Scientia Psychiatrica
Vol. 7, No. 4, 2026, pp. 203–212 · e-ISSN 2715-9736 journal homepage: scientiapsychiatrica.com

ORIGINAL RESEARCH ARTICLE

Transferability of an Antidepressant-Response DNA Methylation Signature to rTMS Outcomes in Treatment-Resistant Depression

Rachmat Hidayat[1*], Abu Bakar[2], Sohyuk Kim[3]

[1]*Department of Medical Biology, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia*
[2]*Department of Radiology, CMHC Research Center, Palembang, Indonesia*
[3]*Department of Biomolecular Medicine, Geum-Jeong Research Center, Busan, South Korea*

ARTICLE INFO

Article history
Received: August 23, 2026
Revised: September 21, 2026
Accepted: October 18, 2026
Available online: October 27, 2026
Published: October 30, 2026

Keywords
Biomarkers
DNA methylation
Epigenetics
Repetitive transcranial magnetic stimulation
Treatment-resistant depression
Treatment response

***Corresponding author**
Rachmat Hidayat
E-mail: dr.rachmat.hidayat@gmail.com
ORCID: 0000-0001-5770-1201

DOI
10.37275/scipsy.v7i4.222

ABSTRACT

Background. Baseline DNA methylation has been linked to antidepressant outcomes. Candidate signals need cross-treatment testing before translational interpretation.

Objective. To test whether five antidepressant-response CpGs in CHN2, JAK2, SCN7A, and IQANK1 transfer to outcomes after repetitive transcranial magnetic stimulation (rTMS) in treatment-resistant depression.

Methods. This secondary analysis used 46 participants of the public GSE300009 cohort with baseline blood methylation and paired Patient Health Questionnaire-9 (PHQ-9) scores. The primary outcome was PHQ-9 improvement. Response (≥50% improvement) and remission (end-treatment PHQ-9 <5) were secondary. Effects per 1-SD higher CpG M-value were estimated with covariate-adjusted linear models (HC3 standard errors) and bias-reduced logistic models. Holm correction, rank correlations, leave-one-out diagnostics, sensitivity analyses, and repeated nested five-fold cross-validation tested robustness and incremental prediction.

Results. Mean PHQ-9 fell from 22.3 to 8.9. Calculated response occurred in 31 participants (67.4%) and remission in 18 (39.1%). After multiplicity correction, no CpG was associated with continuous improvement (Holm-adjusted p = 1.000). Two CHN2 CpGs showed nominal inverse associations with remission, not response: cg06926818, odds ratio 0.36 (95% confidence interval 0.15–0.84), p = 0.018, Holm-adjusted p = 0.089; cg23687322, odds ratio 0.41 (0.19–0.89), p = 0.024, Holm-adjusted p = 0.095. Under one-standard-error regularization, adding the CpGs left root-mean-square error unchanged. The area under the receiver-operating-characteristic curve rose by 0.012 for response and 0.059 for remission, with intervals compatible with no benefit.

Conclusion. The panel did not transfer as a dependable marker of continuous improvement or response. CHN2-remission signals remain hypothesis-generating. Larger, prospectively harmonized external validation must precede clinical use.

Copyright: © 2026 The Author(s). Published by HM Publisher in collaboration with CMHC Research Center on behalf of Scientia Psychiatrica. The authors retain copyright of their work.

Access licence: This is an open access article distributed under the terms of the Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International Licence (CC BY-NC-SA 4.0), which permits non-commercial use, sharing, adaptation and reproduction in any medium, provided the original author and source are credited and any derivative work is distributed under the same licence.

1. Introduction

Major depressive disorder varies widely in symptom profile, course, and treatment response. Symptoms sometimes persist despite adequate antidepressant trials, and treatment-resistant depression then brings prolonged disability and a pressing need for interventions that can be chosen with greater precision. Recent randomized comparisons support repetitive transcranial magnetic stimulation (rTMS) as an effective next-step treatment and keep refining sequential and two-site approaches.[1,2] Multisite and naturalistic cohorts report clinically meaningful benefit in routine practice, but they also document marked variation in response magnitude and symptom trajectory.[3–5] Group-level efficacy cannot tell us who will respond.

Predictive biomarker research has therefore moved from clinical and neurophysiological measures toward molecular features. Peripheral blood markers change alongside rTMS and may capture systemic states related to outcome.[6] Multi-omic studies of depression have combined methylation with transcriptomic or protein-related information, aiming at better biological characterization and risk prediction.[7,8] Locus-specific methylation has been related to cortical structure and to clinically distinct forms of depression, and twin-based

methylome-expression analyses support measurable peripheral epigenetic heterogeneity.[9–11] DNA methylation is attractive for these reasons, but a statistical association in one treatment setting is not automatically a portable treatment-selection marker.

The five candidate cytosine–phosphate–guanine (CpG) sites evaluated here came from an integrated genome-wide methylation and expression study, which identified antidepressant-response signals annotated to CHN2, JAK2, SCN7A, and IQANK1.[12] Later pharmacotherapy investigations reported genome-wide, pathway-level, and targeted methylation correlates of response, involving immune-inflammatory processes, RPS6KA5, ion-channel genes, SLC6A4, HTR1A/HTR1B, and other treatment-sensitive loci.[13–22] One randomized treatment study included methylation analysis. It shows how researchers are trying to connect molecular change with clinical improvement.[23] Other treatment modalities widen this evidence base without making it uniform. Methylome-wide and targeted studies of electroconvulsive therapy found response-related changes, yet no universal panel emerged.[24–26] Longitudinal studies of treatment response and psychotherapy suggest that baseline state, early molecular change, and remission may be different biological questions.[27,28] Methylation scores and multimodal models improve prediction in some development samples. Their usefulness, however, rests on transportability, calibration, and external validation.[29,30]

The public GSE300009 cohort offers an unusually relevant setting for a stringent transfer test; it holds baseline whole-blood methylation and paired Patient Health Questionnaire-9 (PHQ-9) measurements from patients who received an rTMS course for treatment-resistant depression.[31] The parent investigation took an epigenome-scale regional approach to characterize rTMS outcomes. A focused candidate-panel analysis poses a different, related question: does a small antidepressant-derived CpG set keep its outcome information when it crosses treatment modality? A portable marker might index shared response biology. A failure to transfer can prevent premature generalization and sharpen later validation designs.

This study provides a cross-treatment evaluation of an antidepressant-response DNA methylation panel in an independent rTMS cohort. Association, multiplicity-adjusted evidence, robustness, and incremental out-of-sample prediction are explicitly kept apart. The primary aim was to determine whether baseline methylation at five CpGs in CHN2, JAK2, SCN7A, and IQANK1 was associated with continuous PHQ-9 improvement after rTMS; secondary aims were to assess associations with response and remission, and to determine whether the panel added predictive information beyond age, sex, baseline severity, and antidepressant use.

2. Methods

2.1. Study design and setting

Study design and data source. This was a secondary analysis of de-identified public data from the Gene Expression Omnibus (GEO) series GSE300009, in which the candidate CpGs and endpoints were defined before outcome modeling; the parent cohort recruited adults with treatment-resistant depression. They underwent a clinical rTMS course and gave peripheral blood before treatment.[31]

The reanalysis followed an association-and-prediction framework. The association component estimated direction, magnitude, and uncertainty for each candidate CpG-outcome relation. The prediction component compared clinical-only and clinical-plus-CpG models under paired resampling. We drew this distinction on purpose. A locus can be associated with an outcome and still add no useful prediction, particularly in a small cohort. Reporting followed the STROBE statement[32] and the RECORD statement[33], which extends it to routinely collected and publicly deposited data.

2.2. Participants

Analytic population. We used the processed DNA methylation beta-value matrix and the sample-level clinical metadata deposited with the series. Every participant with a baseline methylation profile and complete baseline and end-treatment PHQ-9 measurements entered the analytic population, which gave 46 participants. Selection never depended on methylation value or outcome. The sample size was fixed by the deposited cohort (46 participants with complete PHQ-9 data), and no a priori sample-size calculation was applicable.

2.3. Variables and instruments

Clinical measures and endpoint reconstruction. Depressive-symptom severity was measured with the PHQ-9, a validated nine-item instrument with a total score from 0 to 27;[34] the primary outcome was improvement: baseline PHQ-9 minus end-treatment PHQ-9, so that positive values denote symptom reduction. Response (at least 50% improvement from baseline) and remission (end-treatment PHQ-9 below 5) served as secondary binary outcomes. Both were reconstructed directly from the numeric PHQ-9 fields, so the definitions were applied in the same way to everyone.

DNA methylation variables. The candidate panel consisted of cg06926818 and cg23687322 (annotated to CHN2), cg08584037 (JAK2), cg12958009 (SCN7A), and cg17944171 (IQANK1); these loci were chosen before outcome modeling, because they arose from the integrated antidepressant-response analysis cited above.[12] Deposited processed beta values were used. Repeating upstream normalization was not possible without raw intensity files. Beta values express methylated intensity as a proportion and are intuitive for description. M-values, defined as $\log_2[\beta/(1-\beta)]$, have

distributional properties better suited to regression inference.[35]

Beta values were bounded at 0.001 and 0.999 before M-value transformation, for numerical stability and for no other purpose; each candidate M-value was standardized to its sample standard deviation (SD), so linear coefficients and odds ratios describe a 1-SD higher baseline methylation value. The primary JAK2 locus was cg08584037, because the source study identified it in its results. The source abstract also named cg08339825. That alternate probe was evaluated separately as a sensitivity analysis and did not join the multiplicity family.

2.4. Data collection procedure

Data extraction and missing data. The ordering of selected characteristic labels differed across records of the GEO series matrix. Sample characteristics were therefore parsed at the cell level. Age, sex, baseline PHQ-9, current antidepressant use, and end-treatment PHQ-9 were extracted and checked against numeric ranges and source labels. Two stored response labels disagreed with the response recalculated from PHQ-9 values. The recalculated response was primary. The stored variable was kept for sensitivity analysis. Only participants with complete baseline and end-treatment PHQ-9 measurements were analyzed, and no values were imputed.

2.5. Statistical analysis

Regression models and multiplicity. Participant characteristics are given as count and percentage or as mean and standard deviation; each candidate CpG had its own model, which kept the number of parameters proportionate to the cohort. For continuous PHQ-9 improvement, ordinary least-squares regression included the standardized CpG M-value, age, sex, baseline PHQ-9, and antidepressant use. Heteroskedasticity-consistent HC3 standard errors were chosen because they lower sensitivity to leverage and to small-sample variance misspecification. Coefficients are adjusted mean differences in PHQ-9 improvement per 1-SD higher CpG methylation, with 95% confidence intervals (CIs).

Response and remission were analyzed with bias-reduced logistic regression and the same covariates. Penalized likelihood reduces finite-sample bias. It also limits unstable estimates when events are few or predictors approach separation.[36] We report odds ratios with profile-compatible 95% confidence intervals. p values from the five CpG tests were adjusted within each outcome by the Holm procedure. Evidence was read from effect sizes, confidence intervals, and adjusted p values, and not from nominal thresholds alone.

Robustness checks. Five robustness checks were run: unadjusted Spearman rank correlations for continuous improvement, leave-one-out refitting of every adjusted linear model, Cook's distance, analysis with the stored response label, and analysis of the alternate JAK2 probe. Leave-one-out sign stability was the proportion of refits that kept the full-sample coefficient direction; the checks looked for dependence on any single observation. They also asked whether the endpoint or probe discrepancy materially changed interpretation.

Incremental prediction. Incremental prediction was assessed with 100 repetitions of nested five-fold cross-validation. In each outer training sample, clinical-only and clinical-plus-CpG elastic-net models were tuned in an inner five-fold cross-validation, with paired partitions and random seeds. Nesting separates model selection from performance estimation and limits optimistic error estimates.[37] Clinical predictors were age, sex, baseline PHQ-9, and antidepressant use. The expanded model added all five standardized CpGs. The conservative one-standard-error penalty was primary, since it favors the more parsimonious model when estimated errors are similar; the minimum-error penalty was a sensitivity analysis. Continuous performance was summarized by root-mean-square error (RMSE), mean absolute error, and predictive R^2. Binary performance was summarized by area under the receiver-operating-characteristic curve (AUC) and Brier score. We report mean paired differences with empirical 2.5th–97.5th percentile ranges across repetitions. No final model is offered as clinically deployable, because one cohort necessarily supplied both model development and internal validation.

Software. Analyses ran in R version 4.5.3 with data.table, sandwich, lmtest, brglm2, glmnet, and ggplot2. One analysis script generated every manuscript table and figure from the downloaded data. Two-sided p values are given to three decimals unless extra precision helps interpretation.

2.6. Ethical approval

The CMHC Research Center Ethics Committee Indonesia granted ethical approval (No. 2026/029). The parent study reported institutional review and informed consent. This secondary analysis involved no new participant contact and used de-identified public data.

3. Results

3.1. Cohort characteristics and endpoint reconstruction

The analytic cohort comprised 46 participants; 31 (67.4%) were female, and mean age was 44.8 years (SD 18.3). Mean baseline PHQ-9 was 22.3 (SD 2.5), consistent with severe depressive symptoms. Mean end-treatment PHQ-9 was 8.9 (SD 7.2), corresponding to a mean improvement of 13.3 points (SD 6.9). Calculated response was present in 31 participants (67.4%), remission in 18 (39.1%), and 28 (60.9%) were receiving an antidepressant. These cohort characteristics and outcome frequencies are given in Table 1. Individual baseline-to-end-treatment trajectories are shown in Figure 1 and demonstrate substantial average improvement alongside marked between-participant heterogeneity.

Table 1. Participant characteristics and clinical outcomes.

Characteristic	Value
Analytic sample	46
Female	31 (67.4%)
Age, years	44.8 (18.3)
Baseline PHQ-9	22.3 (2.5)
End-treatment PHQ-9	8.9 (7.2)
PHQ-9 improvement	13.3 (6.9)
Calculated response (≥50% improvement)	31 (67.4%)
Remission (end-treatment PHQ-9 <5)	18 (39.1%)
Antidepressant use	28 (60.9%)

Values are mean (standard deviation, SD) or n (%). PHQ-9, Patient Health Questionnaire-9; PHQ-9 improvement is the baseline score minus the end-treatment score. n = 46.

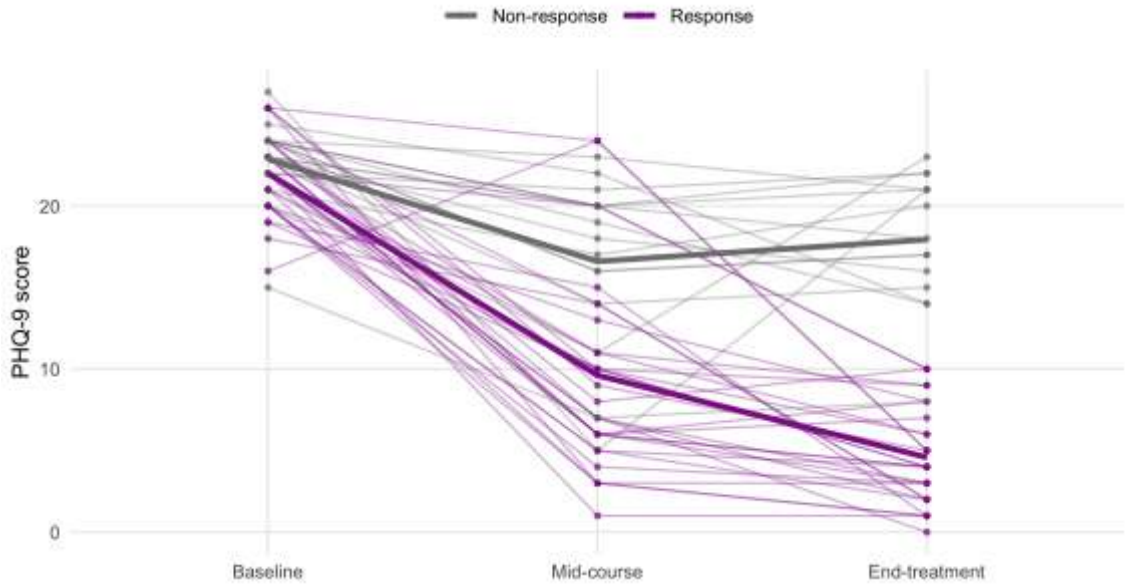


Figure 1. Individual PHQ-9 trajectories from baseline through the mid-course visit to treatment end. Thin lines represent individual participants and heavy lines the group means, purple for participants meeting calculated response and gray for non-response. Lower scores indicate fewer depressive symptoms. n = 46.

Two participants had a stored response label that was inconsistent with the calculation from paired PHQ-9 scores specified in Methods. One participant's PHQ-9 increased from 21 to 23 but was labeled a responder; another improved from 15 to 8, a 46.7% reduction, but was also labeled a responder. Both were classified as nonresponders in the primary recalculated endpoint.

3.2. Associations with continuous PHQ-9 improvement

Neither CHN2 CpG demonstrated a statistically supported association with continuous improvement after adjustment. Per 1-SD higher methylation, cg06926818 was associated with 1.11 points less improvement (95% CI −3.73 to 1.51; p = 0.410), and cg23687322 with 1.08 points less improvement (95% CI −2.94 to 0.78; p = 0.263). The complete adjusted estimates appear in Table 2 and are visualized in Figure 2; the confidence intervals included both a clinically relevant inverse association and little or no association.

Table 2. Adjusted associations between baseline candidate CpGs and PHQ-9 improvement.

CpG (gene)	Adjusted difference (β)	95% CI	p	Holm-adjusted p
cg06926818 (CHN2)	−1.11	−3.73 to 1.51	0.410	1.000
cg23687322 (CHN2)	−1.08	−2.94 to 0.78	0.263	1.000
cg08584037 (JAK2)	−0.16	−2.81 to 2.50	0.909	1.000
cg12958009 (SCN7A)	0.31	−2.87 to 3.50	0.848	1.000
cg17944171 (IQANK1)	−0.18	−3.63 to 3.26	0.918	1.000

Adjusted difference (β) is the regression coefficient: the adjusted difference in PHQ-9 improvement, in points, per 1-SD higher baseline CpG M-value. Models adjusted for age, sex, baseline PHQ-9, and antidepressant use; 95% confidence intervals (CIs) use HC3 standard errors; Holm-adjusted p is the Holm-corrected p value across the five CpGs. n = 46.

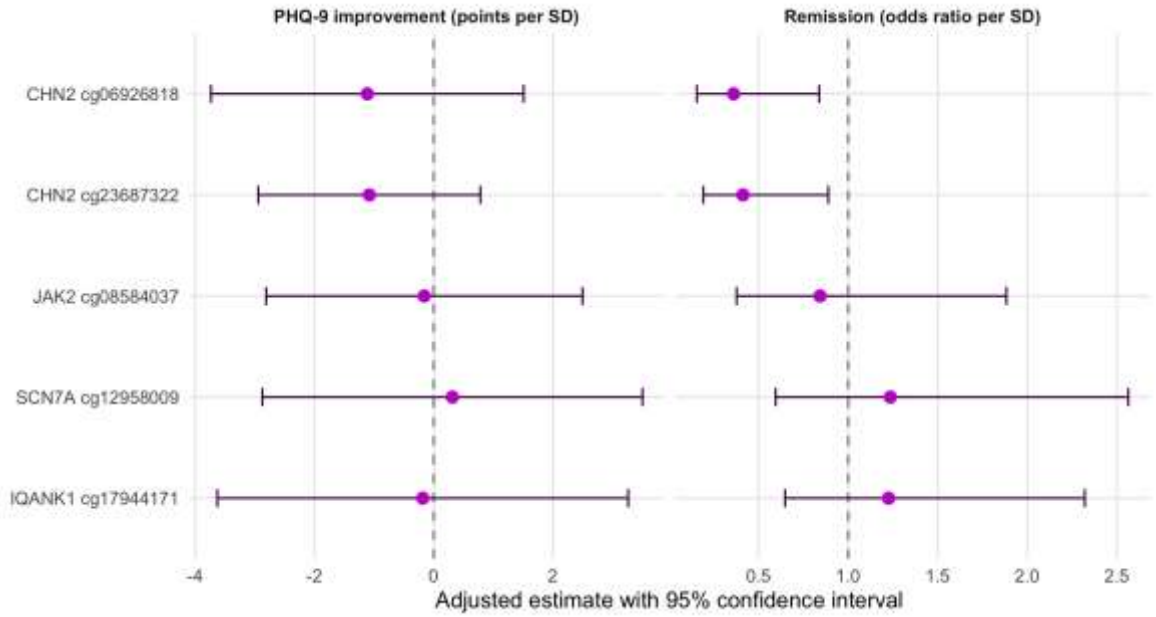


Figure 2. Adjusted associations between baseline candidate CpGs and PHQ-9 improvement (left panel) and remission (right panel). Points are adjusted mean differences or odds ratios per 1-SD higher M-value; bars are 95% confidence intervals from HC3 linear or bias-reduced logistic models. Models adjust for age, sex, baseline PHQ-9, and antidepressant use. Vertical reference lines mark no association (0 for improvement; 1 for remission). n = 46. CI, confidence interval.

The JAK2, SCN7A, and IQANK1 estimates were close to the null and imprecise: cg08584037, −0.16 points (95% CI −2.81 to 2.50; p = 0.909); cg12958009, 0.31 points (95% CI −2.87 to 3.50; p = 0.848); and cg17944171, −0.18 points (95% CI −3.63 to 3.26; p = 0.918). All Holm-adjusted p values were 1.000. Spearman correlations ranged from −0.20 to 0.02 and did not alter the interpretation.

3.3. Associations with response and remission

No candidate CpG was associated with calculated response after covariate adjustment. Adjusted odds ratios ranged from 0.78 for cg06926818 to 1.46 for cg12958009, and every confidence interval included 1.00. Holm-adjusted p values were 1.000. The paired response and remission estimates are summarized in Table 3. Repeating the models with the stored response label also produced no supported association; odds ratios ranged from 0.79 to 1.31, with all adjusted p values equal to 1.000.

Table 3. Adjusted associations between baseline candidate CpGs and response and remission.

CpG (gene)	Response OR (95% CI)	p / Holm-adjusted p	Remission OR (95% CI)	p / Holm-adjusted p
cg06926818 (CHN2)	0.78 (0.38–1.61)	0.507 / 1.000	0.36 (0.15–0.84)	0.018 / 0.089
cg23687322 (CHN2)	0.85 (0.43–1.69)	0.649 / 1.000	0.41 (0.19–0.89)	0.024 / 0.095
cg08584037 (JAK2)	1.06 (0.46–2.47)	0.886 / 1.000	0.84 (0.38–1.88)	0.676 / 1.000
cg12958009 (SCN7A)	1.46 (0.68–3.12)	0.333 / 1.000	1.23 (0.59–2.56)	0.571 / 1.000
cg17944171 (IQANK1)	1.43 (0.71–2.87)	0.317 / 1.000	1.23 (0.65–2.32)	0.533 / 1.000

OR, odds ratio; CI, confidence interval. Odds ratios are per 1-SD higher baseline M-value from bias-reduced logistic models adjusted for age, sex, baseline PHQ-9, and antidepressant use. Each p value is followed by its Holm-adjusted p value within the outcome. Response: 31/46; remission: 18/46.

For remission, the two CHN2 CpGs showed nominal inverse associations. The adjusted odds ratio was 0.36 per 1-SD higher cg06926818 M-value (95% CI 0.15–0.84; p = 0.018) and 0.41 per 1-SD higher cg23687322 M-value (95% CI 0.19–0.89; p = 0.024). After accounting for the five candidate tests, Holm-adjusted p values were 0.089 and 0.095, respectively. JAK2, SCN7A, and IQANK1 estimates were not supported, with odds ratios of 0.84, 1.23, and 1.23 and confidence intervals crossing 1.00. The contrast between the CHN2 estimates and the remaining loci is also shown in Figure 2. Thus, the CHN2 results were retained as hypothesis-generating remission signals rather than multiplicity-adjusted findings.

3.4. Influence, probe sensitivity, and prediction

Leave-one-out analysis preserved the negative direction of both CHN2 continuous coefficients in all 46 refits. The cg06926818 coefficient ranged from −2.04 to −0.54 and the cg23687322 coefficient from −1.53 to −0.69. Direction was less stable for the three near-null loci, as expected when estimates straddled zero. Maximum Cook's distances were below 0.32. These results suggest that the modest inverse CHN2 continuous estimates were not generated by a single participant, while their uncertainty remained substantial. The alternate JAK2 probe cg08339825 yielded an adjusted coefficient of −1.02 points (95% CI

−2.98 to 0.93; p = 0.311), which did not change the panel-level conclusion.

Under primary one-standard-error regularization, the expanded model produced no change in continuous-improvement RMSE relative to clinical predictors alone (both 6.980), and no change in mean absolute error or predictive R^2. The clinical-only and expanded performance estimates appear in Table 4. For response, mean AUC changed from 0.484 to 0.496 (difference 0.012; empirical range −0.017 to 0.095), and Brier score changed from 0.221 to 0.220 (difference −0.001). For remission, mean AUC changed from 0.489 to 0.548 (difference 0.059; range −0.050 to 0.173), while Brier score changed from 0.239 to 0.237 (difference −0.002; range −0.009 to 0.005). Each empirical range remained compatible with no incremental benefit.

Table 4. Incremental prediction under repeated nested cross-validation.

Outcome / metric	Clinical	+ CpGs	Paired Δ (2.5th–97.5th percentile)
Improvement RMSE	6.980	6.980	0.000 (0.000 to 0.000)
Improvement MAE	5.872	5.872	0.000 (0.000 to 0.000)
Improvement predictive R^2	−0.050	−0.050	0.000 (0.000 to 0.000)
Response AUC	0.484	0.496	0.012 (−0.017 to 0.095)
Response Brier score	0.221	0.220	−0.001 (−0.008 to 0.000)
Remission AUC	0.489	0.548	0.059 (−0.050 to 0.173)
Remission Brier score	0.239	0.237	−0.002 (−0.009 to 0.005)

RMSE, root-mean-square error; MAE, mean absolute error; R^2, predictive R-squared; AUC, area under the receiver-operating-characteristic curve. Primary one-standard-error regularization results, averaged over 100 repetitions of nested five-fold cross-validation. Lower RMSE, MAE, and Brier scores are better; higher predictive R^2 and AUC are better. Identical clinical-only and clinical-plus-CpG values for continuous improvement indicate that the elastic net retained no CpG under this rule.

Minimum-error selection, reported as a less conservative sensitivity analysis, also showed negligible continuous performance change and lower response AUC after adding CpGs. Remission AUC increased by 0.156 on average, but the Brier-score difference was small and its empirical range crossed zero. The attenuation under the one-standard-error rule and the absence of consistent calibration improvement indicate that the apparent remission discrimination was sensitive to tuning and should not be interpreted as validated predictive utility.

4. Discussion

4.1. Principal findings

This cross-treatment transferability analysis did not support the five-CpG antidepressant-response panel as a dependable marker of continuous PHQ-9 improvement or calculated response after rTMS. Two CHN2 loci showed nominal inverse associations with remission. Neither met the Holm-adjusted threshold. Conservative repeated nested cross-validation found no incremental value for continuous improvement and little change for response; for remission the AUC change was modest and unstable, with no consistent Brier-score improvement. Taken together, the association and prediction results argue against immediate clinical use. They do identify a focused CHN2-remission hypothesis for larger external evaluation.

Clinical change in this cohort was large. Mean PHQ-9 improved by 13.3 points, and approximately two thirds met calculated response. Even so, broad clinical improvement does not guarantee that a biomarker found for pharmacotherapy will rank rTMS outcomes. Response biology may include shared final pathways, treatment-specific mechanisms, and features that mainly index prognosis rather than treatment effect; the present results carry information even where null. They delimit how far an antidepressant-derived signature can be generalized in its current form.

4.2. Comparison with the existing literature

The present epigenetic result belongs beside a rapidly diversifying rTMS biomarker literature. Randomized mechanistic work has linked intermittent theta-burst stimulation to changes in prefrontal functional connectivity. Connectivity mapping and longitudinal structural-network studies have tested circuit-level correlates of response.[38–40] More recent EEG and machine-learning analyses, trajectory models, and nonlinear prediction pipelines try to estimate individual outcome from baseline or early-treatment information.[41–43] Clinical studies of remission predictors add that performance depends on population, endpoint, and validation design, not on biomarker class alone.[44]

Clinical heterogeneity is also visible in studies of adverse childhood experiences, sleep quality, comorbid anxiety, symptom clusters, early improvement, response trajectories, and residual symptoms after remission;[45–51] these findings offer a reasonable reason for the restrained molecular transportability. A fixed CpG panel may compete with several clinical routes to improvement, and it may behave differently for relative response, absolute remission, or time-dependent trajectories. The current study therefore adds a stringent cross-treatment test and not another within-cohort discovery analysis.

4.3. Interpretation and mechanism

The CHN2 remission pattern. CHN2 encodes beta-chimaerin, a GTPase-activating protein involved in Rac signaling and neuronal circuit development; both CHN2 CpGs share an inverse direction across continuous improvement and remission, which is internally consistent, and leave-one-out estimates kept that direction. Coherence, however, does not equal

multiplicity-adjusted confirmation. The continuous confidence intervals included no association. Response models were null. Remission-adjusted p values stayed above 0.05. Correlation between two loci of one gene may make a five-test correction conservative in a biological sense, but changing the multiplicity framework after seeing the results would overstate the evidence. The appropriate interpretation is a gene-focused signal whose endpoint specificity is uncertain.

Remission may capture a clinical construct different from relative response. A participant can improve by at least 50% and still be symptomatic. Remission requires a low final symptom burden. Baseline molecular features could conceivably relate to the chance of reaching that low state, through inflammatory, synaptic, or stress-response pathways. Here, though, only 18 remission events were available. Even bias-reduced regression cannot create information that the sample lacks, and the fairly wide intervals around the odds ratios leave the magnitude uncertain. Future work should prespecify whether remission is the primary binary endpoint. It should not choose remission because its p value is smaller.

Statistical and clinical importance must also be kept apart. An odds ratio near 0.4 per SD would be meaningful if replicated. Odds ratios do not describe absolute probability directly, though, and they can look large when event counts are modest. Clinical utility would demand risk estimates calibrated across settings, net benefit at reasonable decision thresholds, and an improvement over readily available clinical predictors. None of these conditions was met here.

Blood and brain. Differences between blood and brain remain central to interpretation. Peripheral methylation varies across clinically distinct depression phenotypes, and stress-axis methylation has been related to brain structure. Neither observation establishes locus-specific neural mediation.[52,53] Antidepressant exposure can also change tissue-specific methylomes, so treatment history and sampling compartment may shape the measured signals.[54] Whole-blood methylation further reflects leukocyte mixture, smoking exposure, ancestry, age, medicines, and other environmental or physiological influences. A peripheral signature could still be clinically useful if it predicts reproducibly. It should not be called a direct neural mechanism without functional evidence.

Association, prediction, and transportability. A strength of the analysis is the explicit separation of association from prediction. Regression coefficients ask whether average outcome differs with a CpG after covariate adjustment. Nested cross-validation asks a different thing: does adding the whole panel improve performance for participants who were not used to fit the model? The two questions can diverge, since predictive performance depends on signal-to-noise ratio, correlation among predictors, regularization, and sample size. The CHN2-remission associations coexisted with only modest predictive improvement under conservative tuning. This illustrates why nominal associations alone should not be presented as a usable test.

Minimum-error tuning produced a larger average remission AUC difference. Reported alone, that result could look encouraging. The one-standard-error rule, the paired Brier scores, and the empirical ranges give the essential context; the added discrimination attenuated materially under more parsimonious regularization, and probability accuracy did not improve consistently. Such tuning sensitivity is typical of small validation samples, where a few observations can change the selected coefficients. The appropriate next translational step is not to freeze a classifier from these 46 participants; it is to test a locked panel and model specification in a larger cohort.

Transportability also depends on endpoint harmonization. Reconstructing response revealed two disagreements with the deposited label. The stored-label sensitivity analysis stayed null, so the main inference did not hinge on that choice. Still, the discrepancy shows how a seemingly minor endpoint difference can alter classification in a small study. Future consortia should deposit derivation code next to outcome labels. They should declare how missing visits are handled, and they should distinguish relative response, absolute remission, trajectory, and sustained outcome.

4.4. Implications for practice and policy

Clinical interpretation. For clinical interpretation, the current panel should not be used to recommend, withhold, or sequence rTMS; the analysis found no dependable transfer for continuous improvement or response. The remission result stayed hypothesis-generating after multiplicity correction. This distinction protects patients from premature biomarker claims and keeps a testable lead whose size remains uncertain.

Research implications. For research, the findings support a staged validation strategy. The next study should preregister the CHN2-remission hypothesis and use harmonized response and remission definitions; it should collect detailed rTMS and clinical covariates, estimate leukocyte composition, and reserve a genuinely independent sample for performance assessment. Measurement reliability and between-batch reproducibility should be documented before model fitting. If discrimination replicates, calibration and decision-curve analysis should follow. If it does not, the field should prioritize treatment-specific or multimodal markers over repeated optimization of the same small dataset.

Design priorities for external validation. A credible external validation should start from a written analysis protocol. The protocol should lock the five CpGs, the direction of the CHN2 hypothesis, the outcome hierarchy, and the clinical comparator before any outcome data are analyzed. Recruitment should span more than one rTMS service and should capture stimulation target, frequency, intensity, session number, coil type, treatment interruptions, concomitant psychotherapy, and medication changes; these details

would allow assessment of case-mix and protocol transportability without splitting a modest sample into underpowered subgroups. Repeated PHQ-9 measurements should carry time stamps. End-of-course outcome, early trajectory, and durability could then be told apart.

Molecular procedures need equally careful harmonization. Blood collection time, storage, DNA extraction, array lot, plate position, and laboratory should be recorded prospectively. Technical replicates and blinded quality-control samples can quantify assay reliability. Estimated leukocyte proportions, smoking, body-mass index, ancestry-related structure, inflammatory illness, and relevant medicines should be collected, since each can influence peripheral methylation; they should enter the model according to a prespecified causal and predictive rationale, not because they were screened for significance. The locked candidate model should be applied without cohort-specific probe selection. Any newly discovered features belong in a separate development analysis.

Sample-size planning should rest on precision and model validation, not on an anticipated nominal p value. For association, the target is a confidence interval narrow enough to separate a clinically relevant effect from a trivial one. For prediction, the study needs enough responders and remitters to estimate AUC, calibration intercept and slope, Brier score, and decision-curve net benefit with useful precision. If the external cohort is assembled across centers, internal-external cross-validation can quantify heterogeneity and still preserve a final holdout or a temporally later cohort for confirmation. Reporting should show center-specific case mix and performance. A pooled average alone is not enough.

The decision rule should be defined only after reproducibility is established. A marker that modestly improves remission discrimination may have no clinical value if it changes no treatment decision at a meaningful threshold. Conversely, a small average performance gain might matter if it reliably identifies a group for whom treatment sequencing changes. A future study should therefore state the intended decision, the alternative care pathway, the acceptable false-positive and false-negative consequences, and the incremental cost of methylation measurement; this sequence runs from analytic validity through external performance to decision utility. It would turn the present hypothesis into a clinically interpretable validation program.

4.5. Strengths and limitations

Strengths. The study has several strengths. Candidate loci were specified from external antidepressant evidence. All eligible participants were included. Continuous and binary outcomes were analyzed with methods suited to their distributions and to the sample size, multiplicity was controlled by outcome family, and effect sizes came with uncertainty. Endpoint reconstruction, alternate-probe analysis, rank correlations, influence diagnostics, and the stored-label sensitivity analysis made data-to-result decisions auditable. Repeated nested cross-validation used paired splits for the clinical-only and expanded models, which reduces noise in performance differences; the complete analysis was generated reproducibly from public data.

Limitations. Several limitations define the scope of the conclusions. First, n = 46 limits precision and the number of covariates that can be supported. Remission had only 18 events, so even bias-reduced estimates remain uncertain. Second, this was a retrospective secondary analysis of one cohort. Internal resampling evaluates optimism, but it does not replace external validation. Third, treatment protocol details and clinical histories were not available at the granularity needed to test interactions by coil, stimulation parameters, treatment adequacy, or resistance stage.

Fourth, whole-blood cell composition, smoking, body-mass index, ancestry, and additional medications could not be adjusted comprehensively; each may influence methylation and outcome, so residual confounding remains. Fifth, processed beta values were analyzed because raw array intensities were not part of the working dataset. Intensity-level preprocessing and probe-level quality control could not be repeated. Sixth, annotation to a gene does not establish that the measured CpG regulates that gene or mediates rTMS response. No causal, tissue-specific, or mechanistic inference is made.

Finally, the analysis evaluated five individual CpGs and their joint predictive contribution. It did not examine regional methylation, longitudinal change, gene expression, or multimodal integration. The recent studies cited throughout the Introduction show that methylation associations vary by locus, treatment, tissue, and outcome definition. A broader, biologically informed signature or a treatment-specific rTMS panel may perform differently. Such an expansion belongs in development cohorts of adequate size, followed by testing against a locked external protocol.

5. Conclusion

In conclusion, five antidepressant-response CpGs did not transfer as a dependable baseline signature of PHQ-9 improvement or response in this independent rTMS cohort. Two CHN2 loci showed internally consistent but multiplicity-sensitive associations with remission. Incremental predictive performance was not stable under conservative validation. Larger, prospectively harmonized external studies must show whether CHN2 contributes reproducible remission information, or whether rTMS outcomes need treatment-specific molecular models.

Declarations

Use of Generative Artificial Intelligence. None. The authors declare that no generative artificial intelligence tool was used in the conception, research, analysis, drafting, revision or preparation of this manuscript.

Author Contributions. R.H.: conceptualization, methodology, software, formal analysis, data curation, visualization, writing – original draft. A.B.: validation,

writing – review & editing. S.K.: supervision, writing – review & editing. All authors approved the final manuscript.

Funding. This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest. The authors declare that they have no competing interests, financial or non-financial, in relation to this work.

Ethical Approval. Ethical approval was granted by the CMHC Ethics Committee Indonesia (No. 2026/029).

Informed Consent. The parent study reported institutional review and informed consent. This secondary analysis used de-identified public data and involved no new participant contact. Consent for publication is not applicable; no identifiable individual information is presented.

Data Availability. The source dataset is publicly available from the NCBI Gene Expression Omnibus under accession GSE300009. The reproducible analysis script and derived tables are maintained with the study record.

Acknowledgements. The authors acknowledge the investigators and participants of GSE300009 for making de-identified data available for secondary research.

References

1. Dalhuisen I, van Oostrom I, Spijker J, et al. rTMS as a next step in antidepressant nonresponders: a randomized comparison with current antidepressant treatment approaches. Am J Psychiatry. 2024;181(9):806-814. doi:10.1176/appi.ajp.20230556
2. Zhao YJ, Xiang S, Chen R, et al. A sequential dual-site repetitive transcranial magnetic stimulation for major depressive disorder: a randomized clinical trial. Cell Rep Med. 2025;6(10):102402.
   doi:10.1016/j.xcrm.2025.102402
3. Matsuda Y, Kito S, Hiraki F, et al. A multisite observational real-world study on the effectiveness of repetitive transcranial magnetic stimulation therapy for patients with treatment-resistant depression in Japan. Psychiatry Res. 2024;342:116263.
   doi:10.1016/j.psychres.2024.116263
4. Bouaziz N, Laidi C, Bulteau S, et al. Real world transcranial magnetic stimulation for major depression: a multisite, naturalistic, retrospective study. J Affect Disord. 2023;326:26-35. doi:10.1016/j.jad.2023.01.070
5. Chen X, Blumberger DM, Downar J, et al. Depressive symptom trajectories with prolonged rTMS treatment. Brain Stimul. 2024;17(3):525-532.
   doi:10.1016/j.brs.2024.04.010
6. Ozkan BN, Bozali K, Boylu ME, et al. Altered blood parameters in "major depression" patients receiving repetitive transcranial magnetic stimulation (rTMS) therapy: a randomized case-control study. Transl Psychiatry. 2024;14(1):264. doi:10.1038/s41398-024-02942-8
7. Mokhtari A, Ibrahim EC, Gloaguen A, et al. Using multiomic integration to improve blood biomarkers of major depressive disorder: a case-control study. EBioMedicine. 2025;113:105569.
   doi:10.1016/j.ebiom.2025.105569
8. Diniz BS, Zhao S, Drouard G, et al. Biological aging acceleration in major depressive disorder: a multi-omics analysis. Aging Cell. 2026;25(1):e70310. doi:10.1111/acel.70310
9. Han KM, Choi KW, Kim A, et al. Association of DNA methylation of the NLRP3 gene with changes in cortical thickness in major depressive disorder. Int J Mol Sci. 2022;23(10):5768. doi:10.3390/ijms23105768
10. Yamagata H, Ogihara H, Matsuo K, et al. Distinct epigenetic signatures between adult-onset and late-onset depression. Sci Rep. 2021;11(1):2296. doi:10.1038/s41598-021-81758-8
11. Wang W, Li W, Wu Y, et al. Genome-wide DNA methylation and gene expression analyses in monozygotic twins identify potential biomarkers of depression. Transl Psychiatry. 2021;11(1):416. doi:10.1038/s41398-021-01536-y
12. Ju C, Fiori LM, Belzeaux R, et al. Integrated genome-wide methylation and expression analyses reveal functional predictors of response to antidepressants. Transl Psychiatry. 2019;9(1):254. doi:10.1038/s41398-019-0589-0
13. Lyu N, Zhao Q, Liu D, et al. Genome-wide association and DNA methylation analyses of SSRI treatment response in major depressive disorder. BMC Psychiatry. 2025;25(1):1030. doi:10.1186/s12888-025-07502-y
14. Schiele MA, Crespo Salvador O, Lipovsek J, et al. Epigenome-wide DNA methylation in unipolar depression: predictive biomarker of antidepressant treatment response? int j neuropsychopharmacol. 2024;27(11):pyae045. doi:10.1093/ijnp/pyae045
15. Engelmann J, Zillich L, Frank J, et al. Epigenetic signatures in antidepressant treatment response: a methylome-wide association study in the EMC trial. Transl Psychiatry. 2022;12(1):268. doi:10.1038/s41398-022-02032-7
16. Fuh SC, Fiori LM, Turecki G, et al. Multi-omic modeling of antidepressant response implicates dynamic immune and inflammatory changes in individuals who respond to treatment. PLoS One. 2023;18(5):e0285123. doi:10.1371/journal.pone.0285123
17. Fiori LM, Nagy C, Turecki G, et al. Identification of antidepressant response-related changes to DNA methylation and gene expression. Int J Neuropsychopharmacol. 2026;29(2):pyag001. doi:10.1093/ijnp/pyag001
18. Xu P, Tao Y, Zhang H, et al. RPS6KA5 methylation predict response to 6-week treatment for adolescent MDD patients. BMC Psychiatry. 2022;22(1):561. doi:10.1186/s12888-022-04196-4
19. Martinez-Pinteño A, Rodriguez N, Blázquez A, et al. DNA methylation of fluoxetine response in child and adolescence: preliminary results. Pharmgenomics Pers Med. 2021;14:459-467. doi:10.2147/PGPM.S289480
20. Nascimento LV, Neto FL, Ribeiro Moreira DA, et al. Influence of antidepressant drugs on DNA methylation of ion channels genes in blood cells of psychiatric patients. Epigenomics. 2022;14(14):851-864. doi:10.2217/epi-2022-0089
21. Moon YK, Kim H, Kim S, et al. Influence of antidepressant treatment on SLC6A4 methylation in Korean patients with major depression. Am J Med Genet B Neuropsychiatr Genet. 2023;192(1-2):28-37. doi:10.1002/ajmg.b.32921
22. Xu Z, Chen Z, Shen T, et al. The impact of HTR1A and HTR1B methylation combined with stress/genotype on early antidepressant efficacy. Psychiatry Clin Neurosci. 2022;76(2):51-57. doi:10.1111/pcn.13314
23. Fan L, Zeng P, Wang X, et al. Xiaoyao pills, a Chinese patent medicine, treats mild and moderate depression: a randomized clinical trial combined with DNA methylation analysis. Phytomedicine. 2024;130:155660. doi:10.1016/j.phymed.2024.155660
24. Sirignano L, Frank J, Kranaster L, et al. Methylome-wide change associated with response to electroconvulsive therapy in depressed patients. Transl Psychiatry. 2021;11(1):347. doi:10.1038/s41398-021-01474-9

25. Carvalho Silva R, Martini P, Hohoff C, et al. Unraveling epigenomic signatures and effectiveness of electroconvulsive therapy in treatment-resistant depression patients: a prospective longitudinal study. Clin Epigenetics. 2024;16(1):93. doi:10.1186/s13148-024-01704-z
26. Maier HB, Moschny N, Eberle F, et al. DNA methylation of POMC and NR3C1-1F and its implication in major depressive disorder and electroconvulsive therapy. Pharmacopsychiatry. 2023;56(2):64-72. doi:10.1055/a-2034-6536
27. Van Assche E, Hohoff C, Zang J, et al. Longitudinal early epigenomic signatures inform molecular paths of therapy response and remission in depressed patients. Front Mol Neurosci. 2023;16:1223216.
doi:10.3389/fnmol.2023.1223216
28. Carvalho Silva R, Martini P, Hohoff C, et al. DNA methylation changes in association with trauma-focused psychotherapy efficacy in treatment-resistant depression patients: a prospective longitudinal study. Eur J Psychotraumatol. 2024;15(1):2314913. doi:10.1080/20008066.2024.2314913
29. Barbu MC, Amador C, Kwong ASF, et al. Complex trait methylation scores in the prediction of major depressive disorder. EBioMedicine. 2022;79:104000. doi:10.1016/j.ebiom.2022.104000
30. Ha S, Kang HJ, Lee T, et al. Multimodal machine learning models for predicting remission in major depressive disorder using clinical data, blood biomarkers, and DNA methylation. J Affect Disord. 2026;402:121259. doi:10.1016/j.jad.2026.121259
31. Dahrendorff J, Pages K, Currier G, et al. Epigenetic profiles of response to transcranial magnetic stimulation in treatment resistant depression. BMC Med Genomics. 2025;18(1):181. doi:10.1186/s12920-025-02240-2
32. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457. doi:10.1016/S0140-6736(07)61602-X
33. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885
34. Kroenke K, Spitzer RL, Williams JBW, et al. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x
35. Du P, Zhang X, Huang CC, et al. Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587. doi:10.1186/1471-2105-11-587
36. Kosmidis I, Kenne Pagui EC, Sartori N, et al. Mean and median bias reduction in generalized linear models. Stat Comput. 2020;30(1):43-59. doi:10.1007/s11222-019-09860-6
37. Vabalas A, Gowen E, Poliakoff E, et al. Machine learning algorithm validation with a limited sample size. PLoS One. 2019;14(11):e0224365. doi:10.1371/journal.pone.0224365
38. Struckmann W, Bodén R, Gingnell M, et al. Modulation of dorsolateral prefrontal cortex functional connectivity after intermittent theta-burst stimulation in depression: combining findings from fNIRS and fMRI. Neuroimage Clin. 2022;34:103028. doi:10.1016/j.nicl.2022.103028
39. Elbau IG, Lynch CJ, Downar J, et al. Functional connectivity mapping for rTMS target selection in depression. Am J Psychiatry. 2023;180(3):230-240.
doi:10.1176/appi.ajp.20220306
40. Nestor SM, Mir-Moghtadaei A, Vila-Rodriguez F, et al. Large-scale structural network change correlates with clinical response to rTMS in depression. Neuropsychopharmacology. 2022;47(5):1096-1105. doi:10.1038/s41386-021-01256-3
41. Bailey NW, Fulcher BD, Arns M, et al. Prediction of response to transcranial magnetic stimulation treatment for depression using EEG and machine learning. Clin Neurophysiol. 2025;178:2110937. doi:10.1016/j.clinph.2025.2110937
42. McInnes AN, Olsen ST, Sullivan CRP, et al. Trajectory modeling and response prediction in transcranial magnetic stimulation for depression. Pers Med Psychiatry. 2024;47-48:100135. doi:10.1016/j.pmip.2024.100135
43. Ebrahimzadeh E, Dehghani A, Asgarinejad M, et al. Non-linear processing and reinforcement learning to predict rTMS treatment response in depression. Psychiatry Res Neuroimaging. 2024;337:111764. doi:10.1016/j.pscychresns.2023.111764
44. Göke K, Trevizol AP, Ma C, et al. Predictors of remission after repetitive transcranial magnetic stimulation for the treatment of late-life depression. Psychiatry Res. 2024;334:115822. doi:10.1016/j.psychres.2024.115822
45. Ng E, Wong EHY, Lipsman N, et al. Adverse childhood experiences and repetitive transcranial magnetic stimulation outcomes for depression. J Affect Disord. 2023;320:716-724. doi:10.1016/j.jad.2022.09.153
46. Kweon J, Fukuda AM, Gobin AP, et al. Effect of sleep quality on repetitive transcranial magnetic stimulation outcomes in depression. Front Psychiatry. 2024;15:1458696. doi:10.3389/fpsyt.2024.1458696
47. Kazemi R, Rostami R, Rezaei M, et al. Comorbid anxiety in depression and rTMS treatment response: a retrospective study. J Affect Disord. 2025;376:36-46. doi:10.1016/j.jad.2025.01.148
48. Kaster TS, Downar J, Vila-Rodriguez F, et al. Differential symptom cluster responses to repetitive transcranial magnetic stimulation treatment in depression. EClinicalMedicine. 2023;55:101765.
doi:10.1016/j.eclinm.2022.101765
49. Gill J, De Felice N, Gill J, et al. Repetitive transcranial magnetic stimulation: course and early prediction of response in depression. J Psychiatr Res. 2023;157:108-111. doi:10.1016/j.jpsychires.2022.11.018
50. Adler E, Blumberger DM, Chen X, et al. Trajectories of response to bilateral rTMS in late-life depression. Brain Stimul. 2025;18(6):2067-2074. doi:10.1016/j.brs.2025.10.020
51. Sakurai H, Uribe S, Cirillo P, et al. Residual symptoms after achieving remission with repetitive transcranial magnetic stimulation in depression. J Affect Disord. 2022;301:154-161. doi:10.1016/j.jad.2021.12.115
52. Reszka E, Jabłońska E, Lesicka M, et al. An altered global DNA methylation status in women with depression. J Psychiatr Res. 2021;137:283-289.
doi:10.1016/j.jpsychires.2021.03.003
53. Suh JS, Fiori LM, Ali M, et al. Hypothalamus volume and DNA methylation of stress axis genes in major depressive disorder: a CAN-BIND study report. Psychoneuroendocrinology. 2021;132:105348. doi:10.1016/j.psyneuen.2021.105348
54. Fuchs Weizman N, Wyse BA, Duval C, et al. Major depressive disorder and anti-depressant therapy markedly alters the human follicular niche DNA methylome. Reproduction. 2025;170(3):e240467. doi:10.1530/REP-24-0467