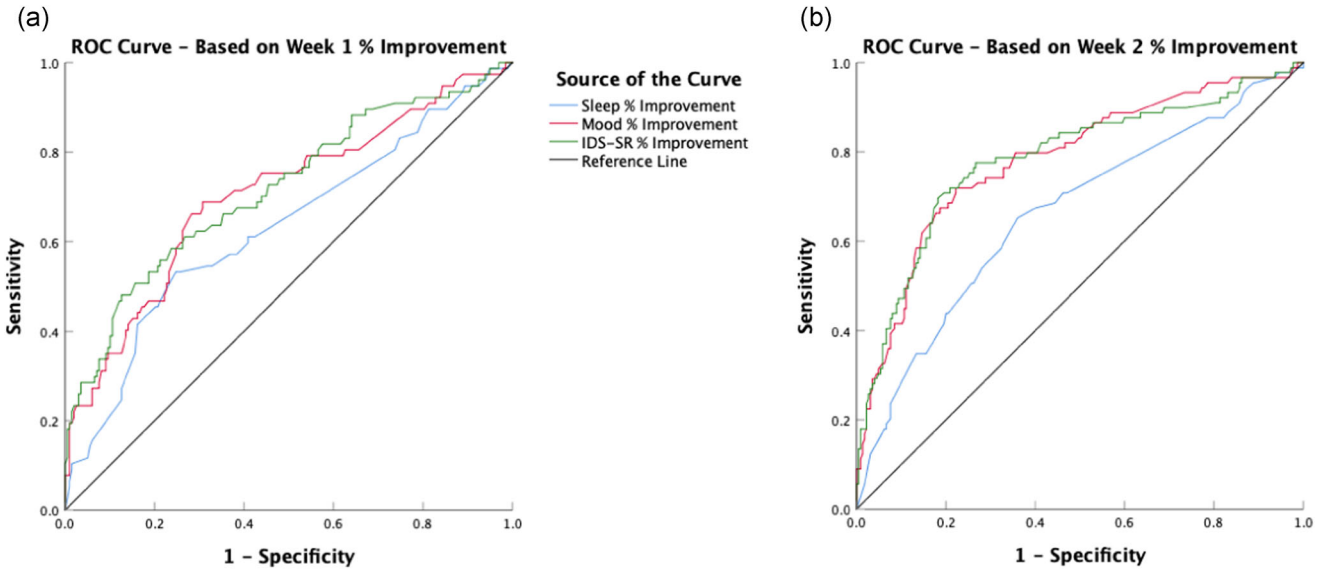

FIGURE 4 ROC curves of response to treatment using percent improvement in *sleep*, *mood*, and overall IDS-SR score after (a) 1 week and (b) 2 weeks of treatment as independent variables

TABLE 3 Lookup table for estimating the likelihood of rTMS nonresponse (<50% improvement) after a full course of treatment based on percent improvement in mood subscale score after one or two weeks of treatment, stratified by baseline IDS-SR severity

	Likelihood of nonresponse to treatment					
	Week 1			Week 2		
	Baseline IDS-SR severity			Baseline IDS-SR severity		
Mood Subscale % Improvement (1 or 2 weeks)	Moderate	Severe	Very severe	Moderate	Severe	Very severe
0%	76.5%	90.0%	88.9%	84.0%	93.8%	100.0%
5%	71.0%	89.7%	90.0%	82.1%	90.0%	95.3%
10%	70.3%	92.7%	92.7%	79.4%	89.2%	92.5%
15%	69.6%	91.1%	91.8%	79.5%	89.6%	93.2%
20%	68.6%	88.9%	84.5%	77.1%	89.3%	92.2%
25%	69.0%	83.3%	84.1%	76.5%	91.2%	89.3%
30%	68.9%	81.6%	84.5%	77.8%	91.7%	85.5%
35%	68.7%	80.0%	82.9%	77.4%	85.2%	83.0%
40%	67.1%	78.6%	82.1%	71.8%	82.0%	81.7%
45%	67.1%	78.4%	81.5%	69.3%	80.2%	81.9%
50%	68.0%	78.4%	81.5%	69.1%	78.8%	81.1%

finding in subsequent analysis (Brakemeier et al., 2008). Our study demonstrates that simultaneously considering baseline IDS-SR score and improvement in *sleep* after 2 weeks can provide a nuanced understanding of each subject's likelihood of nonresponse. Subjects with severe or very severe baseline IDS-SR scores and minimal early *sleep* improvement were most likely to be nonresponders (NPVs between 85% and 90%). However, it should be noted that given the small number of questions comprising the *sleep* score, its dynamic range (and thus utility as a predictor) is limited.

Our findings confirm results of prior research showing that early *mood* changes may predict outcome (Grunhaus et al., 2002). Our findings indicate that *mood* subscale items are a more clinically useful predictor than *sleep*. First, *mood* subscale changes were predictive of nonresponse by 1 week, a week earlier than *sleep* and prior analyses at 2 weeks using composite depression scores (Feffer et al., 2018). Additionally, as in our ROC curve analysis, *mood* closely approximates the overall IDS-SR score as a nonresponse predictor. This finding suggests that *mood* symptoms are a practical and efficient substitute for total IDS-SR,

although AUCs for both models were modest, particularly after the first week. Lastly, NPVs associated with early *mood* changes exceeded those associated with *sleep*, especially in subjects with severe or very severe baseline depression and minimal *mood* improvement. Perhaps the most clinically informative finding in our study is that <20% *mood* improvement by 1 week yields NPVs greater than 90% in those with severe baseline depression. Though this "<20% improvement rule" has traditionally been applied after two weeks of treatment, it appears valid after 1 week of treatment for changes in the *mood* subscale (Feffer et al., 2018; Kudlow et al., 2014).

The predictors in this study were derived from the IDS-SR, a well-validated, widely used self-rated depression scale (Rush et al., 1986, 1996; Wardenaar et al., 2010). One potential issue of this study is the use of subscales (i.e., *mood* and *sleep*) as predictors of the composite variable from which they are derived (i.e., percentage improvement on the final total IDS-SR). To address this concern, we conducted three additional analyses: (1) evaluating for multicollinearity using a VIF analysis; (2) assessing the predictive power of these clinical measures on the IDS-SR scale with the subscale questions removed; and (3) evaluating subscales as predictors of a different composite outcome measure (i.e., the PHQ-9). There was no evidence of multicollinearity as all VIFs were lower than a commonly held threshold (i.e., VIFs > 2.5) denoting significant multicollinearity. Moreover, the total IDS-SR remained a significant predictor of treatment response even when the *mood* and *sleep* items were excluded from the total score. Finally, *mood* improvement predicted outcomes on the PHQ-9 score. These analyses provide substantial evidence that the symptom measures examined here are robust, distinct predictors of clinical outcome, irrespective of outcome measure.

Predicting outcome with an early response is well-documented in the literature for antidepressant medication, psychotherapy, ECT, and rTMS (Lin et al., 2016; Schlagert & Hiller, 2017; Szegedi et al., 2009; van Calker et al., 2009). Notably, early improvements are associated with consistently higher NPVs than PPVs when predicting outcomes from medication or rTMS (Feffer et al., 2018; Hicks et al., 2019); that is, it is considerably easier to predict nonresponse than response. High NPVs could help clinicians identify patients who are unlikely to respond and might benefit from a change in rTMS treatment protocol. Though a high PPV is also useful, those values are consistently below the clinically relevant threshold of 75% (Krepel et al., 2019; Li et al., 2012). Thus, PPV is not discussed at length in this study. Our supplementary lookup table can be used by providers to estimate the likelihood of nonresponse based on early *mood* changes and baseline depression severity (Table 3).

Whether baseline depression severity predicts rTMS outcomes remains unresolved in the literature (Brakemeier et al., 2007, 2008; Fitzgerald et al., 2016; Trevizol et al., 2020). While baseline depression severity as measured on the IDS-SR was not significantly different between responders and nonresponders (Table 1), it was a significant factor contributing to the prediction of outcome when included in the same model as Week 1 but not Week 2 clinical scores (Table 2). Baseline *anxiety*, *sleep*, and *mood* severity alone were not independent predictors of outcome (data not presented). *Mood* and *sleep* findings presented here are both treatment-emergent measures—that is, their predictive significance was only demonstrated after initiating rTMS. These findings indicate that while overall baseline depression severity could be relevant to treatment outcome, treatment-emergent changes in *specific* symptom clusters provide an additional dimension to predictive power.

The results of this study should be interpreted in the context of several limitations. First, rTMS treatment parameters in the current study were adjusted based upon the measurement of depression symptoms and subject tolerability. It is unknown how these changes impacted our results. Second, we examined likelihood of response only after 30 treatments, which is the most commonly used length of treatment in the United States. We did not account for possible "late responders" who might show significant improvement after a greater number of treatments that might yield higher response rates (May & Pridmore, 2019; Silverstein et al., 2015). Future studies should examine subjects receiving >30 treatments to explore this possibility. Third, in our predictions of nonresponse to treatment, we did not take into consideration different response trajectories (Kaster et al., 2019), but instead used the complementary approach of assessing the likelihood of nonresponse as a function of severity. Future studies should consider the rate and timing of response. Fourth, it should be noted that the response rate in this study is lower than that reported in some other clinical study populations (Sackeim et al., 2020). There are several factors that might account for this difference. As discussed above, our population of patients was highly treatment resistant, and our subjects reported severe baseline depressive symptoms (mean baseline IDS-SR score of 43.2). The average subject was in the "severe" category (Rush et al., 1986) with 67% of our subjects having baseline depression severities of "severe" or "very severe." Studies have shown that subjects that start with higher severity illness do not improve to the same level as less severe subjects (Blumberger et al., 2018). Additionally, these subjects were treated in the setting of an academic medical center with the majority of subjects reporting pain or other medical comorbidities that have been shown to increase treatment resistance (Corlier et al., 2021; Leuchter et al., 2010). Academic medical center populations that might be more treatment refractory have largely been excluded from some large registry studies (Sackeim et al., 2020). While the present report therefore may not be entirely representative of community samples, it does focus on a sample of subjects who are in the greatest need of rTMS or other advanced treatments for MDD. Lastly, several different TMS devices were used to deliver treatment (i.e., Magventure, Magstim, and Neurostar). The impact of device type on clinical effectiveness remains unclear, though, our available data did not reveal significant differences in outcome based on device used.

In conclusion, tracking *mood* items on the IDS-SR is a practical and robust method to predict rTMS nonresponse as early as 1 week into the standard 6-week course of treatment typically approved by insurers. Our prior work suggests that in subjects showing limited improvement after 2 weeks of treatment, changing rTMS approach improved outcomes (Lee et al., 2020). Together, these studies provide a rationale for TMS clinicians to consider changing rTMS approach in people with severe baseline depression showing no *mood* improvements after 1 week. This approach could lessen the burden of futile treatment and optimize efficacy in those severely depressed individuals who have most to gain from rTMS.

ACKNOWLEDGMENTS

This project was made possible by the Ryan Family Fund for TMS Research. The authors would like to acknowledge the patients and their families who received treatment in the clinic during the study period, and the dedication of the technicians and administrative team that have made this study possible. They would also like to thank the Ryan Family for their generous support of innovative approaches to depression treatment and of groundbreaking TMS technology. Their contributions have advanced the university's education and research missions.

CONFLICT OF INTERESTS

Andrew C. Wilson has served as a consultant to HeartCloud, Inc., within the past 36 months. Zafiris J. Daskalakis reports grants from Magventure, Inc., and grants from Brainsway, Inc., during the conduct of the study. Andrew F. Leuchter discloses that within the past 36 months, he has received research support from the National Institutes of Health, Department of Defense, CHDI Foundation, and NeuroSigma, Inc. He has served as a consultant to NeoSync, Inc., Ionis Pharmaceuticals, Inc., and ElMindA. He is a Chief Scientific Officer of Brain Biomarker Analytics LLC (BBA). He has equity interest in BBA. Jonathan C. Lee has received in-kind equipment support from Magventure, Inc. The other authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

Research data are not shared.

REFERENCES

Bares, M., Novak, T., Brunovsky, M., Kopecek, M., & Höschl, C. (2017). The comparison of effectiveness of various potential predictors of response to treatment with SSRIs in patients with depressive disorder. *Journal of Nervous and Mental Disease*, *205*(8), 618–626. https://doi.org/10.1097/NMD.0000000000000574

Beam, W., Borckardt, J. J., Reeves, S. T., & George, M. S. (2009). An efficient and accurate new method for locating the F3 position for prefrontal TMS applications. *Brain Stimulation*, *2*, 50–54. https://doi.org/10.1016/j.brs.2008.09.006

Beck, Q. M., Tirrell, E., Fukuda, A. M., Kokdere, F., & Carpenter, L. L. (2020). Can early treatment response serve as a predictor of antidepressant outcome of repetitive transcranial magnetic stimulation? *Brain Stimulation*, *13*(2), 420–421. https://doi.org/10.1016/j.brs.2019.12.002

Blumberger, D. M., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., Knyahnytska, Y., Kennedy, S. H., Lam, R. W., Daskalakis, Z. J., & Downar, J. (2018). Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): A randomised non-inferiority trial. *The Lancet*, *391*(10131), 1683–1692. https://doi.org/10.1016/S0140-6736(18)30295-2

Brakemeier, E. L., Wilbertz, G., Rodax, S., Danker-Hopfe, H., Zinka, B., Zwanzger, P., Grossheinrich, N., Várkuti, B., Rupprecht, R., Bajbouj, M., & Padberg, F. (2008). Patterns of response to repetitive transcranial magnetic stimulation (rTMS) in major depression: Replication study in drug-free patients. *Journal of Affective Disorders*, *108*(1–2), 59–70. https://doi.org/10.1016/j.jad.2007.09.007

Brakemeier, E. L., Luborzewski, A., Danker-Hopfe, H., Kathmann, N., & Bajbouj, M. (2007). Positive predictors for antidepressive response to prefrontal repetitive transcranial magnetic stimulation (rTMS). *Journal of Psychiatric Research*, *41*(5), 395–403. https://doi.org/10.1016/j.jpsychires.2006.01.013

Corlier, J., Tadayonnejad, R., Wilson, A. C., Lee, J. C., Marder, K. G., Ginder, N. D., Wilke, S. A., Levitt, J., Krantz, D., & Leuchter, A. F. (2021). Repetitive transcranial magnetic stimulation treatment of major depressive disorder and comorbid chronic pain: Response rates and neurophysiologic biomarkers. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291721002178

Demitrack, M. A. (2010). Transcranial magnetic stimulation for the treatment of major depression: Clinical, economic, and practical issues. *Psychopharm Review*, *45*(4), 25–30. https://doi.org/10.1097/01.IDT.0000369508.83887.db

Feffer, K., Lee, H. H., Mansouri, F., Giacobbe, P., Vila-Rodriguez, F., Kennedy, S. H., Daskalakis, Z. J., Blumberger, D. M., & Downar, J. (2018). Early symptom improvement at 10 sessions as a predictor of rTMS treatment outcome in major depression. *Brain Stimulation*, *11*(1), 181–189. https://doi.org/10.1016/j.brs.2017.10.010

Fitzgerald, P. B., Benitez, J., de Castella, A., Daskalakis, Z. J., Brown, T. L., & Kulkarni, J. (2006). A randomized, controlled trial of sequential bilateral repetitive transcranial magnetic stimulation for treatment-resistant depression. *American Journal of Psychiatry*, *163*, 88–94. https://doi.org/10.1176/appi.ajp.163.1.88

Fitzgerald, P. B., Hoy, K. E., Anderson, R. J., & Daskalakis, Z. J. (2016). A study of the pattern of response to rTMS treatment in depression. *Depression and Anxiety*, *33*(8), 746–753. https://doi.org/10.1002/da.22503

Gaynes, B. N., Lloyd, S. W., Lux, L., Gartlehner, G., Hansen, R. A., Brode, S., Jonas, D. E., Evans, T. S., Viswanathan, M., & Lohr, K. N. (2014). Repetitive transcranial magnetic stimulation for treatment-resistant depression: A systematic review and meta-analysis. *Journal of Clinical Psychiatry*, *75*(5), 477–489. https://doi.org/10.4088/JCP.13r08815

George, M. S., Lisanby, S. H., Avery, D., McDonald, W. M., Durkalski, V., Pavlicova, M., Anderson, B., Nahas, Z., Bulow, P., Zarkowski, P., Holtzheimer, P. E., Schwartz, T., & Sackeim, H. A. (2010). Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: A sham-controlled randomized trial. *Archives of General Psychiatry*, *67*(5), 507–516. https://doi.org/10.1001/archgenpsychiatry.2010.46

Grunhaus, L., Dolberg, O. T., Polak, D., & Dannon, P. N. (2002). Monitoring the response to rTMS in depression with visual analog scales. *Human Psychopharmacology: Clinical and Experimental*, *17*(7), 349–352. https://doi.org/10.1002/hup.418

Hicks, P. B., Sevilimedu, V., Johnson, G. R., Tal, I., Chen, P., Davis, L. L., Vertrees, J. E., Mohamed, S., & Zisook, S. (2019). Predictability of nonremitting depression after first 2 weeks of antidepressant treatment:

A VAST-D trial report. *Psychiatric Research and Clinical Practice*, *1*(2), 58–67. https://doi.org/10.1176/appi.prcp.20190003

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: A cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality and Quantity*, *52*(4), 1957–1976. https://doi.org/10.1007/s11135-017-0584-6

Kaster, T. S., Downar, J., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., Knyahnytska, Y., Kennedy, S. H., Lam, R. W., Daskalakis, Z. J., & Blumberger, D. M. (2019). Trajectories of response to dorsolateral prefrontal rTMS in major depression: A THREE-D study. *The American Journal of Psychiatry*, *176*(5), 367–375. https://doi.org/10.1176/APPI.AJP.2018.18091096

Krepel, N., Sack, A., Rush, A., & Arns, M. (2019). Clinical and neurophysiological predictors of rTMS response in Major Depressive Disorder: Robustness and clinical relevance. *Brain Stimulation*, *12*(2), 424–272. https://doi.org/10.1016/j.brs.2018.12.374

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kudlow, P. A., McIntyre, R. S., & Lam, R. W. (2014). Early switching strategies in antidepressant non-responders: Current evidence and future research directions. *CNS Drugs*, *28*(7), 601–609. https://doi.org/10.1007/s40263-014-0171-5

Lee, J. C., Wilson, A. C., Corlier, J., Tadayonnejad, R., Marder, K. G., Pleman, C. M., Krantz, D. E., Wilke, S. A., Levitt, J. G., Ginder, N. D., & Leuchter, A. F. (2020). Strategies for augmentation of high-frequency left-sided repetitive transcranial magnetic stimulation treatment of major depressive disorder. *Journal of Affective Disorders*, 277, 964–969. https://doi.org/10.1016/j.jad.2020.09.011

Leuchter, A. F., Husain, M. M., Cook, I. A., Trivedi, M. H., Wisniewski, S. R., Gilmer, W. S., Luther, J. F., Fava, M., & Rush, A. J. (2010). Painful physical symptoms and treatment outcome in major depressive disorder: A STAR*D (Sequenced Treatment Alternatives to Relieve Depression) report. *Psychological Medicine*, *40*(2), 239–251. https://doi.org/10.1017/S0033291709006035

Li, J., Kuk, A. Y. C., & Rush, A. J. (2012). A practical approach to the early identification of antidepressant medication non-responders. *Psychological Medicine*, *42*(2), 309–316. https://doi.org/10.1017/S0033291711001280

Lin, C. H., Chen, M. C., Yang, W. C., & Lane, H. Y. (2016). Early improvement predicts outcome of major depressive patients treated with electroconvulsive therapy. *European Neuropsychopharmacology*, *26*(2), 225–233. https://doi.org/10.1016/j.euroneuro.2015.12.019

May, T., & Pridmore, S. (2019). Impact of transcranial magnetic stimulation on the symptom profile of major depressive episode. *Australasian Psychiatry*, *27*(3), 297–301. https://doi.org/10.1177/1039856219828134

O'Reardon, J. P., Solvason, H. B., Janicak, P. G., Sampson, S., Isenberg, K. E., Nahas, Z., McDonald, W. M., Avery, D., Fitzgerald, P. B., Loo, C., Demitrack, M. A., George, M. S., & Sackeim, H. A. (2007). Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: A multisite randomized controlled trial. *Biological Psychiatry*, *62*(11), 1208–1216. https://doi.org/10.1016/j.biopsych.2007.01.018

Perera, T., George, M. S., Grammer, G., Janicak, P. G., Pascual-Leone, A., & Wirecki, T. S. (2016). The Clinical TMS Society consensus review and treatment recommendations for TMS therapy for major depressive disorder. *Brain Stimulation*, *9*(3), 336–346. https://doi.org/10.1016/j.brs.2016.03.010

Rush, A. J., Bernstein, I. H., Trivedi, M. H., Carmody, T. J., Wisniewski, S., Mundt, J. C., Shores-Wilson, K., Biggs, M. M., Woo, A., Nierenberg, A. A., & Fava, M. (2006). An evaluation of the Quick Inventory Of Depressive Symptomatology and the Hamilton Rating Scale for Depression: A Sequenced Treatment Alternatives to Relieve Depression trial report. *Biological Psychiatry*, *59*(6), 493–501. https://doi.org/10.1016/j.biopsych.2005.08.022

Rush, A. J., Giles, D. E., Schlesser, M. A., Fulton, C. L., Weissenburger, J., & Burns, C. (1986). The inventory for depressive symptomatology (IDS): Preliminary findings. *Psychiatry Research*, *18*(1), 65–87. https://doi.org/10.1016/0165-1781(86)90060-0

Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B., & Trivedi, M. H. (1996). The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychological Medicine*, *26*(3), 477–486. https://doi.org/10.1017/s0033291700035558

Sackeim, H. A., Aaronson, S. T., Bunker, M. T., Conway, C. R., Demitrack, M. A., George, M. S., Prudic, J., Thase, M. E., & Rushi, A. J. (2019). The assessment of resistance to antidepressant treatment: Rationale for the Antidepressant Treatment History Form: Short Form (ATHF-SF). *Journal of Psychiatric Research*, *113*, 125–136. https://doi.org/10.1016/J.JPSYCHIRES.2019.03.021

Sackeim, H. A., Aaronson, S. T., Carpenter, L. L., Hutton, T. M., Mina, M., Pages, K., Verdoliva, S., & West, W. S. (2020). Clinical outcomes in a large registry of patients with major depressive disorder treated with transcranial magnetic stimulation. *Journal of Affective Disorders*, *277*, 65–74. https://doi.org/10.1016/J.JAD.2020.08.005

Schlagert, H. S., & Hiller, W. (2017). The predictive value of early response in patients with depressive disorders. *Psychotherapy Research*, *27*(4), 488–500. https://doi.org/10.1080/10503307.2015.1119329

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(Suppl 20), 22–33.

Silverstein, W. K., Noda, Y., Barr, M. S., Vila-Rodriguez, F., Rajji, T. K., Fitzgerald, P. B., Downar, J., Mulsant, B. H., Vigod, S., Daskalakis, Z. J., & Blumberger, D. M. (2015). Neurobiological predictors of response to dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation in depression: a systematic review. *Depression and Anxiety*, *32*(12), 871–891. https://doi.org/10.1002/da.22424

Szegedi, A., Jansen, W., van Willigenburg, A., van der Meulen, E., Stassen, H., & Thase, M. (2009). Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder. *The Journal of Clinical Psychiatry*, *70*(3), 344–353.

Trevizol, A. P., Downar, J., Vila-Rodriguez, F., Thorpe, K. E., Daskalakis, Z. J., & Blumberger, D. M. (2020). Predictors of remission after repetitive transcranial magnetic stimulation for the treatment of major depressive disorder: An analysis from the randomised non-inferiority THREE-D trial. *EClinicalMedicine*, *22*, 100349. https://doi.org/10.1016/j.eclinm.2020.100349

van Calker, D., Zobel, I., Dykierek, P., Deimel, C. M., Kech, S., Lieb, K., Berger, M., & Schramm, E. (2009). Time course of response to antidepressants: Predictive value of early improvement and effect of additional psychotherapy. *Journal of Affective Disorders*, *114*(1–3), 243–253.

Voigt, J., Carpenter, L., & Leuchter, A. (2019). A systematic literature review of the clinical efficacy of repetitive transcranial magnetic stimulation (rTMS) in non-treatment resistant patients with major depressive disorder. *BMC Psychiatry*, *19*(1), 13. https://doi.org/10.1186/s12888-018-1989-z

Wardenaar, K. J., van Veen, T., Giltay, E. J., den Hollander-Gijsman, M., Penninx, B. W. J. H., & Zitman, F. G. (2010). The structure and dimensionality of the Inventory of Depressive Symptomatology Self Report (IDS-SR) in patients with depressive disorders and healthy controls. *Journal of Affective Disorders*, *125*(1–3), 146–154. https://doi.org/10.1016/j.jad.2009.12.020

Yip, A. G., George, M. S., Tendler, A., Roth, Y., Zangen, A., & Carpenter, L. L. (2017). 61% of unmedicated treatment resistant depression patients who did not respond to acute TMS treatment responded after four weeks of twice weekly deep TMS in the Brainsway pivotal trial. *Brain Stimulation*, *10*(4), 847–849. https://doi.org/10.1016/j.brs.2017.02.013

How to cite this article: Mirman, A. M., Corlier, J., Wilson, A. C., Tadayonnejad, R., Marder, K. G., Pleman, C. M., Krantz, D. E., Wilke, S. A., Levitt, J. G., Ginder, N. D., Ojha, R., Daskalakis, Z. J., Leuchter, A. F., & Lee, J. C. (2022). Absence of early mood improvement as a robust predictor of rTMS nonresponse in major depressive disorder. *Depression and Anxiety*, 39, 123–133. https://doi.org/10.1002/da.23237