BMC Geriatrics

RESEARCH Open Access

Network-based predictive models for artificial intelligence: an interpretable application of machine learning techniques in the assessment of depression in stroke patients

Wenwei Zuo[1] and Xuelian Yang[2*]

Abstract

Background Depression is a common complication after a stroke that may lead to increased disability and decreased quality of life. The objective of this study was to develop and validate an interpretable predictive model to assess the risk of depression in stroke patients using machine learning (ML) methods.

Methods This study included 1143 stroke patients from the NHANES database between 2005 and 2020. First, risk factors for depression in stroke patients were determined by univariate and multivariate logistic regression analysis. Next, five machine learning algorithms were used to construct predictive models, and several evaluation metrics (including area under the curve (AUC)) were used to compare the predictive performance of the models. In addition, the SHAP (Shapley Additive Explanations) method was used to rank the importance of features and to interpret the final model.

Results We screened seven features to construct a predictive model. Among the 5 machine learning models, the XGBoost (extreme gradient boosting) model showed the best discriminative ability, with an AUC of the ROC (receiver operating characteristic curve) in the test set of 0.746 and an accuracy of 0.834. In addition, the prediction results of the XGBoost model were interpreted in detail using the SHAP algorithm. We also developed a web-based calculator that provides a convenient tool for predicting the risk of depression in stroke patients at the following link: https://prediction-model-for-depression.streamlit.app.

Conclusions Our interpretable machine learning model serves as an auxiliary tool for clinical judgment, aimed at early and effective identification of depression risk in stroke patients.

Keywords Stroke, Depression, Machine learning, Interpretable, Predictive model

*Correspondence:
Xuelian Yang
SH3089@126.com
[1]University of Shanghai for Science and Technology, No.516, Jungong Road, Yangpu Area, Shanghai 200093, China
[2]Department of Neurology, Gongli Hospital of Shanghai Pudong New Area, No. 219 Miaopu Road, Pudong New Area, Shanghai 200135, China

Introduction

Depression is a prevalent and serious complication among stroke patients. Studies indicate that it affects approximately 30% of individuals within five years post-stroke, a rate significantly higher than the prevalence in the general population [1, 2]. Stroke not only impairs physical functioning but also poses a substantial challenge to the mental health of patients. Those affected often experience mood swings, loss of self-efficacy, and fear of the future, which collectively contribute to the onset of depression [2]. In stroke patients, depression manifests through a range of symptoms, including persistent sadness, loss of interest, increased fatigue, insomnia or excessive sleepiness, and difficulty concentrating [3, 4]. These symptoms adversely influence the emotional state of patients and may lead to cognitive decline, hindering their recovery process and overall quality of life [5]. Research has established a significant association between depression and poor functional recovery, an increased risk of recurrent stroke, and heightened mortality rates following a stroke [6, 7]. Consequently, the high prevalence of depression among stroke patients represents a public health issue that warrants urgent attention. Furthermore, individuals suffering from depression often exhibit lower adherence and motivation during rehabilitation therapy, undermining treatment effectiveness and exacerbating their suffering. Therefore, early identification and intervention for depressive symptoms in stroke patients is essential.

Nomograms have been extensively utilized in various studies to predict the risk of depression in stroke patients. These studies offer an intuitive risk assessment tool by integrating multiple clinical variables through statistical modeling [8, 9, 10]. Nomograms are user-friendly and provide straightforward risk evaluations. Nevertheless, to enhance the accuracy of depression risk prediction, a machine learning (ML) approach has been employed [11, 12, 13]. ML algorithms can manage numerous variables and identify potential non-linear relationships, thereby demonstrating notable advantages in complex data analysis [14]. Through self-learning, ML models can be continuously refined to improve prediction accuracy. This approach not only boosts the predictive power of the models but also reinforces their clinical applicability. An additional significant advantage is that many ML models can overcome the "black boxes" limitations associated with traditional models [15]. While some ML algorithms face challenges regarding interpretability, feature significance analysis and visualization tools such as SHAP (Shapley Additive Explanations) can elucidate the model's decision-making process, thereby enhancing transparency [16, 17]. This interpretability allows clinicians to comprehend how models derive their predictions, fostering trust in real-world applications.

This study aimed to develop and validate an interpretable ML model for the early and accurate prediction of depression risk in stroke patients using the NHANES 2005–2020 dataset. We employed the SHAP method to clarify the significance of each feature and to elucidate the model's decision-making process. Furthermore, we assessed the model's significance in clinical prognosis to assist healthcare professionals in better identifying and managing the risk of depression in stroke patients, ultimately improving their overall health and quality of life.

Methods

Study design and study population

This study utilized data from the National Health and Nutrition Examination Survey (NHANES) 2005–2020, a comprehensive cross-sectional study of the non-institutionalized civilian population in the United States. NHANES assesses adults' and children's health and nutritional status through household interviews and physical examinations conducted at mobile screening centers. The household interview gathers demographic, socioeconomic, dietary, and health information, while the physical examination includes medical, dental, physiological, and laboratory evaluations. The National Center for Health Statistics (NCHS) Ethics Review Board approved the study, and informed consent was given by each participant. Out of 76,496 participants, 1805 were identified as stroke patients. After excluding those with missing data on depression and other covariates, a total of 1143 participants were included in the analyses (Fig. S1).

Identification of stroke

Stroke patients were identified based on self-reported diagnostic history, as determined by the question: "Has a doctor or other health professional ever told you that you had a stroke?" Participants who answered "yes" to this question were classified as stroke patients.

Determination of depression

The Patient Health Questionnaire-9 (PHQ-9) is a validated self-report instrument for assessing depressive symptoms over the previous two weeks [18]. The PHQ-9 consists of nine items, each rated on a scale of 0 to 3 (0 = 'not at all', 1 = 'a few days', 2 = 'more than half the days', 3 = 'almost every day'). The total score ranges from 0 to 27. In this study, a PHQ-9 score of ≥ 10 was defined as indicating depression, while a score of < 10 was classified as no depression [19].

Predictors

Demographic information collected in this study included gender (male, female), age, race (Mexican American, non-Hispanic white, non-Hispanic black, Hispanic, other

race), education level (less than high school, high school or equivalent, college or above), marital status (married/living with partner, widowed/divorced/separated, never married), and poverty income ratio (PIR). Physical examination data provided body mass index (BMI), calculated as weight (kg) divided by height (m) squared (kg/m^2). Lifestyle factors investigated included smoking behavior (whether participants had smoked at least 100 cigarettes in their lifetime), drinking habits (whether they had consumed at least 12 alcoholic beverages of any type in any given year), and moderate recreational activities. Disease history was also recorded, including diagnoses of hypertension, diabetes, arthritis, congestive heart failure (CHF), coronary heart disease (CHD), heart attack, and cancer. Additionally, data related to sleep duration and sleep disorders were collected through questionnaires, either self-reported by patients or diagnosed by physicians. Important biochemical indicators were obtained from laboratory tests, including total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and creatinine levels.

Data preprocessing

To address the issue of missing values in the dataset, we chose to exclude all individuals with missing data to ensure the accuracy of the analysis and the reliability of the model. While this approach helps minimize potential biases associated with imputing missing values, we acknowledge that excluding participants with missing data may introduce selection bias, especially when the missing data mechanism is not completely random. Additionally, to ensure the fairness and validity of model evaluation, we randomly split the dataset into training and test sets at a 7:3 ratio, with 70% of the data used for model training and 30% for assessing model performance. This procedure ensures the scientific rigor and credibility of the model we developed.

Construction of the model

In the training set, we first analyzed risk factors for depression in stroke patients using univariate logistic regression and selected variables significantly associated with depression ($P<0.05$) for inclusion in the subsequent multivariate logistic regression model. In this analysis, variables with a P-value of less than 0.05 were considered candidates for predicting depression in stroke patients. We calculated odds ratios (OR) and 95% confidence intervals (CI). Furthermore, five different ML models were employed, including random forest (RF), decision tree (DT), extreme gradient boosting (XGBoost), Naïve Bayesian (NB), and support vector machine (SVM), to identify depression in stroke patients. In the hyperparameter settings (Table S1), we provide the corresponding parameter configurations for five models. Taking the XGBoost model as an example, we performed a grid search with 5-fold cross-validation to determine the optimal hyperparameters. Specifically, we explored a range of learning rates (0.001, 0.01, 0.1), maximum depths (3, 5, 7), and the number of estimators (100, 300, 500). Through systematic experiments and evaluation, we selected a learning rate of 0.01, a maximum depth of 3, and 300 estimators as the optimal combination. The choice of a learning rate of 0.01 effectively balances training speed and model performance while reducing the risk of overfitting. A maximum depth of 3 was selected to control model complexity, ensuring that the model captures important features while minimizing overfitting. The 5-fold cross-validation results demonstrated that this depth setting achieved stable performance on both training and validation datasets. Finally, after multiple experiments, we determined that using 300 trees provided the best trade-off between predictive accuracy and computational efficiency. These hyperparameter selections aim to enhance the model's effectiveness and reproducibility, providing a reference for other model research.

Evaluation of the model

Evaluation metrics such as the receiver operating characteristic curve (ROC), area under the curve (AUC), average precision score (APS), accuracy, sensitivity (recall), specificity, negative predictive value (NPV), positive predictive value (PPV), false positive rate (FPR), false negative rate (FNR), and F1 score were used to evaluate the performance of the model. Using the SHAP algorithm, we also visualized the important features influencing the risk of depression in stroke patients, analyzed the importance of individual features on model output, and elucidated the impact of key features on the final model results.

Web deployment tool based on the streamlit framework

To facilitate the clinical application of the model, we developed a user-friendly web application based on the Streamlit Python framework. This application implements our final predictive model, allowing healthcare professionals to assess the risk of depression in stroke patients conveniently. When users input the feature values associated with the final model, the application automatically calculates and returns the probability of depression for that patient and a force map for individual stroke patients based on the pre-trained model.

Statistical analysis

All ML models were constructed and validated within the Python (3.12.0) environment. For statistical analyses, we utilized R software (4.3.2). For continuous data that were not normally distributed, we reported median and interquartile ranges for descriptive purposes and employed

the Mann-Whitney U-test for between-group comparisons. Categorical data were described using frequencies and percentages, with the chi-square test applied to assess differences between groups.

Results

Baseline characteristics of participants

This study included a total of 1143 participants, with the training set comprising 806 individuals and the test set consisting of 337 individuals. The median age of the overall sample was 67 years, with 562 (49.17%) males and 581 (50.83%) females; among these, 200 were diagnosed with depression. In the training set, the median age also stood at 67 years, involving 399 (49.50%) males and 407 (50.50%) females, with 145 patients diagnosed with depression. In contrast, the median age of the test set was 66 years, comprising 163 (48.37%) males and 174 (51.63%) females, with 55 patients diagnosed with depression. There were no statistically significant differences ($p > 0.05$) between the two datasets across all variables (Table 1).

Construction of ML models

In the univariate logistic regression model, a total of 14 variables were associated with the risk of depression in the training group of stroke patients. Following multivariate logistic regression analysis, seven significant predictors were ultimately identified: females had an increased risk of depression compared to men (OR: 1.814, 95% CI: 1.142–2.883); age was negatively associated with the risk of depression (OR: 0.977, 95% CI: 0.961–0.993); and the PIR was also negatively associated with the risk of depression (OR: 0.829, 95% CI: 0.701–0.993). Additionally, drinking increased the risk of depression (OR: 1.976, 95% CI: 1.143–3.415); sleep disorders significantly heightened the risk of depression (OR: 3.390, 95% CI: 2.179–5.272); moderate recreational activities were negatively correlated with the risk of depression (OR: 0.503, 95% CI: 0.299–0.846); and TC levels were positively associated with the risk of depression (OR: 1.215, 95% CI: 1.023–1.445) (Table 2). Subsequently, we applied five ML models to the NHANES dataset using the training dataset that included these seven variables.

Testing the performance of ML models

During the testing phase, we applied the trained model to the test set. The results indicated that the XGBoost model outperformed others in terms of AUC performance (AUC: 0.746; 95% CI: 0.674–0.810). The AUCs for the other models were as follows: 0.711 (95% CI: 0.638–0.778) for RF, 0.719 (95% CI: 0.643–0.792) for DT, 0.671 (95% CI: 0.607–0.736) for NB, and 0.703 (95% CI: 0.627–0.776) for SVM. Figure 1 illustrates the accuracy correction and AUC curves for the five ML models. The accuracy rates for RF (0.825), DT (0.837), XGBoost (0.834), NB (0.837), and SVM (0.837) demonstrated good performance in identifying depression in stroke patients. Table 3 summarizes the estimation performance of each model. Among all five models, XGBoost exhibited the highest average precision score (APS) at 0.353, indicating the best discrimination. The specificity, sensitivity/recall, NPV, PPV, FPR, FNR, FDR, and F1 scores for the five models are presented in Table 3. Fig. S2 provides the confusion matrix for the five models. Comprehensive feature-based analysis confirmed that XGBoost demonstrated the highest precision and robustness in identifying depression in stroke patients.

Visualization of feature importance

Using the SHAP algorithm, we evaluated the importance of each feature for the XGBoost model in predicting the risk of depression among stroke patients. As shown in Fig. 2A, the feature importance plot ranks the most significant features associated with depression in descending order. The horizontal position of each feature in the plot indicates its positive or negative effect on the predicted value, where red reflects a high positive contribution, and blue indicates a low negative contribution (Fig. 2B). The analysis revealed that sleep disorders exhibited the strongest average predictive power among all features, followed by age, PIR, moderate recreational activities, TC, gender, and drinking. In Fig. 2C, the lines represent individual participants in the decision diagram, with characteristics ranked in descending order of importance based on the observed data. Additionally, Fig. 3A and C display individual force diagrams for depressed and non-depressed individuals, respectively. Figure 3B and D illustrate waterfall plots for depressed and non-depressed patients, respectively, to provide insights into the impact of individual characteristics on model predictions. The SHAP values illustrate the predicted characteristics of individual patients and their contributions to the risk of developing depression, where red features indicate an increased risk and blue features indicate a decreased risk. The length and direction of the arrows visualize the degree of influence of each predictive feature.

Facilitating clinical applications

As illustrated in Fig. 4, the XGBoost-based prediction model has been integrated into a web application to facilitate its use in clinical practice. When the actual values of the seven features required by the model are entered, the application automatically predicts the risk of depression in stroke patients. Furthermore, the application displays a graph of the characteristics of a single stroke patient, highlighting the key factors influencing the depression prediction: the blue characteristics on the right contribute to a "non-depression" prediction, while the red

Table 1 Baseline characteristics of the training and test sets

Characteristic	Overall	Training set	Test set	*p*-value
N	1143	806	337	
Age	67.00 (57.00–77.00)	67.00 (57.25-76.00)	66.00 (57.00–77.00)	0.938
Gender				0.726
Male	562 (49.17%)	399 (49.50%)	163 (48.37%)	
Female	581 (50.83%)	407 (50.50%)	174 (51.63%)	
Race				0.808
Mexican American	92 (8.05%)	62 (7.69%)	30 (8.90%)	
Other Hispanic	68 (5.95%)	49 (6.08%)	19 (5.64%)	
Non-Hispanic White	589 (51.53%)	413 (51.24%)	176 (52.23%)	
Non-Hispanic Black	323 (28.26%)	228 (28.29%)	95 (28.19%)	
Other Race	71 (6.21%)	54 (6.70%)	17 (5.04%)	
Education				0.926
Less than high school	144 (12.60%)	100 (12.41%)	44 (13.06%)	
High school or equivalent	538 (47.07%)	382 (47.39%)	156 (46.29%)	
College or above	461 (40.33%)	324 (40.20%)	137 (40.65%)	
Marital				0.574
Married/Living with Partner	604 (52.84%)	434 (53.85%)	170 (50.45%)	
Widowed/Divorced/Separated	469 (41.04)	324 (40.20%)	145 (43.03%)	
Never married	70 (6.12%)	48 (5.96%)	22 (6.53%)	
PIR	1.64 (1.00-2.95)	1.61 (0.99–2.82)	1.71 (1.03–3.20)	0.405
BMI	29.36 (25.37–33.80)	29.20 (25.24–33.70)	29.70 (25.70–33.90)	0.562
Smoking				0.964
No	433 (37.88%)	305 (37.84%)	128 (37.98%)	
Yes	710 (62.12%)	501 (62.16%)	209 (62.02%)	
Drinking				0.599
No	310 (27.12%)	215 (26.67%)	95 (28.19%)	
Yes	833 (72.88%)	591 (73.33%)	242 (71.81%)	
Hypertension				0.552
No	265 (23.18%)	183 (22.70%)	82 (24.33%)	
Yes	878 (76.82%)	623 (77.30%)	255 (75.67%)	
Diabetes				0.213
No	763 (66.75%)	529 (65.63%)	234 (69.44%)	
Yes	380 (33.25%)	277 (34.37%)	103 (30.56%)	
Arthritis				0.217
No	487 (42.61%)	334 (41.44%)	153 (45.40%)	
Yes	656 (57.39%)	472 (58.56%)	184 (54.60%)	
CHF				0.424
No	955 (83.55%)	678 (84.12%)	277 (82.20%)	
Yes	188 (16.45%)	128 (15.88%)	60 (17.80%)	
CHD				0.653
No	945 (82.68%)	669 (83.00%)	276 (81.90%)	
Yes	198 (17.32%)	137 (17.00%)	61 (18.10%)	
Heart attack				0.261
No	909 (79.53%)	634 (78.66%)	275 (81.60%)	
Yes	234 (20.47%)	172 (21.34%)	62 (18.40%)	
Cancer				0.899
No	888 (77.69%)	627 (77.79%)	261 (77.45%)	
Yes	255 (22.31%)	179 (22.21%)	76 (22.55%)	
Sleep duration	7.00 (6.00–8.00)	7.00 (6.00–8.00)	7.07 (6.00–8.00)	0.872
Sleep disorder				0.376
No	647 (56.61%)	463 (57.44%)	184 (54.60%)	
Yes	496 (43.39%)	343 (42.56%)	153 (45.40%)	
Moderate physical activities				0.124

Table 1 (continued)

Characteristic	Overall	Training set	Test set	*p*-value
No	823 (72.00%)	591 (73.33%)	232 (68.84%)	
Yes	320 (28.00%)	215 (26.67%)	105 (31.16%)	
Depression				0.498
No	943 (82.50%)	661 (82.01%)	282 (83.68%)	
Yes	200 (17.50%)	145 (17.99%)	55 (16.32%)	
TC	4.60 (3.90–5.46)	4.60 (3.96–5.43)	4.65 (3.80–5.56)	0.680
HDL-C	1.27 (1.03–1.58)	1.27 (1.03–1.55)	1.27 (1.03–1.58)	0.567
Creatinine	86.63 (70.72-106.96)	86.63 (70.72-106.96)	86.63 (70.72-106.08)	0.698

PIR, poverty income ratio; BMI, body mass index; CHF, congestive heart failure; CHD, coronary heart disease; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol

characteristics on the left indicate a tendency towards "depression". The web application is accessible online at the following link: https://prediction-model-for-depression.streamlit.app.

Discussion

In this study, we selected seven easily accessible clinical variables—gender, age, PIR, drinking, sleep disorders, moderate recreational activities, and TC—to construct a prediction model for the early identification of depression in stroke patients using ML algorithms. Utilizing the XGBoost algorithm, our results demonstrated stable and satisfactory performance, achieving an AUC value of 0.746. This indicated that the predictive model possessed good discriminatory ability, effectively distinguishing between high-risk and low-risk patients. Additionally, we employed the SHAP approach to quantify the importance of each selected feature for model predictions. Finally, we implemented the prediction model as a web application to facilitate its practical application in clinical scenarios.

Our findings revealed a higher prevalence of depression in women compared to men among stroke patients, corroborating previous studies that identified female gender as a significant risk factor for post-stroke depression [20]. Specifically, the annual diagnosis rate for depression was notably higher in women than in men after stroke (HR: 1.53, 95% CI: 1.51–1.55) [21]. Possible explanations for this disparity include the relative psychological vulnerability of women and their comparatively weaker coping mechanisms. The life and work implications of stroke, combined with familial and social pressures, may exacerbate negative emotions. Furthermore, women tend to have poorer prognoses post-stroke, which intensifies both physical and psychological distress along with financial burdens [22]. This study also found a negative correlation between age and depression, indicating that younger stroke patients are more likely to experience depression than older patients. This aligns with a review that noted a significant increase in depression prevalence among adolescents [23]. Young individuals often face greater family and social responsibilities, which can diminish psychological resilience and exacerbate stress [24]. Economic status has a significant impact on reducing the incidence of depression in stroke patients. Patients with better economic conditions generally have access to superior medical resources and social support. One study identified a correlation between the severity of post-stroke depression and patients' economic status (χ^2 = 11.198, P = 0.024) [25]. While our study did not include educational level as a predictor, existing literature suggested that higher education correlated with a reduced risk of depression. For instance, a Chinese population-based study found that stroke patients with a high school education or above had a lower risk of depression compared to those with only primary education (OR: 0.50, 95% CI: 0.28–0.88, P = 0.016) [26]. Future interventions should thus consider enhancing both economic and educational resources to more effectively mitigate depression risk in stroke patients.

We observed a significant association between drinking and depression risk. Specifically, the prevalence of depression was markedly higher among alcohol drinkers compared to non-drinkers. Research-based on the Korean Community Health Survey (KCHS) showed that individuals consuming less than 5 g of alcohol daily had a 20% increased risk of depression (OR: 1.20, 95% CI: 1.07–1.35), while those consuming between 5 and 14.9 g per day had a 39% increased risk (OR: 1.39, 95% CI: 1.13–1.70) [27]. Conversely, a Mendelian randomization study indicated that alcohol consumption did not causally affect depression in older men, possibly due to the stress-relieving and mood-enhancing effects of low to moderate alcohol intake [28, 29]. These discrepancies might arise from differences in study design, sample selection, or genetic variation. Therefore, further research is needed to elucidate the complex relationship between alcohol consumption and depression in this demographic. In addition, stroke patients often contend with multiple comorbidities following acute treatment, making the relationship between sleep disorders and depression particularly important. Studies have indicated that the prevalence of post-stroke depression has been

Table 2 Univariate and multivariate logistic regression analyses of the training set

Variable	Univariate		Multivariate	
	OR (95%CI)	p-value	OR (95%CI)	p-value
Sex				
Male	1.000 (Ref.)		1.000 (Ref.)	
Female	1.840 (1.271,2.663)	**0.001**	1.814 (1.142,2.883)	**0.012**
Age	0.962 (0.950,0.975)	**< 0.001**	0.977 (0.961,0.993)	**0.006**
Race				
Mexican American	1.000 (Ref.)			
Other Hispanic	1.910 (0.731,4.992)	0.187		
Non-Hispanic White	1.223 (0.577,2.592)	0.600		
Non-Hispanic Black	1.253 (0.572,2.746)	0.573		
Other Race	1.867 (0.727,4.792)	0.194		
Education level				
Less than high school	1.000 (Ref.)			
High school or equivalent	1.233 (0.691,2.198)	0.479		
College or above	0.912 (0.500,1.664)	0.764		
Marital status				
Married/Living with Partner	1.000 (Ref.)		1.000 (Ref.)	
Widowed/Divorced/Separated	1.317 (0.897,1.932)	0.160	1.090 (0.693,1.714)	0.708
Never married	3.719 (1.970,7.024)	**< 0.001**	2.046 (0.985,4.250)	0.055
PIR	0.715 (0.612,0.835)	**< 0.001**	0.829 (0.701,0.980)	**0.028**
BMI	1.025 (1.001,1.051)	**0.045**	1.004 (0.977,1.032)	0.791
Smoking				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	1.908 (1.277,2.851)	**0.002**	1.525 (0.955,2.436)	0.077
Drinking				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	1.936 (1.219,3.075)	**0.005**	1.976 (1.143,3.415)	**0.015**
Hypertension				
No	1.000 (Ref.)			
Yes	1.046 (0.678,1.611)	0.840		
Diabetes				
No	1.000 (Ref.)			
Yes	1.166 (0.802,1.693)	0.421		
Arthritis				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	1.541 (1.055,2.250)	**0.025**	1.031 (0.653,1.628)	0.896
CHF				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	1.848 (1.188,2.874)	**0.006**	1.598 (0.906,2.818)	0.106
CHD				
No	1.000 (Ref.)			
Yes	1.280 (0.811,2.020)	0.289		
Heart attack				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	1.666 (1.109,2.502)	**0.014**	1.300 (0.777,2.175)	0.317
Cancer				
No	1.000 (Ref.)			
Yes	1.142 (0.748,1.744)	0.537		
Sleep duration	0.833 (0.759,0.916)	**< 0.001**	0.921 (0.834,1.017)	0.106
Sleep disorder				
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	4.665 (3.136,6.941)	**< 0.001**	3.390 (2.179,5.272)	**< 0.001**
Moderate recreational activities				

Table 2 (continued)

Variable	Univariate		Multivariate	
No	1.000 (Ref.)		1.000 (Ref.)	
Yes	0.461 (0.286,0.741)	**0.001**	0.503 (0.299,0.846)	**0.010**
TC	1.206 (1.039,1.399)	**0.014**	1.215 (1.023,1.445)	**0.027**
HDL-C	0.926 (0.605,1.416)	0.722		
Creatinine	0.996 (0.991,1.000)	0.507		

OR, odd risk; PIR, poverty income ratio; BMI, body mass index; CHF, congestive heart failure; CHD, coronary heart disease; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol

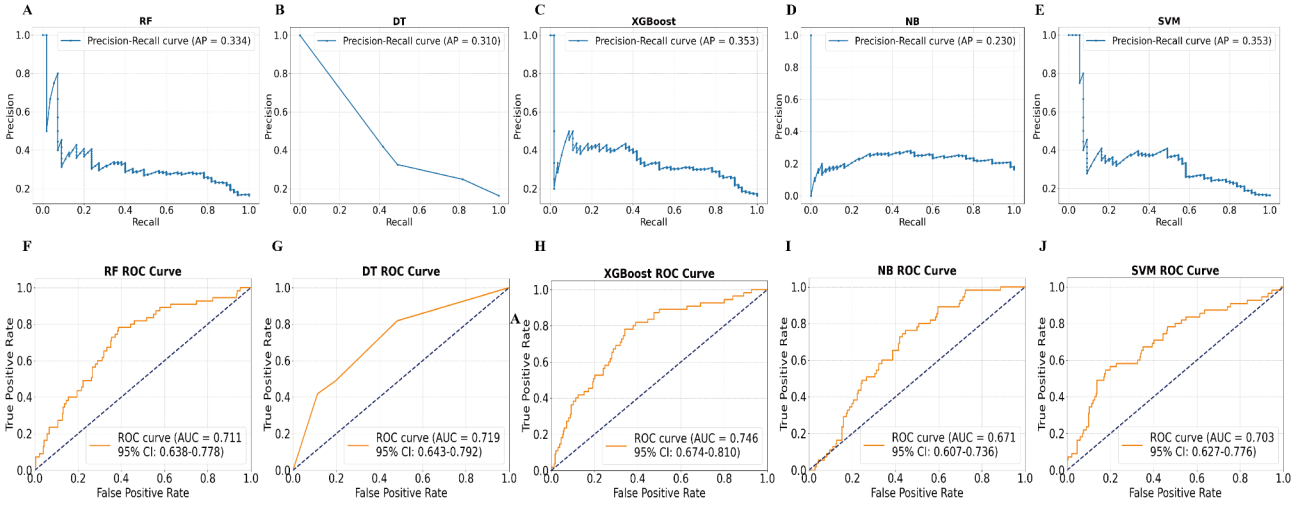

Fig. 1 Precision-recall curves and ROC for ML models. ROC, receiver operating characteristic curve; ML, machine leaning; RF, random forest; DT, decision tree; XGBoost, extreme gradient boosting; NB, Naïve Bayesian; SVM, support vector machine; APS, average precision score; AUC, the area under the curve; CI, confidence interval

Table 3 Comparison of the characteristics of five ML models

Characteristics	RF	DT	XGBoost	NB	SVM
AUC	0.711 (0.638,0.778)	0.719 (0.643,0.792)	0.746 (0.674,0.810)	0.671 (0.607,0.736)	0.703 (0.627,0.776)
APS	0.334	0.310	0.353	0.230	0.353
Accuracy	0.825	0.837	0.834	0.837	0.837
Specificity	0.968	1.000	0.982	1.000	1.000
Sensitivity/Recall	0.091	0.000	0.073	0.000	0.000
NPV	0.845	0.837	0.845	0.837	0.837
PPV	0.357	NA	0.444	NA	NA
FPR	0.032	0.000	0.018	0.000	0.000
FNR	0.909	1.000	0.927	1.000	1.000
F1 score	0.779	0.762	0.780	0.762	0.762

ML, machine learning; RF, random forest; DT, decision tree; XGBoost, extreme gradient boosting; NB, Naïve Bayesian; SVM, support vector machine; AUC, the area under the curve; APS, average precision score; NPV, negative predictive value; PPV, positive predictive value; FPR, false positive rate; FNR, false negative rate; NA: null

higher in patients with poor sleep quality than in those with good sleep quality [30]. Moderate to severe obstructive sleep apnea has been identified as a significant factor influencing post-stroke anxiety during the acute phase [31]. A retrospective study indicated that severe obstructive sleep apnea was significantly associated with an increased risk of post-stroke depression within three months (OR: 4.04, 95% CI: 1.38–9.62) [32]. Consequently, early identification and intervention for sleep disorders in stroke patients are crucial for reducing depression risk.

Our study found that moderate recreational activities significantly lowered the risk of depression among stroke patients. This finding aligns with existing literature, a meta-analysis of nine studies highlighted the potential benefits of home exercise in alleviating post-stroke depression, particularly emphasizing physical and mental exercises, such as tai chi, as effective treatments [2]. Clinical guidelines recommend non-pharmacological interventions, such as physical exercise, for stroke survivors experiencing mild depressive symptoms [33].

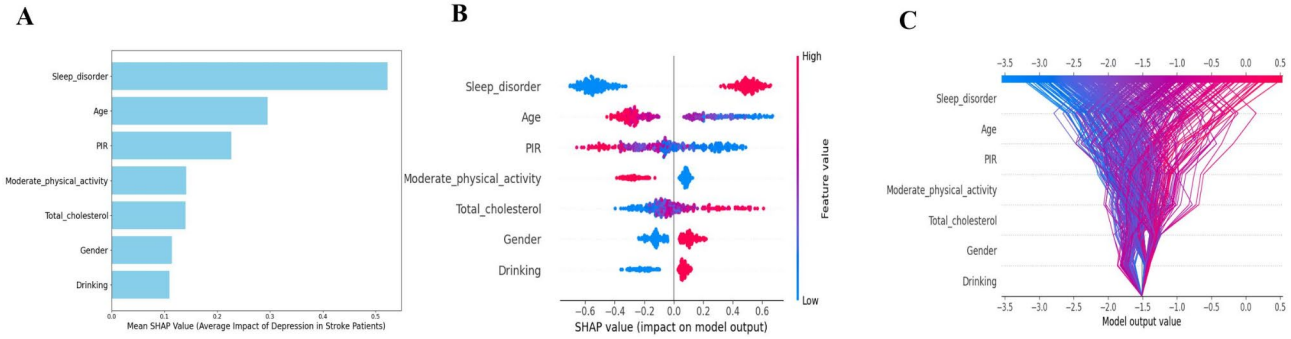

Fig. 2 Global model interpretation using the SHAP method. (**A**) SHAP summary bar plot. (**B**) SHAP summary dot plot. (**C**) SHAP decision plot. In the model, there is a dot for each patient's SHAP value and therefore a dot for each feature for each patient. All variables are in descending order of importance. The color of the dots indicates the actual value of each patient feature, with red indicating a higher feature value and blue indicating a lower feature value

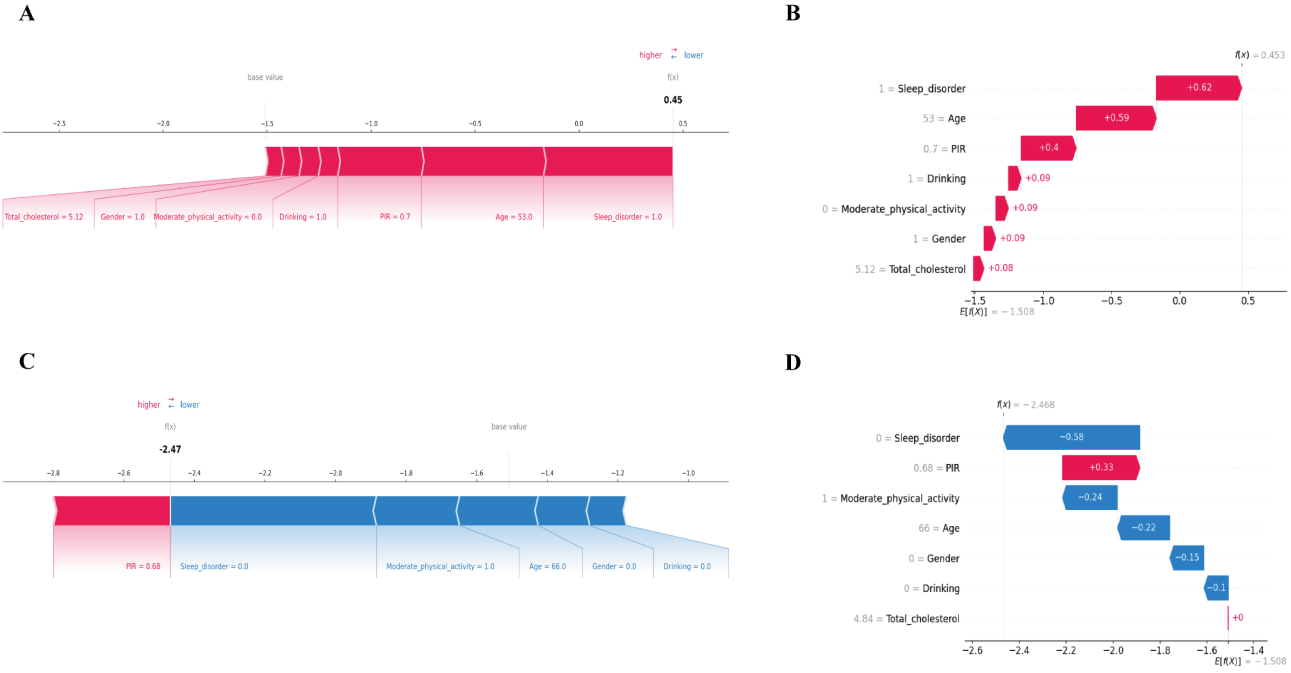

Fig. 3 Local model interpretation using the SHAP method. Fig. 3 A and 3 C show individual force diagrams for depressed and non-depressed patients, respectively. Figure 3B and D, on the other hand, show waterfall plots for depressed and non-depressed patients, respectively. Each patient is represented by the x-axis, while the contribution of features is represented by the y-axis: the larger the red part of each patient, the more likely it is to be judged as 'depression'

Although the precise mechanisms by which exercise alleviates depression are not fully understood, several plausible explanations exist. In the short term, physical activity activates the endorphin system, providing immediate mood enhancement; in the long term, regular exercise promotes neuroplasticity, improves brain function, and enhances the body's stress resistance. Furthermore, exercise positively impacts depression by fostering social interactions and boosting self-esteem [34, 35]. In our analysis, TC emerged as a significant predictor of post-stroke depression. Consistent with existing studies, elevated TC concentrations were associated with a higher risk of depression [36]. Cholesterol plays a critical role in brain function as a vital component of nerve membranes, influencing neurotransmitter synthesis and release, which in turn affects mood and behavior [37]. However, a study conducted on a Japanese population found that elevated cholesterol levels during pregnancy were linked to a reduced risk of postpartum depression [38]. This discrepancy may stem from the unique psychological experiences associated with pregnancy, contrasting with the long-term rehabilitation challenges faced by stroke patients.

In recent years, relatively few studies have employed machine learning methods to predict depression risk in stroke patients. While one study constructed a machine

Predictor of depression in stroke patients

Gender (0=Male, 1=Female):

Male (0)

Age:

40

PIR:

3

Drinking (0=No, 1=Yes):

No (0)

Sleep_disorder (0=No, 1=Yes):

No (0)

Moderate_physical_activity (0=No, 1=Yes):

No (0)

Total_cholesterol:

6.00

Predict

Predicted Class: 0

Prediction Probabilities: [0.91139436 0.08860565]

According to our model, you have a low risk of Depression. The model predicts that your probability of not having Depression is 91.1%. However, maintaining a healthy lifestyle is still very important. I recommend regular check-ups to monitor your heart health, and to seek medical advice promptly if you experience any symptoms.

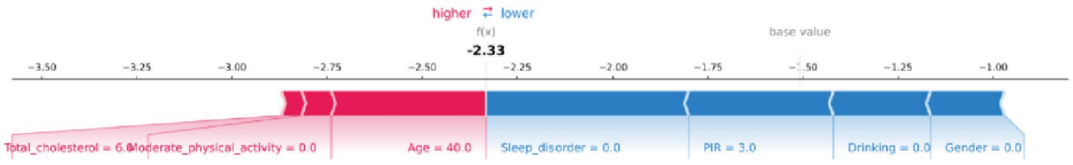

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Application of a web-based predictor on the risk of depression in stroke patients

The final XGBoost model developed in this study is based on seven features that can effectively predict the risk of depression in stroke patients. After inputting the actual values of these seven features, the application automatically calculates and displays the probability that the patient will develop depression. Meanwhile, the force diagram for a single stroke patient shows the features that help determine 'depression': the red features on the left are those that push the prediction into the 'depressed' category, while the blue features on the right are those that push the prediction into the 'non-depressed' category. XGBoost, extreme gradient boosting. The website that predicts the risk of depression in stroke patients is https://prediction-model-for-depression.streamlit.app

learning model utilizing ten features and demonstrated high predictive performance [39], a significant limitation was the lack of application of the SHAP method to explain the model's decision-making process. Machine learning models, particularly deep learning and ensemble methods, are often viewed as "black boxes", rendering their internal mechanisms difficult to interpret. This opacity can lead to confusion among clinicians regarding how predictions are derived, which not only undermines clinical confidence but also hampers the implementation of personalized treatments. To address this challenge, our study employs the SHAP method to provide comprehensive explanations of model outputs. SHAP quantifies the contribution of each feature to the model's predictions, allowing us to identify which factors are pivotal in predicting depression risk in stroke patients. The global interpretation reveals the primary features influencing depression risk, thus offering physicians valuable insights that facilitate the identification of high-risk patients and the development of tailored interventions. Concurrently, local explanations shed light on the predicted outcomes for individual patients, enabling clinicians to understand the specific sources of depression risk for each case—information that is essential for informed clinical decision-making. Additionally, to enhance the usability and convenience of our machine learning model, we developed a tool based on the Streamlit framework, making the predictive model easily accessible via a web interface. This user-friendly platform allows clinicians to conveniently input patient information and instantly retrieve prediction results along with SHAP interpretations. Through these enhancements, we not only improved model interpretability but also promoted knowledge-sharing and collaboration among clinicians. The shareability of this tool enables broader access to and utilization of the predictive model, thereby fostering improved early identification and intervention for depression risk in stroke patients. It is noteworthy that our study did not perform subgroup analyses for different patient populations (such as age, and gender) in terms of model predictions. Therefore, while the model demonstrated good predictive performance in the overall population, its applicability and effectiveness in specific subgroups have yet to be validated. This limitation restricts our comprehensive understanding of the model's practical application and efficacy in various clinical contexts.

There are some limitations of this study that need to be noted. First, the cross-sectional design only identified significant associations between variables, precluding the establishment of causality. Future longitudinal studies or randomized controlled trials could better verify causal relationships. Second, due to the reliance on respondents' subjective judgment and memory in self-reported data, the information reported may be inaccurate or incomplete. This bias can affect the reliability of the study's findings, particularly in the diagnosis of stroke. Therefore, we recommend that future research employ more objective measurement tools to reduce the bias introduced by self-reported data, thereby enhancing the credibility of the research outcomes. Third, the dataset may not comprehensively cover certain key confounding factors, such as the severity of strokes, rehabilitation compliance, and levels of cognitive impairment, which could affect the accuracy of the predictions. Fourth, we recommend considering the use of estimation methods to handle missing data, to make more comprehensive use of the available information, and to reduce potential biases caused by a reduced sample size. This approach will provide more data support for model training, further enhancing the model's generalization ability and performance. Lastly, this study lacks external validation, which may impact the generalizability of the results. When other researchers apply this model to patient populations with different clinical characteristics, the model's performance might differ from the results observed in the NHANES dataset. Therefore, conducting external validation studies is essential to confirm the model's efficacy and reliability in broader populations. Future studies should conduct external validation across diverse independent samples and settings to ascertain the model's reliability and applicability.

Conclusions

We successfully developed an interpretable ML model aimed at predicting depression risk in stroke patients based on clinical data. Following rigorous validation, our XGBoost model demonstrated superior predictive capabilities, establishing a strong foundation for its future application in clinical settings. We hope that this predictive model can serve as an auxiliary tool to help develop more accurate and personalized treatment plans, thereby enhancing the mental health and overall quality of life of stroke patients.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12877-025-05837-5.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements
We thank all the staff in the NHANES for sharing data publicly.

Author contributions
Wenwei Zuo: Writing– original draft, Data curation. Xuelian Yang: Writing– review & editing, Writing– original draft, Formal analysis, Data curation, Conceptualization.

Funding
This research was supported by the General Program of Shanghai Pudong New Area Health Commission (No. PW2022A-12).

Data availability
The data in our study are publicly available online from the NHANES https://www.cdc.gov/nchs/nhanes/index.htm

Declarations

Ethical approval
All study participants gave informed consent following the Institutional Review Board and study ethics guidelines at the Centers for Disease Control and Prevention.

Human Ethics and Consent to Participate declarations
Not applicable.

Competing interests
The authors declare no competing interests.

Clinical trial number
Not applicable.

Received: 14 September 2024 / Accepted: 4 March 2025
Published online: 22 March 2025

References

1. Ayerbe L, Ayis S, Crichton S, Wolfe CD, Rudd AG. The natural history of depression up to 15 years after stroke: the South London stroke register. Stroke. 2013;44(4):1105–10.
2. Chen R, Guo Y, Kuang Y, Zhang Q. Effects of home-based exercise interventions on post-stroke depression: A systematic review and network meta-analysis. Int J Nurs Stud. 2024;152:104698.
3. Carnes-Vendrell A, Deus J, Molina-Seguin J, Pifarré J, Purroy F. Depression and apathy after transient ischemic attack or minor stroke: prevalence, evolution and predictors. Sci Rep. 2019;9(1):16248.
4. Castilla-Guerra L, Fernandez Moreno MDC, Esparrago-Llorca G, Colmenero-Camacho MA. Pharmacological management of post-stroke depression. Expert Rev Neurother. 2020;20(2):157–66.
5. Shin M, Sohn MK, Lee J, Kim DY, Shin YI, Oh GJ, Lee YS, Joo MC, Lee SY, Song MK et al. Post-Stroke depression and cognitive aging: A multicenter, prospective cohort study. J Pers Med 2022, 12(3).
6. Butsing N, Zauszniewski JA, Ruksakulpiwat S, Griffin MTQ, Niyomyart A. Association between post-stroke depression and functional outcomes: A systematic review. PLoS ONE. 2024;19(8):e0309158.
7. Cai W, Stewart R, Mueller C, Li YJ, Shen WD. Poststroke depression and risk of stroke recurrence and mortality: protocol of a meta-analysis and systematic review. BMJ Open. 2018;8(12):e026316.
8. Luo S, Zhang W, Mao R, Huang X, Liu F, Liao Q, Sun D, Chen H, Zhang J, Tian F. Establishment and verification of a nomogram model for predicting the risk of post-stroke depression. PeerJ. 2023;11:e14822.
9. Lan Y, Pan C, Qiu X, Miao J, Sun W, Li G, Zhao X, Zhu Z, Zhu S. Nomogram for persistent Post-Stroke depression and decision curve analysis. Clin Interv Aging. 2022;17:393–403.
10. Zhou L, Chen L, Ma L, Diao S, Qin Y, Fang Q, Li T. A new nomogram including total cerebral small vessel disease burden for individualized prediction of early-onset depression in patients with acute ischemic stroke. Front Aging Neurosci. 2022;14:922530.
11. Huang AA, Huang SY. Use of machine learning to identify risk factors for insomnia. PLoS ONE. 2023;18(4):e0282622.
12. Huang AA, Huang SY. Use of machine learning to identify risk factors for coronary artery disease. PLoS ONE. 2023;18(4):e0284103.
13. Anbarasi J, Kumari R, Ganesh M, Agrawal R. Translational connectomics: overview of machine learning in macroscale connectomics for clinical insights. BMC Neurol. 2024;24(1):364.
14. Ebrahimzadeh E, Fayaz F, Rajabion L, Seraji M, Aflaki F, Hammoud A, Taghizadeh Z, Asgarinejad M, Soltanian-Zadeh H. Machine learning approaches and non-linear processing of extracted components in frontal region to predict rTMS treatment response in major depressive disorder. Front Syst Neurosci. 2023;17:919977.
15. Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, Petzold L, et al. Integrating machine learning and multiscale modeling-perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. NPJ Digit Med. 2019;2:115.
16. Hu J, Xu J, Li M, Jiang Z, Mao J, Feng L, Miao K, Li H, Chen J, Bai Z, et al. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study. EClinicalMedicine. 2024;68:102409.
17. Huang AA, Huang SY. Increasing transparency in machine learning through bootstrap simulation and shapely additive explanations. PLoS ONE. 2023;18(2):e0281922.
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
19. Chen R, Wang K, Chen Q, Zhang M, Yang H, Zhang M, Qi K, Zheng M, Wang Y, He Q. Weekend warrior physical activity pattern is associated with lower depression risk: findings from NHANES 2007–2018. Gen Hosp Psychiatry. 2023;84:165–71.
20. Jørgensen TS, Wium-Andersen IK, Wium-Andersen MK, Jørgensen MB, Prescott E, Maartensson S, Kragh-Andersen P, Osler M. Incidence of depression after stroke, and associated risk factors and mortality outcomes, in a large cohort of Danish patients. JAMA Psychiatry. 2016;73(10):1032–40.
21. Elser H, Caunca M, Rehkopf DH, Andres W, Gottesman RF, Kasner SE, Yaffe K, Schneider ALC. Trends and inequities in the diagnosis and treatment of post-stroke depression: a retrospective cohort study of privately insured patients in the USA, 2003–2020. J Neurol Neurosurg Psychiatry. 2023;94(3):220–6.
22. Yoon CW, Bushnell CD. Stroke in women: A review focused on epidemiology, risk factors, and outcomes. J Stroke. 2023;25(1):2–15.
23. Collishaw S. Annual research review: secular trends in child and adolescent mental health. J Child Psychol Psychiatry. 2015;56(3):370–93.
24. Thapar A, Eyre O, Patel V, Brent D. Depression in young people. Lancet. 2022;400(10352):617–31.
25. Paprocka-Borowicz M, Wiatr M, Ciałowicz M, Borowicz W, Kaczmarek A, Marques A, Murawska-Ciałowicz E. Influence of physical activity and Socio-Economic status on depression and anxiety symptoms in patients after stroke. Int J Environ Res Public Health 2021, 18(15).
26. Cai Q, Qian M, Chen M. Association between socioeconomic status and post-stroke depression in middle-aged and older adults: results from the China health and retirement longitudinal study. BMC Public Health. 2024;24(1):1007.
27. Jeon S, Kang H, Cho I, Cho SI. The alcohol Flushing response is associated with the risk of depression. Sci Rep. 2022;12(1):12569.
28. Almeida OP, Hankey GJ, Yeap BB, Golledge J, Flicker L. The triangular association of ADH1B genetic polymorphism, alcohol consumption and the risk of depression in older men. Mol Psychiatry. 2014;19(9):995–1000.
29. Peele S, Brodsky A. Exploring psychological benefits associated with moderate alcohol use: a necessary corrective to assessments of drinking outcomes? Drug Alcohol Depend. 2000;60(3):221–47.
30. He W, Ruan Y. Poor sleep quality, vitamin D deficiency and depression in the stroke population: A cohort study. J Affect Disord. 2022;308:199–204.

31. Zhu Q, Chen L, Xu Q, Xu J, Zhang L, Wang J. Association between obstructive sleep apnea and risk for post-stroke anxiety: A Chinese hospital-based study in noncardiogenic ischemic stroke patients. Sleep Med. 2023;107:55–63.
32. Li C, Liu Y, Xu P, Fan Q, Gong P, Ding C, Sheng L, Zhang X. Association between obstructive sleep apnea and risk of post-stroke depression: A hospital-based study in ischemic stroke patients. J Stroke Cerebrovasc Dis. 2020;29(8):104876.
33. Lanctôt KL, Lindsay MP, Smith EE, Sahlas DJ, Foley N, Gubitz G, Austin M, Ball K, Bhogal S, Blake T et al. Canadian Stroke Best Practice Recommendations: Mood, Cognition and Fatigue following Stroke, 6th edition update 2019. *Int J Stroke* 2020, 15(6):668–688.
34. Pearce M, Garcia L, Abbas A, Strain T, Schuch FB, Golubic R, Kelly P, Khan S, Utukuri M, Laird Y, et al. Association between physical activity and risk of depression: A systematic review and Meta-analysis. JAMA Psychiatry. 2022;79(6):550–9.
35. Noetel M, Sanders T, Gallardo-Gómez D, Taylor P, Del Pozo Cruz B, van den Hoek D, Smith JJ, Mahoney J, Spathis J, Moresi M, et al. Effect of exercise for depression: systematic review and network meta-analysis of randomised controlled trials. BMJ. 2024;384:e075847.
36. Moreira FP, Jansen K, Cardoso TA, Mondin TC, Magalhães P, Kapczinski F, Souza LDM, da Silva RA, Oses JP, Wiener CD. Metabolic syndrome in subjects with bipolar disorder and major depressive disorder in a current depressive episode: Population-based study: metabolic syndrome in current depressive episode. J Psychiatr Res. 2017;92:119–23.
37. Hu X, Wang T, Luo J, Liang S, Li W, Wu X, Jin F, Wang L. Age-dependent effect of high cholesterol diets on anxiety-like behavior in elevated plus maze test in rats. Behav Brain Funct. 2014;10:30.
38. Mutsuda N, Hamazaki K, Matsumura K, Tsuchida A, Kasamatsu H, Inadera H. Change in cholesterol level during pregnancy and risk of postpartum depressive symptoms: the Japan environment and children's study (JECS). Acta Psychiatr Scand. 2022;145(3):268–77.
39. Chen YM, Chen PC, Lin WC, Hung KC, Chen YB, Hung CF, Wang LJ, Wu CN, Hsu CW, Kao HY. Predicting new-onset post-stroke depression from real-world data using machine learning algorithm. Front Psychiatry. 2023;14:1195586.

Publisher's note