Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Neurotransmitter-Based Model for Identifying Significant Depressive Symptoms in Patients with End-Stage Renal Disease

Fei Huang
Ezhou Central Hospital
Miao Zhu
Xiamen University
Xiang Chen
Ezhou Central Hospital
Lihua Shen
Ezhou Central Hospital
Huan Tian
Ezhou Central Hospital
Feng Cheng
pepsi.g@163.com

Ezhou Central Hospital

Article

Keywords: end-stage renal disease, depressive symptoms, neurotransmitters, machine learning

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11385259/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Neurotransmitter-Based Model for Identifying Significant Depressive Symptoms in Patients with End-Stage Renal Disease

Fei Huang[1#], Miao Zhu[2#], Xiang Chen[3#], Lihua Shen[1], Huan Tian[1*], Feng Cheng[1*]

[1] Department of Medical Laboratory, Ezhou Central Hospital, Ezhou, 436099, China

[2] School of Public Health, Xiamen University, Xiamen, 361005, China

[3] Department of Blood Transfusion, Ezhou Central Hospital, Ezhou, 436099, China

# Fei Huang, Miao Zhu and Xiang Chen contributed equally to this work.

* Correspondence: Huan Tian and Feng Cheng
Email: pepsi.g@163.com; t117vip@163.com;

Abstract

Depressive symptoms are common in patients receiving maintenance hemodialysis, but their identification is complicated by symptoms that overlap with uremia. We examined whether serum neurotransmitters, together with demographic and lifestyle variables, could identify significant depressive symptoms in a single-center, cross-sectional study of 250 adults with end-stage renal disease. Significant depressive symptoms were defined by a score greater than 7 on the depression subscale of the Hospital Anxiety and Depression Scale. Serum neurotransmitters and metabolites were measured by liquid chromatography-tandem mass spectrometry. Least absolute shrinkage and selection operator regression was used to select predictors, and eight machine learning classifiers were compared. Significant depressive symptoms were present in 86 participants (34.4%). The six selected predictors were age, household income, alcohol use, tea consumption, epinephrine, and 5-hydroxyindoleacetic acid. The random forest model achieved the highest area under the receiver operating characteristic curve (0.92), with 84% accuracy, 87% sensitivity, and 80% specificity in the study testing set. These findings support further evaluation of a biomarker-informed approach to identifying depressive symptoms in hemodialysis patients. External validation is required before clinical use.

Keywords: end-stage renal disease; depressive symptoms; neurotransmitters; machine learning

Introduction

Depression is a prevalent and serious issue among end-stage renal disease (ESRD) patients, often going undiagnosed. Studies indicate that depression rates in chronic kidney disease (CKD) and dialysis patients are three to four times higher than in the general population [1]. A global review estimated a 26.5% prevalence of clinical depression in CKD patients, with higher rates in those on maintenance hemodialysis, reaching 40-50% in some regions [1, 2]. Depression in ESRD is linked to adverse outcomes like increased hospitalization and mortality [3]. However, diagnosis is challenging due to the significant overlap between the somatic symptoms of depression (e.g.,

fatigue, anorexia) and uremia itself. This clinical ambiguity underscores the urgent need for objective, biology-based biomarkers to improve diagnostic accuracy.

The pathophysiology of depression in end-stage renal disease (ESRD) is complex and multifactorial, with substantial evidence indicating an interaction among chronic inflammation, uremic toxicity, and neurotransmitter dysregulation [4]. Patients undergoing dialysis often present with a persistent pro-inflammatory state, evidenced by increased levels of C-reactive protein and interleukin-6 (IL-6), the latter of which has been directly associated with depressive symptoms [5]. This chronic inflammatory condition may not only contribute to the elevated mortality observed in depressed ESRD patients through mechanisms such as atherosclerosis but also directly affect central nervous system function. In particular, inflammation can activate the kynurenine pathway, diverting tryptophan metabolism away from serotonin (5-HT) synthesis, a neurotransmitter crucial for mental health regulation [6]. This process, in conjunction with the accumulation of uremic toxins and other physiological disturbances, likely exacerbates the dysregulation of monoamine neurotransmitters, including serotonin (5-HT), norepinephrine (NEP), and dopamine (DA), which is a fundamental aspect of the classical pathophysiology of depression.

Given the high prevalence and severe consequences of depression in patients with end-stage renal disease (ESRD), considerable efforts have been directed toward the development of effective diagnostic and screening strategies. Although self-report scales are commonly employed, their inherent subjectivity constitutes a significant limitation. In response, recent studies have sought to mitigate this issue by constructing predictive models utilizing large-scale databases. For example, Lai et al. and Yan et al. developed nomograms aimed at preventing depression within the general chronic kidney disease (CKD) population, drawing on data from the US National Health and Nutrition Examination Survey (NHANES) [7, 8]. In the context of the Chinese population, a recent study has established a depression risk prediction model specifically tailored for middle-aged and elderly Chinese patients with CKD [9]. Despite these advancements, existing models exhibit two major limitations. Firstly, they predominantly rely on sociodemographic, lifestyle, and comorbidity data, lacking objective biological markers that could elucidate the underlying pathophysiology of depression. Secondly, these models were either developed for general CKD populations or specific age groups (middle-aged and elderly) and have not been specifically tailored to or validated in end-stage renal disease (ESRD) patients

undergoing dialysis. This particular subgroup encounters distinct physiological challenges and metabolic disturbances compared to the general CKD population, thereby limiting the applicability of existing models for this high-risk group.

Consequently, there is a significant need for a diagnostic model tailored specifically to the end-stage renal disease (ESRD) dialysis population, which also incorporates objective biological data. This study was conceived to address this need by integrating advanced bioanalytical techniques with machine learning methodologies. Our objective was to systematically profile an extensive array of serum neurotransmitters and their metabolites utilizing liquid chromatography-tandem mass spectrometry (LC-MS/MS). We aimed to combine this data with clinical variables to develop a robust diagnostic model. We hypothesized that a signature derived from this multi-dimensional dataset could accurately distinguish between depressed and non-depressed ESRD patients, thereby establishing a foundation for a more objective and biologically-informed diagnostic approach.

Results

Population characteristics

In this study, 250 ESRD patients on maintenance hemodialysis were assessed using the depression subscale of the Hospital Anxiety and Depression Scale (HADS-D) and categorized into two groups: the significant depressive symptoms (SDS) group (HADS-D > 7; n=86, 34.4%) and the non-significant depressive symptoms (nSDS) group (HADS-D ≤ 7; n=164, 65.6%).

The demographic and lifestyle characteristics of the study participants are detailed in Table 1. Statistically significant differences between the groups were identified in terms of household income and alcohol consumption. Notably, a greater proportion of individuals in the SDS group reported a monthly household income of ≤2000 RMB (52.3% compared to 26.2%). Additionally, the prevalence of alcohol use was higher in the SDS group (31.4% compared to 17.1%). Conversely, no significant differences were found between the groups regarding smoking status (36.0% compared to 24.4%), employment status (1.2% compared to 6.1%), and tea consumption (7.0% compared to 14.0%). Furthermore, the groups were similar with respect to sex and marital status.

Table 1. Demographic and lifestyle characteristics of the study participants

Variable	nSDS (n=164)	SDS (n=86)	Total (n=250)	P-value
Age (years), Mean ± SD	54.72 ± 11.52	57.98 ± 11.49	55.84 ± 11.57	0.03*
BMI (kg/m ²), Mean ± SD	22.27 ± 3.67	21.94 ± 2.93	22.15 ± 3.43	0.48
Sex, n (%)				0.91
Male	85 (51.8)	46 (53.5)	131 (52.4)	
Female	79 (48.2)	40 (46.5)	119 (47.6)	
Marital Status, n (%)				0.93
Married	143 (87.2)	75 (87.2)	218 (87.2)	
Non-married	11 (6.7)	5 (5.8)	16 (6.4)	
Other	10 (6.1)	6 (7.0)	16 (6.4)	
Household Income (RMB/month), n (%)				<0.01*
≤2000	43 (26.2)	45 (52.3)	88 (35.2)	
2001-4000	55 (33.5)	19 (22.1)	74 (29.6)	
4001-6000	33 (20.1)	9 (10.5)	42 (16.8)	
>6000	33 (20.1)	13 (15.1)	46 (18.4)	
Smoking Status, n (%)				0.07
No	124 (75.6)	55 (64.0)	179 (71.6)	
Yes	40 (24.4)	31 (36.0)	71 (28.4)	
Alcohol Use, n (%)				0.015*
No	136 (82.9)	59 (68.6)	195 (78.0)	
Yes	28 (17.1)	27 (31.4)	55 (22.0)	
Tea Consumption, n (%)				0.15
No	141 (86.0)	80 (93.0)	221 (88.4)	
Yes	23 (14.0)	6 (7.0)	29 (11.6)	
Employment Status, n (%)				0.10
Unemployed	154 (93.9)	85 (98.8)	239 (95.6)	
Employed	10 (6.1)	1 (1.2)	11 (4.4)	

Continuous variables are presented as mean ± standard deviation (SD), and categorical variables as n (%). Groups comprise 164 participants without significant depressive symptoms (nSDS) and 86 with significant depressive symptoms (SDS), defined by Hospital Anxiety and Depression Scale depression subscale scores ≤7 and >7, respectively. P values were calculated using independent-samples t-tests for continuous variables and chi-square or Fisher's exact tests, as appropriate, for categorical variables. *P < 0.05. BMI, body mass index; RMB, renminbi.

Comparison of routine biochemical indicators

The results of the univariate comparisons of 15 routine biochemical indicators between the nSDS and SDS groups are shown in Table 2. The analysis revealed no statistically significant differences between the two groups across all measured biochemical markers. This includes renal function indicators (creatinine, uric acid, urea), lipid profiles (total cholesterol,

triglycerides, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C)), and markers of inflammation/nutritional status (neutrophil-to-lymphocyte ratio [NLR], platelet-to-lymphocyte ratio [PLR], systemic immune-inflammation index [SII], and prognostic nutritional index [PNI]), with all p-values exceeding 0.05.

Table 2. Biochemical indicators in participants with and without significant depressive symptoms

Variable	Unit	nSDS (n=164)	SDS (n=86)	t-statistic	p-value
Creatinine (CRE)	μmol/L	982.14 ± 243.57	968.39 ± 274.93	0.41	0.69
Uric Acid (UA)	μmol/L	418.25 ± 113.20	414.91 ± 95.31	0.23	0.82
Urea (UREA)	mmol/L	22.21 ± 5.39	22.81 ± 5.90	-0.82	0.41
Total Cholesterol (TC)	mmol/L	3.71 ± 0.87	3.92 ± 0.85	-1.82	0.07
Triglycerides (TG)	mmol/L	1.78 ± 0.98	1.90 ± 1.23	-0.88	0.38
HDL-C	mmol/L	0.97 ± 0.29	1.02 ± 0.30	-1.32	0.19
LDL-C	mmol/L	2.12 ± 0.61	2.19 ± 0.67	-0.80	0.42
Glucose (GLU)	mmol/L	7.06 ± 4.45	7.72 ± 5.06	-1.06	0.29
Neutrophils (NEUT)	$\times 10^9$/L	4.16 ± 1.64	4.29 ± 1.39	-0.62	0.54
Lymphocytes (LYM)	$\times 10^9$/L	1.13 ± 0.42	1.07 ± 0.38	1.10	0.27
Platelets (PLT)	$\times 10^9$/L	181.46 ± 66.68	171.80 ± 46.06	1.20	0.23
NLR	ratio	4.12 ± 2.00	4.37 ± 1.85	-0.99	0.33
PLR	ratio	174.86 ± 68.00	174.75 ± 66.72	0.01	0.99
SII	index	750.07 ± 472.54	747.81 ± 363.24	0.04	0.97
PNI	index	43.82 ± 5.02	43.28 ± 3.67	0.88	0.38

Values are mean ± standard deviation. P values and t statistics are from independent-samples t-tests comparing 164 participants without significant depressive symptoms (nSDS) and 86 with significant depressive symptoms (SDS). HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; SII, systemic immune-inflammation index; PNI, prognostic nutritional index.

Profiling of serum neurotransmitters and metabolites

A comprehensive panel of 15 serum neurotransmitters and metabolites was quantified. The comparison of their mean concentrations (via t-test) and their individual associations with depression (via univariate logistic regression) are presented together in Table 3.

Table 3. Serum neurotransmitters and univariate logistic regression by depressive symptom status

Variable	nSDS (n=164)	SDS (n=86)	P-value (t-test)	Odds Ratio (95% CI)	P value (logistic model)
Catecholamines and Metabolites					
Dopamine (DA)	6.03 ± 19.29	7.16 ± 24.35	0.69	1.05 (0.82–1.35)	0.69
Norepinephrine (NEP)	1.66 ± 3.60	1.46 ± 1.57	0.62	0.92 (0.67–1.27)	0.63
Epinephrine (EP)*	0.22 ± 0.22	0.32 ± 0.37	0.01	1.41 (1.08–1.86)	0.01
3-Methoxytyramine (3-MT)	0.92 ± 1.18	0.86 ± 1.16	0.73	0.95 (0.73–1.25)	0.72
L-DOPA	17.52 ± 13.95	20.12 ± 14.91	0.17	1.19 (0.93–1.54)	0.17
Metanephrine (ME)	1.89 ± 1.44	1.97 ± 1.28	0.66	1.06 (0.82–1.37)	0.66
Serotonin and Metabolite					
Serotonin (5-HT)	11.65 ± 8.81	11.42 ± 9.41	0.84	0.97 (0.75–1.27)	0.84
5-HIAA	238.90 ± 180.64	212.16 ± 139.94	0.23	0.85 (0.65–1.11)	0.23
Amino Acid Neurotransmitters					
GABA	9.08 ± 10.24	9.97 ± 10.97	0.52	1.06 (0.84–1.41)	0.52
Glutamate (Glu)	65.75 ± 117.59	92.14 ± 169.30	0.15	1.20 (0.93–1.55)	0.16
Histidine (His)	10.97 ± 12.28	12.79 ± 13.59	0.29	1.15 (0.89–1.49)	0.29
Tyrosine (Tyr)	1.21 ± 1.59	1.32 ± 1.63	0.61	1.07 (0.83–1.38)	0.61
Taurine (Tau)	492.42 ± 1366.48	543.05 ± 943.76	0.76	1.04 (0.81–1.34)	0.76
Cholinergic System					
Acetylcholine (ACh)	4.62 ± 5.11	4.75 ± 5.28	0.85	1.03 (0.79–1.33)	0.85
Choline (Ch)	1186.65 ± 291.26	1200.52 ± 313.26	0.73	1.05 (0.81–1.36)	0.73

Concentrations are presented as mean ± standard deviation. Groups comprise 164 participants without significant depressive symptoms (nSDS) and 86 with significant depressive symptoms (SDS). Group comparisons used independent-samples t-tests; odds ratios and 95% confidence intervals (CI) are from univariate logistic regression. *P < 0.05. EP, epinephrine; 5-HIAA, 5-hydroxyindoleacetic acid; GABA, gamma-aminobutyric acid; L-DOPA, levodopa.

Both analytical methodologies consistently identified epinephrine (EP) as the sole neurotransmitter significantly correlated with depression status within this cohort. The mean serum concentration of EP was notably elevated in the depression group (0.32 ± 0.37 ng/mL) compared to the non-depression group (0.22 ± 0.22 ng/mL; p = 0.01, as determined by a t-test). This finding corresponded to a significantly increased likelihood of depression, as indicated by the logistic regression analysis (Odds Ratio [OR] = 1.41, 95% Confidence Interval [CI]: 1.08–1.86; p = 0.01). Conversely, none of the other 14 quantified neurotransmitters or metabolites demonstrated a significant association with depression in either the group mean comparison or the regression analysis (all p > 0.05). The consistent identification of epinephrine as a key univariate biological marker suggests its potential as a strong candidate for inclusion in subsequent multivariate diagnostic models, where its efficacy can be evaluated in conjunction with other clinical and demographic variables.

Model development and evaluation

To construct a diagnostic model that is both parsimonious and robust, the Least Absolute Shrinkage and Selection Operator (LASSO) regression technique was employed on a comprehensive dataset comprising 39 variables, including demographic, clinical, biochemical, and neurotransmitter data. This method is specifically designed to enhance model simplicity by reducing the coefficients of less significant features to zero.

Utilizing 10-fold cross-validation, an optimal penalty parameter (λ) was identified, culminating in the selection of a final set of six critical features: age, household income, alcohol use, tea consumption, epinephrine (EP), and 5-hydroxyindoleacetic acid (5-HIAA). The performance curve of the cross-validation process is depicted in Fig. 1a. It is noteworthy that, although 5-HIAA did not exhibit a statistically significant difference in the univariate analysis, it was selected by the multivariate LASSO algorithm. This indicates that 5-HIAA, in conjunction with the other five variables, significantly enhances the model's overall diagnostic efficacy.

Utilizing the six features identified through LASSO regression, eight distinct machine learning classifiers were developed to diagnose depression in patients with End-Stage Renal Disease (ESRD). The diagnostic efficacy of each model was rigorously assessed using seven key metrics, with detailed results presented in Table 4. Among the classifiers, the Random Forest (RF) model exhibited significantly superior performance, achieving the highest Area Under the Curve (AUC) of 0.92, along with an accuracy of 0.84, sensitivity of 0.87, and specificity of 0.80. Other models demonstrating high performance included the Gradient Boosting Machine (GBM) and Support Vector Machine (SVM), with AUCs of 0.84 and 0.83, respectively. Conversely, traditional models such as Logistic Regression (LR) and tree-based models like CART displayed only moderate discriminative capabilities (AUC < 0.75). Given its exceptional overall performance, particularly its high AUC and balanced sensitivity and specificity, the Random Forest model was identified as the optimal diagnostic model for this study.

Table 4. Model performance for identifying significant depressive symptoms

Model	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1-score
Random Forest (RF)	0.92	0.84	0.87	0.80	0.89	0.75	0.88
Gradient Boosting Machine (GBM)	0.84	0.79	0.81	0.74	0.86	0.66	0.83
Support Vector Machine (SVM)	0.83	0.78	0.84	0.67	0.85	0.66	0.84
Logistic Regression (LR)	0.74	0.66	0.65	0.68	0.82	0.48	0.72
Linear Discriminant Analysis (LDA)	0.74	0.66	0.65	0.67	0.81	0.48	0.72
Quadratic Discriminant Analysis (QDA)	0.72	0.65	0.65	0.61	0.79	0.49	0.69
Classification and Regression Tree (CART)	0.71	0.68	0.71	0.62	0.80	0.51	0.75
Naïve Bayes	0.69	0.65	0.77	0.43	0.73	0.47	0.74

AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value. Accuracy, sensitivity, specificity, PPV, NPV, and F1-score are reported as proportions. The eight models used the same six selected predictors; values are reproduced from the study evaluation.

The optimal Random Forest (RF) model underwent a thorough assessment of its diagnostic performance, emphasizing its discrimination, calibration, and potential clinical utility. The model exhibited excellent discriminative capacity, as evidenced by the Receiver Operating Characteristic (ROC) curve analysis, which yielded a high mean Area Under the Curve (AUC) of 0.92. The stability of the ROC curves across various data samples, as reflected by the narrow 95% confidence interval, suggests robust generalization capabilities (Fig. 1b). The calibration curve exhibited a robust overall concordance between the predicted probabilities generated by the model and the observed frequencies of depression. While the model demonstrated high accuracy within the high-risk probability range, there was a minor tendency to underestimate risk in the low-probability range, suggesting potential areas for future refinement (Fig. 1c). Decision Curve Analysis (DCA) was conducted to evaluate the model's net benefit in a clinical context. The findings indicated that, across a broad spectrum of threshold probabilities, employing the RF model to inform clinical decisions would yield a greater net benefit compared to the default strategies of treating all patients or treating none (Fig. 1d). This underscores the model's prospective utility as a decision-support tool in clinical practice.

To elucidate the relative contributions of the six predictors within the final Random Forest (RF) model, a feature importance analysis was performed. This analysis, grounded in the mean decrease in Gini impurity, revealed that 5-hydroxyindoleacetic acid (5-HIAA) emerged as the most significant predictor for classifying depression status, followed by age and epinephrine

(EP). Conversely, lifestyle factors, including alcohol and tea consumption, exhibited a comparatively minor impact on the model's diagnostic accuracy within this cohort (Fig. 1e).

Furthermore, to enhance the clinical applicability of these findings, a nomogram was constructed based on a multivariate logistic regression model that integrated the same six key predictors: age, income, alcohol use, tea consumption, EP, and 5-HIAA. This nomogram offers a straightforward graphical interface that enables clinicians to readily calculate an individual patient's probability of depression by aggregating the points assigned to each variable (Fig. 1f).

Figure 1

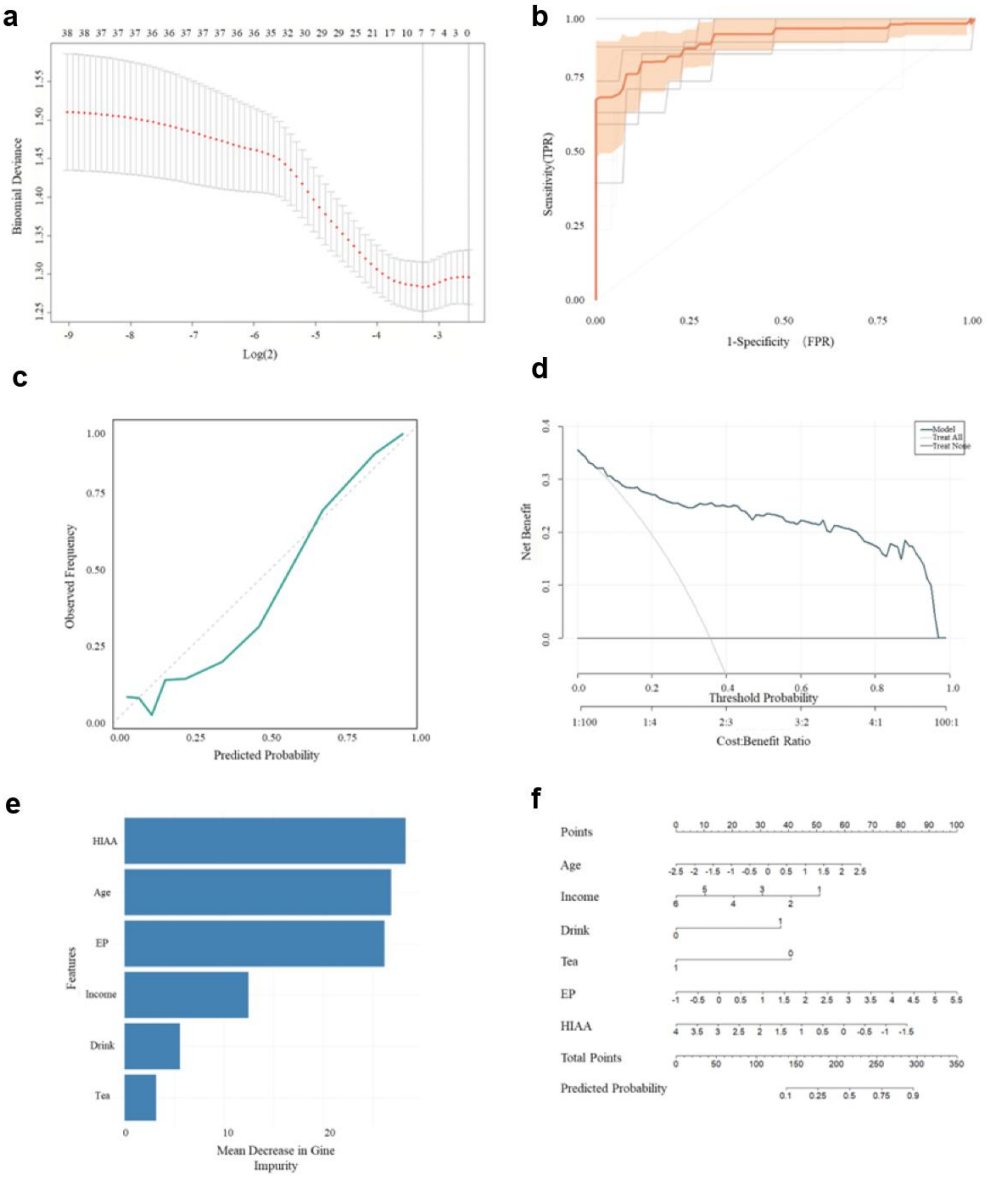


Figure 1. Feature selection and evaluation of models for identifying significant depressive symptoms. (a) Binomial deviance across the LASSO penalty parameter, evaluated by 10-fold cross-validation; six predictors were selected. (b) Receiver operating characteristic curve for the random forest model in the study testing set. The curve represents the mean of 100 bootstrap replicates, and the shaded region denotes the 95% confidence interval. (c) Calibration of predicted probabilities against observed frequencies of significant depressive symptoms. The dashed diagonal indicates perfect calibration. (d) Decision curve analysis comparing the random forest model with the strategies of intervening in all patients or in none across threshold probabilities. (e) Importance of the six predictors in the random forest model, measured by the mean decrease in Gini impurity. (f) Nomogram derived from a separate multivariable logistic regression model using the same six predictors. The study cohort comprised 250 adults receiving maintenance hemodialysis, including 86 with significant depressive symptoms. EP, epinephrine; HIAA, 5-hydroxyindoleacetic acid; Drink, alcohol use; LASSO, least absolute shrinkage and selection operator; TPR, true-positive rate; FPR, false-positive rate. The models identify questionnaire-defined depressive symptoms, not a clinical diagnosis of major depressive disorder.

Discussion

The diagnosis of depression in individuals with end-stage renal disease (ESRD) presents a significant clinical challenge, largely due to the confounding overlap between somatic symptoms and uremia. This study sought to address this issue by developing a quantitative, multi-domain diagnostic model. Our primary finding indicates that a parsimonious six-feature signature, derived from LASSO regression and implemented within a Random Forest algorithm, can accurately classify depressed versus non-depressed hemodialysis patients, achieving a high area under the curve (AUC = 0.92). This signature incorporates demographic, lifestyle, and neurobiological variables, specifically age, income, alcohol use, tea consumption, epinephrine (EP), and 5-HIAA.

A central finding of this research is the significant contribution of specific biological markers to the diagnostic model, particularly EP and 5-HIAA, which reflect the sympathoadrenal and serotonergic systems, respectively.

The observed association between elevated serum epinephrine (EP) and depression within our cohort lends support to the hypothesis of sympathoadrenal hyperactivity as a contributing factor in the pathophysiology of this disorder. This finding is particularly pertinent to the end-stage renal disease (ESRD) population, which is subject to chronic physiological and psychological stressors. Elevated EP may serve as a critical mediator of depressive states, potentially through its downstream effects on autonomic and cardiovascular functions, both of which are often compromised in individuals with ESRD [10]. Previous studies have demonstrated that psychosocial distress in dialysis patients is closely associated with autonomic dysregulation, which subsequently predicts an increased risk of mortality [11]. This observation aligns with existing literature that connects sympathoadrenal hyperactivity with adverse cardiovascular events in depressed populations, thereby providing a potential mechanistic explanation for the heightened mortality observed in this subgroup [12]. The selection of 5-HIAA as the most significant feature by the Random Forest model, despite its lack of significance in univariate analysis, highlights the importance of employing multivariate analytical approaches in the study of complex metabolic diseases. While EP demonstrated a direct univariate association, the significance of 5-HIAA emerged only after accounting for the variance contributed by other

factors within the model. This exemplifies how multivariate methods can uncover critical predictors whose effects may be obscured by other variables in simpler analyses. Systemic 5-HIAA concentration in end-stage renal disease (ESRD) represents a composite measure, reflecting both central serotonin turnover and severely impaired renal clearance [13]. Moreover, chronic inflammation in ESRD activates indoleamine 2,3-dioxygenase (IDO), which diverts tryptophan away from serotonin synthesis towards the kynurenine pathway, a mechanism extensively reviewed by Dantzer et al. [14]. This metabolic shift likely results in subtle yet diagnostically relevant perturbations in the serotonin-5-HIAA axis. Our model indicates that even amidst profound uremic dysmetabolism, these neurochemical signals remain detectable and clinically significant when analyzed in conjunction with confounding variables.

In addition to these biological markers, the model identified advanced age and lower income as significant predictors of depression diagnosis. The correlation with advanced age aligns with findings in both the general population and those with end-stage renal disease (ESRD), likely reflecting an accumulation of risk factors such as a higher burden of comorbidities and the presence of frailty. Frailty is notably prevalent among individuals undergoing hemodialysis and has been recognized as a strong independent predictor of diminished health-related quality of life and depressive symptoms [15]. Low socioeconomic status, as indicated by income, is one of the most robust social determinants of depression [16]. Within the context of ESRD, this is further exacerbated by the concept of "financial toxicity"—the economic burden associated with managing a chronic illness—which serves as a significant psychosocial stressor that exacerbates depressive episodes and diminishes treatment adherence [17].

The integration of lifestyle factors enhances the precision of the model. Within our cohort, alcohol consumption was correlated with an increased risk of depression. This observation is consistent with the well-documented bidirectional relationship between alcohol use and depression, wherein alcohol may serve as a form of self-medication for depressive symptoms or may independently contribute to their onset through neurotoxic effects [18]. In contrast, tea consumption emerged as a potential protective factor, a conclusion corroborated by a large-scale epidemiological study involving older Chinese adults [19]. The bioactive compounds in tea, particularly polyphenols such as epigallocatechin-3-gallate (EGCG), exhibit significant antioxidant and anti-inflammatory properties. These compounds have demonstrated the ability to reduce oxidative stress in kidney tissue and provide neuroprotective benefits, potentially

mitigating the neuroinflammation associated with uremic depression [20]. However, the inclusion of tea consumption as a variable should be approached with caution due to the limited number of tea drinkers in our cohort (n=29), necessitating further validation in larger studies.

The diagnostic performance of our model (AUC = 0.92) represents a significant advancement over recently published prediction tools for depression within the broader chronic kidney disease (CKD) population. For example, Lai et al. and Yan et al. developed nomograms utilizing the NHANES database, which incorporate sociodemographic and lifestyle factors [7, 8]. Additionally, a recent study focused on Chinese middle-aged and elderly CKD patients, developing a risk model based on clinical characteristics [9]. Although these studies offer valuable screening tools for the general CKD population, their dependence on self-reported data and traditional clinical variables constrains their accuracy, especially when applied to the intricate physiology of end-stage renal disease (ESRD).

A significant challenge in this domain is that conventional screening instruments frequently exhibit low specificity in dialysis patients, as the somatic symptoms of uremia—such as fatigue, anorexia, and sleep disturbances—closely resemble those of depression. By integrating specific neurochemical biomarkers, our model addresses this fundamental limitation. In contrast to previous models that may conflate uremic symptoms with depressive characteristics, our biomarker-integrated approach effectively distinguishes this "uremic noise," offering a more objective assessment of the underlying depressive pathology specifically tailored for the dialysis population.

The findings of this study must be interpreted within the framework of several significant limitations. Firstly, the cross-sectional design limits the ability to establish causality between the selected features and depressive symptoms. Secondly, the generalizability of our model is restricted due to the modest sample size and the single-center design. Consequently, future research should focus on the prospective validation of this diagnostic signature within a multi-center cohort. Additionally, the outcome reflects significant depressive symptoms as identified by the HADS screening tool, rather than a formal clinical diagnosis of major depressive disorder. We also recognize that other unmeasured factors, such as specific dialysis parameters and non-psychotropic co-medications, may serve as confounding variables.

In this study, we formulated a comprehensive diagnostic signature for identifying significant depressive symptoms in patients with end-stage renal disease (ESRD), incorporating specific neurochemical biomarkers alongside critical demographic and lifestyle factors. The Random Forest model demonstrated exceptional efficacy, underscoring the central roles of sympathoadrenal hyperactivity, as evidenced by elevated epinephrine (EP) levels, and serotonergic dysregulation, as indicated by 5-hydroxyindoleacetic acid (5-HIAA), in the pathophysiology of depression associated with ESRD. This research establishes a foundation for a more objective and biologically informed approach to screening and risk stratification within the hemodialysis patient population.

Methods

Study design and participants

This cross-sectional study was conducted at Ezhou Central Hospital in Hubei, China, from May 2023 to August 2024. The study was conducted in accordance with the Declaration of Helsinki. All experimental protocols were approved by the Ethics Committee of Ezhou Central Hospital (Approval No. L2022-K-13). Written informed consent was obtained from all patient participants. Eligible participants included adult patients with a confirmed diagnosis of end-stage renal disease (ESRD) who had been undergoing maintenance hemodialysis for a minimum of three months. Additional inclusion criteria required participants to possess the ability to comprehend and complete study questionnaires and to have complete clinical records available. Exclusion criteria encompassed patients with a documented history of major psychiatric or neurological disorders, severe communication impairments, those who had undergone major surgery or experienced significant bleeding within the past month, or individuals currently taking psychotropic medications.

Clinical and demographic data collection

Trained research staff administered a standardized questionnaire to the enrolled hemodialysis patients to collect demographic and lifestyle information, including age, sex, body mass index (BMI), marital status, household income, employment status, smoking history, alcohol consumption, and tea consumption.

Assessment of significant depressive symptoms

Depressive symptoms were evaluated utilizing the depression subscale of the Hospital Anxiety and Depression Scale (HADS-D). The HADS is a 14-item self-report measure, comprising seven items each dedicated to assessing anxiety and depression. Each item is rated on a 4-point Likert scale ranging from 0 to 3, resulting in a subscale score between 0 and 21. In alignment with established literature, a HADS-D score exceeding 7 was employed to categorize patients as exhibiting significant depressive symptoms for the purposes of this study [21].

Routine biochemical measurements

Fasting blood samples were obtained from all participants in the morning prior to their hemodialysis sessions. Routine biochemical parameters were assessed at the hospital's central laboratory utilizing a Beckman Coulter AU5800 chemistry analyzer (Beckman Coulter, Inc., Brea, CA, USA). The parameters measured included fasting glucose (GLU), urea (UREA), creatinine (CRE), uric acid (UA), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and albumin (ALB). A complete blood count, encompassing neutrophil, lymphocyte, monocyte, and platelet counts, was conducted using a Mindray BC-6800Plus auto hematology analyzer (Mindray, Shenzhen, China). Based on these results, four indices related to inflammation and nutrition were calculated: the neutrophil-to-lymphocyte ratio (NLR), monocyte-to-lymphocyte ratio (MLR), systemic immune-inflammation index (SII; calculated as Platelets × Neutrophils / Lymphocytes), and the prognostic nutritional index (PNI; calculated as Albumin [g/L] + 5 × Lymphocyte count [$\times 10^9$/L]).

Quantification of serum neurotransmitters by LC-MS/MS

The serum concentrations of 15 neurotransmitters and their metabolites, including gamma-aminobutyric acid (GABA), glutamate (Glu), dopamine (DA), norepinephrine (NEP), epinephrine (EP), serotonin (5-HT), and 5-hydroxyindoleacetic acid (5-HIAA), were quantified utilizing a validated liquid chromatography-tandem mass spectrometry (LC-MS/MS) method (Waters Xevo TQ-XS). Absolute quantification was achieved through the use of stable isotope-labeled internal standards.

Sample preparation

A 100 μL aliquot of serum was mixed with ascorbic acid to achieve a final concentration of 1000 ng/mL, thereby inhibiting oxidation. A stable isotope-labeled internal standard (IS) solution was then added to obtain a final concentration of 10 ng/mL. Protein precipitation was carried out by adding 400 μL of acetonitrile, followed by vortex mixing for 1 minute. The mixture was centrifuged at 20,000 × g for 10 minutes at 4 °C. The supernatant was carefully transferred to a clean tube and evaporated to dryness under a gentle stream of nitrogen at ambient temperature. The residue was reconstituted in 100 μL of 90% acetonitrile, vortexed, and centrifuged again under the same conditions (20,000 × g, 10 minutes, 4 °C). The resulting supernatant was transferred to an autosampler vial for subsequent LC-MS/MS analysis.

LC-MS/MS analysis

Chromatographic separation was performed using a ZIC-cHILIC column (150 mm × 2.1 mm, 3 μm; Merck) maintained at 30 °C. The mobile phase consisted of (A) 10 mmol/L ammonium formate in water, adjusted to pH 3.5 with formic acid, and (B) 15 mmol/L ammonium formate in 90% acetonitrile, adjusted to pH 4.0 with formic acid. The gradient program was as follows: 0–1.0 min, 90% B; 1.0–4.0 min, 90–80% B; 4.0–7.0 min, 80–60% B; 7.0–7.1 min, 60–50% B; 7.1–9.0 min, 50% B; 9.0–9.1 min, 50–90% B; followed by re-equilibration at 90% B until 12.0 min. The flow rate was 0.3 mL/min, and the column temperature was maintained at 30 °C. The autosampler was kept at 6 °C, and the injection volume was 2 μL. The total run time for each sample was 12 minutes.

The mass spectrometer was operated in positive electrospray ionization (ESI+) mode using multiple reaction monitoring (MRM). Source parameters were optimized to maximize signal intensity and included a capillary voltage of 2.5 kV, a source temperature of 150 °C, and a desolvation temperature of 500 °C. For each analyte and its corresponding internal standard, precursor–product ion transitions, cone voltage, and collision energy were carefully optimized to ensure specificity and sensitivity. The MRM transitions and compound-specific mass spectrometric parameters for all analytes and internal standards are provided in Supplementary Table S1.

Quality control

Each analytical batch comprised quality control (QC) samples at low, medium, and high concentrations, which were prepared by spiking known quantities of standards into a pooled serum matrix. Furthermore, each batch included the analysis of a matrix blank—serum processed with internal standards (IS) but devoid of analyte standards—and a solvent blank, to assess background interference and potential system contamination.

Statistical analysis

Data preprocessing and univariate analysis

Before conducting the analysis, the dataset underwent several preprocessing steps. The proportion of missing values was minimal, at 2.68%, predominantly affecting the LDL-C and income variables. These missing values were imputed using a Random Forest algorithm to maintain data integrity.

For the development of machine learning models, all features were appropriately scaled. Continuous variables, including age, BMI, and the concentrations of all biochemical and neurotransmitter markers, were standardized using a Z-score transformation to achieve a mean of 0 and a standard deviation of 1. Categorical variables, such as employment status, alcohol use, and tea consumption, were converted into numerical format through dummy variable encoding.

Subsequently, univariate comparisons between the depressed and non-depressed groups were conducted. Continuous variables were expressed as mean ± standard deviation (SD) and compared using independent samples t-tests. Categorical variables were expressed as frequencies (percentages) and compared using Chi-square or Fisher's exact tests, as appropriate. A two-sided p-value of less than 0.05 was considered indicative of statistical significance.

Feature selection and model development

The dataset, which includes demographic, biochemical, and neurotransmitter variables, was utilized to determine depression status as a binary outcome (1 indicating depressed, 0 indicating non-depressed). To identify the most pertinent predictors and minimize model complexity, the Least Absolute Shrinkage and Selection Operator (LASSO) regression, coupled with 10-fold cross-validation, was applied.

Model training and evaluation

The feature set identified through the Least Absolute Shrinkage and Selection Operator (LASSO) was employed to train and evaluate eight distinct machine learning classifiers: logistic regression (LR), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), Naïve Bayes (NB), support vector machine (SVM), random forest (RF), gradient boosting machine (GBM), and classification and regression tree (CART). The performance of each model was comprehensively evaluated using a range of metrics, including the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the F1-score.

Data availability

The de-identified datasets and analysis code generated and/or analysed during the current study are not publicly available due to patient privacy and institutional data protection requirements, but are available from the corresponding author on reasonable request and with permission from Ezhou Central Hospital.

References

1. Palmer, S. et al. Prevalence of depression in chronic kidney disease: systematic review and meta-analysis of observational studies. *Kidney Int.* **84,** 179-191 (2013). https://doi.org/10.1038/ki.2013.77

2. Ma, T. K. W. & Li, P. K. T. Depression in dialysis patients. *Nephrology (Carlton)* **21,** 639-646 (2016). https://doi.org/10.1111/nep.12742

3. Farrokhi, F., Abedi, N., Beyene, J., Kurdyak, P. & Jassal, S. V. Association between depression and mortality in patients receiving long-term dialysis: a systematic review and meta-analysis. *Am. J. Kidney Dis.* **63,** 623-635 (2014). https://doi.org/10.1053/j.ajkd.2013.08.024

4. Simões e Silva, A. C., Miranda, A. S., Rocha, N. P. & Teixeira, A. L. Neuropsychiatric disorders in chronic kidney disease. *Front. Pharmacol.* **10,** 932 (2019). https://doi.org/10.3389/fphar.2019.00932

5. Taraz, M., Taraz, S. & Dashti-Khavidaki, S. Association between depression and inflammatory/anti-inflammatory cytokines in chronic kidney disease and end-stage renal disease patients: A review of literature. *Hemodial. Int.* **19,** 11-22 (2015). https://doi.org/10.1111/hdi.12200

6. Cervenka, I., Agudelo, L. Z. & Ruas, J. L. Kynurenines: Tryptophan's metabolites in exercise, inflammation, and mental health. *Science* **357,** eaaf9794 (2017). https://doi.org/10.1126/science.aaf9794

7. Lai, R. L. et al. Development of a nomogram for predicting depression risk in patients with chronic kidney disease: an analysis of data from the US National Health and Nutrition Examination Survey, 2007–2014. *BMJ Open* **15,** e089956 (2025). https://doi.org/10.1136/bmjopen-2024-089956

8. Yan, Q., Liu, G., Wang, R., Li, D. & Wang, D. Development and validation of a nomogram for predicting depression risk in patients with chronic kidney disease based on NHANES 2005–2018. *J. Health Popul. Nutr.* **44,** 136 (2025). https://doi.org/10.1186/s41043-025-00890-7

9. Sun, T., Ye, Q., Li, X. & Wang, D. Development and validation of a depression risk-predicting model for Chinese middle-aged and elderly Chronic Kidney Disease patients. Preprint at Research Square https://doi.org/10.21203/rs.3.rs-4545265/v1 (2024).

10. Converse, R. L. et al. Sympathetic overactivity in patients with chronic renal failure. *N. Engl. J. Med.* **327,** 1912-1918 (1992). https://doi.org/10.1056/nejm199212313272704

11. Kimmel, P. L., Cukor, D., Cohen, S. D. & Peterson, R. A. Depression in end-stage renal disease patients: a critical review. *Adv. Chronic Kidney Dis.* **14,** 328-334 (2007). https://doi.org/10.1053/j.ackd.2007.07.007

12. Joynt, K. E., Whellan, D. J. & O'Connor, C. M. Depression and cardiovascular disease: mechanisms of interaction. *Biol. Psychiatry* **54,** 248-261 (2003). https://doi.org/10.1016/s0006-3223(03)00568-7

13. Zeng, L. et al. Rapid LC-MS/MS-based quantitative serum tryptophan metabolomics platform identifies and monitors novel biomarkers for diabetic kidney disease. *Anal. Bioanal. Chem.* **418,** 285-297 (2026). https://doi.org/10.1007/s00216-025-06221-3

14. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W. & Kelley, K. W. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat. Rev. Neurosci.* **9,** 46-56 (2008). https://doi.org/10.1038/nrn2297

15. McAdams-DeMarco, M. A. et al. Frailty and health-related quality of life in end stage renal disease patients of all ages. *J. Frailty Aging* **5,** 174–179 (2016). https://doi.org/10.14283/jfa.2016.106

16. Ridley, M., Rao, G., Schilbach, F. & Patel, V. Poverty, depression, and anxiety: causal evidence and mechanisms. *Science* **370,** eaay0214 (2020). https://doi.org/10.1126/science.aay0214

17. Bhanvadia, S. K. et al. Financial toxicity among patients with prostate, bladder, and kidney cancer: a systematic review and call to action. *Eur. Urol. Oncol.* **4,** 396-404 (2021). https://doi.org/10.1016/j.euo.2021.02.007

18. Boden, J. M. & Fergusson, D. M. Alcohol and depression. *Addiction* **106,** 906-914 (2011). https://doi.org/10.1111/j.1360-0443.2010.03351.x

19. Yao, Y. et al. Type of tea consumption and depressive symptoms in Chinese older adults. *BMC Geriatr.* **21,** 331 (2021). https://doi.org/10.1186/s12877-021-02203-z

20. Pervin, M., Unno, K., Takagaki, A., Isemura, M. & Nakamura, Y. Function of green tea catechins in the brain: epigallocatechin gallate and its metabolites. *Int. J. Mol. Sci.* **20,** 3630 (2019). https://doi.org/10.3390/ijms20153630

21. Zigmond, A. S. & Snaith, R. P. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* **67,** 361-370 (1983). https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Acknowledgements

The authors thank all participants and the staff of the Hemodialysis Center and Clinical Laboratory of Ezhou Central Hospital for their assistance with participant recruitment, sample collection and laboratory testing.

Funding

This work was supported by the Joint Fund Project of the Hubei Provincial Natural Science Foundation of China (Grant Nos. 2026AFC0639 and 2026AFC1458) and the Hubei Provincial Central-Guided Local Science and Technology Development Project (Grant No. 2022BCE064).

Author contributions

FH, MZ and XC contributed equally to this work. FC and HT conceived and designed the study. FH, XC and LS contributed to participant recruitment, clinical data collection and sample processing. FH, LS and HT performed the laboratory measurements and quality control. MZ performed the statistical analysis and machine learning model development. FH, MZ and XC drafted the manuscript. XC, LS, HT and FC contributed to data interpretation and critically revised the manuscript. FC and HT supervised the study. All authors read and approved the final manuscript.

Additional information

Competing interests

The authors declare that they have no competing interests.

Abbreviations

ESRD: end-stage renal disease;
SDS: significant depressive symptoms;
nSDS: non-significant depressive symptoms;
HADS-D: depression subscale of the Hospital Anxiety and Depression Scale;
LC-MS/MS: liquid chromatography-tandem mass spectrometry;
LASSO: least absolute shrinkage and selection operator;
RF: random forest;
AUC: area under the receiver operating characteristic curve;
ROC: receiver operating characteristic;
DCA: decision curve analysis;
BMI: body mass index;
NLR: neutrophil-to-lymphocyte ratio;
PLR: platelet-to-lymphocyte ratio;
SII: systemic immune-inflammation index;
PNI: prognostic nutritional index;
EP: epinephrine;

5-HIAA: 5-hydroxyindoleacetic acid;

5-HT: serotonin;

GABA: gamma-aminobutyric acid;

DA: dopamine;

NEP: norepinephrine;

PPV: positive predictive value;

NPV: negative predictive value.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.docx