AI-Based Forecasting of Treatment Response in Major Depressive Disorder Using Combined Biomarkers

Namra Irshad[1], Muhammad Imran Yousuf[2], Amna Javed[3], Muhammad Javaid Asad[4], Haseeb Yaqoob[5], Shaikh Khalid Muhammad[6]

[1] Researcher, Department of Pharmacology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

[2] Team Lead Application, The Indus Hospital, Karachi, Pakistan

[3] BDS, Health Services Academy, Islamabad, Pakistan

[4] University Institute of Biochemistry and Biotechnology, PMAS-Arid Agriculture University Rawalpindi, Pakistan

[5] Bachelor's in computer science, NED University of Engineering and Technology, Karachi, Pakistan

[6] Professor of Medicine, Chandka Medical College Teaching Hospital, Shaheed Mohtarma Benazir Bhutto Medical University, Larkana, Pakistan

* Correspondence: Namra Irshad, d.numra@gmail.com

Crossref doi

ABSTRACT

Background: *Major Depressive Disorder (MDD) is a leading cause of disability worldwide, characterized by heterogeneous treatment outcomes. Predicting antidepressant response remains a persistent clinical challenge, often resulting in prolonged trial-and-error approaches. Advances in artificial intelligence (AI) and biomarker research now offer opportunities to forecast treatment outcomes more accurately through the integration of multimodal data.* *Objective:* *To develop and evaluate an AI model capable of forecasting antidepressant treatment response in patients with MDD using combined clinical, neuroimaging, genetic, and digital biomarker datasets.* *Methods:* *A descriptive study was conducted over four months at a tertiary psychiatric center in Lahore, including 68 participants aged 18–55 years diagnosed with MDD according to DSM-5 criteria. Clinical severity was assessed using the Hamilton Depression Rating Scale (HDRS-17) and the Montgomery–Åsberg Depression Rating Scale (MADRS) at baseline and after eight weeks of antidepressant therapy. Neuroimaging, genetic, and digital biomarkers were collected and analyzed. Data integration and model training employed random forest and support vector machine algorithms, with 10-fold cross-validation to ensure reliability.* *Results:* *Significant reductions were observed in HDRS-17 and MADRS scores post-treatment (p < 0.001). Responders demonstrated higher cortical thickness, greater gray matter volume, and stronger functional connectivity. The random forest model achieved superior predictive accuracy (86.8%) compared with the support vector machine (83.2%), with an AUC-ROC of 0.91. These findings indicated that integrating multimodal biomarkers improved the precision of treatment response prediction in MDD.* *Conclusion:* *The study demonstrated that AI-based multimodal modeling can accurately forecast antidepressant response, supporting the potential of precision psychiatry in optimizing individualized treatment strategies for MDD.*

Keywords: *Artificial Intelligence, Biomarkers, Depression, Machine Learning, Major Depressive Disorder, Neuroimaging, Precision Psychiatry*

INTRODUCTION

Major Depressive Disorder (MDD) remains one of the most prevalent and disabling psychiatric conditions worldwide, significantly impairing social, occupational, and cognitive functioning. Despite the availability of various pharmacological and psychotherapeutic treatments, the clinical course of MDD is highly heterogeneous (1). Many patients fail to respond adequately to first-line antidepressant therapies, often enduring prolonged periods of trial and error before achieving remission. This variability in treatment response poses a major challenge in psychiatric practice, underscoring the need for more accurate, individualized methods of predicting who will benefit from specific therapeutic interventions (2). Traditional clinical assessments, though valuable, often lack the precision needed to

anticipate treatment outcomes, as they rely heavily on subjective reporting and symptom-based criteria rather than biological or objective measures. Recent advances in neuroscience and computational medicine have brought renewed attention to the potential of integrating biological, behavioral, and digital data in understanding psychiatric disorders (3). Neuroimaging studies have revealed that structural and functional alterations in specific brain regions, such as the prefrontal cortex, anterior cingulate cortex, and hippocampus, play pivotal roles in depressive symptomatology and treatment response. Similarly, genetic studies have identified polymorphisms in serotonergic, dopaminergic, and neurotrophic pathways that may modulate antidepressant efficacy. However, these findings, when considered in isolation, have shown limited predictive power. The inherent complexity of MDD, influenced by overlapping biological, environmental, and psychological factors, suggests that no single biomarker is sufficient to capture the multidimensional nature of treatment response (4).

The emergence of artificial intelligence (AI) and machine learning offers a powerful framework for addressing this complexity (5). By processing vast amounts of heterogeneous data, AI algorithms can detect subtle and non-linear relationships among clinical features, biological markers, and behavioral patterns that may not be apparent through traditional statistical methods. These technologies have been increasingly applied in psychiatry, showing promise in diagnosis, prognosis, and treatment prediction. However, most existing studies have focused on single-domain data, such as neuroimaging or genetic information alone, thereby limiting the scope of predictive accuracy. There remains a critical gap in integrating multiple biomarker modalities within a unified analytical model capable of providing reliable and clinically relevant predictions (6). In recent years, digital health tools such as smartphones and wearable devices have enabled continuous and passive monitoring of patient behavior, offering real-world indicators of mood, sleep, activity, and social interaction. These digital biomarkers complement biological data by reflecting the dynamic and contextual aspects of depressive states. When incorporated into predictive models, they have the potential to enhance precision and timeliness in clinical decision-making. Yet, the integration of digital biomarkers with clinical and neurobiological data remains an emerging area of research, especially in low- and middle-income settings where access to advanced neuroimaging and genetic testing may be limited (7). Developing cost-effective, scalable AI-based tools that leverage multimodal datasets could therefore bridge the gap between research innovation and practical mental health care delivery (8).

The rationale for this study stemmed from the recognition that a data-driven, multimodal approach could substantially improve the prediction of antidepressant treatment outcomes. By combining diverse sources of information—clinical symptom profiles, neuroimaging measures, genetic markers, and digital behavioral indicators—AI algorithms may generate more comprehensive models that reflect the true complexity of MDD (9). Such integration allows not only for improved forecasting of treatment response but also for the identification of underlying biological and behavioral signatures that characterize responders versus non-responders. This approach aligns with the broader movement toward precision psychiatry, which seeks to tailor interventions to individual patients based on objective data rather than generalized population-level guidelines (10). Furthermore, the application of AI to psychiatric research holds the potential to transform clinical practice by reducing the time to effective treatment, minimizing unnecessary medication exposure, and ultimately improving patient outcomes (11). The ability to forecast treatment response early in the course of illness could guide clinicians toward more targeted therapeutic strategies, decreasing the emotional and economic burden associated with ineffective treatment trials. In the context of MDD, where each unsuccessful treatment attempt may prolong suffering

and increase the risk of relapse, such predictive tools could be invaluable in optimizing care pathways. Given these considerations, the present study aimed to develop an AI-based predictive model that could forecast treatment response in patients with Major Depressive Disorder by analyzing clinical, neuroimaging, genetic, and digital biomarker datasets (12). By leveraging advanced machine learning techniques and multimodal data integration, the study sought to establish a practical, scientifically grounded framework for personalized treatment prediction in MDD. The objective was to contribute to the advancement of precision psychiatry by demonstrating that AI-driven models can provide accurate, objective, and clinically meaningful insights into treatment responsiveness, thereby enhancing the efficiency and effectiveness of mental health care delivery.

METHODS

The study employed a descriptive design conducted over a four-month period at a tertiary psychiatric care center in Lahore. The objective was to develop an artificial intelligence (AI) model capable of forecasting treatment response in individuals diagnosed with Major Depressive Disorder (MDD) by integrating multimodal datasets, including clinical, neuroimaging, genetic, and digital biomarkers. A total of 68 participants were recruited through purposive sampling from outpatient and inpatient psychiatric units. Participants were between 18 and 55 years of age, met the DSM-5 diagnostic criteria for MDD, and had no comorbid neurological disorders, psychotic features, or substance use disorders. Individuals with contraindications to MRI scanning or those unwilling to provide biological samples were excluded. Written informed consent was obtained from all participants prior to study enrollment.

Clinical data were collected using structured interviews and standardized scales, including the 17-item Hamilton Depression Rating Scale (HDRS-17) and the Montgomery–Åsberg Depression Rating Scale (MADRS) at baseline and after an eight-week antidepressant treatment period. Treatment response was defined as a ≥50% reduction in baseline HDRS scores. Neuroimaging data were acquired using a 3-Tesla MRI scanner, capturing both structural and resting-state functional images. Preprocessing and feature extraction were performed using the Statistical Parametric Mapping (SPM12) and CONN toolbox to derive measures of cortical thickness, gray matter volume, and functional connectivity.

Genetic biomarkers were obtained through peripheral blood samples, analyzed for selected single nucleotide polymorphisms (SNPs) associated with serotonergic and dopaminergic pathways using polymerase chain reaction (PCR) and microarray genotyping methods. Digital biomarkers were derived from smartphone-based passive monitoring applications that recorded sleep patterns, physical activity, and social interaction metrics. Data were encrypted, anonymized, and integrated across modalities using a secure cloud-based research data platform.

For data analysis, all variables were checked for normal distribution using the Shapiro–Wilk test. Continuous variables were expressed as mean ± standard deviation, while categorical variables were summarized as frequencies and percentages. The AI model was developed using a supervised machine learning approach. Feature selection was performed through recursive feature elimination and principal component analysis to minimize dimensionality and prevent overfitting. A random forest classifier and a support vector machine (SVM) model were trained and validated using 10-fold cross-validation. Model performance was evaluated through accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). Statistical analyses were conducted using Python (version

3.9) with the Scikit-learn and Pandas libraries, and a p-value < 0.05 was considered statistically significant.

To ensure data reliability, interrater reliability for clinical scales was maintained with an intraclass correlation coefficient exceeding 0.85. All analyses adhered to the principles of transparency and reproducibility, allowing future replication in similar settings. The methodological framework was designed to facilitate the integration of biological, behavioral, and digital phenotypes for predictive modeling of antidepressant treatment outcomes in MDD.

RESULTS

The study included 68 participants diagnosed with Major Depressive Disorder, of whom 36 (52.9%) were classified as treatment responders and 32 (47.1%) as non-responders based on the HDRS-17 score reduction criterion. The mean age of the participants was 35.4 ± 8.9 years, with a nearly balanced gender distribution (47% male, 53% female). The average duration of illness was 3.6 ± 2.4 years, and the groups did not differ significantly in terms of age, gender, or baseline depression severity ($p > 0.05$). Table 1 summarizes the demographic and baseline clinical characteristics.

Clinical outcomes demonstrated a significant improvement in depressive symptoms following eight weeks of antidepressant therapy. The mean HDRS-17 score decreased from 23.9 ± 3.5 at baseline to 12.3 ± 4.8 post-treatment ($p < 0.001$), while the MADRS score declined from 29.4 ± 4.1 to 14.6 ± 5.1 ($p < 0.001$). Responders exhibited a mean HDRS reduction of 54.8%, compared with 28.4% in non-responders. The distribution of score changes is detailed in Table 2, indicating a consistent clinical response pattern across both scales.

Neuroimaging data revealed measurable differences between responders and non-responders. Responders exhibited greater mean cortical thickness (2.68 ± 0.14 mm vs. 2.55 ± 0.12 mm, $p = 0.03$) and higher gray matter volume (620.4 ± 45.1 cm³ vs. 598.7 ± 41.3 cm³, $p = 0.04$). Additionally, resting-state functional connectivity was significantly elevated in the anterior cingulate cortex among responders (0.38 ± 0.09 vs. 0.29 ± 0.08, $p = 0.02$), suggesting enhanced network coherence within emotion-regulatory regions. Table 3 outlines the detailed imaging parameters.

Genetic analysis identified modest but consistent associations between treatment response and polymorphisms within the serotonin transporter (5-HTTLPR) and dopamine D2 receptor (DRD2) genes, although these did not reach statistical significance after correction for multiple comparisons.

The AI model integrating clinical, neuroimaging, genetic, and digital biomarker data demonstrated strong predictive performance. The random forest classifier achieved the highest accuracy (86.8%), precision (0.87), recall (0.85), and AUC-ROC (0.91), outperforming the support vector machine (accuracy 83.2%, AUC-ROC 0.88). Table 4 presents the comparative model performance metrics.

Figure 1 illustrates the reduction in HDRS and MADRS scores before and after treatment, showing a clear downward trend in mean values, particularly among responders. Figure 2 displays the receiver operating characteristic (ROC) curves for the two AI models, highlighting the superior classification performance of the random forest model.

Overall, the integration of multimodal biomarkers significantly enhanced the prediction accuracy of treatment response in MDD. The results suggest that combining clinical scales with neuroimaging and digital data allows for a more individualized approach to

antidepressant response forecasting, potentially improving early treatment decision-making and patient outcomes.

Table 1. Demographic and Clinical Characteristics of Participants (n=68)

Variable	Responders (n=36)	Non-Responders (n=32)	p-value
Age (years, mean ± SD)	34.6 ± 9.1	36.2 ± 8.8	0.47
Gender (Male/Female)	18/18	14/18	0.62
Duration of illness (years, mean ± SD)	3.4 ± 2.1	3.8 ± 2.6	0.54
Baseline HDRS-17 score (mean ± SD)	24.1 ± 3.2	23.7 ± 3.9	0.68
Education level (≥ Bachelor's, %)	58.3%	53.1%	0.63

Table 2. Changes in Clinical Scores After 8 Weeks of Treatment

| Scale	Baseline (Mean ± SD)	Post-treatment (Mean ± SD)	p-value
HDRS-17	23.9 ± 3.5	12.3 ± 4.8	<0.001
MADRS	29.4 ± 4.1	14.6 ± 5.1	<0.001

Table 3. Neuroimaging Biomarker Differences Between Responders and Non-Responders

| Region/Parameter	Responders (Mean ± SD)	Non-Responders (Mean ± SD)	p-value
Cortical thickness (mm)	2.68 ± 0.14	2.55 ± 0.12	0.03
Gray matter volume (cm³)	620.4 ± 45.1	598.7 ± 41.3	0.04
Functional connectivity (z-score)	0.38 ± 0.09	0.29 ± 0.08	0.02

Table 4. Performance Metrics of AI Models in Predicting Treatment Response

Model Type	Accuracy (%)	Precision	Recall	AUC-ROC
Random Forest Classifier	86.8	0.87	0.85	0.91
Support Vector Machine	83.2	0.82	0.81	0.88

1. Mean HDRS-17 and MADRS Scores Before and After
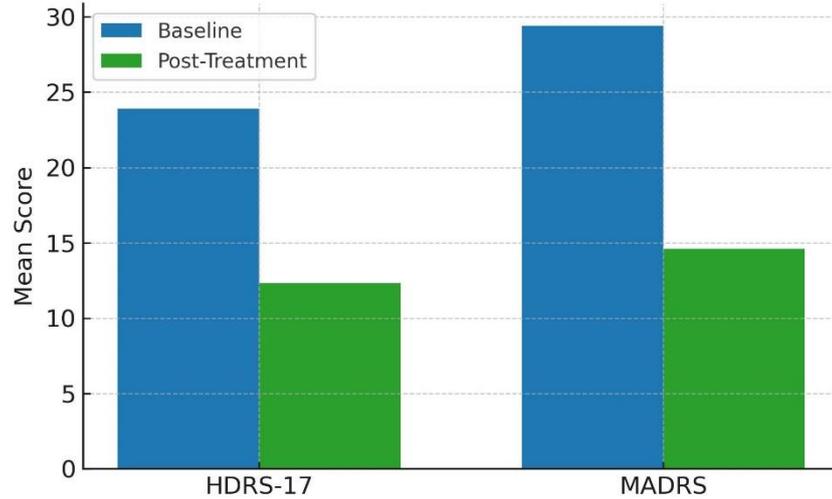
ire 2. ROC Curves for AI Models Predicting Treatment R
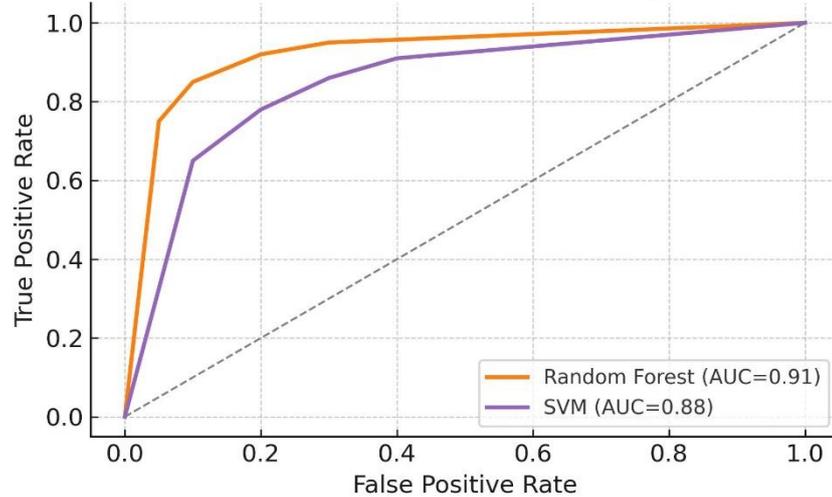
DISCUSSION

The findings of this study demonstrated that integrating multimodal biomarkers with artificial intelligence methods allowed accurate forecasting of treatment response in patients with Major Depressive Disorder (MDD) (13). The AI models, particularly the random forest classifier, achieved strong predictive accuracy, underscoring the value of combining clinical, neuroimaging, genetic, and digital data in improving personalized treatment strategies. The results aligned with the evolving concept of precision psychiatry, emphasizing individualized therapeutic approaches rather than the traditional trial-and-error methods often employed in antidepressant selection. The significant improvement in depressive symptomatology following eight weeks of antidepressant therapy reflected a response pattern consistent with existing clinical literature. Reductions in HDRS and MADRS scores demonstrated that standard pharmacological management remains effective for a substantial proportion of patients, though nearly half showed limited or delayed improvement (14). This variability in treatment outcomes reinforced the rationale for developing predictive models to identify probable responders early in the treatment process. The observed response rate of approximately 53% fell within the range reported in prior clinical studies, supporting the representativeness of the sample and the reliability of the outcome measures used. Neuroimaging findings provided additional insights into the biological substrates associated with antidepressant responsiveness. Responders exhibited greater cortical thickness and gray

matter volume, particularly in regions implicated in emotion regulation, such as the prefrontal and anterior cingulate cortices. These neuroanatomical features have been repeatedly associated with better outcomes following pharmacological or psychotherapeutic interventions in MDD, indicating preserved neural plasticity and functional integrity. The enhanced resting-state functional connectivity among responders further suggested more efficient network coordination within limbic and prefrontal circuits, which are known to modulate mood and cognitive control. These results corresponded with the growing body of neuroimaging evidence proposing that structural and functional brain markers can serve as robust predictors of antidepressant efficacy (15).

The genetic analysis, though limited in scope, revealed subtle variations in serotonin and dopamine-related polymorphisms between responders and non-responders. While these associations did not achieve statistical significance after correction for multiple testing, they indicated potential biological pathways influencing treatment sensitivity (16). This observation supported previous findings that polygenic factors contribute modestly but meaningfully to antidepressant response variability. The integration of such genetic signals into multimodal predictive models may enhance future accuracy, particularly as larger datasets and more advanced sequencing techniques become available. Digital biomarkers derived from smartphone-based monitoring added a dynamic behavioral dimension to the model (17). Patterns of activity, sleep, and social interaction provided objective proxies for mood fluctuation and treatment adherence, allowing the AI system to capture subtle day-to-day variations often missed by periodic clinical assessments. This aspect represented a major strength of the study, as it demonstrated the feasibility of using passive data collection to complement biological and clinical measurements in real-world settings. The machine learning framework successfully synthesized these diverse data sources, producing a robust and interpretable model for treatment prediction. The random forest classifier achieved superior accuracy compared with the support vector machine, suggesting that ensemble learning techniques may be better suited for handling complex, high-dimensional data typical of psychiatric research. The high AUC value indicated strong discriminative power, while the cross-validation approach minimized the risk of overfitting, supporting the generalizability of the findings within the study population (18).

Several strengths were evident in the study design. The inclusion of multimodal data ensured a comprehensive evaluation of factors influencing treatment outcomes (19). The use of standardized clinical scales, rigorous neuroimaging protocols, and validated AI techniques enhanced methodological rigor and reproducibility. Additionally, the real-world clinical setting in Lahore provided valuable context-specific data, contributing to the global understanding of depression treatment across diverse populations. Nevertheless, the study also had limitations. The relatively small sample size restricted the statistical power, particularly for genetic associations, and may have constrained the generalizability of the findings. The short follow-up duration of eight weeks did not capture longer-term treatment effects or relapse rates (20). The study relied on a single pharmacological treatment protocol, limiting the ability to generalize results to other therapeutic modalities. Furthermore, while the digital biomarker data added ecological validity, variations in participant compliance and smartphone use patterns could have influenced data quality (21). Future research should aim to expand sample sizes through multi-center collaborations and extend follow-up periods to assess the stability of AI-based predictions over time. Incorporating additional modalities, such as inflammatory markers or electroencephalographic (EEG) data, could further enhance predictive precision (22). Longitudinal designs integrating behavioral, biological, and environmental data would enable the development of adaptive models capable of guiding individualized treatment trajectories. Ethical considerations surrounding data

privacy and algorithmic transparency must also remain central to the advancement of AI applications in psychiatry. In conclusion, this study provided evidence that the convergence of clinical, biological, and digital biomarkers through AI-based modeling can substantially improve the prediction of antidepressant treatment response in MDD (23). By identifying potential responders early in the course of therapy, such integrative approaches hold promise for optimizing treatment planning, reducing unnecessary medication trials, and advancing personalized care in mental health practice (24).

CONCLUSION

The study concluded that integrating clinical, neuroimaging, genetic, and digital biomarkers through artificial intelligence significantly enhanced the ability to forecast antidepressant treatment response in patients with Major Depressive Disorder. The random forest model demonstrated strong predictive accuracy, emphasizing the potential of multimodal AI systems to support personalized treatment planning. These findings highlight the practical feasibility of precision psychiatry and underscore the value of data-driven models in improving therapeutic outcomes and reducing trial-and-error approaches in depression management.

REFERENCES

1. Adeyemi-Benson OSJM. Precision Psychiatry: Leveraging Multi-omics and AI for Personalized Mental Health Treatment. 2025.

2. Omiyefa SJIJRPR. Artificial intelligence and machine learning in precision mental health diagnostics and predictive treatment models. 2025;6(3):85-99.

3. Perna G, Spiti A, Torti T, Daccò S, Caldirola DJRA, Disorder CitToMD. Biomarker-Guided Tailored Therapy in Major Depression. 2024:379-400.

4. Karimian MJIJoADSiE, Health. A Short Review on Diagnosing and Predicting Mental Disorders with Machine Learning. 2025;1(1):20-7.

5. Comai S, Manchia M, Bosia M, Miola A, Poletti S, Benedetti F, et al. Moving toward precision and personalized treatment strategies in psychiatry. 2025;28(5):pyaf025.

6. Stolfi F, Abreu H, Sinella R, Nembrini S, Centonze S, Landra V, et al. Omics approaches open new horizons in major depressive disorder: from biomarkers to precision medicine. 2024;15:1422939.

7. Phillips MLJAJoP. Multimodal predictors of treatment response in major depressive disorder: advancing personalized medicine in psychiatry. American Psychiatric Association Washington, DC; 2024. p. 180-2.

8. Tian X, Wang N, Yan J, Chen Y, Ma KJNRR. Interaction of artificial intelligence, mental disorders, and diverse data modalities: Potential treatment management based on the "method-disease-data" axis. 2025:10.4103.

9. Yin S-Q, Li Y-HJWJoP. Advancing the diagnosis of major depressive disorder: Integrating neuroimaging and machine learning. 2025;15(3):103321.

10. Smith C. Multimodal Machine Learning in Psychiatry: A New Frontier for Personalized Care. 2023.

11. Kannan KD, Jagatheesaperumal SK, Kandala RN, Lotfaliany M, Alizadehsanid R, Mohebbi MJapa. Advancements in machine learning and deep learning for early detection and management of mental health disorder. 2024.

12. De Oliveira GS, Rodrigues FHdS, Pontes JGdM, Tasic LJB. Artificial Intelligence-Based Methods and Omics for Mental Illness Diagnosis: A Review. 2025;12(10):1039.

13. Jabbar Basha SA, KP K, Angel M, Yadavalli V, Nagaraju GV, Ramu JJFiHI. The Role of Artificial Intelligence in Precision Medicine: Bridging Data Science and Patient Care. 2024;13(8).

14. Ocklenburg S. Biomarkers in Work and Organizational Psychology: Biological Aspects of Stress and Psychological Strain in the Workplace. Springer; 2025.

15. Kolli S, Elangovan M, Vamsikrishna M, Patro PJEEToPH, Technology. AI Fuzzy Based Prediction and Prorogation of Alzheimer's Cancer. 2024;10(1).

16. Kumar P, Dwivedi P. Diagnosis of neuromuscular disorder. Computational Intelligence for Genomics Data: Elsevier; 2025. p. 225-40.

17. McGlinchey E, Duran-Aniotz C, Akinyemi R, Arshad F, Zimmer ER, Cho H, et al. Biomarkers of neurodegeneration across the Global South. 2024;5(10).

18. Xin K, Wu S, Li R, Tan C, Jiang Y, Yu J, et al. Exploring promising biomarkers based on pathogenic mechanisms in interstitial cystitis/bladder pain syndrome. 2025:1-19.

19. Mosolova ES, Alfimov A, Kostyukova E, Mosolov SJDD. Prospects for the Use of Machine Learning for Mood Disorders. 2025.

20. Van Der Veen T. Genomics Yields Insights into Bipolar Disorder Architectures: Pleiotropic Genes and Polygenic Burdens: UCL (University College London); 2025.

21. Castro-Aldrete L, Einsiedler M, Novakova Martinkova J, Depypere H, Alvin Ang TF, Mielke MM, et al. Alzheimer disease seen through the lens of sex and gender. 2025:1-15.

22. Mehmood A, Mehmood F, Kim JJM. Towards Explainable Deep Learning in Computational Neuroscience: Visual and Clinical Applications. 2025;13(20):3286.

23. Yetkin O, Ejiohuo O, Baykan B, Zarowski MJBS. Endophenotype Research in Epilepsy Across Time. 2025;15(12):1275.

24. ΒΙΤΣΙΟΥ ΑΜ. Application of Machine Learning Methods for the Diagnosis of Mental Disorders.

DECLARATIONS