Brazilian Journal of Psychiatry
bjp
Revista Brasileira de Psiquiatria

CC BY

LETTER TO THE EDITORS

Artificial intelligence for depression: early detection and personalized treatment

Braz J Psychiatry. 2026;48:e20254543
doi:10.47626/1516-4446-2025-4543

CC BY

Clinical decision-making in depression still struggles with blurred diagnostic boundaries, heterogeneous treatment response, and nontrivial relapse risk – areas where artificial intelligence and machine learning have been proposed as pragmatic complements to clinical judgment, including in discussions previously hosted by the *Brazilian Journal of Psychiatry*.[1] Recent work has shown that individualized response prediction is feasible in major depressive disorder through multimodal models[2] and that attention to algorithmic fairness and transparent evaluation is indispensable for safe implementation in mental health.[3] Here, we outline near-term opportunities for early detection and personalized treatment in depression using artificial intelligence/machine learning (Box 1), and we offer practical implementation guidance tailored to Brazil's Unified Health System (Sistema Único de Saúde).

Early screening and risk stratification

Although early recognition improves prognosis, many cases remain undetected in primary care. In a nationwide cohort of perinatal patients, models trained on electronic health records predicted postpartum depression with an area under the receiver operating characteristic curve of approximately 0.712.[4] Among adolescents, ensembles including random forests and support vector machines achieved areas under the curve > 0.90 for depression and suicidal ideation.[5] As a rule of thumb, area under the curve values around 0.70 indicate moderate discrimination, whereas values $\geqslant 0.90$ are excellent. However, clinical usefulness still depends on calibration, decision thresholds, and external validation. In Brazil's Unified Health System, these tools can facilitate triage, resulting in earlier and more consistent clinical assessment.

Personalized treatment selection

The trial-and-error cycle can be shortened when predictions guide initial choices. In treatment-resistant depression, a random-forest model predicted response to intranasal esketamine with roughly 68.5% accuracy over three months.[6] Multimodal approaches combining symptoms, sociodemographics, and magnetic resonance imaging connectivity features predicted sertraline response with an area under the curve of approximately 0.73.[2] While advanced imaging is not routine in the Unified Health System, the underlying logic of fusing diverse, routinely available data can be implemented using structured clinical and laboratory variables from electronic health records.

Subtype differentiation with therapeutic impact

Misclassifying bipolar depression as unipolar depression can precipitate antidepressant-induced switching, while the reverse may cause unnecessary exposure to mood stabilizers. Using functional connectivity, machine learning achieved an accuracy of about 87.5% in distinguishing unipolar from bipolar depression.[7] Deep learning applied to resting-state electroencephalography reached 90-100% accuracy when separating major depressive disorder, bipolar disorder and schizophrenia.[8] In Brazilian referral settings, electroencephalography is a cost-effective means of augmenting complex diagnostics; models should complement, not replace, longitudinal clinical assessment.

Digital phenotyping and longitudinal monitoring

Smartphones and wearables allow continuous capture of behavioral and physiological data (speech patterns, mobility, sleep architecture, heart-rate variability). Smartphones and wearables can passively capture speech patterns, mobility, and sleep architecture in real time, providing signals that precede full relapse and are amenable to machine learning analysis.[9] Electroencephalography-based deep learning has also differentiated treatment-resistant depression from treatment-responsive depression with high accuracy, illustrating biomarker-guided monitoring.[10] The goal is not surveillance, but earlier, patient-centered action.

Implementation priorities for Brazil

Three pillars should guide implementation: interpretability (e.g., Shapley additive explanations; feature importance; transparent decision paths), external validation across Brazil's diverse populations and care levels, and governance with fairness – including bias detection/correction and equity metrics in routine evaluation.[3] Capacity-building through continuing education will ensure that artificial intelligence augments, rather than displaces, clinical expertise.

The first pillar is interpretability. In psychiatry, treatment choices and risk communication require reasoning that clinicians can investigate. Model-agnostic tools (e.g., Shapley additive explanations for local/global feature attribution), summaries of feature importance, and transparent decision paths (e.g., shallow trees, rule lists, or monotonic constraints for high-stakes triggers) make outputs auditable and support shared decision-making at the point of care.

The second is external validation for Brazil. Beyond internal cross-validation, we recommend multi-site testing

Box 1 Near-term opportunities for artificial intelligence in depression - applications, methods, performance, and key references

Application	Data source(s) & method(s)	Representative performance (metric)	Key reference(s)
Screening in primary & perinatal care	EHR-based features; gradient boosting/random forest	Postpartum depression AUC $\cong$0.712	Hochman et al.[4]
Adolescent risk stratification	Surveys + school data; RF, SVMs	Depression/ideation AUC > 0.90	Huang et al.[5]
Personalized pharmacotherapy	Clinical + MRI multimodal model	Sertraline response AUC ~0.73	Poirot et al.[2]
TRD esketamine allocation	Baseline clinical profiles; RF	Response accuracy ~68.5%	Pettorruso et al.[6]
Subtype differentiation	MRI connectivity; SVM	Unipolar vs bipolar accuracy ~87.5%	Zhang et al.[7]
EEG-based differential diagnosis	Resting EEG; deep learning	MDD/bipolar/schizophrenia accuracy 90-100%	Ravan et al.[8]
Digital phenotyping for monitoring	Smartphone/wearable signals; ML	Early relapse detection (proof-of-concept)	Zhang et al.,[9] Metin et al.[10]
Governance & fairness	Bias detection/correction pipelines	Improved parity across datasets	Dang et al.[3]

AUC = area under the curve; EEG = electroencephalography; EHR = electronic health record; MDD: major depressive disorder; ML = machine learning; MRI = magnetic resonance imaging; RF = random forest; SVM = support vector machine; TRD = treatment-resistant depression.

across states and levels of care (primary, community, tertiary); temporal-split validation (train on earlier cohorts, test on later periods); calibration assessment (calibration plots/Brier score); and adherence to Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis and the Prediction model Risk of Bias Assessment Tool. Where feasible, reserve a held-out hospital/municipality as a true external test to mitigate dataset shift.

The third is fairness and governance. Mitigate bias via representative sampling and reweighting, consider in-processing constraints (e.g., equalized odds) or post-processing calibration when disparities appear, and monitor fairness metrics (e.g., gaps in true positive rate/false positive rate) longitudinally. Use model cards/data sheets, ensure General Personal Data Protection Law-compliant de-identification, and favor privacy-preserving data sharing (e.g., federated learning) to enable reproducible, equitable artificial intelligence in the Unified Health System.

Conclusions

Evidence now supports feasible uses for artificial intelligence in early detection and personalized treatment for depression. With rigorous validation, transparency and equity, Brazil can translate these tools into better outcomes in real-world services.

Naian Acá Oliveira **Mathias**, Filipe Paiva **Rodrigues**, Luiz Guilherme Coimbra de **Brito**, Antonio E. **Nardi**
Instituto de Psiquiatria, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

Submitted Sep 09 2025, accepted Dec 14 2025.

Disclosure

The authors report no conflicts of interest.

Data availability statement

Not applicable.

Author contributions

NAOM: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing.
FPR: Investigation, Data curation, Writing – review & editing.
LGCB: Methodology, Validation, Writing – review & editing.
AEN: Supervision, Resources, Validation, Writing – review & editing.
All authors have read and approved of the final version to be published.

Handling Editor: Thiago Henrique Roza

How to cite this article: Mathias NAO, Rodrigues FP, Brito LGC, Nardi AE. Artificial intelligence for depression: early detection and personalized treatment. Braz J Psychiatry. 2026;48:e20254543. Epub 2026 Jan 7. http://doi.org/10.47626/1516-4446-2025-4543

References

1 Laskoski PB, Serralta FB, Passos IC, Hauck S. Machine-learning approaches in psychotherapy: a promising tool for advancing the understanding of the psychotherapeutic process. Braz J Psychiatry. 2019;41:568-9.
2 Poirot MG, Ruhe HG, Mutsaerts HMM, Maximov II, Groote IR, Bjørnerud A, et al. Treatment response prediction in major depressive disorder using multimodal MRI and clinical data: secondary analysis of a randomized clinical trial. Am J Psychiatry. 2024;181:223-33.
3 Dang VN, Cascarano A, Mulder RH, Cecil C, Zuluaga MA, Hernández-González J, et al. Fairness and bias correction in machine learning for depression prediction across four study populations. Sci Rep. 2024;14:7848.
4 Hochman E, Feldman B, Weizman A, Krivoy A, Gur S, Barzilay E, et al. Development and validation of a machine learning-based

postpartum depression prediction model: a nationwide cohort study. Depress Anxiety. 2021;38:400-11.
5 Huang Y, Zhu C, Feng Y, Ji Y, Song J, Wang K, et al. Comparison of three machine learning models to predict suicidal ideation and depression among Chinese adolescents: a cross-sectional study J Affect Disord. 2022;319:221-8.
6 Pettorruso M, Guidotti R, d'Andrea G, De Risio L, D'Andrea A, Chiappini S, et al. Predicting outcome with intranasal esketamine treatment: a machine-learning, three-month study in treatment-resistant depression (ESK-LEARNING). Psychiatry Res. 2023;327:115378.
7 Zhang A, Qiao D, Wang Y, Yang C, Wang Y, Sun N, et al. Distinguishing between bipolar and unipolar depression based on reward circuit activities and clinical characteristics: a machine-learning analysis. J Affect Disord. 2023;327:46-53.
8 Ravan M, Noroozi A, Sanchez MM, Borden L, Alam N, Flor-Henry P, et al. Diagnostic deep-learning algorithms that use resting EEG to distinguish major depressive disorder, bipolar disorder, and schizophrenia from each other and from healthy volunteers. J Affect Disord. 2024;346:285-98.
9 Zhang Y, Folarin AA, Dineley J, Conde P, de Angel V, Sun S, et al. Identifying depression-related topics in smartphone-collected free-response speech recordings using an automatic speech recognition system and a deep learning topic model. J Affect Disord. 2024; 355:40-9.
10 Metin SZ, Uyulan Ç, Farhad S, Ergüzel TT, Türk Ö, Metin B, et al. Deep learning-based artificial intelligence can differentiate treatment-resistant and responsive depression cases with high accuracy. Clin EEG Neurosci. 2025;56:119-30.