Mental Health and Lifestyle Journal

Year 2027
Volume 5
Issue 5

Predicting Treatment Resistance in Obsessive–Compulsive Disorder Using Deep Learning Networks

Mohammadjavad. Shahmiri[1], Mahboobeh. Javanmard Ahandani[2*], Mehrnoosh. Heydari Goojani[3], Sogand. Asghari[4]

1 Department of Psychology, ST.C., Islamic Azad University, Tehran, Iran.
2 Department of Psychology, Ka.C., Islamic Azad University, Karaj, Iran
3 Department of Psychology, Feyzoleslam University, Khomeyni Shahr, Iran
4 Department of Psychology, Qu.C., Islamic Azad University, Quchan, Iran
*Correspondence: alimohamadrz2011@gmail.com

Article type:
Original Research

Article history:
Received 20 July 2026
Revised 05 October 2026
Accepted 16 October 2026
Initial Publish 16 November 2026
Published online 01 September 2027

ABSTRACT

This study aimed to develop and evaluate an explainable deep learning model for predicting treatment resistance in patients with obsessive–compulsive disorder based on baseline demographic, clinical, psychopathological, and treatment-history characteristics. This retrospective longitudinal predictive study included 612 adults with obsessive–compulsive disorder who received psychiatric and psychological services in specialized centers in Tehran, Iran, between 2021 and 2026. Treatment resistance was defined as failure to achieve clinically meaningful improvement after at least two adequate serotonin reuptake inhibitor trials. Baseline predictors included Yale–Brown Obsessive Compulsive Scale scores, Obsessive–Compulsive Inventory–Revised scores, depressive and anxiety symptoms, insight, psychiatric comorbidities, illness duration, age at onset, previous hospitalization, medication adherence, exposure and response prevention, and prior augmentation strategies. The dataset was divided into training, validation, and independent test sets. A multilayer deep neural network was developed and compared with logistic regression, random forest, and gradient boosting models. Model performance was assessed using AUROC, AUPRC, accuracy, sensitivity, specificity, F1 score, and Brier score, and SHAP analysis was applied for interpretability. The deep neural network achieved an AUROC of .891, AUPRC of .844, accuracy of 82.6%, sensitivity of 81.1%, specificity of 83.6%, F1 score of .789, and Brier score of .139 in the independent test set. It significantly outperformed logistic regression (p = .012) and random forest (p = .041), while the difference from gradient boosting was not significant (p = .174). SHAP analysis identified baseline Y-BOCS severity, illness duration, poor insight, depressive symptom severity, earlier age at onset, previous antipsychotic augmentation, and completion of adequate exposure and response prevention as the most influential predictors. Deep learning demonstrated strong and clinically interpretable performance in predicting treatment resistance in obsessive–compulsive disorder and may support earlier identification of high-risk patients and more individualized treatment planning.

Keywords: Obsessive–compulsive disorder, treatment resistance, deep learning, neural networks, machine learning, explainable artificial intelligence, prediction

How to cite this article:
Shahmiri, M., Javanmard Ahandani, M., Heydari Goojani, M., & Asghari, S. (2027). Predicting Treatment Resistance in Obsessive–Compulsive Disorder Using Deep Learning Networks. *Mental Health and Lifestyle Journal, 5*(5), 1-21. https://doi.org/10.61838/mhlj.340

Introduction

Obsessive–compulsive disorder (OCD) is a chronic and frequently disabling psychiatric condition characterized by recurrent intrusive thoughts, images, or impulses and repetitive behaviors or mental acts performed to reduce distress or prevent feared outcomes. Although the phenomenology of OCD is highly heterogeneous, its central features commonly include persistent doubt, contamination concerns, checking,

© 2027 the authors. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

symmetry-related needs, aggressive or taboo obsessions, and ritualized responses that may occupy substantial portions of daily life. The disorder often follows a prolonged course and is associated with marked functional impairment, deterioration in occupational and academic performance, disruption of interpersonal and marital relationships, and reduced quality of life. The burden of OCD extends beyond symptom severity because recurrent obsessions and compulsions may interfere with autonomy, decision-making, social functioning, and family relationships, while associated avoidance and reassurance-seeking behaviors can further reinforce disability (1, 2). Cognitive mechanisms such as thought–action fusion and maladaptive attempts to suppress intrusive thoughts have also been implicated in the persistence of obsessive–compulsive symptoms, supporting the view that OCD reflects a complex interaction between intrusive cognition, threat appraisal, affect regulation, and repetitive behavioral responses rather than a unitary symptom process (3).

Current evidence-based treatment for OCD relies primarily on serotonin reuptake inhibitor pharmacotherapy and cognitive-behavioral interventions, especially exposure and response prevention. For many patients, these approaches produce clinically meaningful reductions in symptom severity and improve functioning, and cognitive-behavioral therapy remains a central component of contemporary OCD management (4, 5). Nevertheless, a substantial subgroup of patients experiences incomplete response, persistent residual symptoms, or minimal improvement despite apparently adequate treatment. Pharmacoresistant OCD therefore remains a major clinical challenge, particularly when patients have severe baseline symptoms, chronic illness, psychiatric comorbidity, poor insight, low treatment engagement, or previous failure of multiple therapeutic strategies (6). Treatment resistance is clinically important because it often leads to repeated medication trials, augmentation strategies, prolonged psychotherapy, specialist referral, hospitalization, and, in the most severe cases, consideration of invasive or experimental interventions. The distinction between treatment response and treatment resistance is thus not simply descriptive; it has direct implications for prognosis, treatment sequencing, resource allocation, and the timing of escalation to more intensive care.

The problem of treatment resistance has stimulated the development of specialized and multimodal treatment programs. Intensive tertiary OCD services can improve outcomes among patients with complex or treatment-resistant presentations, particularly when pharmacological optimization, specialized psychotherapy, and multidisciplinary management are integrated (7). Similarly, intensive inpatient treatment has demonstrated clinically meaningful long-term benefit in patients with severe and resistant OCD, although substantial heterogeneity remains in the magnitude and durability of response (8). For patients who have not responded sufficiently to conventional exposure-based treatment, alternative psychological interventions have also been explored. Inference-based therapy has shown potential across symptom subtypes and in some treatment-resistant cases by targeting dysfunctional inferential processes that contribute to obsessional doubt (9). Acceptance and commitment-based interventions have likewise been investigated as approaches to reducing obsessive symptoms in resistant OCD, particularly through modification of experiential avoidance and psychological inflexibility (10). More intensive schema-oriented approaches combined with cognitive-behavioral day treatment have also been examined in treatment-resistant anxiety disorders and OCD, reflecting growing recognition that entrenched maladaptive schemas and broader personality-related processes may interfere with standard treatment response (11).

Contemporary reviews therefore increasingly emphasize treatment innovation beyond conventional exposure protocols, including interventions designed to address cognitive rigidity, metacognitive processes, self-related beliefs, and broader transdiagnostic mechanisms (12).

Treatment resistance in OCD is likely to arise from multiple interacting sources rather than from a single mechanism. Clinical heterogeneity is one major contributor. Patients differ markedly in age at onset, illness duration, symptom dimensions, insight, family history, psychiatric comorbidity, affective burden, and previous treatment exposure. Genetic and hereditary influences may also contribute to vulnerability, illness persistence, and variability in therapeutic response, with emerging evidence linking OCD to complex polygenic mechanisms and interacting neurobiological pathways (13). Early environmental adversity may further shape neurobiological vulnerability through stress-related alterations in neural development, emotional regulation, and cortico-striatal functioning (14). At the systems level, contemporary neurocircuit models emphasize dysfunction across cortico-striato-thalamo-cortical networks, although these models are increasingly recognized as incomplete representations of a highly distributed and heterogeneous disorder (15). Repetitive behavior may also be understood within predictive-processing frameworks involving abnormal learning, uncertainty reduction, and basal ganglia mechanisms, providing a theoretical bridge between compulsive repetition, habit formation, and maladaptive prediction-error processing (16). Reward-related processes and dopaminergic modulation may also contribute to cognitive control abnormalities and dysfunctional working memory processes relevant to compulsive behavior (17).

Neurobiological research has further shown that obsessive and compulsive states may be associated with individualized neural signatures rather than a uniform disorder-specific biomarker. Intracranial recordings from the ventral striatum have demonstrated patient-specific neural patterns associated with obsessions and compulsions, highlighting the possibility that treatment resistance may partly reflect individualized circuit-level pathophysiology (18). Such findings are consistent with broader evidence that psychiatric heterogeneity and comorbidity may be better characterized through data-driven analysis of core neural regions and patterns rather than through categorical diagnosis alone (19). The increasing availability of neuroimaging and neurophysiological data has therefore created opportunities for machine-learning approaches capable of detecting latent patterns not readily identifiable through traditional clinical reasoning. Reviews of machine learning in OCD neuroimaging indicate that computational models have already been applied to diagnostic classification, symptom prediction, and characterization of structural and functional brain abnormalities (20). Likewise, hybrid deep learning architectures based on electroencephalographic data have shown promising capacity to identify individuals with OCD, suggesting that deep neural networks may effectively capture complex nonlinear relationships embedded within high-dimensional psychiatric data (21).

These computational developments are particularly relevant to the prediction of treatment resistance. Traditional prognostic methods generally depend on a limited number of variables analyzed using linear statistical models. Although such approaches are interpretable and clinically familiar, they may not adequately represent nonlinear relationships, threshold effects, higher-order interactions, and heterogeneous combinations of risk factors. In OCD, for example, severe baseline symptoms may have different prognostic significance depending on illness duration, insight, depressive comorbidity, previous treatment exposure, adherence, and symptom subtype. Deep learning models are theoretically well suited to

this problem because they can learn multilayered representations from complex predictor sets and identify interactions without requiring all relationships to be specified a priori. Precision psychiatry approaches increasingly propose that deep learning could support individualized treatment selection and risk estimation in OCD by integrating demographic, clinical, cognitive, neurobiological, and treatment-history variables (22). This shift from group-level prediction to individual-level probability estimation is clinically important because two patients with superficially similar symptom severity may have substantially different probabilities of responding to standard treatment.

At the same time, predictive modeling in OCD should be interpreted within the broader evolution of treatment strategies for severe and refractory illness. Deep brain stimulation has emerged as an option for highly selected patients with severe treatment-resistant OCD, although evidence regarding optimal targets, patient selection, long-term benefit, and adverse effects remains under continuing investigation (23). Neurosurgical treatments for psychiatric disorders are becoming increasingly informed by connectomics, network targeting, and individualized neuroanatomical mapping, creating a movement toward more precise and technologically integrated interventions (24). However, psychiatric neurosurgery raises substantial ethical and safety concerns, including informed consent, reversibility, autonomy, adverse effects, vulnerability of severely ill patients, and the balance between experimental innovation and established evidence (25). Reviews of psychosurgical evolution similarly emphasize that modern neuromodulation must be understood within a historical trajectory from destructive procedures toward adjustable and targeted circuit interventions (26). Deep brain stimulation is now being investigated across several severe psychiatric disorders, and its efficacy and safety in treatment-resistant conditions continue to be refined through better patient selection and technological advances (27). Research in other severe psychiatric conditions, including deep brain stimulation targeting the nucleus accumbens, further illustrates the broader translational relevance of reward-circuit modulation and the importance of careful feasibility assessment when applying invasive neuromodulatory methods (28).

Newer noninvasive and minimally invasive neuromodulation strategies may also expand future options for refractory OCD. Temporal interference stimulation of deep neural structures such as the nucleus accumbens has recently been investigated as a potential approach for modulating circuits implicated in obsessive–compulsive pathology (29). At the preclinical level, transgenic animal models are being used to clarify the neurobiological basis of repetitive behavior and to evaluate novel treatment mechanisms, including circuit-specific and molecular interventions (30). Such developments underscore a broader movement toward mechanism-based and individualized care. Nevertheless, the availability of advanced interventions also increases the importance of identifying treatment resistance earlier in the clinical trajectory. If patients with a high probability of nonresponse to routine treatment can be identified before multiple sequential failures occur, clinicians may be able to intensify monitoring, optimize adherence, prioritize specialized psychotherapy, use augmentation strategies more efficiently, or refer patients earlier to expert services.

The clinical complexity of OCD is further amplified by somatic and behavioral comorbidities. Chronic pain, for example, may coexist with OCD and contribute to greater functional burden, cognitive preoccupation, and treatment complexity (31). Other unusual clinical presentations involving somatic fixation or perceived bodily abnormalities may also overlap phenomenologically with obsessive, compulsive,

or delusional processes and require careful diagnostic differentiation (32). These examples illustrate why treatment outcome cannot be predicted reliably from obsessive–compulsive severity alone. Instead, prognostic modeling should consider interacting dimensions of psychopathology, comorbidity, behavioral patterns, and treatment history.

Recent neurobiological intervention research also suggests that treatment-resistant OCD may involve modifiable functional network abnormalities. Mindfulness-based cognitive therapy has been associated with changes in resting-state networks among patients with treatment-resistant OCD, indicating that even refractory presentations may retain neuroplastic potential when targeted interventions engage relevant regulatory systems (33). Such findings support the idea that resistance is not necessarily a fixed patient characteristic but may reflect the interaction between illness biology, treatment mechanism, treatment adequacy, and individual clinical context. Consequently, predictive models should not be interpreted as deterministic tools. Rather, they may function as probabilistic decision-support systems that identify patients whose combination of characteristics places them at elevated risk for inadequate response and who may therefore benefit from modified or intensified care.

Despite growing interest in precision psychiatry and computational modeling, several important gaps remain. Much of the machine-learning literature in OCD has focused on diagnosis, neuroimaging classification, or biological signal analysis rather than prediction of clinically defined treatment resistance using routinely available baseline clinical information. Studies that rely exclusively on neuroimaging, EEG, or highly specialized biomarkers may also have limited applicability in routine psychiatric services where such data are not consistently available. In contrast, demographic variables, symptom severity, insight, psychiatric comorbidity, treatment history, and adherence are commonly documented in clinical settings and may offer a more immediately translatable foundation for predictive modeling. Furthermore, many previous predictive studies have involved relatively small samples, limited external validation, or insufficient attention to overfitting, calibration, and explainability. For a predictive model to be clinically useful, it should not only discriminate between resistant and responsive patients but also generate well-calibrated probabilities and provide interpretable information regarding the factors that most strongly influence individual predictions.

Deep learning offers a promising framework for addressing these limitations because it can model complex nonlinear dependencies across heterogeneous clinical predictors. However, the use of neural networks in psychiatry introduces a parallel need for interpretability. Clinicians are unlikely to rely on a model that produces accurate predictions without clarifying why a particular patient is categorized as high risk. Explainable artificial intelligence methods such as SHapley Additive exPlanations can help address this issue by quantifying the contribution of individual variables to model predictions and distinguishing global predictor importance from patient-specific effects. This is particularly relevant in OCD, where combinations of severe symptoms, poor insight, early onset, prolonged duration, depressive comorbidity, previous augmentation, and incomplete psychotherapy may interact in ways that are difficult to capture using conventional regression. An explainable deep learning model may therefore provide both predictive accuracy and clinically meaningful information regarding the profile of treatment resistance.

The increasing convergence of clinical psychiatry, neurobiology, neuromodulation, and artificial intelligence creates a strong rationale for developing data-driven tools capable of identifying high-risk OCD

patients earlier in the treatment pathway. Such prediction may support more efficient treatment planning, reduce prolonged exposure to ineffective interventions, facilitate earlier specialist referral, and provide a quantitative basis for individualized risk stratification. At the same time, predictive models based on routine clinical data may be more feasible for real-world implementation than approaches requiring expensive imaging or invasive biomarkers. Given the substantial burden of persistent OCD, the heterogeneity of treatment response, and the growing availability of computational methods capable of modeling complex multidimensional data, there is a clear need to evaluate whether deep neural networks can reliably distinguish patients who will develop treatment resistance from those who will respond to standard treatment.

Accordingly, the aim of this study was to develop and evaluate an explainable deep learning model for predicting treatment resistance in patients with obsessive–compulsive disorder using baseline demographic, clinical, psychopathological, and treatment-history characteristics.

Methods and Materials

Study Design and Participants

This study employed a retrospective longitudinal predictive design to develop and evaluate a deep learning model for identifying treatment resistance among patients with obsessive–compulsive disorder (OCD). The study population consisted of adults with a confirmed diagnosis of OCD who received psychiatric and psychological services at specialized psychiatric hospitals and outpatient mental health clinics affiliated with major medical universities in Tehran, Iran, between January 2021 and March 2026. Medical records were initially screened to identify patients who had undergone a sufficiently documented course of pharmacological treatment and for whom baseline clinical, demographic, symptom-related, and treatment variables were available. A total of 684 potentially eligible patient records were identified. After exclusion of records with substantial missing baseline information, uncertain diagnostic status, insufficient treatment duration, incomplete follow-up assessment, or major neurological and medical conditions that could substantially affect psychiatric evaluation, the final analytical sample comprised exactly 612 patients with OCD. Participants were aged 18–60 years and had OCD as their primary psychiatric diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria. The diagnosis documented in the clinical records was additionally required to have been established or confirmed through a structured diagnostic interview conducted by a psychiatrist or trained clinical psychologist.

Eligibility criteria included a primary diagnosis of OCD, age between 18 and 60 years, a baseline Yale–Brown Obsessive Compulsive Scale (Y-BOCS) total score of at least 16, availability of pretreatment clinical information, completion of at least one adequately documented pharmacological treatment course, and availability of treatment-response information sufficient to determine the final response status. Patients were excluded when they had a current or lifetime diagnosis of schizophrenia-spectrum or other psychotic disorders, bipolar I disorder, a severe substance use disorder requiring independent clinical intervention, intellectual disability, major neurocognitive disorder, epilepsy or another major neurological disorder, or a serious uncontrolled medical condition likely to interfere with psychiatric treatment or symptom evaluation. Records were also excluded when more than 20% of the candidate predictor variables were missing or when treatment adherence could not be adequately established from clinical documentation. Comorbid depressive

and anxiety disorders were not exclusion criteria because these conditions commonly occur in patients with OCD and were considered potentially informative predictors of treatment resistance.

The primary outcome was treatment resistance, operationalized on the basis of longitudinal treatment history and symptom change. Patients were classified as treatment resistant when they had failed to achieve a clinically meaningful response following at least two adequate trials of serotonin reuptake inhibitor treatment administered at therapeutic or maximally tolerated doses for an appropriate duration, with each trial lasting at least 12 weeks and including an adequate period at the target dose. An adequate treatment response was defined as a reduction of at least 35% in the Y-BOCS total score from pretreatment baseline together with a Clinical Global Impression–Improvement score indicating substantial clinical improvement. Patients who did not satisfy these response criteria despite two adequate treatment trials were assigned to the treatment-resistant group, whereas those demonstrating the specified improvement during adequately delivered treatment were classified as treatment responsive. The final outcome was coded dichotomously for predictive modeling. To prevent information leakage, only demographic, clinical, psychopathological, and treatment-history characteristics available before determination of the final resistance status were entered as predictors. The study protocol was conducted in accordance with the principles of the Declaration of Helsinki. All extracted data were de-identified before analysis, and access to patient information was restricted to the research team in accordance with institutional requirements for the secondary analysis of clinical records.

Data Collection Tools

The Structured Clinical Interview for DSM-5 Disorders–Clinician Version (SCID-5-CV) was used within the clinical diagnostic process to confirm OCD and identify major psychiatric comorbidities. Diagnostic information obtained from structured psychiatric assessments was reviewed to ensure that OCD constituted the primary disorder relevant to the treatment episode included in the study. The diagnostic dataset additionally incorporated the presence or absence of major depressive disorder, anxiety disorders, tic-related disorders, and other clinically relevant psychiatric comorbidities. Age at onset of OCD, duration of illness, family history of OCD or related psychiatric disorders, previous psychiatric hospitalization, number of previous treatment courses, history of psychotherapy, medication adherence, and the presence of chronic or episodic symptom patterns were obtained from standardized clinical records. Sociodemographic predictors included age, sex, marital status, educational level, employment status, and selected indicators of socioeconomic status where these variables were available in sufficiently consistent form across records.

The Yale–Brown Obsessive Compulsive Scale was used as the principal measure of OCD symptom severity. The Y-BOCS is a clinician-administered instrument consisting of 10 core severity items that assess the time occupied by obsessive thoughts and compulsive behaviors, associated interference, subjective distress, resistance, and perceived control. Five items assess obsessions and five assess compulsions, with individual items rated from 0 to 4 and total scores ranging from 0 to 40. Higher scores indicate greater OCD severity. Baseline Y-BOCS total scores and, where available, obsession and compulsion subscale scores were incorporated into the prediction dataset. Longitudinal Y-BOCS scores recorded after treatment were used exclusively to establish treatment-response status and were not included as baseline predictors in the deep learning model. Information concerning predominant OCD symptom dimensions, including contamination

and cleaning, symmetry and ordering, forbidden or intrusive thoughts, checking, and other repetitive behavioral or mental rituals, was also extracted when documented through structured assessment.

The Obsessive–Compulsive Inventory–Revised (OCI-R) was used as an additional measure of obsessive–compulsive symptom expression. The instrument includes 18 items covering major dimensions of obsessive–compulsive symptomatology, including washing, checking, ordering, obsessing, hoarding-related symptoms, and neutralizing. Higher scores indicate greater severity of obsessive–compulsive symptoms. OCI-R total and dimensional scores available at baseline were considered candidate predictors because variation in symptom configuration may be associated with differential treatment outcomes. Depressive symptom severity was assessed using the Beck Depression Inventory–II (BDI-II), a 21-item self-report measure assessing cognitive, affective, and somatic manifestations of depression. Anxiety severity was assessed using the Beck Anxiety Inventory (BAI), which contains 21 items measuring common subjective and somatic manifestations of anxiety. Total baseline scores from these instruments were included as potential predictors when available. In addition, clinical variables reflecting suicidal ideation, level of psychosocial impairment, insight into obsessive–compulsive symptoms, and psychiatric comorbidity were extracted from the clinical documentation when recorded using standardized procedures.

Medication-related information was systematically extracted from psychiatric treatment records. These variables included the number and type of previous serotonin reuptake inhibitor trials, use of selective serotonin reuptake inhibitors or clomipramine, maximum prescribed dosage, treatment duration, augmentation with antipsychotic medication, medication switching, reported adverse effects, treatment discontinuation, and documented adherence. Exposure to cognitive behavioral therapy, particularly exposure and response prevention, was recorded in terms of whether treatment had been received, approximate number of sessions, and completion status whenever sufficiently reliable information was available. The resulting candidate predictor set therefore represented multiple clinically meaningful domains, including demographic characteristics, illness course, OCD severity and dimensions, psychiatric comorbidity, depressive and anxiety symptoms, insight, previous treatment exposure, and pharmacological treatment characteristics. Variables directly generated after the treatment-response endpoint were excluded from the predictor set to preserve the temporal validity of the prediction model.

Data Analysis

Data analysis was performed using Python and standard machine-learning and deep-learning libraries. Prior to predictive modeling, the dataset underwent systematic preprocessing and quality control. Continuous variables were examined for implausible values and extreme observations, whereas categorical variables were evaluated for inconsistent coding and sparse categories. Patients with more than 20% missing candidate predictor information had already been excluded during eligibility screening. For the remaining dataset, missing continuous values were imputed using statistics estimated exclusively from the training data, and missing categorical observations were handled using the most appropriate category-based imputation strategy. All preprocessing transformations were estimated within the training dataset and subsequently applied without refitting to validation and test datasets in order to avoid data leakage. Continuous predictors were standardized to a mean of zero and a standard deviation of one, and categorical variables were transformed into machine-readable representations using appropriate binary or one-hot

encoding procedures. Highly redundant variables and variables with near-zero variance were removed before model development. Correlations among continuous predictors and clinically overlapping variables were examined to reduce unnecessary duplication while retaining variables with plausible prognostic significance.

The complete dataset of 612 patients was randomly partitioned at the patient level into training, validation, and independent test subsets using stratification according to treatment-resistance status. Seventy percent of the sample was allocated to the training set, corresponding to 428 patients, 15% to the validation set, corresponding to 92 patients, and the remaining 15% to the independent test set, corresponding to 92 patients. The test dataset was isolated before model optimization and was not used during variable selection, architecture development, hyperparameter tuning, threshold selection, or early stopping. Because treatment resistance was not expected to be equally distributed across outcome categories, class frequencies were evaluated within each subset. Class imbalance in the training data was addressed through class-weighted loss functions so that errors involving the less frequent outcome received proportionately greater weight during optimization. Resampling procedures, when examined during sensitivity analyses, were restricted exclusively to the training dataset.

A multilayer feed-forward deep neural network was developed to estimate the probability of treatment resistance. The input layer represented the complete set of preprocessed baseline predictors, followed by multiple fully connected hidden layers employing rectified linear unit activation functions to capture nonlinear relationships and higher-order interactions among demographic, clinical, symptom-related, and treatment-history characteristics. Batch normalization was incorporated to improve optimization stability, and dropout layers were introduced between dense layers to reduce overfitting. The final output layer consisted of a single neuron with a sigmoid activation function generating a probability between 0 and 1 for treatment resistance. Binary cross-entropy was used as the primary loss function, and network parameters were optimized using the Adam optimizer. Hyperparameters, including the number of hidden layers, number of neurons within each layer, dropout rate, batch size, learning rate, and regularization strength, were selected according to performance in the validation data. Early stopping was applied by monitoring validation loss, with the model weights corresponding to the best validation performance restored before final testing.

Predictive performance was evaluated in the independent test sample using the area under the receiver operating characteristic curve as the principal discrimination measure. Accuracy, sensitivity, specificity, positive predictive value, negative predictive value, F1 score, and the area under the precision–recall curve were additionally calculated to provide a comprehensive assessment of classification performance. Because clinically useful prediction requires not only discrimination but also correspondence between estimated and observed probabilities, model calibration was evaluated using the Brier score and calibration plots. Ninety-five percent confidence intervals for major performance indices were estimated using bootstrap resampling of the independent test set. The optimal classification threshold was determined using the validation data rather than the test data and was subsequently fixed before final evaluation.

To determine whether the additional complexity of deep learning produced clinically meaningful predictive improvement, the performance of the final neural network was compared with conventional and machine-learning benchmark models, including logistic regression, random forest, and gradient-boosting

classifiers, using the same training and test partitions and the same baseline predictor set. Model interpretability was examined using SHapley Additive exPlanations (SHAP), which quantified the contribution of individual predictors to the model's estimated probability of treatment resistance. Global SHAP importance values were used to rank the predictors that contributed most strongly to the overall model, while patient-level SHAP values were inspected to clarify how particular combinations of clinical characteristics influenced individual predictions. This procedure made it possible to evaluate nonlinear effects and interactions while retaining a clinically interpretable representation of the deep learning model. Sensitivity analyses were conducted to examine the stability of predictive performance under alternative class-weighting procedures and model specifications. All statistical tests used a two-sided significance level of .05 where inferential comparisons were performed, and predictive performance was interpreted primarily according to effect magnitude, confidence intervals, discrimination, calibration, and out-of-sample generalizability rather than statistical significance alone.

Findings and Results

The final analytical sample consisted of 612 patients with obsessive–compulsive disorder, of whom 246 (40.2%) met the predefined criteria for treatment resistance and 366 (59.8%) were classified as treatment responsive. The mean age of the total sample was 34.72 years (SD = 9.61; range = 18–60 years). A total of 322 participants (52.6%) were women and 290 (47.4%) were men. Regarding marital status, 331 participants (54.1%) were married, 255 (41.7%) were single, and 26 (4.2%) were divorced or widowed. In terms of educational attainment, 96 patients (15.7%) had completed secondary education or less, 214 (35.0%) held a diploma or associate degree, 225 (36.8%) held a bachelor's degree, and 77 (12.6%) had postgraduate education. Overall, 387 participants (63.2%) were employed or studying at the time of baseline assessment, whereas 225 (36.8%) were unemployed, homemakers, retired, or otherwise economically inactive. The mean age was 35.11 ± 9.48 years in the treatment-resistant group and 34.46 ± 9.70 years in the treatment-responsive group, and this difference was not statistically significant (p = .416). Women constituted 53.7% of the treatment-resistant group and 51.9% of the treatment-responsive group, with no significant difference in sex distribution between the two outcome groups (p = .659). Likewise, marital status, education, and employment status did not differ significantly between groups (all p > .05). These findings indicated that the subsequent predictive differences were not attributable to major imbalances in basic sociodemographic composition. Stratified allocation of the sample resulted in 428 participants in the training set, 92 in the validation set, and 92 in the independent test set. The proportions of treatment-resistant patients were 40.2%, 40.2%, and 40.2% in the training, validation, and test datasets, respectively, demonstrating successful preservation of the outcome distribution across the three analytical subsets.

Table 1. Baseline Clinical and Treatment Characteristics According to Treatment-Response Status

Clinical characteristic	Total sample (N = 612)	Treatment-resistant (n = 246)	Treatment-responsive (n = 366)	p value	Effect size
Age at OCD onset, years, M ± SD	23.68 ± 8.02	21.39 ± 7.54	25.22 ± 7.98	<.001	d = 0.49

Duration of OCD, years, M ± SD	9.24 ± 7.06	11.71 ± 7.63	7.58 ± 6.11	<.001	d = 0.61
Baseline Y-BOCS total score, M ± SD	24.99 ± 5.92	27.79 ± 5.46	23.10 ± 5.42	<.001	d = 0.86
OCI-R total score, M ± SD	29.23 ± 9.47	32.71 ± 9.28	26.89 ± 8.91	<.001	d = 0.64
BDI-II total score, M ± SD	17.85 ± 9.62	21.84 ± 9.76	15.17 ± 8.67	<.001	d = 0.73
BAI total score, M ± SD	16.81 ± 9.14	18.93 ± 9.21	15.38 ± 8.84	<.001	d = 0.39
Poor or absent insight, n (%)	132 (21.6)	78 (31.7)	54 (14.8)	<.001	φ = .20
Major depressive disorder, n (%)	200 (32.7)	107 (43.5)	93 (25.4)	<.001	φ = .19
Comorbid anxiety disorder, n (%)	228 (37.3)	105 (42.7)	123 (33.6)	.023	φ = .09
Tic-related disorder, n (%)	61 (10.0)	35 (14.2)	26 (7.1)	.004	φ = .12
Positive family history of OCD, n (%)	181 (29.6)	89 (36.2)	92 (25.1)	.003	φ = .12
Previous psychiatric hospitalization, n (%)	76 (12.4)	45 (18.3)	31 (8.5)	<.001	φ = .15
Previous antipsychotic augmentation, n (%)	171 (27.9)	96 (39.0)	75 (20.5)	<.001	φ = .20
Previous structured ERP treatment, n (%)	301 (49.2)	107 (43.5)	194 (53.0)	.021	φ = .09
Completed an adequate ERP course, n (%)	293 (47.9)	88 (35.8)	205 (56.0)	<.001	φ = .20
Documented medication adherence ≥80%, n (%)	469 (76.6)	176 (71.5)	293 (80.1)	.014	φ = .10

As shown in Table 1, treatment-resistant patients exhibited a substantially more severe and clinically complex baseline profile than treatment-responsive patients. The largest between-group difference was observed for baseline obsessive–compulsive symptom severity, with the resistant group recording a mean Y-BOCS score of 27.79 compared with 23.10 in the responsive group ($p < .001$, $d = 0.86$), representing a large standardized effect. Treatment-resistant patients also had a significantly earlier mean age of OCD onset and a substantially longer mean duration of illness, indicating that greater chronicity and earlier emergence of the disorder were associated with subsequent nonresponse. Significant differences were also observed for broader psychopathology. Mean OCI-R, BDI-II, and BAI scores were consistently higher among resistant patients, with depressive symptom severity showing a moderate-to-large difference between groups. Poor or absent insight was present in almost one-third of treatment-resistant patients compared with approximately 15% of responsive patients, while major depressive disorder was observed in 43.5% and 25.4% of the two groups, respectively. Tic-related disorders, positive family history of OCD, and previous psychiatric hospitalization were also significantly more frequent in resistant patients. Treatment history further differentiated the outcome groups. Previous antipsychotic augmentation was nearly twice as frequent among treatment-resistant patients, whereas completion of an adequate course of exposure and response prevention was substantially less common in the resistant group. Documented medication adherence of at least 80% was also lower among resistant patients. Collectively, these baseline findings suggested that treatment resistance was associated with a multidimensional pattern involving greater OCD severity, earlier and more chronic illness, more severe depressive and anxiety symptoms, poorer insight, greater psychiatric complexity, and a more complicated prior treatment trajectory rather than with any single isolated clinical characteristic.

Table 2. Performance of the Final Deep Neural Network Across Training, Validation, and Independent Test Datasets

Performance index	Training set (n = 428)	Validation set (n = 92)	Independent test set (n = 92)
AUROC	.941	.901	.891
AUPRC	.913	.861	.844

Accuracy	.876	.837	.826
Sensitivity	.884	.811	.811
Specificity	.870	.855	.836
Positive predictive value	.820	.789	.769
Negative predictive value	.919	.870	.868
F1 score	.851	.800	.789
Brier score	.096	.128	.139
Classification threshold	.430	.430	.430

The final deep neural network demonstrated strong predictive discrimination and maintained relatively stable performance when transferred from development data to previously unseen patients. As reported in Table 2, the model achieved an AUROC of .941 in the training sample, .901 in the validation sample, and .891 in the independent test sample. The modest reduction from training to validation and test performance indicated some expected decrease in discrimination outside the development data but did not suggest substantial overfitting. The area under the precision–recall curve was .844 in the test sample, indicating that the model retained favorable performance even when evaluated in relation to the less prevalent treatment-resistant class. At the probability threshold of .430 selected exclusively from validation data, test-set accuracy was 82.6%. Sensitivity was 81.1%, indicating that approximately four out of five genuinely treatment-resistant patients were correctly identified by the model, while specificity was 83.6%, demonstrating a similarly high ability to correctly identify treatment-responsive patients. Positive predictive value was 76.9%, negative predictive value was 86.8%, and the F1 score was .789, indicating an appropriate balance between detection of resistant patients and avoidance of excessive false-positive classifications. In the 92-patient independent test dataset, which contained 37 treatment-resistant and 55 treatment-responsive patients, the network correctly classified 30 of the 37 resistant patients and 46 of the 55 responsive patients. Accordingly, there were seven false-negative and nine false-positive predictions. The test-set Brier score of .139 further indicated acceptable probabilistic accuracy, suggesting that the output probabilities provided meaningful information beyond binary classification. The relatively narrow differences between validation and independent test indices supported the stability and generalizability of the fitted network.

Table 3. Comparison of the Deep Neural Network With Benchmark Prediction Models in the Independent Test Set

Model	AUROC (95% CI)	AUPRC	Accuracy	Sensitivity	Specificity	F1 score	Brier score
Deep neural network	.891 (.821–.951)	.844	.826	.811	.836	.789	.139
Gradient boosting	.862 (.782–.930)	.808	.793	.784	.800	.763	.154
Random forest	.836 (.748–.909)	.774	.772	.757	.782	.737	.169
Logistic regression	.797 (.699–.879)	.721	.728	.703	.745	.684	.188

Comparison with benchmark algorithms demonstrated that the deep learning model provided the strongest overall prediction of treatment resistance. As shown in Table 3, the final neural network obtained the highest AUROC (.891), AUPRC (.844), accuracy (.826), sensitivity (.811), specificity (.836), and F1 score (.789), while also producing the lowest Brier score (.139). Gradient boosting was the strongest comparator, reaching an AUROC of .862 and an accuracy of 79.3%, but its performance remained below that of the neural network for all principal discrimination and calibration indices. Random forest produced an AUROC of .836, whereas conventional multivariable logistic regression achieved the weakest discrimination with an AUROC

of .797. The absolute AUROC improvement of the deep neural network over logistic regression was .094, suggesting that modeling nonlinear associations and interactions among clinical variables provided appreciable additional predictive information beyond a conventional linear probability model. Bootstrap comparison of discrimination indicated that the deep neural network significantly outperformed logistic regression (p = .012) and random forest (p = .041). The difference between the neural network and gradient boosting was smaller and did not reach conventional statistical significance (p = .174), although the deep network retained numerically superior discrimination, classification accuracy, and calibration. The Brier score pattern was particularly consistent, progressively improving from .188 for logistic regression to .169 for random forest, .154 for gradient boosting, and .139 for the neural network. These findings indicated that the advantage of deep learning was not restricted to ranking patients according to risk but also extended to the accuracy of the predicted treatment-resistance probabilities.

Table 4. Most Influential Predictors of Treatment Resistance According to SHAP Analysis

Rank	Predictor	Mean absolute SHAP value	Predominant relationship with predicted treatment resistance
1	Baseline Y-BOCS total score	.184	Higher scores increased predicted risk
2	Duration of OCD	.139	Longer duration increased predicted risk
3	Poor or absent insight	.118	Poorer insight increased predicted risk
4	BDI-II total score	.104	Greater depressive severity increased predicted risk
5	Age at OCD onset	.091	Earlier onset increased predicted risk
6	Previous antipsychotic augmentation	.079	Previous augmentation increased predicted risk
7	Completion of adequate ERP	.073	Completion decreased predicted risk
8	Major depressive disorder	.066	Presence increased predicted risk
9	OCI-R total score	.058	Higher symptom burden increased predicted risk
10	BAI total score	.049	Greater anxiety severity increased predicted risk
11	Previous psychiatric hospitalization	.044	Previous hospitalization increased predicted risk
12	Documented medication adherence ≥80%	.039	Higher adherence decreased predicted risk
13	Positive family history of OCD	.034	Positive history modestly increased predicted risk
14	Tic-related disorder	.030	Presence modestly increased predicted risk
15	Comorbid anxiety disorder	.027	Presence modestly increased predicted risk

SHAP analysis demonstrated that the predictions generated by the deep neural network were driven primarily by clinically interpretable indicators of OCD severity, chronicity, insight, affective comorbidity, and previous treatment history. Baseline Y-BOCS severity was the most influential predictor, with a mean absolute SHAP value of .184. Higher baseline scores consistently shifted predictions toward treatment resistance, although the magnitude of this effect increased nonlinearly at the upper end of the severity distribution. Duration of OCD was the second most influential predictor, indicating that progressively longer illness duration contributed substantially to predicted nonresponse. Poor or absent insight ranked third, followed by depressive symptom severity measured using the BDI-II. Earlier age at onset also contributed meaningfully to resistance predictions, particularly when early onset co-occurred with long illness duration and high baseline symptom severity. Previous antipsychotic augmentation was associated with increased predicted resistance, most likely reflecting a more difficult prior treatment course, whereas completion of an adequate exposure and response prevention program shifted predictions toward treatment responsiveness and therefore functioned as a protective predictor within the model. Major depressive disorder, higher OCI-R and BAI scores, previous psychiatric hospitalization, positive family history, and tic-related disorder contributed additional predictive information, although their average effects were smaller

than those of the leading variables. Higher medication adherence was generally associated with lower predicted probability of resistance. Inspection of individual-level SHAP profiles further showed that risk was rarely determined by one variable alone. The highest predicted probabilities were typically generated by combinations of severe baseline OCD symptoms, prolonged illness, poor insight, high depressive symptom burden, early onset, and evidence of previous treatment complexity, illustrating the capacity of the network to integrate interacting clinical characteristics into individualized estimates of treatment resistance.

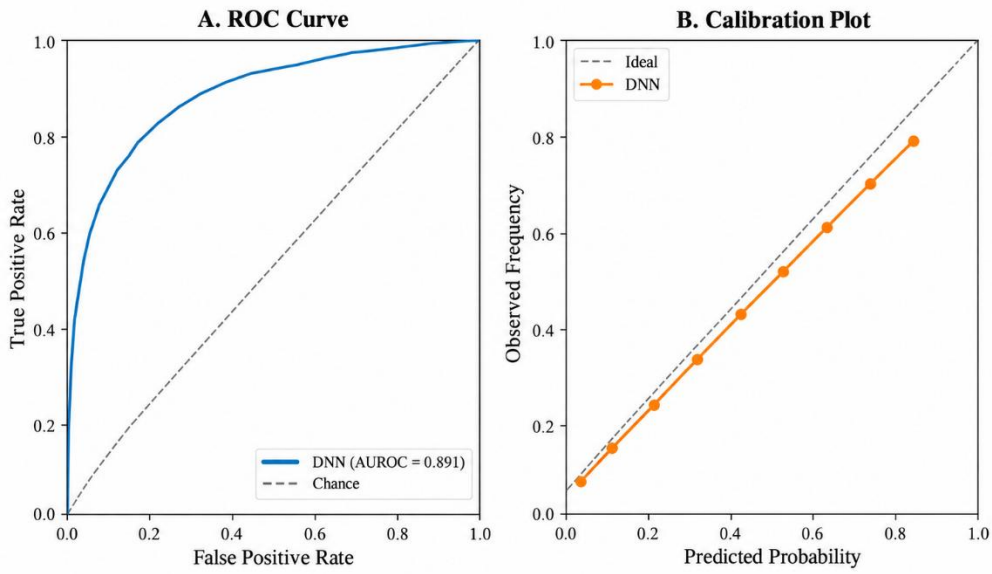


Figure 1. Receiver Operating Characteristic and Calibration Curves of the Final Deep Neural Network for Predicting Treatment Resistance in the Independent Test Sample

Figure 1 demonstrates the discrimination and calibration characteristics of the final deep neural network in the independent test dataset. The receiver operating characteristic curve remained consistently above the chance diagonal over the clinically relevant range of classification thresholds and yielded an AUROC of .891, confirming strong separation between treatment-resistant and treatment-responsive patients. The curve indicated that clinically useful sensitivity could be achieved without a disproportionate decline in specificity, consistent with the sensitivity of 81.1% and specificity of 83.6% observed at the validation-derived classification threshold of .430. The calibration component of the figure showed close correspondence between predicted and observed treatment-resistance probabilities across most of the probability range. Predictions in the low- and intermediate-risk strata were particularly well calibrated, while mild overestimation of resistance was observed among patients assigned the highest predicted probabilities. Nevertheless, the overall calibration pattern remained close to the ideal reference line and was consistent with the test-set Brier score of .139. Taken together, the discrimination and calibration results indicate that the network was capable not only of distinguishing patients who subsequently demonstrated treatment resistance but also of providing clinically meaningful estimates of the magnitude of individual risk. The combination of high AUROC, balanced sensitivity and specificity, favorable negative predictive value, and acceptable calibration supports the potential utility of the model as a risk-stratification instrument for identifying patients who may require closer monitoring, earlier optimization of treatment, or more intensive multimodal intervention.

Discussion and Conclusion

The present study developed and evaluated an explainable deep learning model for predicting treatment resistance in obsessive–compulsive disorder using baseline demographic, clinical, psychopathological, and treatment-history characteristics. The principal findings showed that treatment-resistant patients had a more severe and clinically complex profile than treatment-responsive patients, particularly in relation to baseline OCD severity, illness chronicity, earlier age at onset, poorer insight, depressive and anxiety symptoms, psychiatric comorbidity, previous hospitalization, and prior treatment complexity. The deep neural network demonstrated strong predictive discrimination, with an AUROC of .891 in the independent test sample, an accuracy of 82.6%, sensitivity of 81.1%, specificity of 83.6%, and an F1 score of .789. The model also showed acceptable calibration, with a Brier score of .139, and outperformed logistic regression, random forest, and gradient boosting across most performance indicators. SHAP analysis further showed that baseline Y-BOCS score, duration of OCD, poor or absent insight, depressive severity, age at onset, previous antipsychotic augmentation, completion of adequate exposure and response prevention, major depressive disorder, and overall obsessive–compulsive symptom burden were the most influential predictors. Taken together, the results suggest that treatment resistance in OCD is not determined by a single clinical characteristic but instead emerges from a multidimensional interaction among symptom severity, chronicity, insight, affective burden, comorbidity, and treatment history.

The finding that greater baseline OCD severity was the strongest predictor of treatment resistance is consistent with the broader clinical literature on severe and refractory OCD. Patients who enter treatment with more intense obsessions and compulsions generally require more prolonged, specialized, and multimodal intervention, and those with severe illness are disproportionately represented in tertiary services and intensive inpatient programs (7, 8). The large difference in baseline Y-BOCS scores between resistant and responsive patients in the present study indicates that initial symptom burden may be a clinically meaningful marker of later nonresponse. This finding is also compatible with work on pharmacoresistant OCD showing that conventional interventions can be less effective when severe symptoms are accompanied by entrenched behavioral patterns, poor insight, and multiple previous treatment failures (6). Although cognitive-behavioral therapy remains one of the most effective interventions for OCD, severe presentations often require greater treatment intensity, more carefully structured exposure, augmentation strategies, or alternative therapeutic approaches (4, 5). Thus, the strong importance assigned to baseline Y-BOCS severity by the neural network is clinically plausible and reinforces the relevance of early symptom quantification when evaluating the likely trajectory of treatment response.

Longer illness duration and earlier age at onset were also among the most influential predictors of treatment resistance. This pattern is consistent with the view that prolonged illness may be associated with greater behavioral consolidation, stronger avoidance and ritualization, more extensive functional impairment, and reduced flexibility in responding to conventional interventions. Early-onset OCD may also reflect a more developmentally entrenched or biologically loaded form of the disorder, particularly when symptoms emerge during childhood or adolescence and persist into adulthood (1). Genetic and familial research also supports the possibility that early and persistent forms of OCD may be associated with stronger hereditary vulnerability and complex neurobiological mechanisms (13). Similarly, research on early life stress suggests that developmental adversity may alter neurobiological systems implicated in compulsivity,

emotional regulation, and stress responsivity, potentially increasing the risk of a chronic and difficult-to-treat course (14). From a predictive perspective, the combined importance of age at onset and illness duration suggests that temporal characteristics of the disorder should not be treated as secondary background variables; rather, they may capture clinically significant information regarding disease persistence and cumulative treatment burden.

Poor or absent insight emerged as the third most influential predictor in the model and was considerably more common among treatment-resistant patients. This result is theoretically consistent with cognitive models of OCD in which dysfunctional appraisals, certainty seeking, rigid beliefs, and excessive confidence in obsession-related inferences sustain compulsive behavior. Thought–action fusion and thought suppression have long been associated with the persistence of obsessive–compulsive symptoms, illustrating how maladaptive cognitive processes can strengthen symptom maintenance (3). Poor insight may reduce willingness to challenge obsessive beliefs, interfere with exposure-based learning, and increase reliance on compulsive neutralization. The relevance of insight is also compatible with the inference-based approach, which specifically targets faulty reasoning processes that contribute to obsessional doubt and has shown usefulness across symptom subtypes and in some treatment-resistant patients (9). Accordingly, the strong predictive role of insight in the present model may reflect the extent to which patients can recognize the excessive or unreasonable nature of their symptoms and engage flexibly with corrective therapeutic experiences.

Depressive symptom severity and major depressive disorder were also important predictors of treatment resistance. Patients in the resistant group had substantially higher BDI-II scores and a greater prevalence of depressive comorbidity, indicating that affective burden may complicate the treatment course of OCD. Depression may influence outcome through several pathways, including reduced motivation, diminished behavioral activation, hopelessness, lower adherence, impaired concentration, and reduced ability to engage consistently in exposure and response prevention. Depression may also represent greater overall psychiatric severity rather than a separate additive condition. The broader literature on difficult-to-treat OCD has repeatedly emphasized the importance of complex comorbidity and the need for more individualized treatment planning when standard interventions are insufficient (5, 12). The present findings further support this perspective by showing that depression was not merely associated with resistance at the group level but contributed directly to the model's individualized risk estimates. The observed role of anxiety severity may similarly reflect a more generalized threat-processing burden that amplifies distress and avoidance, although its predictive contribution was smaller than that of depression.

The model also identified previous antipsychotic augmentation and psychiatric hospitalization as predictors of greater treatment resistance. These variables are best interpreted as markers of prior illness severity and treatment complexity rather than as evidence that these interventions themselves produce resistance. Patients who have already required augmentation or hospitalization are likely to represent a subgroup with more severe, persistent, or functionally impairing illness. This interpretation is supported by studies of specialized and tertiary OCD care, in which treatment-resistant patients often have extensive histories of failed pharmacological and psychological interventions before entering higher-intensity services (7, 8). At the more severe end of the treatment spectrum, deep brain stimulation and other neurosurgical or neuromodulatory approaches are generally considered only after multiple evidence-based treatments have

failed (23, 27). The increasing development of precision neurosurgical and connectomic approaches further reflects the clinical need to distinguish ordinary nonresponse from highly refractory illness requiring advanced intervention (24, 25). The present model may therefore have practical value in recognizing patients who are moving toward this more complex treatment trajectory before numerous sequential failures occur.

In contrast, completion of an adequate exposure and response prevention course and better medication adherence were associated with a lower probability of treatment resistance. This finding reinforces the importance of distinguishing true treatment resistance from apparent nonresponse caused by inadequate dose, insufficient treatment duration, poor adherence, or incomplete behavioral therapy. Cognitive-behavioral therapy, particularly exposure and response prevention, remains a core evidence-based intervention for OCD, and its effectiveness depends on sufficient intensity, engagement, and treatment fidelity (4, 6). Alternative psychological approaches have also shown promise in patients who do not respond fully to conventional therapy, including acceptance and commitment-based approaches, inference-based interventions, and schema-focused programs (9-11). The present findings therefore support a stepped-care interpretation of resistance: patients should not be regarded as genuinely refractory until the adequacy of previous evidence-based treatment has been established carefully.

The superior performance of the deep neural network compared with logistic regression, random forest, and gradient boosting is also noteworthy. The difference between the neural network and logistic regression suggests that treatment resistance may be partly determined by nonlinear and interacting relationships that are not readily captured by conventional linear models. In OCD, the prognostic meaning of one variable may depend on the presence of others; for example, severe baseline symptoms may be especially problematic when combined with long illness duration, poor insight, depressive comorbidity, and repeated previous treatment failure. Deep learning architectures are well suited to detecting such higher-order interactions, and recent work has highlighted the potential of deep learning within precision psychiatry for individualized prediction in OCD (22). The present findings are also consistent with evidence that computational models can successfully extract complex information from neuroimaging and EEG data in OCD (20, 21). Moreover, machine-learning approaches have increasingly been used to characterize psychiatric heterogeneity and comorbidity through multidimensional patterns rather than isolated predictors (19). The current results extend this logic to clinically defined treatment resistance using routinely available clinical variables.

The explainability findings are especially important because predictive accuracy alone is insufficient for clinical implementation. SHAP analysis showed that the network's predictions were driven by clinically interpretable variables that are already recognized as relevant to OCD severity and chronicity. This enhances confidence that the model was not relying primarily on obscure or spurious statistical artifacts. The prominence of symptom severity, chronicity, insight, and depression is compatible with current neurobiological models that conceptualize OCD as a heterogeneous disorder involving multiple interacting neural circuits rather than a single uniform pathway (15). Patient-specific neural signatures of obsessions and compulsions identified in ventral striatal recordings further support the idea that clinically similar patients may differ substantially in the mechanisms maintaining their symptoms (18). Predictive-processing accounts of repetitive behavior and biophysical models involving dopaminergic reward and working-memory systems likewise suggest that compulsivity emerges from distributed and interacting mechanisms rather

than a single deficit (16, 17). The multidimensional predictive pattern identified in the present study is therefore broadly aligned with contemporary mechanistic views of OCD.

The relevance of individualized prediction is further underscored by emerging neurobiological and neuromodulatory interventions. Resting-state network changes associated with mindfulness-based cognitive therapy in treatment-resistant OCD suggest that refractory symptoms may remain biologically modifiable and that treatment resistance should not be viewed as a fixed or irreversible patient trait (33). Experimental neuromodulation approaches, including temporal interference stimulation of the nucleus accumbens, represent additional efforts to target dysfunctional circuits in severe OCD (29). Preclinical research using transgenic animal models is likewise advancing understanding of molecular and circuit-level mechanisms and may eventually inform more individualized treatment selection (30). These developments make early and accurate risk stratification increasingly relevant. A patient identified as being at high risk of resistance might benefit from more intensive monitoring, earlier optimization of pharmacotherapy, specialist psychotherapy, or referral to tertiary services before years of ineffective sequential treatment accumulate.

The findings should also be interpreted in light of the broader clinical complexity of OCD. Chronic pain and other comorbid somatic conditions may contribute to functional impairment and increase the complexity of treatment planning (31). Some patients may present with unusual somatic preoccupations or phenomenological overlap with related psychiatric conditions, requiring careful differential diagnosis (32). Such complexity may partly explain why treatment outcome is difficult to predict through symptom severity alone. The present model's ability to combine multiple clinical domains is therefore one of its principal strengths. Rather than treating resistance as the product of a single adverse factor, the model generated individualized risk estimates from combinations of illness history, symptom severity, insight, comorbidity, and previous treatment exposure.

Several limitations should be considered when interpreting these findings. First, the study was retrospective and relied on clinical records, which may contain variability in documentation quality, treatment adherence estimates, psychotherapy exposure, and the timing of symptom assessments. Second, all patients were recruited from psychiatric services in Tehran, which may limit generalizability to other regions, healthcare systems, cultural contexts, and community-based populations. Third, although the sample of 612 patients was adequate for the present model, larger multicenter datasets would provide greater stability for deep learning and allow more extensive subgroup analyses. Fourth, some potentially relevant predictors, including detailed neurocognitive measures, genetic markers, neuroimaging features, digital phenotyping variables, therapist characteristics, and socioeconomic stressors, were not available consistently and therefore could not be included. Fifth, the treatment-resistance outcome was dichotomized, which may obscure clinically important gradations of partial response, remission, relapse, and time to response. Finally, although an independent internal test set was used, external validation in a completely separate sample was not performed.

Future research should validate the model prospectively in larger and more diverse multicenter cohorts and should examine whether predictive performance remains stable across different treatment settings and patient populations. Studies should also compare purely clinical models with multimodal models integrating neuroimaging, EEG, genetics, neurocognitive measures, digital behavior, and ecological data. Longitudinal prediction frameworks may be especially valuable because treatment resistance evolves over time and may

be better represented as a dynamic trajectory than as a fixed binary outcome. Future work should also examine separate prediction models for pharmacological resistance, psychotherapy resistance, partial response, remission, and relapse. Greater attention should be given to external validation, model recalibration, fairness across demographic subgroups, and clinically meaningful decision thresholds. In addition, prospective impact studies are needed to determine whether model-assisted care actually improves outcomes, reduces ineffective treatment cycles, or accelerates referral to specialized services.

In clinical practice, the findings suggest that systematic assessment of baseline OCD severity, illness duration, age at onset, insight, depressive symptoms, comorbidity, previous augmentation, psychotherapy completion, and medication adherence may help identify patients at elevated risk for treatment resistance. A predictive model of this type should be used as a decision-support instrument rather than as a substitute for clinical judgment. Patients classified as high risk could receive closer follow-up, earlier review of treatment adequacy, stronger adherence support, more intensive exposure and response prevention, and timely multidisciplinary consultation. Clinicians should also distinguish true resistance from pseudo-resistance caused by insufficient dose, inadequate duration, incomplete psychotherapy, or poor adherence. Integrating interpretable predictive tools into routine psychiatric assessment may ultimately support more individualized treatment planning and reduce the delay between recognition of a difficult clinical course and delivery of more intensive or specialized care.

Acknowledgments

The authors express their deep gratitude to all participants who contributed to this study.

Authors' Contributions

All authors equally contributed to this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants. This article was derived from the first author's doctoral dissertation in psychology at the Central Tehran Branch of Islamic Azad University. The study received ethical approval under code IR.IAU.CTB.REC.1404.191.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

References

1. Krebs G, Heyman I. Obsessive-compulsive disorder in children and adolescents. Arch Dis Child. 2015;100(5):495-9. Epub 20141114. doi: 10.1136/archdischild-2014-306934. PubMed PMID: 25398447; PubMed Central PMCID: PMC4413836.
2. Sylvia MI, Vidya N. Obsessive-Compulsive Disorder and Marital Relationship Harmony – A Qualitative Analysis. International Journal of Health Sciences and Pharmacy. 2023:1-14. doi: 10.47992/ijhsp.2581.6411.0105.
3. Rassin E, Muris P, Schmidt H, Merckelbach H. Relationships between thought–action fusion, thought suppression and obsessive–compulsive symptoms: a structural equation modeling approach. Behaviour Research and Therapy. 2000;38(9):889-97. doi: 10.1016/S0005-7967(99)00104-7.
4. Vieira FFP. Cognitive-Behavioral Therapy for Obsessive-Compulsive Disorder: A Comprehensive Synthesis of the Evidence. 2025. doi: 10.21203/rs.3.rs-7835365/v1.
5. Roh D, Jang K-W, Kim CH. Clinical Advances in Treatment Strategies for Obsessive-Compulsive Disorder in Adults. Clinical Psychopharmacology and Neuroscience. 2023;21(4):676-85. doi: 10.9758/cpn.23.1075.
6. Vyskocilova J, Prasko J, Sipek J. Cognitive behavioral therapy in pharmacoresistant obsessive-compulsive disorder. Neuropsychiatric Disease and Treatment. 2016;12:625-39. doi: 10.2147/NDT.S101721.
7. Pilunthanakul T, Yim J, Jayaraman H, Lau BJ, Ang D, Wong S, et al. Effectiveness of a tertiary specialist obsessive-compulsive disorder (OCD) service in improving clinical outcomes for patients with treatment-resistant/complex OCD. Singapore Medical Journal. 2024. doi: 10.4103/singaporemedj.SMJ-2024-033.
8. Balachander S, Bajaj A, Hazari N, Kumar A, Anand N, Manjula M, et al. Long-term outcomes of intensive inpatient care for severe, resistant obsessive-compulsive disorder. The Canadian Journal of Psychiatry. 2020;65(11):779-89. doi: 10.1177/0706743720927830.
9. Aardema F, O'Connor KP, Delorme ME, Audet JS. The inference-based approach (IBA) to the treatment of obsessive-compulsive disorder: An open trial across symptom subtypes and treatment-resistant cases. Clinical Psychology & Psychotherapy. 2017;24(2):289-301. doi: 10.1002/cpp.2024.
10. Izadi R, Abedi MR. Alleviation of obsessive symptoms in treatment-resistant obsessive-compulsive disorder using acceptance and commitment-based therapy. Feyz Medical Sciences Journal. 2013;17(3):275-86.
11. Remmerswaal KCP, Cnossen TEA, van Oppen P, van Balkom AJLM. Schema Therapy with Cognitive Behaviour Day-Treatment in Patients with Treatment-Resistant Anxiety Disorders and Obsessive-Compulsive Disorder: An Uncontrolled Pilot Study. Behavioural and Cognitive Psychotherapy. 2023;51(2):174-9. doi: 10.1017/S1352465822000625.
12. Coughtrey AE, Melli G. Beyond Exposure: Innovations in the Treatment of Obsessive–Compulsive Disorder. Journal of Clinical Psychology. 2026;82(5):647-50. doi: 10.1002/jclp.70099.
13. Dhiman A, Mehan S, Khan Z, Tiwari A, Gupta GD, Narula AS. Hereditary Patterns and Genetic Associations in Obsessive-Compulsive Disorder (OCD): Neuropsychiatric Insights, Genetic Influences, and Treatment Perspectives. Current Gene Therapy. 2025;25(3):257-316. doi: 10.2174/0115665232316708240828063527.
14. Debnath A, Sarkar S. Impact of Early Life Stress on Neurobiological Vulnerability to Obsessive-Compulsive Disorder: Early Life Stress &Amp; Neurobiological OCD Risk. Brain Science Advances. 2024;10(2). doi: 10.26599/bsa.2024.905005.
15. Shephard E, Batistuzzo MC, Hoexter MQ, Stern ER, Zuccolo PF, Ogawa CY, et al. Neurocircuit Models of Obsessive-Compulsive Disorder: Limitations and Future Directions for Research. Brazilian Journal of Psychiatry. 2022;44(2):187-200. doi: 10.1590/1516-4446-2020-1709.
16. Spee BTM, Sladky R, Fingerhut J, Laciny A, Kraus C, Carls-Diamante S, et al. Repeating Patterns: Predictive Processing Suggests an Aesthetic Learning Role of the Basal Ganglia in Repetitive Stereotyped Behaviors. Frontiers in Psychology. 2022;13. doi: 10.3389/fpsyg.2022.930293.
17. Yin L, Han F, Wang Q. A Biophysical Model for Dopamine Modulating Working Memory Through Reward System in Obsessive–compulsive Disorder. Cognitive Neurodynamics. 2023;18(4):1895-911. doi: 10.1007/s11571-023-09999-z.
18. Friðgeirsson EA, Bais MN, Eijsker N, Thomas RM, Smit DJA, Bergfeld IO, et al. Patient Specific Intracranial Neural Signatures of Obsessions and Compulsions in the Ventral Striatum. Journal of Neural Engineering. 2023;20(2):026008. doi: 10.1088/1741-2552/acbee1.

19. Lv Q, Zeljic K, Zhao S, Zhang J, Zhang J, Wang Z. Dissecting Psychiatric Heterogeneity and Comorbidity With Core Region-Based Machine Learning. Neuroscience Bulletin. 2023;39(8):1309-26. doi: 10.1007/s12264-023-01057-2.

20. Li X, Gu H. A Comprehensive Review for Machine Learning on Neuroimaging in Obsessive-Compulsive Disorder. Frontiers in Human Neuroscience. 2023;17. doi: 10.3389/fnhum.2023.1280512.

21. Farhad S, Metin SZ, Uyulan Ç, Makouei STZ, Metin B, Erguzel TT, et al. Application of Hybrid DeepLearning Architectures for Identification of Individuals With Obsessive Compulsive Disorder Based on EEG Data. Clinical Eeg and Neuroscience. 2024;55(5):543-52. doi: 10.1177/15500594231222980.

22. Zaboski BA, Bednarek L. Precision Psychiatry for Obsessive-Compulsive Disorder: Clinical Applications of Deep Learning Architectures. Journal of Clinical Medicine. 2025;14(7):2442. doi: 10.3390/jcm14072442.

23. Vicheva P, Butler M, Shotbolt P. Deep brain stimulation for obsessive-compulsive disorder: A systematic review of randomised controlled trials. Neuroscience & Biobehavioral Reviews. 2020;109:129-38. doi: 10.1016/j.neubiorev.2020.01.007.

24. Torres CV, Almunia ML, Granados GO, García MN, Pulido P, Pedro M, et al. Complications and Ethical Challenges in Neurosurgery for Psychiatric Disorders. Brain Sciences. 2025;15(12):1303. doi: 10.3390/brainsci15121303.

25. Torres CV, García MN, Pulido P, Almunia ML, Alén JF. Shaping the Future of Psychiatric Neurosurgery: From Connectomic Precision to Technological Integration. Brain Sciences. 2025;15(6):647. doi: 10.3390/brainsci15060647.

26. Sakthiyendran NA, Padala VJ, Seide M, See JW, Sabu N, Sharma A, et al. Past and Present Role of Neurosurgical Interventions in the Management of Psychiatric Disorders: A Literature Review on the Evolution of Psychosurgery. Cureus. 2025. doi: 10.7759/cureus.79022.

27. Khosravi M. Deep Brain Stimulation in Treatment-Resistant Psychiatric Disorders: Efficacy, Safety, and Future Directions. Brain Sciences. 2025;15(11):1244. doi: 10.3390/brainsci15111244.

28. Scaife J, Eraifej J, Green AL, Petric B, Aziz TZ, Park RJ. Deep Brain Stimulation of the Nucleus Accumbens in Severe Enduring Anorexia Nervosa: A Pilot Study. Frontiers in Behavioral Neuroscience. 2022;16. doi: 10.3389/fnbeh.2022.842184.

29. Huang B, Cheng Y, Cao H, Li D, Shi D, Wang Z. Temporal Interference Stimulation of the Nucleus Accumbens in Obsessive-compulsive Disorder. General Psychiatry. 2026;39(5). doi: 10.1002/gps3.70059.

30. Huang X, Xiao L, Wang M, Wu Y, Deng H, Wang W. Advancing Obsessive–Compulsive Disorder Research: Insights From Transgenic Animal Models and Innovative Therapies. Brain Sciences. 2025;15(1):43. doi: 10.3390/brainsci15010043.

31. Sloley C, Bell C, Shipton EA, Williman J. Chronic Pain and Obsessive–Compulsive Disorder: A Scoping Review. Pain and Therapy. 2025;14(2):513-87. doi: 10.1007/s40122-024-00704-4.

32. Fiorese GS, Uribe NC, Trüeb RM. Morgellons-Like Disease of the Scalp. International Journal of Trichology. 2025;17(2):133-6. doi: 10.4103/ijt.ijt_72_24.

33. De la Peña-Arteaga V, Cano M, Porta-Casteràs D, Vicent-Gil M, Miquel-Giner N, Martínez-Zalacaín I, et al. Mindfulness-based cognitive therapy neurobiology in treatment-resistant obsessive-compulsive disorder: A domain-related resting-state networks approach. European Neuropsychopharmacology. 2024;82:72-81.