Journal of Affective Disorders 405 (2026) 121631

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Utilizing stimulation-evoked hemodynamic activity to predict antidepressant response to intermittent theta-burst stimulation in adults with major depression

Rebecca L.D. Kan [a], Alvin H.P. Tang [a], Penny P. Qin [a], Minxia Jin [a,b], Adam W.L. Xia [a], Bella B.B. Zhang [a], Tim T.Z. Lin [a], Sharie X. Wang [a], Jessie J. Lin [a], Michael K. Yeung [c,d], Sherry K.W. Chan [e,f], Fan Li [g], Fidel Vila-Rodriguez [h,i], Kenneth N.K. Fong [a], Frank Padberg [j,k,l], Georg S. Kranz [a,m,n,*]

[a] *Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China*
[b] *Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Shanghai, China*
[c] *Department of Psychology, The Education University of Hong Kong, Hong Kong, China*
[d] *University Research Facility of Human Behavioral Neuroscience, The Education University of Hong Kong, Hong Kong, China*
[e] *Department of Psychiatry, School of Clinical Medicine, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China*
[f] *Department of Psychiatry, Queen Mary Hospital, Hong Kong, China*
[g] *Department of Biostatistics, Yale School of Public Health, Yale University, New Haven, Connecticut, USA*
[h] *Non-invasive Neurostimulation Therapies Laboratory, Department of Psychiatry, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada*
[i] *School of Biomedical Engineering, Faculty of Applied Science/Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada*
[j] *Department of Psychiatry and Psychotherapy, LMU University Hospital, LMU Munich, Munich, Germany*
[k] *DZPG (German Center for Mental Health), Partner Site Munich-Augsburg, Germany*
[l] *Center for Non-invasive Brain Stimulation Munich-Augsburg, Munich, Germany*
[m] *Mental Health Research Center (MHRC), The Hong Kong Polytechnic University, Hong Kong, China*
[n] *University Research Facility in Behavioral and Systems Neuroscience (UBSN), The Hong Kong Polytechnic University, Hong Kong, China*

ABSTRACT

Background: The instantaneous neural response to prefrontal theta burst stimulation (TBS) may serve as predictive marker for antidepressant treatment success. This study aimed to (1) assess whether baseline theta burst stimulation (TBS)-induced prefrontal hemodynamic responses can predict treatment outcome of four weeks of TBS in adults with major depressive disorder (MDD); (2) assess the test-retest reliability of TBS-induced hemodynamic responses.
Methods: Forty-four MDD participants were recruited and underwent two consecutive-day concurrent TBS/functional near-infrared spectroscopy (fNIRS) measurements. Participants then received four weeks intermittent TBS (iTBS) treatment. An additional 45 healthy controls (HCs) were recruited for the test-retest reliability analysis. Baseline TBS-induced hemodynamic responses were utilized to classify treatment responders via logistic regression and supervised machine learning. Intraclass correlation coefficients (ICCs) were calculated using a two-way mixed-effects model with absolute agreement to assess the test-retest reliability of TBS-induced hemodynamic responses.
Results: A logistic regression model distinguished responders from non-responders ($R^2 = 0.617$, $p < 0.001$) and a support-vector machine classifier achieved an accuracy of 82.9% and an AUC of 0.902 in identifying responders. The test-retest reliability of TBS-induced prefrontal hemoglobin responses (Single Measures ICCs) ranged from 0.301 to 0.752, suggesting poor to excellent reliability.
Conclusions: TBS-induced prefrontal hemodynamic response provides valuable information for predicting antidepressant treatment response, highlighting its potential as predictive imaging marker. The reliability of TBS-induced prefrontal response is comparable to previous neuroimaging marker studies, although better control of external factors is needed to enhance reliability.

1. Introduction

Repetitive transcranial magnetic stimulation (rTMS) targeting the left dorsolateral prefrontal cortex (DLPFC) has been established as an effective treatment for treatment resistant major depressive disorder (MDD) (Hsu et al., 2024; Milev et al., 2016; Perera et al., 2016).

---

* Corresponding author.
E-mail addresses: georg.kranz@polyu.edu.hk, georg.kranz@meduniwien.ac.at (G.S. Kranz).

https://doi.org/10.1016/j.jad.2026.121631
Received 7 December 2025; Received in revised form 11 March 2026; Accepted 12 March 2026

Intermittent theta burst stimulation (iTBS), a patterned form of rTMS, has shown comparable effects on depressive symptoms in a significant shorter daily treatment time compared to conventional high-frequency rTMS (3 vs 37.5 mins) (Blumberger et al., 2018). Despite these advancements, the response and remission rates of iTBS are still not satisfying, standing at 49% and 32%, respectively (Blumberger et al., 2018). This indicates that a substantial number of patients do not significantly benefit from this treatment. The lack of response not only imposes economic burden and wastes resources (Smith, 2011), but also diminishes patients' confidence in future treatments. Hence, the development of a reliable and robust predictive marker for assessing rTMS treatment response in depression is crucial.

Significant progress has been made in identifying markers for predicting rTMS treatment outcomes using electroencephalography (EEG) or magnetic resonance imaging (MRI). Our recent meta-analysis revealed that classification models based on neurophysiological and neuroimaging data could effectively differentiate between antidepressant rTMS treatment responders (R) and non-responders (NR), yielding a pooled AUC of 0.87 across studies (Jin et al., 2024). However, the lack of interpretability of these predictive imaging marker candidates hinders our understanding of the underlying neural mechanisms of rTMS treatment. Regarding the closest to a proof-of-concept predictive marker, frontal theta cordance, studies have shown that cognitive task-induced frontal theta activity holds greater predictive potential than resting state EEG data (Li et al., 2016). This suggests that, compared to a resting state, an excited brain – such as one stimulated by a task or external perturbation – may more effectively reveal pathophysiological changes and the brain's potential for recovery, making it a stronger predictor of treatment response.

Similarly, acute cortical hemodynamic responses evoked by a standard daily iTBS session (i.e., 600 pulses) provide immediate feedback on the neurophysiological effects of therapeutic stimulation. Thus, these real-time responses may serve as ideal prognostic markers for depression treatment responses of iTBS. In recent years, TBS-evoked acute cortical hemodynamic responses have been assessed using interleaved functional MRI (fMRI) (Chang et al., 2024a; Chang et al., 2024b). However, the high cost and sophisticated designs that are necessary for TBS/fMRI limits its clinical application. In contrast, functional near-infrared spectroscopy (fNIRS) is a cost-effective imaging method for assessing cortical activation, offering a more affordable alternative to fMRI. It seamlessly integrates with TMS because the far-infrared light in fNIRS and the magnetic field generated by TMS do not interfere with each other. FNIRS has been employed in multiple studies to examine the association between changes in frontal lobe activation and symptom improvement in depression following rTMS treatment (Chou et al., 2023; Lin et al., 2025; Tsuji et al., 2025). Notably, these investigations primarily assessed frontal cortex activation during cognitive task performance, which may not accurately capture the neural activity specifically modulated by therapeutic rTMS. In contrast, direct measurements of rTMS-induced activation in the frontal lobe may provide a more precise and clinically relevant neuroimaging marker for evaluating the antidepressant efficacy of rTMS.

Yet, another concern is the large variability observed in TMS-induced brain activation, as reported in numerous studies (Hordacre et al., 2017; Pell et al., 2011). The high inter- and intra-individual variability in TBS-induced prefrontal responses were also observed in our previous concurrent TBS/fNIRS studies (Kan et al., 2023; Kan et al., 2024). This variability presents a critical challenge in identifying reliable and reproducible neuroimaging markers. Therefore, assessing the test-retest reliability of TBS-induced prefrontal hemoglobin response changes is equally essential. Test-retest reliability refers to the consistency or reproducibility of a measurement when repeated under the same conditions over time (Koo and Li, 2016). High test-retest reliability ensures that neuroimaging markers reflect stable physiological responses rather than being influenced by external noise or random fluctuations.

Summarized, the present study had two key aims: (1) to determine whether TBS-induced prefrontal hemoglobin responses (i.e., the response occurring during and immediately after TBS) before treatment can serve as a predictor of treatment response following four weeks of antidepressant iTBS therapy, and (2) to assess the test-retest reliability of TBS-induced brain activation, as measured by fNIRS.

2. Methods and materials

2.1. Study design

This study was an open-label trial. Participants were recruited from March 2023 to April 2024. The study conformed with the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of The Hong Kong Polytechnic University (HSEARS20200120005) and of the University of Hong Kong Hospital Authority Hong Kong West Cluster (UW 22-326). Consent forms have been obtained from all participants. This report presents the prognostic potential of concurrent TMS/fNIRS for antidepressant treatment success of iTBS therapy. Further aims and outcomes of this study will be reported separately. The study protocol was preregistered at ClinicalTrials.gov (identifier: NCT04526002) and published (Kan et al., 2022).

2.2. Participants

Participants with MDD were recruited for the study. Inclusion criteria were diagnosed with unipolar depression according to DSM-5 criteria by a psychiatrist and were further screened using the 17-items Hamilton Depression Rating Scale (HAMD-17), with a cut-off score of 18 (Rush Jr et al., 2009). Patients with comorbid personality disorders or psychotic features were excluded from the study. Other exclusion criteria were active suicidal intent, severe somatic comorbidities, or a history of seizures. Patients were either treatment naïve or on stable psychopharmacological medication for at least 4 weeks before study inclusion, in which case they were required to maintain the same medication throughout the study period.

Additionally, a group of healthy control (HCs) participants were recruited to be included in the test-retest reliability analysis. For HCs, individuals were excluded if they had a current or previous diagnosis of any psychiatric or neurological disorder or if they had first-degree relatives with a history of psychiatric disorders. None of the participants had any contraindications to TMS (Rossi et al., 2021). For details, see the published protocol (Kan et al., 2022).

2.3. Study procedure

For MDD patients, basic demographic information and clinical data were collected. Each MDD patient underwent two sessions of concurrent TBS/fNIRS measurements over two days (Kan et al., 2023; Kan et al., 2024). On each of the two days, iTBS was applied to the left DLPFC (MNI: -38, +44, +26) using neuronavigation system, followed one hour later by cTBS applied to the right DLPFC (MNI: +38, +44, +26) with neuronavigation, see Figure 1. FNIRS measurements were conducted concurrently with TBS, recording hemoglobin concentrations in the DLPFC during stimulation, lasting either 3 minutes and 8 seconds (i.e., iTBS) or 40 seconds (i.e., cTBS), as well as during pre- and post-stimulation periods, each lasting 3 minutes. To this end, the fNIRS probe was placed directly below the TMS coil on the DLPFC, separated by a 3D-printed polylactic acid bridge (Kan et al., 2024) (also see Supplementary figure S1 the concurrent setup). For stimulation, a figure-of-eight cooling coil (Cool-B65) and the MagPro magnetic stimulator (MagVenture, Denmark) were used. TBS was delivered at 90% of each participant's resting motor threshold (rMT) given that subthreshold intensity is more tolerable for TMS-naïve individuals, and 90% rMT is sufficient to induce individualized brain responses to therapeutic rTMS (Huang et al., 2005). The intensity was adjusted to account for the increased distance between the TMS coil and the skull due to the fNIRS

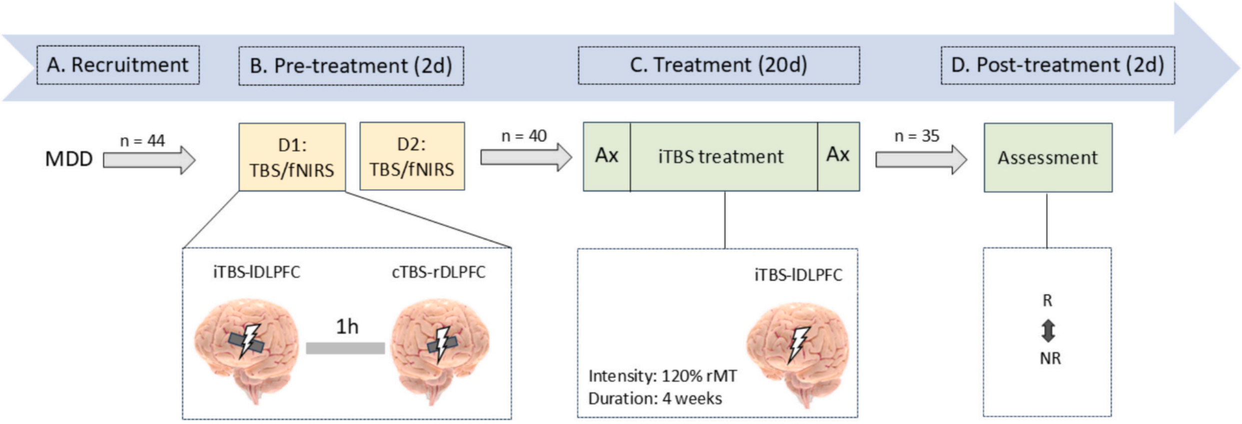

Figure 1. Schematic diagram of the study. After recruitment (A), participants underwent two consecutive days of concurrent TBS/fNIRS measurements (B). This was followed by four weeks of iTBS treatment for MDD participants (C), and for MDD participants who completed the treatment, the same concurrent TBS/fNIRS measurements were repeated post-treatment (D).

optode beneath the coil (Stokes et al., 2007). Throughout the concurrent TBS/fNIRS sessions, participants remained seated comfortably in a chair to ensure optimal conditions for accurate data collection.

After completing the concurrent TMS/fNIRS measurements, MDD patients underwent a standard course of daily neuronavigated iTBS targeting the left DLPFC (MNI: -38, +44, +26). The treatment was administered at an intensity of 120% (FDA approved standard stimulation intensity) of each individual's rMT over a four-week period, consisting of five sessions per week for a total of 20 sessions (Blumberger et al., 2018). Psychometric evaluations were conducted by an experienced rater, including the HAMD-17 to evaluate symptom severity for inclusion criteria, in line with most clinical trials. The Montgomery-Asberg depression rating scale (MADRS) was used to measure symptom changes over time, as it is considered more sensitive to treatment-related changes (Montgomery and Asberg, 1979). The MADRS was administered at baseline, mid-treatment and post-treatment. Response rate was defined as a minimum 50% reduction in MADRS score by the end of the 4-week treatment. For MDD participants who completed the treatment, the same concurrent TBS/fNIRS measurements were conducted again, also twice on consecutive days. See Figure 1.

HC participants received the same two-session concurrent TBS/fNIRS measurement as described above.

2.4. *fNIRS measurement*

The concurrent TBS/fNIRS setup utilized a frequency-domain NIRS system (OxiTs; ISS Inc, Champaign, Illinois USA) to measure the absolute concentrations of oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) at the stimulation site, for details see (Kan et al., 2023; Kan et al., 2024). Absolute hemoglobin concentrations were calculated based on the absolute properties (absorption and reduced scattering coefficients) by measuring the change in the intensity modulation and phase shift (Fantini and Sassaroli, 2020). The data were collected at 25 Hz, where the manufacturer's software downsampled the signals to 0.5 Hz; the downsampled data at 0.5 Hz was then used for analysis. The downsampling method effectively mitigated high-frequency noise (e.g., heartbeat), while for low-frequency noise (e.g., signal drift), bad signals were artificially eliminated. The fNIRS probe consisted of eight emitters at two wavelengths (687 nm & 830 nm) and one detector (for details, see Supplementary figure S2 NIRS geometry). The emitters were 2.0 to 3.5 cm from the detector. Experiments on gelatin-based models have demonstrated that frequency-domain NIRS with emitter-detector distances between 1.5 and 4.5 cm pick up signals representative of the underlying block when the superficial layer is thinner than 0.4 cm (Franceschini et al., 1998). Thus, while superficial muscle or blood are a common source of noise in conventional fNIRS data acquisition, the frequency-domain NIRS used in this study is less affected by such confounding effects. Stimulation-induced oxygen-hemoglobin concentration responses (ΔHbO, ΔHbR) during and after TBS served as the primary imaging endpoints and were calculated by subtracting the corresponding mean of the absolute baseline values (Kan et al., 2023).

2.5. *Statistical analyses*

For objective 1, baseline stimulation-induced responses in HbO and HbR during and after stimulation, as well as demographic information and clinical data (e.g., age, episodes), were used as predictive features (for details, see Supplementary 1.3), while the four-week treatment response was categorized into responders (R) and non-responders (NR). To simplify classification, the Least Absolute Shrink and Selection Operator (LASSO) method was applied to identify the most relevant features. Final features were examined for collinearity using the variance inflation factor and correlation analyses to ensure they were statistically independent. Following our protocol (Kan et al., 2022), a binary logistic regression was conducted first to examine whether these final features could predict treatment response following iTBS therapy. Subsequently, machine learning analyses were performed by employing multiple algorithms and cross-validation procedures to maximize predictive accuracy and generalizability (see supplementary 1.3 for details). Additionally, a straightforward descriptive comparison was made between R and NR in terms of significant features. This analysis served to complement the prediction model by providing additional insights into the characteristics of TBS-evoked prefrontal activation. The combined approach aimed to achieve a more comprehensive assessment of its effectiveness in classifying individuals as responders or non-responders to TBS treatment.

For objective 2, the test-retest reliability of absolute prefrontal hemoglobin concentrations (HbO, HbR) during and after stimulation were first assessed using the interclass correlation coefficient (ICC) analysis, applying a two-way mixed-effects model with absolute agreement and average measurements (Koo and Li, 2016). Only when excellent reliability was achieved for the absolute values at each phase (i.e., before stimulation, during stimulation, after stimulation) was the test-retest reliability of TBS-induced prefrontal responses (ΔHbO, ΔHbR) further assessed. This approach ensured that no additional variability was introduced when assessing the reliability of TBS-induced prefrontal responses. We followed the ICC classification criteria proposed by Li et al. (Li et al., 2015), where ICC value < 0.40 was considered as "poor", 0.40 to 0.59 as "fair", 0.60 to 0.74 as "good", and > 0.75 as "excellent". For the test-retest reliability of TBS-induced prefrontal responses, ICCs were calculated using a two-way mixed-effects model with absolute agreement. Single Measures ICCs were reported, as each session provided an independent measurement (Koo and Li, 2016). In the context of stimulus-induced cortical hemoglobin changes measured by fMRI or fNIRS, an ICC of fair level (≥0.4) is generally considered as acceptable reliability (Huang et al., 2017; Li et al., 2015). Furthermore, the reproducibility of TBS-induced prefrontal hemoglobin responses (ΔHbO, ΔHbR) was evaluated using Pearson's correlation coefficients (Huang et al., 2017). Statistical significance was set at $p < 0.05$. For the identification and handling of potential outliers, please refer to supplementary 1.4. All statistical analyses were performed using SPSS Version 26.

3. Results

3.1. Participants

Of the 44 MDD participants initially enrolled, 40 received iTBS treatment. Among them, 34 completed the full 4-week treatment protocol (for dropout details, see Supplementary 2.1). One additional participant, classified as an early responder, discontinued treatment after 2 weeks due to an unrelated fall. In total, data from 35 patients were included in the treatment outcome prediction analysis (Aim 1), with 24 classified as responders (R) and 11 as non-responders (NR). To assess the impact of including this early responder, we conducted a sensitivity analysis (see Supplementary 2.1). For the test-retest reliability analysis (Aim 2), a total of 119 datasets from both MDD participants and HCs were included. For details, see supplementary 2.1. Additionally, the demographic and clinical information of all participants is provided in Table 1 and Supplementary Table S1.

3.2. The utility of baseline TBS-induced prefrontal hemodynamic responses in predicting treatment response

Four baseline features including age, the ΔHbO after iTBS, the ΔHbO during cTBS and family history of a psychiatric disorder, were selected after performing LASSO. No collinearity was observed among the selected features. The overall model of the binary logistic regression was statistically significant, $\chi^2(4) = 20.25$, $p < 0.001$, indicating that these predictors reliably distinguished between R and NR. The model explained 61.7% of the variance in treatment response (Nagelkerke $R^2 = 0.617$) and correctly classified 85.7% of cases (sensitivity: 91.7%; specificity: 72.7%). Age, the ΔHbO after iTBS, and family history were statistically significant predictors ($p < 0.05$). Results were consistent in the sensitivity analysis after excluding the early responder (n=34, see Supplementary 2.1). For more details, see Supplementary 2.6.2 and Figure 2A.

Using MATLAB's Machine Learning toolbox, linear SVM was determined to be the most effective classification model, achieving an accuracy of 82.9% (sensitivity: 91.7%; specificity: 63.6%) and an AUC of 0.902 (see Figure 2B). The descriptive comparison between R and NR showed that patients who exhibit greater decrease in HbO change following iTBS indicated a better treatment response (R: -0.5 ± 1.238μM; NR: 0.308 ± 1.195μM). Furthermore, older individuals and those without a family history of psychiatric disorders were more likely to respond positively to four weeks iTBS treatment. Detailed results from the descriptive comparison between the R and NR can be found in Table 2.

3.3. Test-retest reliability

Due to missing data or poor signal quality, the final datasets included in the test-retest reliability analysis of prefrontal hemoglobin concentrations comprised 112 iTBS-related datasets and 104 cTBS-related analysis (see Supplementary 2.3).

The test-retest reliability results for absolute prefrontal hemoglobin concentrations (HbO, HbR) at three phases (before, during and after stimulation) are presented in Supplementary Table S2. The ICC values ranged from 0.874 to 0.973 (mean = 0.918), indicating excellent test-retest reliability. For the test-retest reliability of TBS-induced prefrontal activation (ΔHbO, ΔHbR), the Single Measures ICC ranged from 0.301 to 0.752 (mean = 0.477), suggesting poor to excellent reliability (see Table 3). Notably, the significant hemodynamic predictor (i.e., ΔHbO after iTBS) demonstrated acceptable reliability (ICC = 0.469).

Table 1

Demographics and clinical information for all participants.

A	Participants (All)			Statistics
	All (n = 89) mean (SD)	MDD (n = 44) mean (SD)	HC (n = 45) mean (SD)	
Demographics				
Age (years)	39.65 (11.37)	38.82 (10.52)	40.47 (12.22)	p = 0.497
Sex, M: F (% Female)	31:58 (65.17%)	13: 31 (70.45%)	18:27 (60%)	p = 0.375
Education (years)	15.39 (2.89)	14.45 (2.33)	16.31 (3.10)	p = 0.002*
RMT (%MSO)	51% (0.11)	51% (0.12)	52% (0.11)	p = 0.712

B	MDD Participants with completed iTBS treatment			
	All (n = 35)	Responders (n = 24)	Non-responders (n = 11)	
Demographics				
Age (years)	39.23 (9.92)	41.83 (8.90)	33.55 (10.01)	p = 0.019*
Sex, M: F (% Female)	13:22 (62.86%)	9:15 (62.50%)	4:7 (63.64%)	p = 0.950
Education (years)	14.6 (2.35)	14.58 (2.41)	14.64 (2.34)	p = 0.952
Job, Full: Part: no (% Full)	19: 6: 10 (54.29%)	11: 5: 8 (45.83%)	8: 1: 2 (72.73%)	p = 0.197
RMT(%MSO)	51% (0.13)	51% (0.13)	49% (0.08)	p = 0.639
Clinical characteristics				
Onset age of the first episode (years)	32.54 (8.99)	34.46 (7.97)	28.36 (10.03)	p = 0.061
Illness duration since first episode (years)	6.77 (7.69)	7.5 (8.27)	5.18 (6.29)	p = 0.416
Episodes (no.)	1.51 (1.17)	1.63 (1.35)	1.27 (0.65)	p = 0.417
Comorbidities (no.)	0.83 (0.66)	0.79 (0.66)	0.91 (0.70)	p = 0.634
Family history, yes: no (% yes)	17: 18 (48.57%)	10: 14 (41.67%)	7: 4 (63.63%)	p = 0.240
Clinical symptoms				
Pre MADRS	28.11 (7.62)	28.25 (8.71)	27.82 (4.77)	p = 0.879
Post MADRS	11.89 (6.86)	8.42 (4.79)	19.45 (3.91)	p < 0.001*

RMT: Resting motor threshold; MSO: Maximum Stimulator Output.

Pearson's correlation coefficients for TBS-induced prefrontal responses across two days are displayed in Supplementary (Figure S4). Significant correlations were observed in all eight paired comparisons (all $p < 0.05$), with correlation coefficient ranging from 0.264 to 0.703, indicating negligible to strong correlations.

4. Discussion

This study examined whether baseline stimulation-induced prefrontal hemoglobin responses could predict treatment response after

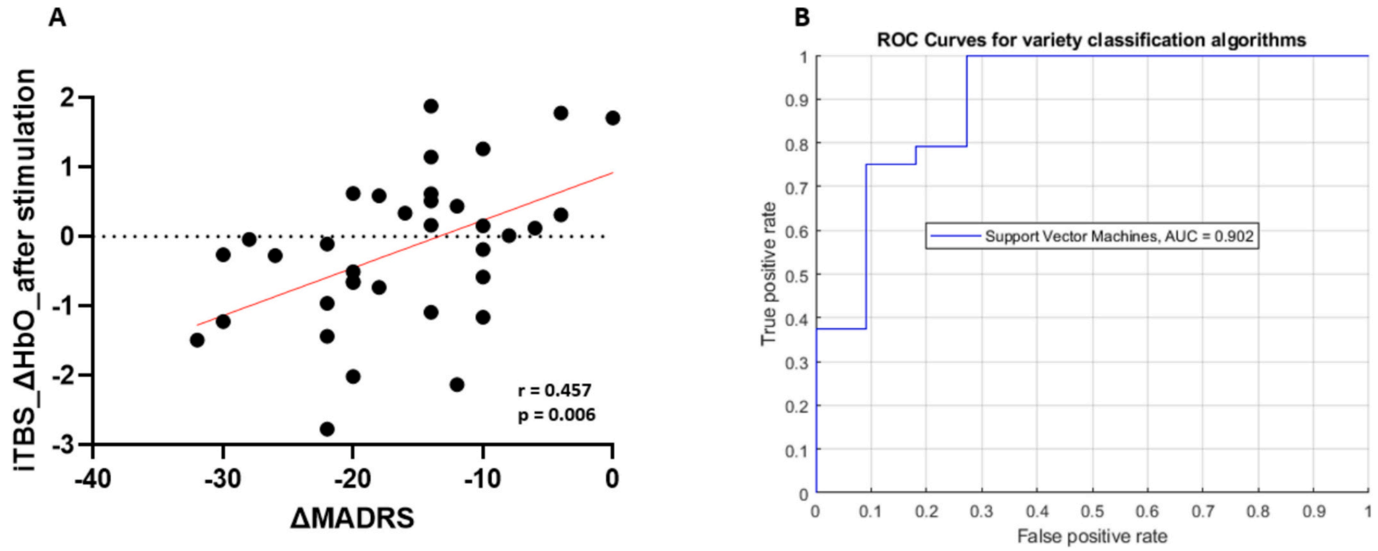

Figure 2. Baseline TBS-induced prefrontal hemodynamic response predicts treatment response. (A) Baseline ΔHbO after iTBS was a significant predictor of antidepressant response. A greater decrease in prefrontal HbO after iTBS was associated with better treatment response (i.e., larger reduction in symptom severity scores). (B) Receiver operating characteristic (ROC) curve analysis. The linear SVM model, based on four selected baseline features, achieved an area under the curve (AUC) of 0.902 in predicting clinical response to iTBS treatment for depression.

Table 2
Descriptive comparison between Responders and Non-responders.

Features	Participants (MDD)		
	All (n = 35) mean (SD)	Responders (n = 24) mean (SD)	Non-responders (n = 11) mean (SD)
iTBS_post_HbO (μmol/L)	-0.24 (1.27)	-0.50 (1.24)	0.31 (1.20)
Age (years)	39.23 (9.92)	41.83 (8.9)	33.55 (10.01)
Family history, yes: no (% yes)	17: 18 (48.57%)	10: 14 (41.67%)	7: 4 (63.63%)

four weeks of daily iTBS in MDD and assessed the test-retest reliability of TBS-induced prefrontal hemoglobin responses. The results indicated (1) that two imaging features from baseline TBS/fNIRS measurements and two demographic features together can reliably distinguish responders from non-responders to four weeks of iTBS treatment; (2) TBS-induced prefrontal hemoglobin responses demonstrated a range of reliability, from poor to excellent.

4.1. Predicting treatment response by baseline TBS-induced hemoglobin responses

Based on the above-mentioned imaging and demographic features, logistic regression accounted for 61.7% of the variance in the classification of individuals as R and NR following four weeks of iTBS treatment. Furthermore, a SVM model achieved an accuracy of 82.9%. This is comparable to previous published classifiers using prefrontal theta cordance as a feature derived from EEG data (Erguzel et al., 2015; Hasanzadeh et al., 2019; Jin et al., 2024). Given that demographic, as well as imaging characteristics were selected as features in our study, the differential contribution of these features in predicting antidepressant outcomes requires further investigation. A recent study involving 518 MDD patients utilized pretreatment symptoms and EEG features to predict antidepressant treatment outcomes. The findings indicated that the baseline symptom score was the most crucial predictive feature, while EEG features showed smaller but still meaningful associations with specific symptom improvements (Rajpurkar et al., 2020). In our study, the prediction based on both fNIRS and demographic features achieved an accuracy of 82.9%, higher than the prediction based solely on fNIRS or on demographic features (see supplementary 2.6.1), justifying our decision to include both kinds of features. Moreover, we argue that TBS-induced brain activity changes, with TBS being an objective stimulus, may provide additional information on the brain's capacity for neuroplasticity, beyond the disease-related brain activity observed during resting state or cognitive tasks. Future replication studies are needed to clarify whether demographic features, internal brain activity, external (objective) stimulus-induced cortical activation, or a combination of these will prove to be a better predictor of treatment outcomes.

The descriptive comparison results together with additional exploratory analysis (see supplementary 2.6.2) suggest that patients whose brains showed a greater decrease in HbO following iTBS had a better treatment response, in line with previous fMRI-BOLD findings (Smith et al., 2014). Previous fNIRS studies assessing task-evoked frontal lobe activation in both pre- and post rTMS treatment have demonstrated that

Table 3
Test-retest reliability of TBS-induced prefrontal hemoglobin responses (ΔHbO, ΔHbR) across two consecutive days.

Measures			Day 1 (mean ± SD)	Day 2 (mean ± SD)	ICC (95% CI) (single measures)	P value
iTBS (n=112)	ΔHbO (μmol/L)	During	-0.23±1.00	-0.16±0.94	0.622 (0.494, 0.723)	P<0.001
		Post	-0.29±1.12	-0.22±1.09	0.469 (0.311, 0.602)	P<0.001
	ΔHbR (μmol/L)	During	0.38±0.48	0.21±0.47	0.439 (0.267, 0.582)	P<0.001
		Post	0.25±0.46	0.12±0.54	0.397 (0.231, 0.541)	P<0.001
cTBS (n=104)	ΔHbO (μmol/L)	During	-0.35±1.33	-0.18±1.00	0.362 (0.185, 0.518)	P<0.001
		Post	-0.16±1.04	0.07±0.84	0.301 (0.120, 0.464)	P=0.001
	ΔHbR (μmol/L)	During	0.63±0.94	0.56±0.91	0.752 (0.655, 0.825)	P<0.001
		Post	0.23±0.43	0.13±0.38	0.402 (0.230, 0.550)	P<0.001

increased frontal lobe activity is associated with improvements in depressive symptoms (Chou et al., 2023; Lin et al., 2025). However, whether TBS-induced hemodynamic responses similarly increase following treatment remains to be elucidated and warrants further investigation. Additionally, our analysis shows that particularly older patients, as well as patients without a family history of psychiatric disorders are more likely to respond to the treatment. This aligns with earlier studies that reported a higher likelihood of response to TMS treatment among elderly individuals (Fitzgerald et al., 2016; Hopman et al., 2021).

4.2. Test-retest reliability of TBS-induced prefrontal hemoglobin responses

We observed excellent test-retest reliability for absolute prefrontal hemoglobin concentrations (HbO, HbR). However, for TBS-induced prefrontal activation (ΔHbO, ΔHbR), the reliability ranged from poor to excellent reliability (based on Single Measures ICC). This result was expected, as change scores tend to have lower variance and accumulate measurement error from both conditions, thereby reducing reliability. Our findings align with previous fMRI studies, which have reported large variability in reliability, ranging from poor to excellent (Holiga et al., 2018; Plichta et al., 2012; Zuo and Xing, 2014). Additionally, task-based fMRI measures generally exhibit lower reliability compared to resting-state fMRI, and reliability outcomes tend to be task-dependent (Elliott et al., 2020; Holiga et al., 2018). Interestingly, studies have found that greater task-induced activation magnitudes are associated with higher reliability (Holiga et al., 2018; Korucuoglu et al., 2020), a pattern that is also reflected in our results (see Table 3). Similarly, in EEG-based reliability studies, including those involving resting-state EEG measurements and TMS-induced brain activation assessed via EEG, the findings have been variable, with good reliability only observed in some specific features (Bertazzoli et al., 2025; Hirano et al., 2020; Kerwin et al., 2018; Tang et al., 2025).

Importantly, the sample size used for reliability validation in our study was significantly larger than in previous neuroimaging reliability studies, where typical sample sizes ranged from 10 to 30 participants. This larger sample size strengthens the robustness of our findings. However, many previous studies have demonstrated that certain external factors, such as sleep, caffeine intake, and time-of-day, influence neuroimaging reliability (Kan et al., 2023; Wang et al., 2017). Unfortunately, in current study, due to logistical constraints, a tightly scheduled laboratory shared with other research groups, we were unable to systematically control for these external factors, which may have influenced the reliability of our findings. Future studies investigating neuroimaging markers should be implement strictly controlled conditions to minimize these potential confounds.

4.3. Limitations

There were several limitations in our study. First, the sample size of the MDD group was relatively small. Although the study provides important preliminary evidence supporting the use of stimulation-evoked hemoglobin responses as predictive factors, larger studies are needed to validate and extend these findings. In particular, the small number of participants may increase the risk of overfitting in machine learning models. Second, the sensor of ISS NIRS only covered a small brain area, restricting the investigation to local hemodynamic effects and preventing an assessment of functional connectivity between the prefrontal cortex and other brain regions. Thirdly, potential confounding factors that could influence TBS-induced hemoglobin concentration responses were not adequately controlled.

4.4. Conclusion

In conclusion, this study assessed the test-retest reliability of TBS-induced prefrontal hemoglobin responses and found that its reliability was comparable to previous studies on neuroimaging markers. However, future research should aim to better control external factors to enhance reliability. Additionally, our study showed baseline TBS-induced hemodynamic responses were found to predict antidepressant treatment outcomes. However, given the relatively small sample size of the MDD group, these predictive results should be interpreted with caution. Further studies with larger sample size are needed to validate and expand upon these findings.

CRediT authorship contribution statement

Rebecca L.D. Kan: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Alvin H.P. Tang:** Methodology, Investigation. **Penny P. Qin:** Writing – review & editing, Investigation. **Minxia Jin:** Investigation. **Adam W.L. Xia:** Writing – review & editing, Investigation. **Bella B.B. Zhang:** Methodology, Investigation. **Tim T.Z. Lin:** Software, Investigation, Formal analysis. **Sharie X. Wang:** Investigation. **Jessie J. Lin:** Writing – review & editing, Supervision, Resources, Investigation. **Michael K. Yeung:** Writing – review & editing, Supervision, Formal analysis. **Sherry K.W. Chan:** Writing – review & editing, Resources. **Fan Li:** Writing – review & editing, Software, Methodology, Formal analysis. **Fidel Vila-Rodriguez:** Writing – review & editing, Methodology. **Kenneth N.K. Fong:** Writing – review & editing, Supervision. **Frank Padberg:** Writing – review & editing, Methodology. **Georg S. Kranz:** Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Funding

This work was supported by the General Research Fund (grant numbers 15100120 and 15106222) under the University Grands Committee of the HKSAR, as well as the Mental Health Research Center (grant numbers 0048822 and 0040786), The Hong Kong Polytechnic University to Georg S. Kranz.

Declaration of competing interest

FP is a member of the European Scientific Advisory Board of BrainsWay Inc. (Jerusalem, Israel), and the International Scientific Advisory Board of Sooma (Helsinki, Finland); he has received speaker honoraria from Mag&More GmbH, the neuroCare Group (Munich, Germany), and BrainsWay Inc.; his lab has received support with equipment from neuroConnGmbH (Ilmenau, Germany), Mag&More GmbH, and BrainsWay Inc. **FVR** has received research support from CIHR, Brain Canada, Michael Smith Foundation for Health Research, Vancouver Coastal Health Research Institute, and Weston Brain Institute for investigator-initiated research. Philanthropic support from Seedlings Foundation. In-kind equipment support for investigator-initiated trial from MagVenture. He has received honoraria for participation in an advisory board for Allergan. FVR is a volunteer director on the board of directors of the British Columbia Schizophrenia Society. He is a member of the Educational Committee of the Clinical TMS Society (unpaid). **GSK** received travel grants and/or speaker honoraria from Roche, Pfizer, Orphan Pharmaceuticals AG, HealthLink, Academy of Brain Stimulation and Storz Medical. The other authors do not report any conflict of interest.

Acknowledgements

We are deeply grateful to all participants for their time, trust, and willingness to contribute to our study. We would further like to thank Arthur D.P. Mak and Raymond C.K. Chung for clinical and statistical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2026.121631.

References

Bertazzoli, G., Dognini, E., Fried, P.J., Miniussi, C., Julkunen, P., Bortoletto, M., 2025. Bridging the gap to clinical use: A systematic review on TMS-EEG test-retest reliability. Clin. Neurophysiol. 171, 133–145. https://doi.org/10.1016/j.clinph.2025.01.002.

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. Lancet 391, 1683–1692. https://doi.org/10.1016/s0140-6736(18)30295-2.

Chang, K.Y., Tik, M., Mizutani-Tiebel, Y., Schuler, A.L., Taylor, P., Campana, M., Vogelmann, U., Huber, B., Dechantsreiter, E., Thielscher, A., Bulubas, L., Padberg, F., Keeser, D., 2024a. Neural response during prefrontal theta burst stimulation: Interleaved TMS-fMRI of full iTBS protocols. Neuroimage 291, 120596. https://doi.org/10.1016/j.neuroimage.2024.120596.

Chang, K.Y., Tik, M., Mizutani-Tiebel, Y., Taylor, P., van Hattem, T., Falkai, P., Padberg, F., Bulubas, L., Keeser, D., 2024b. Dose-Dependent Target Engagement of a Clinical Intermittent Theta Burst Stimulation Protocol: An Interleaved Transcranial Magnetic Stimulation-Functional Magnetic Resonance Imaging Study in Healthy People. Biol. Psychiatry: Cognit. Neurosci. Neuroimaging. https://doi.org/10.1016/j.bpsc.2024.08.009.

Chou, P.H., Liu, W.C., Wang, S.C., Lin, W.H., Chung, Y.L., Chang, C.H., Su, K.P., 2023. Associations between frontal lobe activity and depressive symptoms in patients with major depressive disorder receiving rTMS treatment: a near-infrared spectroscopy study. Front Psychiatry. 14, 1235713. https://doi.org/10.3389/fpsyt.2023.1235713.

Elliott, M.L., Knodt, A.R., Ireland, D., Morris, M.L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T.E., Caspi, A., Hariri, A.R., 2020. What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. Psychol. Sci. 31, 792–806. https://doi.org/10.1177/0956797620916786.

Erguzel, T.T., Ozekes, S., Gultekin, S., Tarhan, N., Hizli Sayar, G., Bayram, A., 2015. Neural Network Based Response Prediction of rTMS in Major Depressive Disorder Using QEEG Cordance. Psychiatry Investig. 12, 61–65. https://doi.org/10.4306/pi.2015.12.1.61.

Fantini, S., Sassaroli, A., 2020. Frequency-Domain Techniques for Cerebral and Functional Near-Infrared Spectroscopy. Front. Neurosci. 14, 300. https://doi.org/10.3389/fnins.2020.00300.

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. Depress. Anxiety 33, 746–753. https://doi.org/10.1002/da.22503.

Franceschini, M.A., Fantini, S., Paunescu, L.A., Maier, J.S., Gratton, E., 1998. Influence of a superficial layer in the quantitative spectroscopic study of strongly scattering media. Appl. Opt. 37, 7447–7458. https://doi.org/10.1364/ao.37.007447.

Hasanzadeh, F., Mohebbi, M., Rostami, R., 2019. Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal. J. Affect. Disord. 256, 132–142. https://doi.org/10.1016/j.jad.2019.05.070.

Hirano, Y., Nakamura, I., Tamura, S., Onitsuka, T., 2020. Long-Term Test-Retest Reliability of Auditory Gamma Oscillations Between Different Clinical EEG Systems. Front Psychiatry. 11, 876. https://doi.org/10.3389/fpsyt.2020.00876.

Holiga, Š., Sambataro, F., Luzy, C., Greig, G., Sarkar, N., Renken, R.J., Marsman, J.C., Schobel, S.A., Bertolino, A., Dukart, J., 2018. Test-retest reliability of task-based and resting-state blood oxygen level dependence and cerebral blood flow measures. PLoS One 13, e0206583. https://doi.org/10.1371/journal.pone.0206583.

Hopman, H.J., Chan, S.M.S., Chu, W.C.W., Lu, H., Tse, C.Y., Chau, S.W.H., Lam, L.C.W., Mak, A.D.P., Neggers, S.F.W., 2021. Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. J. Affect. Disord. 290, 261–271. https://doi.org/10.1016/j.jad.2021.04.081.

Hordacre, B., Goldsworthy, M.R., Vallence, A.M., Darvishi, S., Moezzi, B., Hamada, M., Rothwell, J.C., Ridding, M.C., 2017. Variability in neural excitability and plasticity induction in the human cortex: A brain stimulation study. Brain Stimul. 10, 588–595. https://doi.org/10.1016/j.brs.2016.12.001.

Hsu, T.W., Yeh, T.C., Kao, Y.C., Thompson, T., Brunoni, A.R., Carvalho, A.F., Hsu, C.W., Tu, Y.K., Liang, C.S., 2024. The dose-effect relationship of six stimulation parameters with rTMS over left DLPFC on treatment-resistant depression: A systematic review and meta-analysis. Neurosci. Biobehav. Rev. 162, 105704. https://doi.org/10.1016/j.neubiorev.2024.105704.

Huang, Y., Mao, M., Zhang, Z., Zhou, H., Zhao, Y., Duan, L., Kreplin, U., Xiao, X., Zhu, C., 2017. Test-retest reliability of the prefrontal response to affective pictures based on functional near-infrared spectroscopy. J. Biomed. Opt. 22, 16011. https://doi.org/10.1117/1.Jbo.22.1.016011.

Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P., Rothwell, J.C., 2005. Theta burst stimulation of the human motor cortex. Neuron 45, 201–206. https://doi.org/10.1016/j.neuron.2004.12.033.

Jin, M.X., Qin, P.P., Xia, A.W.L., Kan, R.L.D., Zhang, B.B.B., Tang, A.H.P., Li, A.S.M., Lin, T.T.Z., Giron, C.G., Pei, J.J., Kranz, G.S., 2024. Neurophysiological and neuroimaging markers of repetitive transcranial magnetic stimulation treatment response in major depressive disorder: A systematic review and meta-analysis of predictive modeling studies. Neurosci. Biobehav. Rev. 162, 105695. https://doi.org/10.1016/j.neubiorev.2024.105695.

Kan, R.L.D., Lin, T.T.Z., Zhang, B.B.B., Giron, C.G., Jin, M., Qin, P.P.I., Xia, A.W.L., Chan, S.K.W., Chau, B.K.H., Kranz, G.S., 2023. Moderators of stimulation-induced neural excitability in the left DLPFC: A concurrent iTBS/fNIRS case study. Brain Stimul. https://doi.org/10.1016/j.brs.2023.09.015.

Kan, R.L.D., Mak, A.D.P., Chan, S.K.W., Zhang, B.B.B., Fong, K.N.K., Kranz, G.S., 2022. Protocol for a prospective open-label clinical trial to investigate the utility of concurrent TBS/fNIRS for antidepressant treatment optimisation. BMJ Open 12, e053896. https://doi.org/10.1136/bmjopen-2021-053896.

Kan, R.L.D., Zhang, B.B.B., Lin, T.T.Z., Tang, A.H.P., Xia, A.W.L., Qin, P.P.I., Jin, M., Fong, K.N.K., Becker, B., Yau, S.Y., Kranz, G.S., 2024. Sex differences in brain excitability revealed by concurrent iTBS/fNIRS. Asian J. Psychiatr. 96, 104043. https://doi.org/10.1016/j.ajp.2024.104043.

Kerwin, L.J., Keller, C.J., Wu, W., Narayan, M., Etkin, A., 2018. Test-retest reliability of transcranial magnetic stimulation EEG evoked potentials. Brain Stimul. 11, 536–544. https://doi.org/10.1016/j.brs.2017.12.010.

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Korucuoglu, O., Harms, M.P., Astafiev, S.V., Kennedy, J.T., Golosheykin, S., Barch, D.M., Anokhin, A.P., 2020. Test-retest reliability of fMRI-measured brain activity during decision making under risk. Neuroimage 214, 116759. https://doi.org/10.1016/j.neuroimage.2020.116759.

Li, C.T., Hsieh, J.C., Huang, H.H., Chen, M.H., Juan, C.H., Tu, P.C., Lee, Y.C., Wang, S.J., Cheng, C.M., Su, T.P., 2016. Cognition-Modulated Frontal Activity in Prediction and Augmentation of Antidepressant Efficacy: A Randomized Controlled Pilot Study. Cereb. Cortex 26, 202–210. https://doi.org/10.1093/cercor/bhu191.

Li, L., Zeng, L., Lin, Z.J., Cazzell, M., Liu, H., 2015. Tutorial on use of intraclass correlation coefficients for assessing intertest reliability and its application in functional near-infrared spectroscopy-based brain imaging. J. Biomed. Opt. 20, 50801. https://doi.org/10.1117/1.Jbo.20.5.050801.

Lin, C.E., Chen, L.F., Chung, C.H., Sack, A.T., Chang, H.A., 2025. Functional near-infrared spectroscopy as a biomarker of TMS efficacy in treatment-resistant depression. Psychiatry Clin. Neurosci. 79, 765–775. https://doi.org/10.1111/pcn.13890.

Milev, R.V., Giacobbe, P., Kennedy, S.H., Blumberger, D.M., Daskalakis, Z.J., Downar, J., Modirrousta, M., Patry, S., Vila-Rodriguez, F., Lam, R.W., MacQueen, G.M., Parikh, S.V., Ravindran, A.V., 2016. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 4. Neurostimulation Treatments. Can. J. Psychiatry. 61, 561–575. https://doi.org/10.1177/0706743716660033.

Montgomery, S.A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. Br. J. Psychiatry 134, 382–389. https://doi.org/10.1192/bjp.134.4.382.

Pell, G.S., Roth, Y., Zangen, A., 2011. Modulation of cortical excitability induced by repetitive transcranial magnetic stimulation: influence of timing and geometrical parameters and underlying mechanisms. Prog. Neurobiol. 93, 59–98. https://doi.org/10.1016/j.pneurobio.2010.10.003.

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 Stimul. 9, 336–346. https://doi.org/10.1016/j.brs.2016.03.010.

Plichta, M.M., Schwarz, A.J., Grimm, O., Morgen, K., Mier, D., Haddad, L., Gerdes, A.B., Sauer, C., Tost, H., Esslinger, C., Colman, P., Wilson, F., Kirsch, P., Meyer-Lindenberg, A., 2012. Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. Neuroimage 60, 1746–1758. https://doi.org/10.1016/j.neuroimage.2012.01.129.

Rajpurkar, P., Yang, J., Dass, N., Vale, V., Keller, A.S., Irvin, J., Taylor, Z., Basu, S., Ng, A., Williams, L.M., 2020. Evaluation of a Machine Learning Model Based on Pretreatment Symptoms and Electroencephalographic Features to Predict Outcomes of Antidepressant Treatment in Adults With Depression: A Prespecified Secondary Analysis of a Randomized Clinical Trial. JAMA Netw. Open 3, e206653. https://doi.org/10.1001/jamanetworkopen.2020.6653.

Rossi, S., Antal, A., Bestmann, S., Bikson, M., Brewer, C., Brockmöller, J., Carpenter, L.L., Cincotta, M., Chen, R., Daskalakis, J.D., Di Lazzaro, V., Fox, M.D., George, M.S., Gilbert, D., Kimiskidis, V.K., Koch, G., Ilmoniemi, R.J., Lefaucheur, J.P., Leocani, L., Lisanby, S.H., Miniussi, C., Padberg, F., Pascual-Leone, A., Paulus, W., Peterchev, A. V., Quartarone, A., Rotenberg, A., Rothwell, J., Rossini, P.M., Santarnecchi, E., Shafi, M.M., Siebner, H.R., Ugawa, Y., Wassermann, E.M., Zangen, A., Ziemann, U., Hallett, M., 2021. Safety and recommendations for TMS use in healthy subjects and patient populations, with updates on training, ethical and regulatory issues: Expert Guidelines. Clin. Neurophysiol. 132, 269–306. https://doi.org/10.1016/j.clinph.2020.10.003.

Rush Jr., A.J., First, M.B., Blacker, D., 2009. Handbook of psychiatric measures. American Psychiatric Pub.

Smith, K., 2011. Trillion-dollar brain drain. Nature 478, 15. https://doi.org/10.1038/478015a.

Smith, R., Allen, J.J.B., Thayer, J.F., Fort, C., Lane, R.D., 2014. Increased association over time between regional frontal lobe BOLD change magnitude and cardiac vagal control with sertraline treatment for major depression. Psychiatry Res. Neuroimaging 224, 225–233. https://doi.org/10.1016/j.pscychresns.2014.08.015.

Stokes, M.G., Chambers, C.D., Gould, I.C., English, T., McNaught, E., McDonald, O., Mattingley, J.B., 2007. Distance-adjusted motor threshold for transcranial magnetic stimulation. Clin. Neurophysiol. 118, 1617–1625. https://doi.org/10.1016/j.clinph.2007.04.004.

Tang, X., Wang, S., Xu, X., Luo, W., Zhang, M., 2025. Test-retest reliability of resting-state EEG intrinsic neural timescales. Cereb. Cortex 35. https://doi.org/10.1093/cercor/bhaf034.

Tsuji, A., Nishida, K., Imazu, S.I., Kawabata, Y., Toyoda, K., Matsumoto, K., Onishi, N., Kubo, Y., Yamane, T., Kinoshita, S., Nishizawa, Y., Kanazawa, T., 2025. Hemodynamic characteristics at baseline and following repeated transcranial magnetic stimulation treatment. J. Affect. Disord. 381, 459–466. https://doi.org/10.1016/j.jad.2025.03.168.

Wang, J., Han, J., Nguyen, V.T., Guo, L., Guo, C.C., 2017. Improving the Test-Retest Reliability of Resting State fMRI by Removing the Impact of Sleep. Front. Neurosci. 11, 249. https://doi.org/10.3389/fnins.2017.00249.

Zuo, X.N., Xing, X.X., 2014. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. Neurosci. Biobehav. Rev. 45, 100–118. https://doi.org/10.1016/j.neubiorev.2014.05.009.