scientific reports

Check for updates

OPEN

Machine learning helps predict early onset psychosis with serum protein biomarkers, neuropsychometry, and clinicodemographic data

Przemyslaw T. Zakowicz[1,9], Maksymilian A. Brzezicki[1,2,9]✉, Joanna Pawlak[4], Maria Skibinska[4], Szymon Jurga[3], Aleksandra Lewandowska[5], Benedikt Vogel[6], Emily Ungermann[7] & Barbara Remberk[8]

Early-onset psychosis presents diagnostic challenges due to overlapping clinical presentations and complex comorbidities, typically requiring specialized tertiary care with extensive neuroimaging, neuropsychometric testing, and multidisciplinary evaluation. This case-control study investigated whether machine learning could integrate multiple diagnostic modalities to create an objective diagnostic framework for early-onset psychosis. We recruited 45 patients with early-onset psychosis and 34 healthy controls from a tertiary referral centre. Participants underwent comprehensive assessment including serum protein biomarker analysis (brain-derived neurotrophic factor, proBDNF, p75 neurotrophin receptor, S100B), neuropsychometric testing (Iowa Gambling Task, Simple Response Time, Zabor Verbal Task), and demographic evaluation. Four machine learning algorithms (logistic regression, support vector machine, random forest, XGBoost) were trained on five feature combinations using nested cross-validation with hyperparameter optimization. XGBoost demonstrated superior performance, achieving optimal classification with the complete multimodal dataset (accuracy: 0.91 ± 0.08, precision: 0.92 ± 0.08, area under curve: 0.97 ± 0.04). Feature importance analysis revealed cognitive measures, particularly Zabor Verbal Task errors and response time parameters, as most discriminative, with brain-derived neurotrophic factor pathway components showing highest biomarker importance. Machine learning effectively integrated neuropsychometric and protein biomarker data for high-accuracy early-onset psychosis classification, with multimodal approaches outperforming single-domain assessments.

Keywords Biomarkers, Adolescents, Neuropsychometry, XGBoost, Classification, Neurotrophins

Primary psychotic disorder is a significant and growing contributor to the global burden of neurological diseases. It leads to a decline in disability-adjusted life years as well as a wide variety of occupational, familial, and social problems[1,2]. In early onset psychosis (EOS) patients often present with less demarcated clinical syndromes[3] and may have various developmental and socioeconomic comorbidities, thus lowering diagnostic certainty. This leads to delays in care and up to four times longer duration of time of untreated psychosis[4], which results in worse prognosis[5–7], especially if the independent living, gaining employment and satisfactory school performance are to be considered as priorities of care.

[1]Department of Neural Engineering and Space Medicine, Institute of Medical Sciences, University of Zielona Góra, Zielona Góra, Poland. [2]Jesus College, University of Oxford, Oxford, UK. [3]Department of Neurology, University Hospital, Zielona Góra, Poland. [4]Department of Genetics in Psychiatry, Poznan University of Medical Sciences, Poznan, Poland. [5]Department of Child and Adolescent Psychiatry, Babinski Hospital, Lodz, Poland. [6]Querschnittgelähmten-Zentrum, BG Klinikum, Hamburg, Germany. [7]Institute of Forensic and Traffic Medicine, University Hospital Heidelberg, Heidelberg, Germany. [8]Institute of Neurology and Psychiatry, Warsaw, Poland. [9]Przemyslaw T. Zakowicz and Maksymilian A. Brzezicki contributed equally to this work. ✉email: mbrzezicki@neurologicalsociety.org

To that end several diagnostic panels were proposed. The neurocognitive element could be measured through standardised protocols using tests such as Simple Response Time, Iowa Gambling Task[8], Zabor Verbal Task[9], Wechsler Intelligence Measurement[10] and others (for a comprehensive review of neuropsychometric testing, see Zakowicz et al., 2025[11]). Neuropsychometric testing may not be appropriate for all populations, however, as they require patient engagement and appropriate skill in interpretation. Several studies suggested identifying neuroimaging abnormalities in EOS patients, e.g.: measuring cortical grey matter thickness, and assessing hippocampal, parahippocampal and cingulate gyrus morphology[12,13]. While these studies are encouraging, a comprehensive MRI protocol may require general anaesthesia and neuroradiology expertise, both necessitating adequate resources. Genetic analysis of EOS patients identified several copy number variations (CNVs) that may be linked to EOS[14]. Similarly, investigation of polygenic risk scores (PRS) provides an important insight into cognitive and neurodevelopmental background of psychosis[15]. Even the classical genetic methods enabled identification of chromosomal microdeletions associated with an increased risk for psychosis, like 22q11.2[16]. However, whilst extremally useful in selected cases, genetic testing takes months to report and may not capture cases of EOS with sporadic aetiology.

Investigation into the neurobiological underpinnings of EOS has increasingly focused on the role of neurotrophins and related proteins due to their critical involvement in neuronal development, differentiation, and synaptic plasticity. Brain-derived neurotrophic factor (BDNF), proBDNF, S100B, p75NTR (p75 Neurotrophin Receptor) are often mentioned as biomarker candidates for early rather than adult-onset psychosis due to their involvement in neurodevelopmental disruptions underlying this condition[17–20]. These proteins are crucial in adolescent brain maturation processes, neuroplasticity, glial-neuronal interactions, and neuroinflammation that are severely disrupted in EOS[3,17,19,21]. Furthermore, their accessibility through peripheral blood sampling, dynamic responsiveness to developmental changes, correlation with clinical severity and cognitive dysfunction, and predictive value for treatment outcomes make them interesting targets for further study as follows.

BDNF is a well-established target[18,22]. It is crucial for neuronal development[23] and response to stress[24]. It also modulates various monoaminergic[25], GABAergic[26], and cholinergic[27] systems, all implicated in pathology of psychotic disorders[18]. Clinical studies consistently report reduced BDNF levels in patients with schizophrenia and first-episode psychosis when compared to healthy controls[18,22,24,28]. The BDNF precursor, proBDNF, and its receptor p75NTR are also critical in physiological neurosignalling[29–31]. Increased proBDNF/p75NTR signalling during development can disrupt the normal GABAergic excitatory/inhibitory balance, leading to depolarizing and excitatory actions of GABA in adulthood, which may contribute to neurodevelopmental conditions[30–34]. In the context of psychotic disorders specifically, however, the exact role of p75NTR remains to be elucidated.

Another candidate is the S100B calcium-binding protein. This protein is produced by astrocytes[35], serves as an activator of glial metabolism[36], and promotes the stability of brain-blood barrier[37]. Thus, its plasma levels have been proposed as a potential clinical marker of delirium[38], traumatic brain injuries[37], and neurodegeneration[39]. Elevated S100B levels have also been observed in psychotic disorders, correlating with neuroinflammatory processes and cognitive dysfunction[40–42]. Whilst these protein biomarkers demonstrate significant promise in elucidating the pathophysiology of EOS, the absence of standardised reference ranges and definitive normative data currently limits their utility as reliable clinical diagnostic tools.

We hypothesize that the multifactorial aetiology of EOS precludes diagnosis through any single pathognomonic test. A comprehensive diagnostic panel, therefore, necessitates the integration of multiple diagnostic modalities, often in various data formats and dimensionalities. Machine learning has emerged as a powerful tool for neuropsychiatric classification[43–46], with applications ranging from EEG-based identification of emotional dysfunctions and rumination patterns[47–49] to neuroimaging-based diagnosis of psychotic disorders. Four algorithms have demonstrated particular utility across diverse neuropsychiatric applications: logistic regression, which has successfully predicted symptomatic improvement in early psychosis from fMRI data[50] and distinguished schizophrenia patients from healthy controls using structural MRI[51]; support vector machines (SVM), which achieved 88% accuracy discriminating schizophrenia from major depressive disorder using brain network connectivity[52] and 80% accuracy in distinguishing schizophrenia from autistic spectrum disorders based on neuroimaging data[53]; random forests, which demonstrated 88% accuracy classifying social media posts of schizophrenia patients versus controls[54], and 98% accuracy in multiclass classification of depression and schizophrenia from motor activity data[55]; and gradient boosting methods (XGBoost), which predicted mental states in schizophrenia from smartphone sensor data[56] and achieved superior performance in genetic prediction of schizophrenia risk[57]. Several of these approaches successfully used multimodal data inputs, e.g. integrating fMRI data with neurocognitive assessments, combining neuroimaging with genetic data or smartphone outputs with clinical ratings. Despite these advances in adult populations, machine learning applications to early-onset psychosis remain limited, and integration of cognitive assessments with molecular biomarkers has been underexplored. The present study systematically evaluates these four established algorithms across multimodal feature sets combining neuropsychometric assessments, demographics, and BDNF pathway proteins to optimize early detection of adolescent-onset psychotic disorders.

Methods

Subjects and protocol

The data utilised in this investigation were derived from the Comparison of Biomarkers in Schizophrenia and Bipolar Affective Disorder Study. All experiments were performed in accordance with relevant guidelines and regulations. This case-control study was conducted with approval from the Ethics Committee at Poznan University of Medical Sciences (reference number 1066/19), in accordance with European regulatory standards. Each participant and/or their legal representatives received comprehensive written documentation regarding the study parameters prior to enrolment. Informed consent was obtained from all subjects and/or their legal guardian(s). The recruitment methodology employed a systematic approach whereby all patients presenting

with psychotic symptoms upon admission to the Zabor Centre for Child and Adolescent Psychiatry (a tertiary referral centre for psychiatry for the West of Poland) were invited to participate. For full research protocol including blood and analysis see[8,58]. The diagnosis of EOS was set according to DSM-4 criteria. In short, for EOS patients, the baseline (t0) blood samples were taken within one week of admission, and the post-treatment samples (t1) were obtained 6–8 weeks after admission, when the stabilisation of symptoms was achieved. Zabor Verbal Task (ZVT) was performed at t0 and t1 since it is a task that has previously been proven to be well-tolerated even by actively psychotic individuals[9]. The remainder of neuropsychometric testing was performed at t1. For healthy controls (HC) patients, both blood samples and cognitive tests were obtained at a single (t0) visit. A summary of clinicodemographic data is available in Table 1, and the recruitment protocol is illustrated in Fig. 1. The final diagnosis of EOS was confirmed by two independent physicians specialising in child and adolescent psychiatry.

Blood analysis

Ten millilitres of venous blood were collected into BD Vacutainer PPT Plasma Preparation Tubes between 07:30 and 09:30 following overnight fasting. Plasma was immediately separated by centrifugation, aliquoted into storage tubes, and stored at − 80 °C until analysis.

Enzyme-linked immunosorbent assay analyses were performed using DuoSet ELISA Development Kits from R&D System (Minneapolis, MN, USA), specifically BDNF (cat. No.: DY248), proBDNF (cat. No.: DY3175), p75NTR (cat. No.: DY367), and S100B (cat. No.: DY1820-05). All assays were conducted according to manufacturer instructions with minor modifications as previously described[8,59].

The standard curve ranges were 1500–23.4 pg/mL for BDNF, 5000–78.1 pg/mL for proBDNF, 3000–46.9 pg/mL for NGFR, and 2000–31.2 pg/mL for S100B. Dilutions were used where necessary. Quality control measures showed intra-assay variability of less than 5% coefficient of variation and inter-assay variability of less than 10% coefficient of variation for each studied protein.

	EOS		HC		
Parameter	**Mean**	**SD**	**Mean**	**SD**	***p***
N	45		34		
Age	14.7	1.7	15.2	1.8	0.26
Gender (female/male)	25/20		19/15		0.98
Years of education	8.7	1.7	9.2	1.8	0.26
Age of onset	14.5	1.6			
Family history	21		0		> 0.99
Disease duration	0.2	0.8			
PANSS T0	109.3	20.5			
PANSS T1	57.4	18.9			
Suicide attempts	1.6	0.9			
Chlorpromazine equivalent (mg)	336.3	172.2			
Mean SRT	463.31	123.71	364.26	49.57	< 0.01
Median SRT	411.02	86.57	338.40	38.79	< 0.01
Minimum SRT	197.29	42.57	224.55	45.57	0.04
Maximum SRT	1756.75	786.60	1096.85	486.55	< 0.01
SD SRT	202.34	122.35	100.73	44.96	< 0.01
IGT Advantageous cards	45.79	7.60	46.15	10.36	0.89
IGT Disadvantageous cards	54.21	7.60	53.85	10.36	0.89
BDNF_T0	1664.46	1406.13	3000.92	1586.74	< 0.01
BDNF_T1	1784.70	1505.20	3000.92	1586.74	< 0.01
PROBDNF_T0	3081.78	1928.87	3494.89	2043.23	0.39
PROBDNF_T1	2792.30	1615.09	3494.89	2043.23	0.12
P75NTR_T0	703.36	1743.08	995.91	1677.31	0.48
P75NTR_T1	621.61	1664.63	995.91	1677.31	0.38
S100B_T0	162.48	301.47	318.66	490.89	0.14
S100B_T1	250.72	606.61	318.66	490.89	0.61
ZVT_wrong_sum_1	5.61	6.07	1.57	2.64	< 0.01
ZVT_wrong_sum_2	5.41	5.88			

Table 1. Demographics of the studied population. *p* values for student's t test and Chi square tests (where applicable) biomarkers in pg/mL. EOS = Early Onset Schizophrenia, HC = Healthy Control, SD = Standard deviation, PANSS = positive and negative symptoms scale, SRT = Simple Reaction Time, IGT = Iowa Gambling Task, BDNF = Brain-derived neurotrophic factor, p75NTR = p75 neurotrophin receptor, ZVT = Zabor Verbal Task.

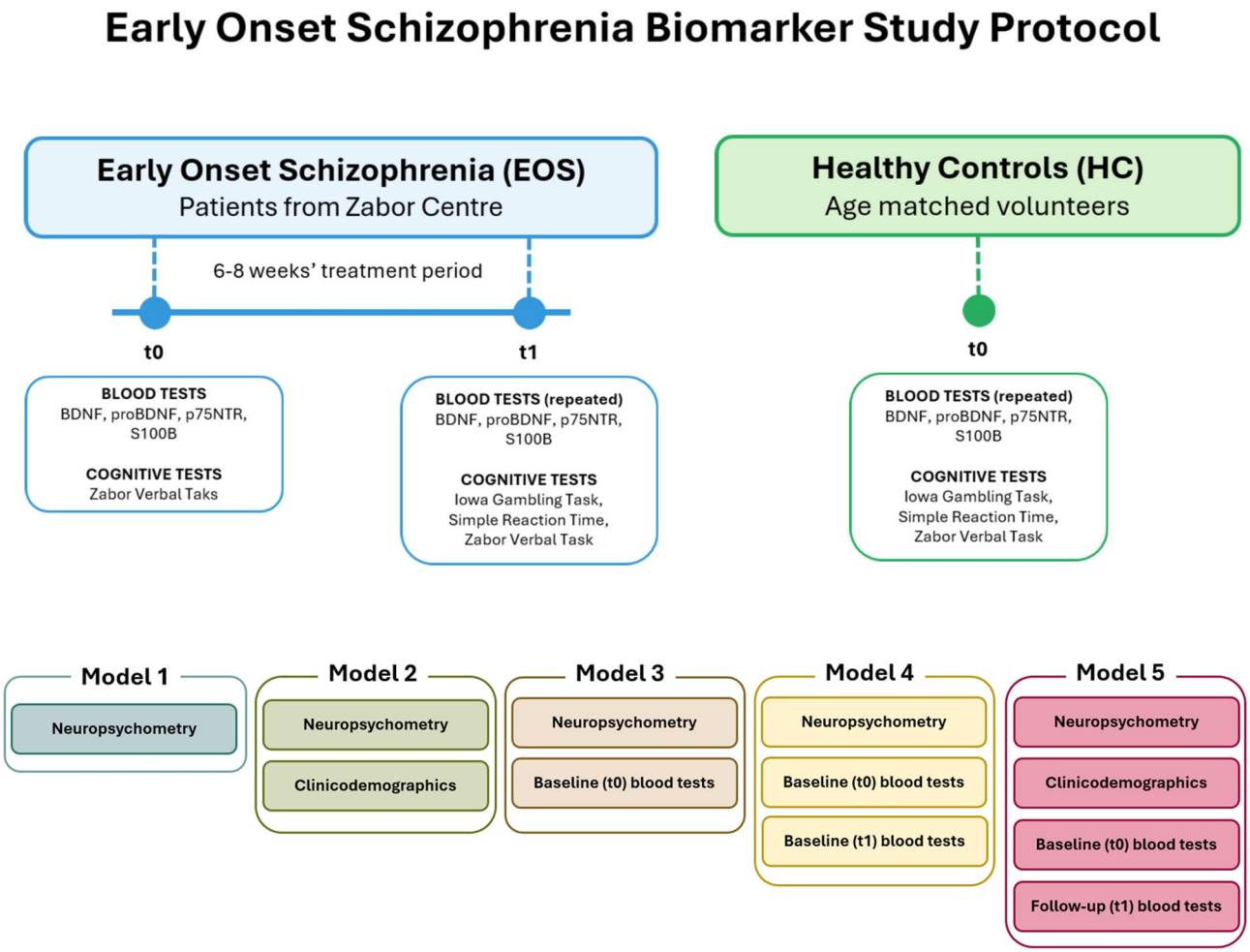

Fig. 1. Summary of study protocol and feature selection for models trained in different experimental conditions.

Neuropsychometry analysis

Two cognitive tests were administered in this study to assess executive function: the Iowa Gambling Task and the Simple Response Time task. Both tests were selected for their widespread use in executive function assessment and were delivered using Psychological Experiment Building Language software.

The Iowa Gambling Task was administered using standard protocol[60]. Participants completed 100 card selections, and performance was quantified by calculating the total number of advantageous cards chosen and the total number of disadvantageous cards chosen from the full set of 100 trials.

The Simple Response Time task required participants to press the left mouse button as quickly as possible when a visual stimulus (the letter X) appeared on screen. The task consisted of 200 trials divided into four blocks of 50 trials each, with rest breaks between blocks. Performance measures included the mean reaction time, standard deviation of reaction times, minimum reaction time, and maximum reaction time across all trials.

ZVT[9] is a clinical assessment tool examining language embodiment by evaluating how individuals organize words based on their spatial meaning. The task employs a 60 cm × 30 cm board positioned centrally in the participant's visual field, containing seven rectangular boxes arranged in three spatial tiers: two boxes at the top, three in the middle, and two at the bottom. Participants place word cards into boxes according to their intuitive spatial associations. Scoring is based on expected real-world spatial relationships, where words like "cloud" should be placed in upper boxes, "chair" in middle boxes, and "ground" in lower boxes. Performance is quantified by counting incorrectly placed words, with higher scores reflecting greater difficulty in spatial-conceptual word organization. This test was chosen for its capability of assessing linguistic and visuospatial domains of disorganised thought in EOS.

Machine learning approach

Four machine learning algorithms were employed for classification: logistic regression, support vector machine, random forest, and XGBoost. These models were selected for their effectiveness with small sample sizes and were trained on four different feature combinations to evaluate the relative contribution and optimal integration of different data modalities. The training conditions included blood biomarkers alone at baseline

(t0), neuropsychometric measures alone, combined neuropsychometric measures and baseline biomarkers (t0), and combined neuropsychometric measures with biomarkers from both time points (t0, t1) using imputation methods for missing values. This systematic approach allowed for comparison of unimodal versus multimodal classification performance and assessment of whether longitudinal biomarker data improved predictive accuracy. All features were standardized using StandardScaler to ensure comparable ranges across variables.

Model development followed a standardized scikit-learn[61] pipeline incorporating feature standardization and model fitting to optimize performance and prevent overfitting. Given the relatively balanced class distribution (approx. 57% to 43% split), no class rebalancing techniques were applied. This ratio (1.32:1) is considered mild imbalance; most ML practitioners only consider rebalancing when the ratio exceeds 3:1 or 4:1. No feature selection was performed beyond the predefined feature combinations, as the small sample size relative to the number of features made it preferable to retain all available information and rely on the inherent regularization properties of the selected algorithms to handle potential overfitting. To assess feature stability and mitigate potential risks of overfitting, we conducted bootstrap resampling analysis with 100 iterations on the optimal configurations across all four algorithms (LR, SVM, RF, XGBoost) using the Model 5 condition (See Supplementary Material).

Hyperparameter optimization was performed using nested cross-validation with exhaustive grid search to identify optimal parameter settings for each algorithm. The outer loop employed 5-fold stratified cross-validation to assess model performance, while the inner loop performed grid search across predefined hyperparameter spaces to select optimal configurations based on area under the receiver operating characteristic curve (ROC-AUC). This systematic optimization was conducted independently for each of the five experimental conditions (Fig. 1) to ensure optimal performance across different feature combinations. The optimal values identified for each condition are presented in Table 2. All optimizations were performed using scikit-learn 1.7.0, with feature scaling applied within each cross-validation fold to prevent data leakage.

Model performance was assessed using multiple metrics including accuracy, precision, recall, F1-score, and area under the curve. Confusion matrices were generated to evaluate classification performance. All classification metrics were calculated using the default probability threshold of 0.5, which was not adjusted during the training process. The feature importance for the winning model was calculated using standard feature weights and Shapley value explanation (SHAP)[62].

To rigorously assess whether performance differences between algorithms and feature sets were statistically significant rather than due to sampling variation, we employed two complementary statistical tests: McNemar's Test for Prediction Pattern Comparison and DeLong's Test for ROC AUC Comparison. McNemar's test evaluates whether two classifiers make systematically different errors by analysing the contingency table of correct/incorrect predictions on the same test samples. The test follows a chi-square distribution with 1 degree of freedom under the null hypothesis that both classifiers have equal error rates. We applied continuity correction for small sample sizes (Yates' correction). McNemar's test was performed on predictions obtained from the outer cross-validation loop to ensure independence from hyperparameter optimization. DeLong's test compares the ROC AUC between two models by estimating the covariance of AUC values using the theory of generalized U-statistics. The test works under the null hypothesis of equal AUCs. DeLong's method accounts for the correlation between AUC estimates from the same test samples, providing more accurate p-values than independent t-tests. This test was applied to probability predictions from the outer cross-validation loop. For algorithm comparisons (6 pairwise tests: LR vs. SVM, LR vs. RF, LR vs. XGB, SVM vs. RF, SVM vs. XGB, RF vs. XGB), we report both uncorrected p-values and Bonferroni-corrected significance thresholds ($\alpha = 0.05/6 = 0.0083$). For feature set comparisons using XGBoost (10 pairwise tests among 5 models), Bonferroni correction yields $\alpha = 0.005$. Statistical tests were implemented using Python's statsmodels package (McNemar's test) and scipy.stats package (DeLong's test via Mann-Whitney U-statistic approximation for AUC comparison). All tests were two-tailed with $\alpha = 0.05$ as the nominal significance level before multiple comparison correction.

Results

Sociodemographic and clinical features

The study included 45 EOS patients and 34 healthy controls with similar age distributions (14.7 ± 1.7 vs. 15.2 ± 1.8 years). EOS patients demonstrated a mean age of onset at 14.5 ± 1.6 years, with 46.7% reporting positive family history of psychiatric illness. Clinical severity was reflected in baseline PANSS scores of 109.3 ± 20.5, which improved to 57.4 ± 18.9 following treatment stabilisation. Patients received antipsychotic medication with mean chlorpromazine equivalent doses of 336.3 ± 172.2 mg/day and had comparable level of intellectual capability as evidenced by only one participant meeting criteria for at least mild disability according to the psychologists' gestalt assessment.

Machine learning model performance

Four machine learning algorithms were evaluated across five different feature combinations to assess their effectiveness in classifying participants into EOS and HC (Table 3). The experimental conditions included: neuropsychometric measures alone (Model 1), neuropsychometric measures with demographics (Model 2), neuropsychometric measures with baseline protein biomarkers at t0 (Model 3), neuropsychometric measures with protein biomarkers at t1 (Model 4), and the complete dataset including neuropsychometry, demographics, and protein biomarkers from both time points (Model 5) (Fig. 2).

XGBoost demonstrated the highest performance across all experimental conditions, achieving the largest accuracy, precision, and area under the curve (AUC) values. The algorithm's performance was particularly robust with the complete dataset (Model 5), reaching an accuracy of 0.91 ± 0.08, precision of 0.92 ± 0.08, recall of 0.93 ± 0.05, and AUC of 0.97 ± 0.04. Notably, XGBoost maintained consistently high performance across all feature combinations, with accuracy ranging from 0.86 to 0.91 and AUC values consistently above 0.94. Random Forest showed a comparably good overall performance, with optimal results achieved using neuropsychometric

Algorithm	Condition	Key optimal hyperparameters
LR	Model 1	C = 0.01, penalty = l2, solver = lbfgs
	Model 2	C = 0.1, penalty = l2, solver = lbfgs
	Model 3	C = 0.01, penalty = l2, solver = lbfgs
	Model 4	C = 0.01, penalty = l2, solver = lbfgs
	Model 5	C = 0.01, penalty = l2, solver = lbfgs
SVM	Model 1	C = 1, kernel = linear, gamma = scale
	Model 2	C = 1, kernel = rbf, gamma = scale
	Model 3	C = 0.1, kernel = linear, gamma = scale
	Model 4	C = 0.1, kernel = linear, gamma = scale
	Model 5	C = 0.1, kernel = linear, gamma = scale
RF	Model 1	n_est = 50, max_depth = 5, min_split = 2
	Model 2	n_est = 50, max_depth = 5, min_split = 2
	Model 3	n_est = 50, max_depth = 5, min_split = 2
	Model 4	n_est = 100, max_depth = 5, min_split = 10
	Model 5	n_est = 50, max_depth = 5, min_split = 2
XGBoost	Model 1	n_est = 100, max_depth = 3, learning_rate = 0.05, subsample = 0.6
	Model 2	n_est = 100, max_depth = 3, learning_rate = 0.05, subsample = 0.5
	Model 3	n_est = 50, max_depth = 2, learning_rate = 0.05, subsample = 0.6
	Model 4	n_est = 100, max_depth = 3, learning_rate = 0.01, subsample = 0.5
	Model 5	n_est = 100, max_depth = 3, learning_rate = 0.01, subsample = 0.5

Table 2. Optimal hyperparameters identified for each algorithm across experimental conditions. Values shown achieved highest ROC-AUC during 5-fold stratified cross-validation within the following search space. For Logistic Regression (LR), we optimised regularisation strength (C: np.logspace(-4, 4, 20)), penalty type (['l1', 'l2']), and solver (['liblinear', 'saga']). Support Vector Machine (SVM) hyperparameters included regularization parameter (C: [0.1, 1, 10]), kernel type (['rbf', 'poly', 'linear', 'sigmoid']), and kernel coefficient (gamma: ['scale', 'auto', 1e-2, 1e-1, 1, 10]). Random Forest (RF) optimization explored number of trees ([50, 100, 200]), maximum depth ([None, 10, 20]), minimum samples for split ([2, 5, 10]). For eXtreme Gradient Boosting (XGBoost), we optimized number of boosting rounds ([50, 100]), maximum tree depth ([2, 3]), learning rate ([0.01, 0.05]), subsample ratio ([0.5, 0.6]). Model 1 = neuropsychometry only, Model 2 = neuropsychometry and clinicodemographic data, Model 3 = neuropsychometry and protein biomarkers at baseline (t0), Model 4 = Model 3 with addition of biomarkers after stabilisation of treatment (t1), Model 5 = Model 4 with addition of clinicodemographic data (complete multimodal model). Grid search with 5-fold stratified cross-validation was performed across all hyperparameter combinations shown.

measures alone (Model 1), yielding an accuracy of 0.81 ± 0.10 and AUC of 0.86 ± 0.15. However, Random Forest performance appeared to decline with the addition of biomarker data, suggesting potential overfitting or feature interaction issues in this algorithm.

Logistic Regression demonstrated moderate and stable performance across conditions, with the best results observed in the Model 4 condition (accuracy: 0.76 ± 0.04, AUC: 0.82 ± 0.10). The algorithm showed consistent performance with relatively low standard deviations, indicating stable predictions across cross-validation folds. SVM showed the most variable performance, with results ranging from 0.60 to 0.71 accuracy across different feature combinations. The algorithm performed best with the Model 3 condition (accuracy: 0.71 ± 0.06, AUC: 0.82 ± 0.07) but showed concerning overfitting patterns in some conditions, particularly Model 1, where perfect recall (1.00) was achieved at the cost of reduced precision.

Standard deviation values across cross-validation folds indicated varying degrees of model stability. XGBoost demonstrated the most consistent performance with standard deviations typically below 0.10 for accuracy measures. Logistic Regression also showed good stability, particularly in the Model 3 and Model 4 conditions. Random Forest and SVM exhibited higher variability, suggesting potential sensitivity to training data composition or hyperparameter selection.

Statistical testing using McNemar's and DeLong's tests revealed that XGBoost significantly outperformed logistic regression (McNemar: $p = 0.032$; DeLong: $p = 0.041$) and support vector machine ($p < 0.001$ for both tests), but did not significantly exceed Random Forest ($p = 0.118$ and $p = 0.089$), indicating comparable performance between ensemble methods (Supplementary Table S1). Critically, comparisons across feature sets using XGBoost showed that Model 5 (multimodal integrated, 0.91 accuracy) did not significantly outperform simpler models: Model 5 vs. Model 1 (neurocognitive only, 0.89): $p = 0.772$ (McNemar), $p = 0.584$ (DeLong); Model 5 vs. Model 3 (baseline biomarkers, 0.86): $p = 0.505$ and $p = 0.412$ (Supplementary Table S2). No pairwise feature set comparisons reached statistical significance (all $p > 0.3$), suggesting that the 2–5% point accuracy differences likely reflect sampling variation rather than genuine incremental value from progressive feature addition. These findings indicate that simpler models (1 to 3) provide statistically equivalent performance to complex, imputed, multimodal models.

Algorithm	Feature set	*n* features	Accuracy	Precision	Recall	ROC AUC	F1 score
Logistic regression	Model 1: Neuropsychometry	8	0.74 ± 0.20	0.76 ± 0.20	0.73 ± 0.26	0.80 ± 0.18	0.74 ± 0.24
	Model 2: + Clinicodemographic	10	0.66 ± 0.12	0.70 ± 0.16	0.82 ± 0.17	0.76 ± 0.18	0.73 ± 0.09
	Model 3: + Baseline biomarkers (t0)	12	0.71 ± 0.05	0.76 ± 0.06	0.73 ± 0.05	0.78 ± 0.08	0.74 ± 0.04
	Model 4: + Follow-up biomarkers (t1)[a]	16	0.76 ± 0.04	0.87 ± 0.11	0.71 ± 0.15	0.82 ± 0.10	0.76 ± 0.08
	Model 5: Complete integration[a]	18	0.72 ± 0.07	0.73 ± 0.09	0.87 ± 0.08	0.80 ± 0.10	0.78 ± 0.03
Support vector machine	Model 1: Neuropsychometry	8	0.60 ± 0.07	0.59 ± 0.05	1.00 ± 0.00	0.81 ± 0.12	0.74 ± 0.04
	Model 2: + Clinicodemographic	10	0.61 ± 0.10	0.67 ± 0.14	0.73 ± 0.22	0.68 ± 0.13	0.67 ± 0.09
	Model 3: + Baseline biomarkers (t0)	12	0.71 ± 0.06	0.74 ± 0.04	0.76 ± 0.13	0.82 ± 0.07	0.74 ± 0.08
	Model 4: + Follow-up biomarkers (t1)[a]	16	0.62 ± 0.15	0.65 ± 0.14	0.73 ± 0.23	0.78 ± 0.11	0.68 ± 0.16
	Model 5: Complete integration[a]	18	0.67 ± 0.08	0.69 ± 0.05	0.76 ± 0.15	0.79 ± 0.13	0.72 ± 0.09
Random forest	Model 1: Neuropsychometry	8	0.81 ± 0.10	0.86 ± 0.10	0.80 ± 0.13	0.86 ± 0.15	0.82 ± 0.10
	Model 2: + Clinicodemographic	10	0.77 ± 0.14	0.80 ± 0.13	0.80 ± 0.13	0.85 ± 0.14	0.80 ± 0.13
	Model 3: + Baseline biomarkers (t0)	12	0.70 ± 0.12	0.72 ± 0.09	0.73 ± 0.18	0.75 ± 0.15	0.72 ± 0.13
	Model 4: + Follow-up biomarkers (t1)[a]	16	0.70 ± 0.14	0.75 ± 0.14	0.71 ± 0.15	0.77 ± 0.14	0.73 ± 0.13
	Model 5: Complete integration[a]	18	0.67 ± 0.10	0.72 ± 0.11	0.71 ± 0.09	0.80 ± 0.09	0.71 ± 0.09
XGBoost	Model 1: Neuropsychometry	8	0.89 ± 0.12	0.92 ± 0.12	0.89 ± 0.10	0.95 ± 0.06	0.90 ± 0.10
	Model 2: + Clinicodemographic	10	0.89 ± 0.12	0.90 ± 0.11	0.91 ± 0.08	0.95 ± 0.08	0.90 ± 0.10
	Model 3: + Baseline biomarkers (t0)	12	0.86 ± 0.13	0.85 ± 0.12	0.93 ± 0.09	0.94 ± 0.08	0.89 ± 0.11
	Model 4: + Follow-up biomarkers (t1)[a]	16	0.90 ± 0.08	0.90 ± 0.09	0.93 ± 0.05	0.97 ± 0.04	0.92 ± 0.06
	Model 5: Complete integration[a]	18	0.91 ± 0.08	0.92 ± 0.08	0.93 ± 0.05	0.97 ± 0.04	0.92 ± 0.06

Table 3. Binary classification performance across feature sets and algorithms. Binary classification task: Early-Onset Psychosis (EOS, $n = 45$) vs. Healthy Controls (HC, $n = 34$). Performance metrics calculated using nested 5-fold stratified cross-validation (outer loop). Values shown as mean ± standard deviation across 5 folds. [a]Models 4 and 5 include follow-up biomarker measurements (t1) which required KNN imputation (k = 5) for healthy controls who did not return for follow-up visits. Performance estimates for these models may be inflated due to imputation artifacts (see Discussion). XGBoost = eXtreme Gradient Boosting.

Feature analysis

Feature importance analysis revealed that cognitive measures were the most discriminative variables, with ZVT error scores showing the highest importance at baseline (0.103) and follow-up (0.15), indicating that verbal cognitive dysfunction is a robust marker that strengthens over time (Fig. 2). Response time parameters collectively ranked highly (median RT: 0.072, mean RT: 0.047, SD: 0.046), reflecting processing speed deficits and response inconsistency in early psychosis. Among biomarkers, BDNF pathway components demonstrated the greatest importance (BDNF_T0: 0.069, P75NTR_T1: 0.066, BDNF_T1: 0.062), supporting altered neuroplasticity signalling in the disorder. Follow-up measurements generally outperformed baseline values, suggesting temporal changes provide superior discriminative power, while demographic variables (age: 0.02, gender: 0.036) showed minimal importance compared to neurocognitive and neurobiological factors.

Correlation analysis revealed strong within-domain associations and modest cross-domain relationships. Reaction time measures demonstrated high intercorrelations (mean-median RT: $r = 0.97$, mean RT-SD: $r = 0.91$), while protein biomarkers showed strong temporal stability (P75NTR: $r = 0.93$, BDNF: $r = 0.84$, proBDNF: $r = 0.83$) and significant cross-protein correlations, particularly proBDNF_T0 with P75NTR_T0 ($r = 0.71$). ZVT performance was consistent across trials ($r = 0.79$) and showed moderate correlations with reaction time measures ($r = 0.2$–0.4). Age demonstrated negative associations with reaction time variables (r = -0.34 to -0.41) and positive correlation with advantageous decision-making ($r = 0.17$). Cross-domain correlations remained generally modest, supporting the complementary nature of different measurement modalities.

Diagnostic performance

The confusion matrix of the winning model (Fig. 2) revealed that the best-perfoming XGBoost model demonstrated a good classification capability, with 33 healthy controls correctly identified (true negatives), 44 EOS patients correctly classified (true positives), and only 2 misclassifications total—1 false positive (healthy control misclassified as EOS) and 1 false negative (EOS patient misclassified as healthy control).

A survey of medical records was conducted by two independent physicians, in an attempt to eludciate possible causes for this mismatch. No obvious reasons, including typical risk factors or known socioeconomic status was found. We may not exclude the chance for future morbidity for this individual.

Discussion

To our knowledge, this is the first study utilising machine learning in an attempt to classify EOS from HC, based on an ensemble method of integrating protein biomarkers, neuropsychometry and clinicodemographic data. The comparison of unimodal versus multimodal approaches revealed important insights into contribution of individual features and group of features. Neuropsychometric measures alone (Model 1) provided a strong

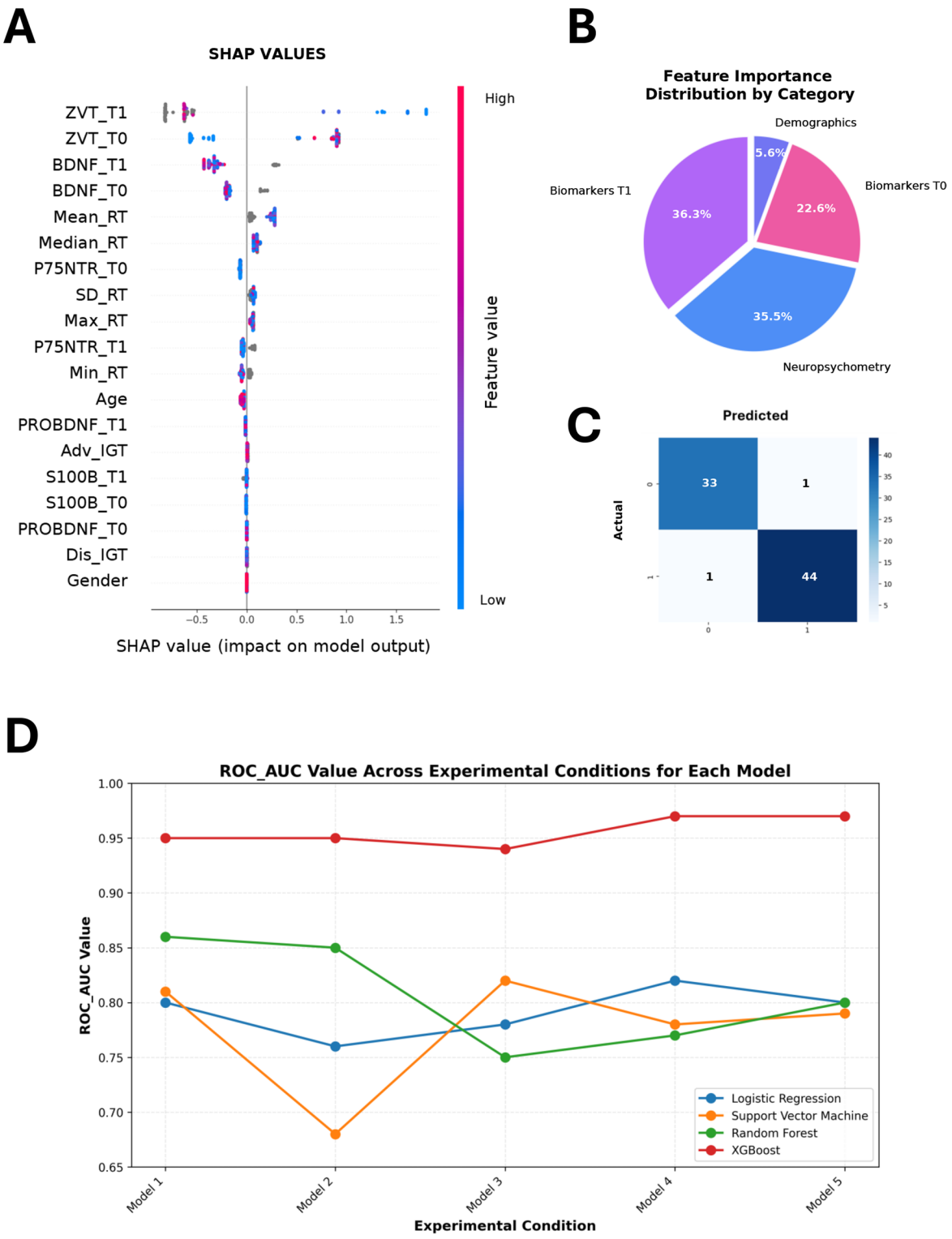

Fig. 2. Feature importance analysis for the winning model analysed through SHAP values (**A**) and simple feature importance values of the algorithm (**B**), ROC_AUC value across different models and experimental conditions (**C**), Confusion matrix for the winning classification model (**D**). ROC_AUC = Receiver Operating Characteristic Area Under Curve, Model 1 = neuropsychometry only, Model 2 = neuropsychometry and clinicodemographic data, Model 3 = neuropsychometry and protein biomarkers at baseline (t0), Model 4 = Model 3 with addition of biomarkers after stabilisation of treatment (t1), Model 5 = Model 4 with addition of clinicodemographic data. SD = standard deviation, SVM = Support Vector Machines, XGBoost = eXtreme Gradient Boosting. ZVT = Zabor Verbal Task. RT = Reaciton Time, Adv IGT = Advantageous Cards at Iowa Gambling Task, Dis IGT = Disadvantageous Cards at Iowa Gambling Task, T0 = Baseline, T1 = Follow-up.

foundation for classification, with XGBoost achieving 0.89 accuracy using cognitive testing data only. The addition of demographic data (Model 2) did not substantially improve performance across most algorithms, suggesting limited additional predictive value from demographic variables alone. Inclusion of baseline protein biomarkers at t0 (Model 3) generally maintained performance, with XGBoost showing marginal decline (0.86 accuracy) while other algorithms showed mixed results. Interestingly, using protein biomarkers from the post-treatment timepoint t1 (Model 4) yielded improved results for XGBoost (0.90 accuracy, 0.97 AUC), suggesting that protein levels after clinical stabilization may provide stronger diagnostic signals than baseline measurements. The complete dataset combining all modalities (Model 5) provided the best overall classification performance for XGBoost, achieving the highest accuracy (0.91) and maintaining excellent AUC (0.97). This suggests that the integration of neuropsychometric measures, demographic factors, and longitudinal protein biomarker data creates synergistic diagnostic value beyond any single modality. This, however, did not reach statistical significance, which warrants caution as regards increasing the complexity of the feature inputs. The lack of statistical significance between models likely reflects insufficient power ($n = 79$); nonetheless, even if differences are small, integrating baseline biomarkers (Model 3) provides clinical value through convergent validation, thus ensuring diagnoses are not solely dependent on cognitive performance, which may be confounded by anxiety, education, or fatigue in individual cases, whilst avoiding imputation requirements of Model 5.

The superior performance of the XGBoost algorithm in our study highlights the complex, non-linear relationships underlying early onset psychosis that may not be captured through traditional univariate statistical approaches. While individual biomarkers such as BDNF, proBDNF, P75NTR, and S100B showed mixed patterns of statistical significance between healthy controls and patients[8,63–67], XGBoost was able to identify meaningful combinations and interactions among these neurobiological variables alongside clinical, cognitive, and demographic factors. This ensemble learning approach excels at detecting subtle multi-dimensional signatures where, for instance, specific combinations of protein expression levels (even when individually non-significant), cognitive performance metrics from the IGT and ZVT clinical severity scores, and demographic variables collectively contribute to diagnostic classification. The algorithm's ability to weight and combine features dynamically allows it to capture the heterogeneous nature of early psychosis, where the disorder may manifest through diverse pathophysiological pathways rather than consistent directional changes in individual biomarkers.

The feature importance analysis provides crucial insights into the neurobiological and cognitive architecture underlying early-onset psychosis. The prominence of ZVT error scores as the most discriminative feature (importance: 0.103–0.15) underscores the central role of formal thought disorder and language-spatial integration deficits in EOS pathophysiology. This finding aligns with established theories of psychosis as fundamentally a disorder of cognitive coherence[68–72] and semantic processing[73–76], where patients struggle to maintain logical connections between concepts and their spatial representations[77,78]. The collective high importance of response time parameters (median RT: 0.072, mean RT: 0.047, SD: 0.046) further supports the hypothesis that information processing speed deficits represent a core endophenotype of psychosis[79–81], reflecting compromised neural efficiency across distributed brain networks[82]. The dominance of BDNF pathway components (BDNF_T0: 0.069, P75NTR_T1: 0.066, BDNF_T1: 0.062) among biomarkers suggests that altered neurotrophin signaling—critical for synaptic plasticity[83], neuronal survival[84], and circuit refinement[85]—may serve as a reliable molecular signature of the disorder. Notably, the superior discriminative power of follow-up measurements compared to baseline values indicates that EOS classification may benefit more from capturing dynamic neurobiological changes during treatment response rather than static baseline assessments. The minimal contribution of demographic variables (age: 0.02, gender: 0.036) suggests that the pathophysiological signature of EOS transcends traditional demographic risk factors, emphasizing the importance of neurobiological and neurocognitive markers in disease identification and potentially supporting a more biologically informed diagnostic approach.

An intriguing finding warranting discussion is that post-treatment biomarker measurements (t1) demonstrated higher feature importance than baseline levels (t0) for several proteins, particularly p75NTR and BDNF. This observation raises important questions about the biological and clinical significance of temporal biomarker dynamics. Several non-mutually exclusive explanations merit consideration. First, baseline measurements may be confounded by state-dependent factors associated with acute psychosis, including stress axis activation, systemic inflammation, sleep disruption, and metabolic dysregulation, that obscure underlying trait-level neurobiological signatures. Post-stabilization measurements, obtained after symptom resolution and clinical improvement, may capture more stable disease-related alterations once these acute confounds have resolved. Second, the trajectory of biomarker change during treatment may itself be diagnostically informative. Differential patterns of BDNF/p75NTR response to antipsychotic treatment could reflect fundamental differences in neuroplasticity capacity, synaptic reorganization potential, or compensatory mechanisms between EOS patients and healthy individuals. In this interpretation, t1 measurements capture not merely disease signatures but also individual variability in treatment response and recovery trajectories that enhance diagnostic discrimination. Third, antipsychotic-induced neurobiological changes may paradoxically improve classification accuracy if medication effects differ systematically between diagnostic groups due to underlying pathophysiological differences in receptor systems or downstream signalling cascades. Interpretation of these temporal dynamics is constrained by our asymmetric study design as discussed in more detail below.

Another potential benefit of establishing an algorithm based on biomarkers could be post-mortem diagnosis of EOS. In cases where the individual is dead the diagnosis of psychosis based on the commonly used neuropsychometric tools becomes impossible. Thus, diagnosing a first or subtle onset of psychosis postmortem is based on the medical history and reports of relatives or friends while remaining impossible to diagnose via commonly used clinical diagnostic tools. Determining serum levels of biomarkers post-mortem comes with the additional challenge of changing serum levels based on haemolysis and proteolysis[86]. Studies investigating the post-mortem serum levels of S100B and BDNF in individuals where the time of death was known (failed CPR

attempts in ER setting or witnessed collapse and death) without head trauma showed an increase of S100B[87] concentration with an increasing post-mortem interval whereas BDNF showed stability up to the examined time period of 48 h post-mortem[88]. Taking the known post-mortem changes in serum levels into account another way of diagnosis EOS post-mortem could arise. Further research on the post-mortem serum levels of the listed biomarkers in known EOS patients compared to HC could lead to a better way of approaching post-mortem EOS diagnosis.

Our machine learning classification model can also be used as a screening in workplaces that require a high level of reliability under prolonged periods of stress (e.g. soldiers on overseas missions, submarine crews, researchers on expeditions in isolation or astronauts on space missions). A diagnostic panel which is fast, reliable, interviewer-independent and objective, could be integrated into routine examinations to screen for cognitive deterioration. Using biomarkers in diagnostic tools for early-onset psychosis might also help to reduce the social stigma associated with neurological diseases. In workplaces where psychosis or other neurological disorder might be career-ending (e.g. pilots), there is an incentive to conceal symptoms and diseases from the employer and regulatory authorities. Using diagnostic models including biomarkers could offer a new way to approach neurological diseases and their long-term management, as they could be used for monitoring after an initial psychotic episode.

Real-world deployment would likely follow a tiered approach rather than universal application. Initial diagnostic evaluation would proceed through standard clinical assessment (psychiatric interview, symptom rating scales, developmental history), with the multimodal model reserved for cases presenting diagnostic uncertainty, such as patients with ambiguous symptomatology, atypical presentations, or comorbidities that complicate differential diagnosis. This targeted approach would optimize resource utilization while maintaining accessibility.

The practical workflow might involve: (1) brief cognitive screening using tablet- or computer-based assessments (ZVT, reaction time tasks requiring ~ 20–30 min total administration time); (2) blood sampling for biomarker analysis via standard venipuncture; (3) automated model prediction generation within 24–48 h once laboratory results are available; and (4) integration of model outputs with clinical judgment to inform diagnostic formulation and treatment planning. Critically, machine learning predictions should be presented as decision support tools rather than autonomous diagnoses, providing clinicians with probabilistic risk estimates and feature importance profiles to inform rather than replace clinical reasoning.

Cost-benefit considerations present both opportunities and challenges. The direct costs include neuropsychometric assessment (approximately 30 min of trained personnel time), biomarker immunoassay panels (estimated $200–400 for four-plex ELISA in research settings, though clinical laboratory costs may be significantly less in mass-production), and computational infrastructure for model deployment. These expenses must be weighed against the substantial costs associated with diagnostic uncertainty in early psychosis: prolonged hospitalization during diagnostic workup (average $1,000–2,000 per day), delayed initiation of appropriate treatment with associated functional deterioration, potential exposure to inappropriate medications with attendant side effects and costs, and the long-term socioeconomic burden of suboptimal early intervention. In tertiary referral centres and specialized early intervention services, where diagnostic complexity is highest and consequences of misdiagnosis most severe, the cost-effectiveness profile may be favourable. However, in resource-limited settings or primary care contexts, the added expense may not be justified compared to standard clinical evaluation.

Several methodological limitations should be considered when interpreting these findings. The relatively small sample size ($n = 79$) may limit the generalizability of our machine learning models and increase susceptibility to overfitting, particularly given the high-dimensional feature space employed. The mild class imbalance (ratio 1.32:1) may affect model generalizability despite multiple mitigation strategies including stratified cross-validation, conservative regularization, and feature stability analysis. While internal analyses suggest imbalance effects are minimal as evidenced by balanced class-specific performance (sensitivity 0.93 vs. specificity 0.88) and performance well above majority-class baseline (86–91% vs. 57%), external validation in larger, balanced cohorts is essential before clinical deployment. Until such validation confirms transportability, the current results should be interpreted as hypothesis-generating evidence supporting the feasibility of multimodal machine learning approaches to early psychosis detection, rather than as clinically validated diagnostic tools ready for implementation. Future work will prioritize simplified models (neurocognitive-only or stable-feature subsets) that demonstrate more favourable sample-to-feature ratios and may prove more robust to sampling variation.

That notwithstanding, our findings suggest that even without longitudinal biomarker data, clinically meaningful classification performance can be achieved using readily available baseline measurements and cognitive assessments. The consistency of results across different feature combinations, including those free from imputation artifacts, strengthens confidence in the overall methodological approach and supports the robustness of our core findings.

It is important to consider the universal medication status of EOS patients during assessment (mean chlorpromazine equivalent 336.3 mg/day), which introduces potential confounding of disease-related signatures with treatment effects. Antipsychotic medications influence both cognitive performance and serum biomarker levels, raising the question of whether our model detects core disease pathophysiology, medication-induced changes, or both. Meta-analytic evidence indicates that antipsychotics have time-dependent effects on BDNF levels[89,90], though our baseline measurements (t0) were obtained within one week of admission, potentially before substantial medication-induced changes occurred. Regarding cognitive effects, the relationship between antipsychotics and neurocognitive performance appears dose-dependent, with detrimental effects primarily emerging at higher doses and with significant anticholinergic burden[91,92]. Given that our patients received relatively moderate doses, medication-induced cognitive impairment is likely to be limited. Several considerations suggest substantial disease-related signal persists: the strongest discriminative features were cognitive measures

(ZVT errors, response time parameters) reflecting core pathophysiological processes in psychosis rather than medication side effects; feature importance patterns align with established theories of psychosis as a disorder of (meta)cognitive coherence[69,93] and linguistic processing[73,94]. BDNF pathway components are primarily implicated in neurodevelopmental disruptions underlying EOS rather than being exclusively treatment-related[18]. While medication status limits direct generalizability to medication-naïve populations, this reflects clinical reality where EOS diagnosis typically occurs after treatment initiation in tertiary care settings. From a clinical utility perspective, a model performing robustly in medicated patients may be more practically valuable than one requiring medication-free status, as most diagnostic decisions occur in treatment-engaged populations. Nevertheless, future studies could employ medication-naïve cohorts or longitudinal pre-post treatment designs to disentangle disease-specific signatures from treatment effects and examine dose-response relationships between antipsychotic exposure (both current and cumulative) and model predictions to clarify the extent to which classification depends on medication-related factors.

In summary, it is important to acknowledge that our model evaluation relies entirely on cross-validation within a single dataset from one tertiary referral centre. External validation on completely independent, unseen datasets from different institutions and populations is the gold standard for establishing diagnostic accuracy and represents an essential next step before any clinical implementation could be considered. The current results should be interpreted as preliminary and hypothesis-generating rather than clinically validated, demonstrating proof-of-concept for multimodal machine learning approaches to early psychosis detection but requiring prospective external validation to confirm generalizability, robustness to population variability, and true clinical utility.

Conclusions

This study demonstrates that machine learning algorithms, particularly XGBoost, can effectively integrate neuropsychometric assessments and protein biomarkers to achieve high-accuracy classification of early-onset psychosis patients from healthy controls. The superior performance of multimodal approaches over single-domain assessments underscores the multifactorial nature of EOS and supports the development of comprehensive diagnostic panels that capture both cognitive and neurobiological dimensions of the disorder. The prominence of formal thought disorder measures (ZVT) and BDNF pathway proteins as key discriminative features provides valuable insights into the core pathophysiological mechanisms underlying early psychosis and suggests specific targets for future biomarker development.

While longitudinal biomarker data enhanced classification performance, the substantial diagnostic value achieved using baseline measurements alone indicates the potential for developing clinically feasible screening tools that do not require extended follow-up periods. These findings represent an important step towards an objective, data-driven approach for the early identification of psychosis in adolescents, potentially enabling more timely intervention and improved clinical outcomes. Future validation studies in larger, independent cohorts will be essential to establish the clinical utility and generalisability of this machine learning-based diagnostic framework.

Data availability

The data code and materials necessary to reproduce the analyses presented here are not publicly accessible, however they will be made available from the first author upon reasonable request.

Received: 16 October 2025; Accepted: 22 December 2025
Published online: 30 December 2025

References

1. Solmi, M. et al. Incidence, prevalence, and global burden of schizophrenia - data, with critical appraisal, from the global burden of disease (GBD) 2019. *Mol. Psychiatry*. **28**, 5319–5327 (2023).
2. Tripathi, A., Kar, S. K. & Shukla, R. Cognitive deficits in schizophrenia: Understanding the biological correlates and remediation strategies. *Clin. Psychopharmacol. Neurosci.* **16**, 7–17 (2018).
3. Pelizza, L. et al. Diagnostic shift in adolescents with first episode psychosis: findings from the 2-year follow-up of the 'Parma early psychosis' program. *Soc. Psychiatry Psychiatr Epidemiol.* **60**, 375–385 (2025).
4. Di Luzio, M. et al. Clinical features and comorbidity in very early-onset schizophrenia: a systematic review. *Front. Psychiatry* **14** https://doi.org/10.3389/fpsyt.2023.1270799 (2023).
5. Coulon, N. et al. Early and very early-onset schizophrenia compared with adult-onset schizophrenia: French FACE-SZ database. *Brain Behav.* **10**, e01495 (2020).
6. Liu, Z. et al. Comparison of clinical symptoms and symptom structure across different onset ages in schizophrenia inpatients. *Schizophr Res.* **277**, 177–184 (2025).
7. Baykal, S. et al. A comparison of clinical characteristics and course predictors in early- and childhood-onset schizophrenia. *Early Interv Psychiatry*. **19**, e13594 (2025).
8. Zakowicz, P. et al. Plasma biomarkers in adolescents with schizophrenia-spectrum disorder. *Early Interv Psychiatry*. **17**, 1154–1161 (2023).
9. Zakowicz, P., Skibińska, M. & Pawlak, J. Disembodied Language in Early-Onset schizophrenia. *Front. Psychiatry*. **13**, 888844 (2022).
10. Ueland, T. et al. Cognitive functioning in adolescents with schizophrenia spectrum disorders. *Psychiatry Res.* **126**, 229–239 (2004).
11. Zakowicz, P. et al. Detection of formal thought disorders in child and adolescent psychosis using machine learning and neuropsychometric data. *Front. Psychiatry* **16** https://doi.org/10.3389/fpsyt.2025.1550571 (2025).
12. White, T. et al. Limbic structures and networks in children and adolescents with schizophrenia. *Schizophr Bull.* **34**, 18–29 (2008).
13. Gogtay, N. Cortical brain development in schizophrenia: insights from neuroimaging studies in childhood-onset schizophrenia. *Schizophr Bull.* **34**, 30–36 (2008).
14. Brownstein, C. A. et al. Similar rates of deleterious copy number variants in Early-Onset psychosis and autism spectrum disorder. *AJP* **179**, 853–861 (2022).

15. Chang, S. E. et al. Attention-mediated genetic influences on psychotic symptomatology in adolescence. *Nat. Mental Health*. **2**, 1518–1531 (2024).
16. Momtazmanesh, S. et al. Brain microstructural abnormalities in 22q11.2 deletion syndrome: A systematic review of diffusion tensor imaging studies. *Eur. Neuropsychopharmacol.* **52**, 96–135 (2021).
17. Chen, B-Y. et al. Neurodevelopment regulators miR-137 and miR-34 family as biomarkers for early and adult onset schizophrenia. *Npj Schizophr*. **7**, 35 (2021).
18. Nieto, R. R. et al. BDNF as a biomarker of cognition in Schizophrenia/Psychosis: an updated review. *Front. Psychiatry*. **12**, 662407 (2021).
19. van de Kerkhof, N. W. A. et al. BDNF and S100B in psychotic disorders: evidence for an association with treatment responsiveness. *Acta Neuropsychiatr*. **26**, 223–229 (2014).
20. Chukaew, P. et al. Correlation of BDNF, VEGF, TNF-α, and S100B with cognitive impairments in chronic, medicated schizophrenia patients. *Neuropsychopharmacol. Rep.* **42**, 281–287 (2022).
21. Lundin, N. B. et al. Identification of psychosis risk and diagnosis of First-Episode psychosis: advice for clinicians. *Psychol. Res. Behav. Manag.* **17**, 1365–1383 (2024).
22. Buckley, P. F. et al. Brain derived neurotropic factor in first –Episode psychosis. *Schizophr Res.* **91**, 1–5 (2007).
23. Cohen-Cory, S. et al. Brain-derived neurotrophic factor and the development of structural neuronal connectivity. *Dev. Neurobiol.* **70**, 271–288 (2010).
24. Gören, J. L. Brain-derived neurotrophic factor and schizophrenia. *Ment Health Clin.* **6**, 285–288 (2016).
25. Juric, D. M., Miklic, S. & Carman-Krzan, M. Monoaminergic neuronal activity up-regulates BDNF synthesis in cultured neonatal rat astrocytes. *Brain Res.* **1108**, 54–62 (2006).
26. Waterhouse, E. G. et al. BDNF promotes differentiation and maturation of adult-born neurons through GABAergic transmission. *J. Neurosci.* **32**, 14318–14330 (2012).
27. da Penha Berzaghi, M. et al. Cholinergic regulation of brain-derived neurotrophic factor (BDNF) and nerve growth factor (NGF) but not neurotrophin-3 (NT-3) mRNA levels in the developing rat hippocampus. *J. Neurosci.* **13**, 3818–3826 (1993).
28. Toll, A. & Mané, A. Brain-derived neurotrophic factor levels in first episode of psychosis: A systematic review. *World J. Psychiatry*. **5**, 154–159 (2015).
29. Yang, C-R. et al. Upregulation of proBDNF/p75NTR signaling in immune cells and its correlation with inflammatory markers in patients with major depression. *FASEB J.* **38**, e23312 (2024).
30. Riffault, B. et al. Pro-Brain-Derived neurotrophic factor (proBDNF)-Mediated p75NTR activation promotes depolarizing actions of GABA and increases susceptibility to epileptic seizures. *Cereb. Cortex*. **28**, 510–527 (2018).
31. Alnoaman, H. et al. Dysregulation of proBDNF/p75NTR and BDNF/TrkB signaling in acute ischemic stroke: different sides of the same coins. *Brain Res. Bull.* **226**, 111338 (2025).
32. Jourdi, G. et al. Soluble p75 neurotrophic receptor as a reliable biomarker in neurodegenerative diseases: what is the evidence? *Neural Regen Res.* **19**, 536–541 (2023).
33. Fujii, T. & Kunugi, H. p75NTR as a therapeutic target for neuropsychiatric diseases. *Curr. Mol. Pharmacol.* **2**, 70–76 (2009).
34. Shu, Y-H. et al. Update on the role of p75NTR in neurological disorders: A novel therapeutic target. *Biomed. Pharmacother*. **76**, 17–23 (2015).
35. Brozzi, F. et al. S100B protein regulates astrocyte shape and migration via interaction with Src kinase: implications for astrocyte development, activation, and tumor growth. *J. Biol. Chem.* **284**, 8797–8811 (2009).
36. Lam, A. G. et al. Mechanism of glial activation by S100B: involvement of the transcription factor NFkappaB. *Neurobiol. Aging*. **22**, 765–772 (2001).
37. Koh, S. X. T. & Lee, J. K. W. S100B as a marker for brain damage and blood-brain barrier disruption following exercise. *Sports Med.* **44**, 369–385 (2014).
38. Khan, B. A. et al. S100 calcium binding protein B as a biomarker of delirium duration in the intensive care unit - an exploratory analysis. *Int. J. Gen. Med.* **6**, 855–861 (2013).
39. Steiner, J. et al. S100B protein in neurodegenerative disorders. *Clin. Chem. Lab. Med.* **49**, 409–424 (2011).
40. Kozlowski, T. et al. Peripheral S100B protein levels in five major psychiatric disorders: A systematic review. *Brain Sci.* **13**, 1334 (2023).
41. Schümberg, K. et al. Serum S100B is related to illness duration and clinical symptoms in schizophrenia—a meta-regression analysis. *Front. Cell Neurosci.* **10** https://doi.org/10.3389/fncel.2016.00046 (2016).
42. Steiner, J. et al. S100B serum levels in schizophrenia are presumably related to visceral obesity and insulin resistance. *Cardiovasc. Psychiatry Neurol.* **2010**, 480707. (2010).
43. Li, G. et al. Identification of binge drinkers via convolutional neural network and support vector machine. In: *2021 IEEE International Conference on Mechatronics and Automation (ICMA)* 715–720. (2021).
44. Zhang, Z. et al. Identifying biomarkers of subjective cognitive decline using graph convolutional neural network for fMRI analysis. In: *2022 IEEE International Conference on Mechatronics and Automation (ICMA)* 1306–1311. (2022).
45. Li, G. et al. Cognitive challenges are better in distinguishing binge from Nonbinge drinkers: an exploratory Deep-Learning study of fMRI data of multiple behavioral tasks and resting state. *J. Magn. Reson. Imaging*. **57**, 856–868 (2023).
46. Zhang, Z. et al. Application of artificial intelligence in the MRI classification task of human brain neurological and psychiatric diseases: A scoping review. *Diagnostics (Basel)*. **11**, 1402 (2021).
47. Aydın, S., Demirtaş, S. & Yetkin, S. Cortical correlations in wavelet domain for Estimation of emotional dysfunctions. *Neural Comput Applic*. **30**, 1085–1094 (2018).
48. Aydın, S. Cross-validated adaboost classification of emotion regulation strategies identified by spectral coherence in Resting-State. *Neuroinform* **20**, 627–639 (2022).
49. Aydın, S. & Akın, B. Machine learning classification of maladaptive rumination and cognitive distraction in terms of frequency specific complexity. *Biomed. Signal Process. Control.* **77**, 103740 (2022).
50. Smucny, J., Lesh, T. A. & Carter, C. S. Baseline frontoparietal Task-Related BOLD activity as a predictor of improvement in clinical symptoms at 1-Year Follow-Up in Recent-Onset psychosis. *AJP* **176**, 839–845 (2019).
51. Greenstein, D. et al. Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Front. Psychiatry* **3**. https://doi.org/10.3389/fpsyt.2012.00053 (2012).
52. Schnack, H. G. et al. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *NeuroImage* **84**, 299–306 (2014).
53. Yassin, W. et al. Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl Psychiatry*. **10**, 278 (2020).
54. Birnbaum, M. L. et al. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *J. Med. Internet Res.* **19**, e289 (2017).
55. Rodríguez-Ruiz, J. G. et al. Classification of depressive and schizophrenic episodes using Night-Time motor activity signal. *Healthc. (Basel)*. **10**, 1256 (2022).
56. Jean, T., Hottin, R. G. & Orban, P. Forecasting mental States in schizophrenia using digital phenotyping data. *PLOS Digit. Health*. **4**, e0000734 (2025).
57. Bracher-Smith, M. et al. Machine learning for prediction of schizophrenia using genetic and demographic factors in the UK biobank. *Schizophr. Res.* **246**, 156–164 (2022).

58. Zakowicz, P. T. et al. Detection of formal thought disorders in child and adolescent psychosis using machine learning and neuropsychometric data. *Front. Psychiatry*. **16**, 1550571 (2025).
59. Rajewska-Rager, A. et al. Longitudinal assessment of S100B serum levels and clinical factors in youth patients with mood disorders. *Sci. Rep.* **11**, 11973 (2021).
60. Mueller, S. T. & Piper, B. J. The psychology experiment Building Language (PEBL) and PEBL test battery. *J. Neurosci. Methods*. **222**, 250–259 (2014).
61. Pedregosa, F. et al. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
62. Lundberg, S. M. & Lee, S-I. A Unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems*. (Accessed 1 July 2025). https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html (Curran Associates, Inc., 2017).
63. Simsek, S. et al. Lower Brain-Derived neurotropic factor levels in untreated adolescents with First-Episode psychosis. *J. Clin. Psychopharmacol.* **35**, 596 (2015).
64. Kelsven, S. et al. Immuno-inflammatory changes across phases of early psychosis: the impact of antipsychotic medication and stage of illness. *Schizophr Res.* **226**, 13–23 (2020).
65. Yee, J. Y., Lee, T-S. & Lee, J. A longitudinal study of serum Brain-Derived neurotrophic factor levels in First-Episode schizophrenia. *J. Clin. Psychopharmacol.* **39**, 639–643 (2019).
66. Akyol, E. S. et al. Decreased serum levels of brain-derived neurotrophic factor in schizophrenic patients with deficit syndrome. *Neuropsychiatr Dis. Treat.* **11**, 865–872 (2015).
67. Milleit, B. et al. Serum S100B protein is specifically related to white matter changes in schizophrenia. *Front Cell. Neurosci.* **10**. https://doi.org/10.3389/fncel.2016.00033 (2016).
68. Lysaker, P. H. et al. Metacognitive deficits in schizophrenia: presence and associations with psychosocial outcomes. *J. Nerv. Ment. Dis.* **203**, 530 (2015).
69. Kahn, R. S. et al. Schizophrenia. *Nat. Rev. Dis. Primers*. **1**, 15067 (2015).
70. Keefe, R. S. E. & Harvey, P. D. Cognitive impairment in schizophrenia. *Handb. Exp. Pharmacol.* 11–37. (2012).
71. Gladsjo, J. A. et al. A Six-Factor model of cognition in schizophrenia and related psychotic disorders: relationships with clinical symptoms and functional capacity. *Schizophr. Bull.* **30**, 739–754 (2004).
72. Carter, C. S. & Barch, D. M. Cognitive Neuroscience-Based approaches to measuring and improving treatment effects on cognition in schizophrenia: the CNTRICS initiative. *Schizophr. Bull.* **33**, 1131–1137 (2007).
73. Corcoran, C. M. et al. Language as a biomarker for psychosis: A natural Language processing approach. *Schizophr. Res.* **226**, 158–166 (2020).
74. de Boer, J. N. et al. Language in schizophrenia: relation with diagnosis, symptomatology and white matter tracts. *Npj Schizophr*. **6**, 10 (2020).
75. de Boer, J. N. et al. Anomalies in Language as a biomarker for schizophrenia. *Curr. Opin. Psychiatry*. **33**, 212 (2020).
76. Andreasen, N. C. Scale for the assessment of Thought, Language, and communication (TLC). *Schizophr. Bull.* **12**, 473–482 (1986).
77. Shergill, S. S. et al. Evidence for sensory prediction deficits in schizophrenia. *AJP* **162**, 2384–2386 (2005).
78. Thakkar, K. N. & Park, S. Impaired passive maintenance and spared manipulation of internal representations in patients with schizophrenia. *Schizophr. Bull.* **38**, 787–795 (2012).
79. Knowles, E. E. M., David, A. S. & Reichenberg, A. Processing speed deficits in schizophrenia: reexamining the evidence. *Am. J. Psychiatry*. **167**, 828–835 (2010).
80. Cadenhead, K. S. et al. Information processing deficits of schizophrenia patients: relationship to clinical ratings, gender and medication status. *Schizophr Res.* **28**, 51–62 (1997).
81. Nieman, D. H. et al. Psychosis prediction: stratification of risk Estimation with Information-Processing and premorbid functioning variables. *Schizophr. Bull.* **40**, 1482–1490 (2014).
82. Sun, X. & Xia, M. Schizophrenia and neurodevelopment: insights from connectome perspective. *Schizophr. Bull.* **51**, 309–324 (2025).
83. Lu, B., Nagappan, G. & Lu, Y. BDNF and synaptic plasticity, cognitive function, and dysfunction. *Handb. Exp. Pharmacol.* **220**, 223–250 (2014).
84. Lipsky, R. H. & Marini, A. M. Brain-derived neurotrophic factor in neuronal survival and behavior-related plasticity. *Ann. N Y Acad. Sci.* **1122**, 130–143 (2007).
85. Wollet, M. & Kim, J. H. Brain-Derived neurotrophic factor is involved in Activity-Dependent tonotopic refinement of MNTB neurons. *Front. Neural Circuits*. **16**, 784396 (2022).
86. Maeda, H. et al. Significance of postmortem biochemistry in determining the cause of death. *Leg. Med.* **11**, S46–S49 (2009).
87. Ondruschka, B. et al. S100B and NSE as useful postmortem biochemical markers of traumatic brain injury in autopsy cases. *J. Neurotrauma*. **30**, 1862–1871 (2013).
88. Ondruschka, B. et al. Post-mortem in situ stability of serum markers of cerebral damage and acute phase response. *Int. J. Legal Med.* **133**, 871–881 (2019).
89. Green, M. J. et al. Brain-derived neurotrophic factor levels in schizophrenia: a systematic review with meta-analysis. *Mol. Psychiatry*. **16**, 960–972 (2011).
90. Fernandes, B. S. et al. Peripheral brain-derived neurotrophic factor (BDNF) as a biomarker in bipolar disorder: a meta-analysis of 52 studies. *BMC Med.* **13**, 289 (2015).
91. Haddad, C. et al. Effects of antipsychotic and anticholinergic medications on cognition in chronic patients with schizophrenia. *BMC Psychiatry*. **23**, 61 (2023).
92. Rehse, M. et al. Influence of antipsychotic and anticholinergic loads on cognitive functions in patients with schizophrenia. *Schizophr. Res. Treatment* **2016** 8213165. (2016).
93. Lysaker, P. H. et al. Metacognitive deficits in schizophrenia: presence and associations with psychosocial outcomes. *J. Nerv. Ment Dis.* **203**, 530–536 (2015).
94. de Boer, J. N. et al. Anomalies in Language as a biomarker for schizophrenia. *Curr. Opin. Psychiatry*. **33**, 212–218 (2020).

Author contributions

Przemyslaw T. Zakowicz Ideas; formulation or evolution of overarching research goals and aims, creation and drafting, editing of the manuscript. Maksymilian A. Brzezicki Ideas; formulation or evolution of overarching research goals and aims, creation and drafting, editing of the manuscript. PT Zakowicz and MA Brzezicki contributed to this work equally. Joanna Pawlak creation and drafting, editing of the manuscript. Maria Skibinska Ideas; formulation or evolution of overarching research goals and aims. Szymon Jurga creation and drafting, editing of the manuscript. Aleksandra Lewandowska creation and drafting, editing of the manuscript. Benedikt Vogel creation and drafting, editing of the manuscript. Emily Ungermann creation and drafting, editing of the manuscript. Barbara Remberk Ideas; formulation or evolution of overarching research goals and aims, Oversight and leadership responsibility for the research activity planning and execution, including mentorship.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-33765-2.

Correspondence and requests for materials should be addressed to M.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025