Development and interpretation of a machine learning predictive model for early cognitive impairment in hypertension associated with environmental factors

Xia Zhong[1], Tianen Zhao[2], Shimeng Lv[3], Guangheng Zhang[3], Jing Li[1]*, Donghai Liu[4]* and Huachen Jiao[5]*

[1]Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China, [2]Purchasing Department, Jinan Lixia Dezhengtang Hospital of Traditional Chinese Medicine, Jinan, Shandong, China, [3]Institute of First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China, [4]College of Laboratory Animal Science, Shandong First Medical University, Jinan, Shandong, China, [5]Institute of Cardiology Department, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China

Background and objective: Risk-based predictive models are a reliable tool for early identification of hypertensive cognitive impairment. However, the evidence of the combination of individual factors and natural environmental factors is still insufficient. The aim of this study was to establish a well-performing machine learning (ML) model based on personal and natural environmental factors to help assess the risk of early cognitive impairment in hypertension.

Methods: In this study, a total of 757 Chinese hypertensive patients from from different regions of Shandong Province, China (aged 31–95, male 49.01%) were randomly divided into training group (70%) and verification group (30%). Modelling variables were determined by a 5-fold cross-validated least absolute shrinkage and selection operator (LASSO) regression analysis. Five ML classifiers, XGB (extreme gradient boosting), LR (logistic regression), AdaBoost (adaptive boosting), GNB (gaussian naive bayes), and SVM (support vector machines), have been developed. Area under the ROC curve (AUC), accuracy, sensitivity, specificity, and F1 scores were used to access the model performance. Shape Additive explanation (SHAP) models reveal the feature importance. The clinical performance of the model was evaluated by Decision Curve Analysis (DCA).

Results: Cognitive impairment was diagnosed in 17.44% ($n = 132$). LASSO regression analyses suggested that age, waist circumference, urban green coverage, educational levels, annual sunshine hours, and area whole-day average noise were considered significant predictors of early cognitive impairment in hypertension. The obtained XGBoost model yielded good predictive performance with the AUC (0.893), F1 score (0.627), accuracy (0.837), sensitivity (0.780), and specificity (0.853). The predictive model's clinical net benefit was confirmed through DCA analysis.

Conclusion: The XGBoost model developed based on personal factors and natural environmental factors can predict early cognitive impairment of hypertension with superior predictive performance. Larger population cohorts are needed in the future to validate these findings and potentially enhance the ability to identify the occurrence of early cognitive impairment in people with hypertension.

KEYWORDS

environmental exposure factor, cognitive impairment, hypertension, predictive model, machine learning

1 Introduction

Dementia, an increasingly prevalent and challenging neurodegenerative disease, affects about 50 million people worldwide (1), a number that is expected to triple by 2050 (2). Cognitive impairment is the preclinical stage of dementia, and active prevention can reduce the likelihood of developing dementia. Hypertension is one of the major risk factors for cognitive impairment (3) and is associated with a 1.62-fold increase in the risk of cognitive impairment (4). Currently, there is no conclusive evidence supporting the role of pharmacological therapies in preventing cognitive decline (5). In recent years, there has been a rapid increase in the number of patients with hypertensive cognitive disorders seeking medical help and counselling. The specific mechanisms that trigger and promote cognitive decline are unclear, and treatment options are limited, so there is an urgent need to develop more precise treatment strategies for hypertensive cognitive disorders.

Risk factors for cognitive impairment have not been fully explored. Most studies have focused on individual factors such as aging, smoking, physical inactivity, low levels of education, obesity, vitamin D deficiency, diabetes, and hypertension (6, 7), but these cannot fully explain cognitive impairment. Increasing evidence supports the need to focus the risk of cognitive impairment on environmental factors. In addition, it is increasingly recognized that environmental factors may hold promise for future predictive methods and are more universal than brain imaging markers and cerebrospinal fluid (CSF). Several studies have reported the relationship between natural environmental factors and cognitive impairment, including climate (8), greenness (9), air pollution (10), noise exposure (11), etc. These modifiable natural environmental factors are considered promising predictors of cognitive impairment, despite varying reported results.

Recently, related risk factors and predictive models for hypertensive cognitive impairment developed based on these risk factors have been preliminarily explored, especially in the Chinese population. Zhang et al. (12) showed that plateau environment, age, abdominal circumference, and serum uric acid (SUA) were independent risk factors for hypertensive cognitive impairment. Li et al. (13) reported that duration of hypertension, systolic blood pressure (SBP), homocysteine (Hcy), and SUA were risk factors for cognitive dysfunction, and duration of education was negatively correlated with cognitive dysfunction. Ma et al. (14) also revealed that low levels of education, elevated body mass index (BMI) and waist- to- height ratio (WHR) were independent risk factors for hypertensive cognitive impairment. Lu et al. (15) developed a predictive model of hypertensive cognitive impairment based on a number of influencing factors including hypertension grade, smoking, sleep disorder, and duration of hypertension, and the AUC, sensitivity, and specificity of the model were 0.765, 0.630, and 0.877, respectively. Currently recognized cognitive impairment is more likely to involve a combination of factors (5), and prevention strategies based on modifiable factors appear to be more important. To the best of our knowledge, a predictive model based on personal and natural environmental factors for early cognitive impairment in hypertensive patients has not been developed.

Machine learning (ML) can be based on important information modeling, which helps reveal the relationship between factors and diseases in complex data environments (16). ML techniques have shown benefits in developing risk prediction models for cardiovascular and cerebrovascular events (17–20). Furthermore, predictive models for cognitive impairment based on ML algorithms have also been reported (21–24). However, evidence for ML-based risk prediction models for early cognitive impairment in hypertension remains limited, especially when environmental factors are involved.

Therefore, we aimed to develop for the first time a superior ML prediction model that considers modifiable personal and environmental exposure factors to predict the risk of early cognitive impairment in hypertension, with the hope of providing optimal strategies for the early diagnosis and management of hypertensive cognitive impairment.
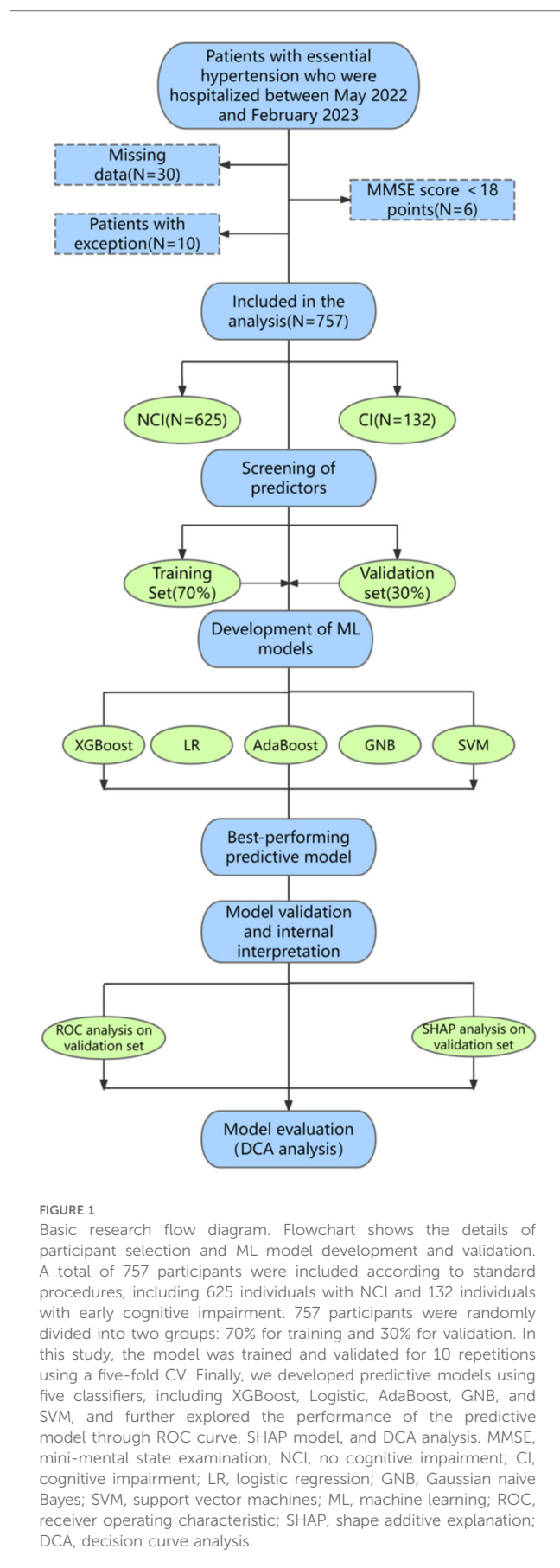
2 Methods

2.1 Participants and study design

This study received approval from the Institutional Review Board (IRB) of the Affiliated Hospital of Shandong University of Traditional Chinese Medicine [Approval Number: (2023) Ethics Review No. (109)-KY] with the informed consent of all participants. In this study, the research area is Shandong Province, China (115–120°E and 35–38°N). We collected 803 hypertension patients from 8 hospitals in 5 prefecture-level cities by stratified sampling between May 2022 and February 2024. Finally, a total of 757 hypertensive patients recruited from four prefecture-level hospitals in Yantai, Jinan, Weifang and Dongying in Shandong Province were eligible for the development and validation of the prediction model. Participants were included if they were (1) 30 years of age or older; and (2) essential hypertension. Participants were excluded if they (1) had a history of neurological disease (e.g., Parkinson's disease, stroke, epilepsy, brain tumor, brain trauma, mental or psychiatric illness, and dementia) or (2) had a history of cardiovascular disease (e.g., severe arrhythmia, heart failure, and cardiac surgery) or (3) had severe vision or hearing impairment. A detailed study flow diagram of the selection of participants is shown in Figure 1. For better analysis, we excluded 30 individuals with missing data, 10 individuals with abnormal data, and 6 individuals with the Mini-Mental State Examination (MMSE) scores below 18 points. Data collection and quality control were standardized across all centers through electronic medical records and centralized the Standard Operating Procedure (SOP) training, with key variables (e.g., MMSE) demonstrating high inter-site reliability (Kappa >0.8, ICC >0.9).

2.2 Assessment of hypertension and early cognitive impairment

All patients were diagnosed by experienced cardiologists using the following criteria. Diagnostic criteria for hypertension are (25):

FIGURE 1
Basic research flow diagram. Flowchart shows the details of participant selection and ML model development and validation. A total of 757 participants were included according to standard procedures, including 625 individuals with NCI and 132 individuals with early cognitive impairment. 757 participants were randomly divided into two groups: 70% for training and 30% for validation. In this study, the model was trained and validated for 10 repetitions using a five-fold CV. Finally, we developed predictive models using five classifiers, including XGBoost, Logistic, AdaBoost, GNB, and SVM, and further explored the performance of the predictive model through ROC curve, SHAP model, and DCA analysis. MMSE, mini-mental state examination; NCI, no cognitive impairment; CI, cognitive impairment; LR, logistic regression; GNB, Gaussian naive Bayes; SVM, support vector machines; ML, machine learning; ROC, receiver operating characteristic; SHAP, shape additive explanation; DCA, decision curve analysis.

(1) systolic blood pressure (SBP) ≥ 140 mmHg, or (2) diastolic blood pressure (DBP) ≥ 90 mmHg, (3) and/or the use of antihypertensive medications. In addition, early cognitive impairment was determined by the Chinese version of the MMSE assessment questionnaire. MMSE scores range from 0 to 30, with low scores representing poor cognitive function (26). MMSE scores above 27 are considered normal cognitive function, and MMSE scores above 18 and below 27 are identified as early cognitive impairment (27, 28).

2.3 Predictors and feature selection

Based on previously reported cognitive impairment factors, personal factors (age, sex, estimated duration of hypertension, levels of education, type of work, smoking, drinking, BMI, waist circumference, waist-to-hip ratio) and natural environmental factors (air temperature, annual sunshine hours, urban precipitation, relative humidity, area whole-day average noise value, traffic average noise value, per capita public green areas, urban green coverage) were analyzed. Basic information about the participants, including age, sex, estimated duration of hypertension, education levels, type of work, smoking and drinking, was obtained through questionnaires. Anthropometric indicators including weight (kg), height (cm), hip circumference (cm) and waist circumference (cm) were standardized measurements. BMI was obtained by dividing weight into kilograms by height in meters squared in kg/m^2 (29). WHR was obtained by dividing waist circumference (cm) by hip circumference (cm). We obtained annual average data from the Shandong Meteorological Bureau and local authoritative environmental monitoring stations for 2018–2022, including air temperature, annual sunshine hours, urban precipitation, relative humidity, area whole-day average noise value, traffic average noise value, per capita public green areas, urban green coverage. Air temperature (°C) is the temperature of the air and is calculated by dividing the average monthly temperature over 12 months by 12. Annual sunshine hours (h) are the number of hours the sun actually hits the ground and are calculated by adding up the number of hours of sunshine over 12 months. Precipitation (mm) is the depth of liquid or solid water (after melting) that falls from the sky to the ground and accumulates on the ground without evaporation, penetration, or loss and is calculated by the cumulative amount of monthly precipitation over 12 months. Relative humidity (%) is the ratio of the actual water pressure in the air to saturated water pressure at the prevailing temperature and is calculated by adding up the average monthly humidity over 12 months and dividing by 12. Monitoring sites used to measure noise include major urban traffic intersections and residential areas. Area whole-day average noise value (db) mainly includes industrial noise, traffic noise, construction noise, social noise. Traffic average noise value (db) depends mainly on the load or traffic flow on the road (30). Traffic flow is determined by the number of vehicles passing

through the center of the intersection in a unit of time. Per capita public green areas (m^2) in built-up areas refers to the average area of public green space occupied by each resident of a city. Urban green coverage rate (%) is obtained by the ratio of green coverage area to built-up area in urban built-up area.

We chose LASSO regression for feature selection because of its ability to handle high-dimensional, small sample datasets through L1 regularization, which automatically eliminates irrelevant features by shrinking their coefficients to zero while retaining critical predictors (31). This approach balances sparsity and computational efficiency, addressing overfitting risks inherent in limited data. Though recursive feature elimination (RFE) and Shape Additive explanation (SHAP) -based methods offer interpretability (32, 33), their limitations outweighed their benefits: Computational Costs: RFE's iterative process and SHAP's prediction dependency become infeasible with large feature sets. Noise Sensitivity: Both methods struggle with redundant/noisy features common in small samples, risking false-positive selections. Scalability Issues: RFE's recursive elimination and SHAP's permutation evaluations degrade performance as feature counts increase. LASSO's integrated feature selection and regularization mechanism eliminated iterative processing needs, making it uniquely suited for our dataset.

2.4 Outcomes

In this study, a total of 757 patients with hypertension were included for analysis, of which 132 (17.44%) participants were identified as having early cognitive impairment. Finally, we screened 6 core predictors from 18 variables, including 3 personal parameters (age, waist circumference, educational levels) and 3 natural environmental parameters (urban green coverage, annual sunshine hours, and area whole-day average noise value), to develop and validate the predictive model of cognitive impairment in hypertension.

2.5 Development and validation of Ml models

The flow of machine learning model development and validation is shown in Figure 1. First, we selected predictors and randomly assigned participants to two groups: 70% for model development and training, and 30% for model evaluation and validation to prevent model overfitting. A five-fold CV was used to train and verify the model. Then, ROC curve analysis was performed to evaluate and compare the predictive performance of five ML classifiers [XGBoost, logistic regression (LR), AdaBoost, GNB, and SVM], including AUC, sensitivity, specificity, accuracy, and F1 scores. Finally, we further explored the significance of predictors and the clinical applicability of the predictive model through SHAP and DCA analysis. According to the classification confusion matrix (34), the ML model is defined as true positive (TP) and true negative (TN) if it can correctly predict cognitive impairment; conversely, the ML model was defined as false positive (FP) or

false negative (FN). Sensitivity is defined as the percentage of samples that test positive and is calculated by the formula: Sensitivity = TP/(TP + FN). Specificity is defined as the percentage of samples that are actually negative that are judged to be negative, and the formula is calculated as follows: specificity = TN/(TN + FP). Accuracy refers to the proportion of samples correctly classified for a given data in the total samples, and its calculation formula is: Accuracy = (TP + TN)/(TP + TN + FP + FN). F1 score is used to measure the overall performance of the classifier, and its formula is as follows: F1 score = 2×precision × recall/(precision + recall).

2.6 Statistical analyses

Continuous data were represented by mean ± standard deviation (SD) or median [25th, 75th], while categorical data were expressed by number (%). We performed the t-test, Mannwhitney-U test, Analysis of Variance (ANOVA), and Chi-square test to select variables with comparative differences between groups for further LASSO regression dimension reduction, and screened predictors for modeling through five-fold cross-validation (CV). All patients were randomly assigned to a training or validation group (7:3), and predictive models were developed using five classifiers (XGBoost, LR, AdaBoost, GNB, and SVM). We developed ROC curves and compared the AUC, accuracy, sensitivity, specificity, and F1 scores of the five models. Finally, we used SHAP analysis to determine the importance of features in the predictive model, and developed DCA curves to evaluate the clinical applicability of the predictive model. All statistical analysis results were obtained using Python version 3.7 and R 3.6.3. A $p < 0.05$ was considered statistically significant, and a 2-sided test was performed.

3 Results

3.1 Comparison of personal and environmental factors between early cognitive impairment and controls

Table 1 shows the personal and environmental factors for all patients. The mean age of all participants was 67.11 ± 11.47 years, with 48.98% male. Of the 757 participants, 132 (17.44%) had early cognitive impairment. Compared to NCI, individuals with early cognitive impairment were older, had a longer estimated duration of hypertension, lower educational attainment, and larger waist circumference, and were more likely to perform manual labor (all $p < 0.05$). In addition, participants with early cognitive impairment lived in areas with lower temperatures and urban green coverage, fewer hours of annual sunshine, higher relative humidity, and area whole-day average noise value (all $p < 0.05$). However, there were no significant differences between the two groups in gender, smoking, drinking, BMI, waist-to-hip ratio, urban precipitation, traffic average noise value, and per capita public green areas (all $p > 0.05$).

3.2 Predictors were screened by LASSO regression analysis

For the simplicity of the model, we performed LASSO regression analysis to reduce the dimension of 10 indicators with statistical differences in Table 1, including age, estimated duration of hypertension, educational levels, type of work, waist circumference, air temperature, annual sunshine hours, relative humidity, area whole-day average noise value, and urban green coverage. As shown in Figure 2, as the log (λ) value increases, the normalization coefficients of the 10 candidate parameters are compressed to varying degrees until all changes are zero (35). Finally, we selected four predictors for model development, including age, education levels, waist circumference, annual sunshine hours, area whole-day average noise, and urban green coverage.

3.3 Development of predictive models based on five classifiers

Figure 3 shows the performance of predictive models based on different classifiers. Current results suggest that the best performance model was the XGBoost model, with an AUC of 0.893, accuracy of 0.837, sensitivity of 0.780, specificity of 0.853, and F1 score of 0.627. The AdaBoost model ranked second in performance, with AUC of 0.854, accuracy of 0.715, sensitivity of

0.902, specificity of 0.672, and F1 score of 0.519. Compared with the other four models, the SVM model has poor performance, with AUC of 0.645, accuracy of 0.628, sensitivity of 0.750, specificity of 0.605, and F1 score of 0.427. Figure 4 shows the training and validation ROC curves for the five classifiers. The AUC of the XGBoost model in the training set was 0.893 (Figure 4A), and the AUC of the validation set was 0.770 (Figure 4B). The predictive performance (including accuracy, sensitivity, specificity, and F1 score) of the XGBoost model was significantly better than that of the other four classifiers (all P-values < 0.05).

3.4 Outcomes of XGBoost model evaluation

Figure 5 shows the evaluation of the XGBoost model. Based on a five-fold CV, the mean AUC for the XGBoost training and validation model was 0.880 (Figure 5A) and 0.804 (Figure 5B). The predicted probability of early cognitive impairment was positively correlated with the actual probability, and the XGBoost model had excellent calibration ($P > 0.05$).
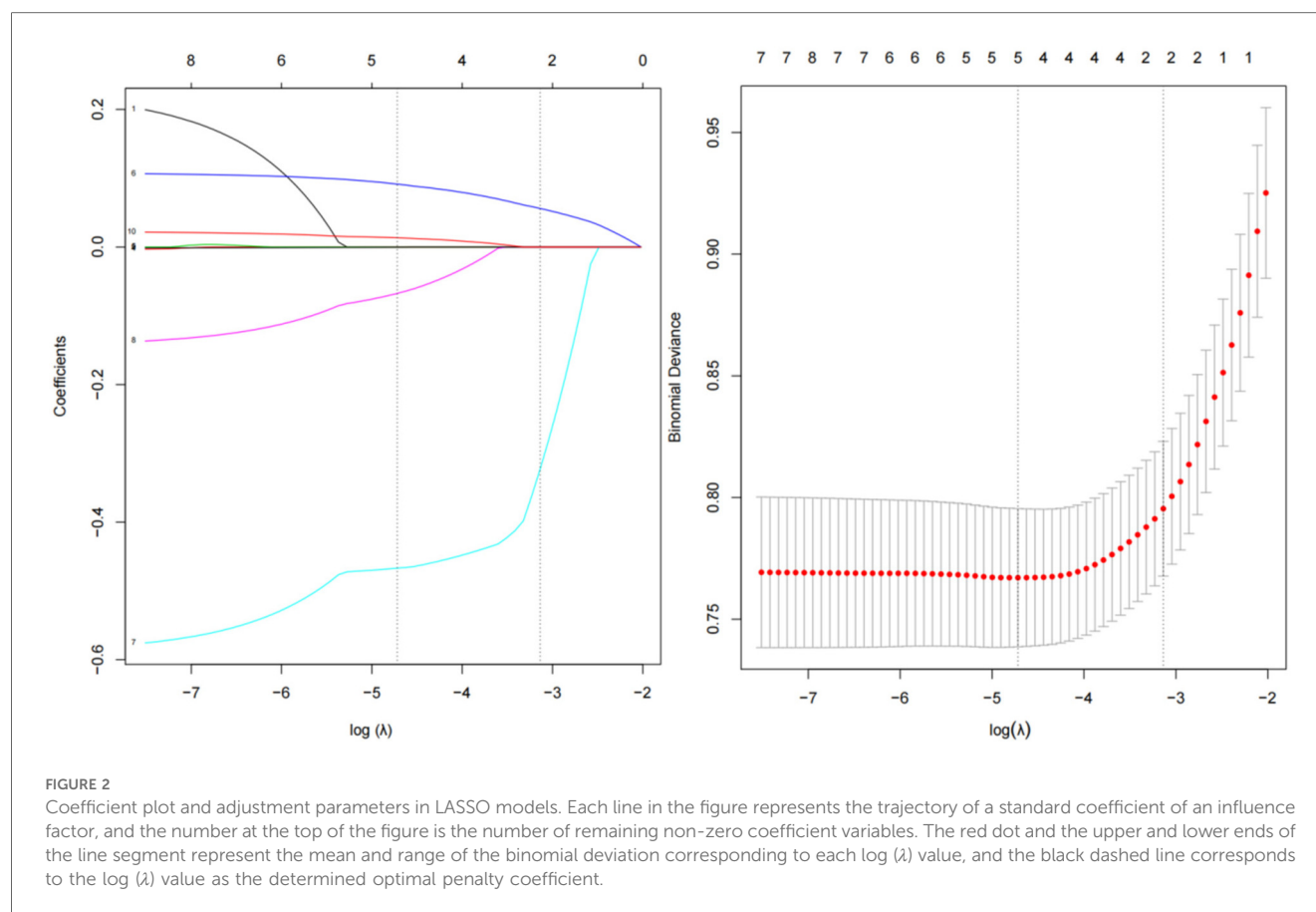
3.5 DCA analysis of predictive model

To evaluate the clinical application of the model, we performed DCA analysis. Figure 6 show the threshold probability for XGBoost

TABLE 1 Comparison of personal and environmental factors between cognitive impairment and controls.

Indicators	Total (N = 757)	NCI (N = 625)	CI (N = 132)	P value
Personal factors				
Age, years	67.11 ± 11.47	65.27 ± 11.21	75.83 ± 8.22	<0.001*
Sex (men), n (%)	371 (49.01)	307 (49.12)	64 (48.49)	0.894
Estimated duration of hypertension, months, median [IQR]	117.00 [59.00, 179.00]	109.00 [58.00, 176.00]	134.00 [67.00, 212.00]	0.017*
Educational levels, n (%)				<0.001*
Primary school or below	299 (39.50)	209 (33.44)	90 (68.18)	
Junior high school or senior high school	411 (54.29)	369 (59.04)	42 (31.82)	
University or above	47 (6.21)	47 (7.52)	0 (0.00)	
Type of work, n (%)				<0.001*
Manual	406 (53.63)	306 (48.96)	100 (75.76)	
Mental	104 (13.74)	100 (16.00)	4 (3.03)	
Both manual and mental	247 (32.63)	219 (35.04)	28 (21.21)	
Current drinker, n (%)	195 (25.76)	168 (26.88)	27 (20.46)	0.125
Current smoker, n (%)	231 (30.52)	199 (31.84)	32 (24.24)	0.085
BMI, kg/m^2, mean ± SD	24.97 ± 3.30	24.87 ± 3.31	25.45 ± 3.19	0.068
Waist circumference, cm, mean ± SD	84.02 ± 15.38	83.28 ± 15.76	87.53 ± 12.83	0.001*
Waist-to-hip ratio, mean ± SD	0.86 ± 0.12	0.86 ± 0.13	0.87 ± 0.09	0.494
Environmental factors				
Air temperature, °C, mean ± SD	13.92 ± 0.83	13.96 ± 0.85	13.71 ± 0.74	<0.001*
Annual sunshine hours, h, mean ± SD	2,374.66 ± 39.90	2,376.05 ± 40.84	2,368.08 ± 34.35	0.021*
Urban precipitation, mm, mean ± SD	775.17 ± 23.04	775.35 ± 23.95	774.28 ± 18.09	0.562
Relative humidity, %, mean ± SD	64.10 ± 3.17	63.97 ± 3.22	64.73 ± 2.85	0.007*
Area whole-day average noise value, db, mean ± SD	54.46 ± 0.93	54.42 ± 0.97	54.65 ± 0.69	0.001*
Traffic average noise value, db, mean ± SD	67.38 ± 0.71	67.40 ± 0.72	67.28 ± 0.61	0.051
Per capita public green areas, sq.m/person, mean ± SD	16.34 ± 1.64	16.30 ± 1.69	16.53 ± 1.37	0.092
Urban green coverage, %, mean ± SD	41.02 ± 0.43	41.03 ± 0.44	40.92 ± 0.32	<0.001*

NCI, no cognitive impairment; CI, cognitive impairment; BMI, body mass index.
*Statistically significant ($P < 0.05$).

FIGURE 2
Coefficient plot and adjustment parameters in LASSO models. Each line in the figure represents the trajectory of a standard coefficient of an influence factor, and the number at the top of the figure is the number of remaining non-zero coefficient variables. The red dot and the upper and lower ends of the line segment represent the mean and range of the binomial deviation corresponding to each log (λ) value, and the black dashed line corresponds to the log (λ) value as the determined optimal penalty coefficient.

Model. Primary analysis demonstrated optimal performance within 15% (Δnet benefit = 0.041) - 25% (Δnet benefit = 0.068) threshold range. The maximum net benefit of 0.085 (95% CI: 0.062–0.109) was obtained at the 20% threshold. Specifically, at a threshold probability of 0.20 (representing 20% risk of early cognitive impairment), the XGBoost model demonstrated superior clinical utility with a maximum net benefit of 0.085 (95% CI: 0.062–0.109). This corresponds to identifying 8.5 true positive cases per 100 patients while avoiding unnecessary interventions in 14.7% of cases (false positives), translating to a risk-benefit ratio of 1:1.72.

3.6 Feature importance of predictive model

Figures 7A, B show the feature importance of the XGBoost model based on SHAP analysis. SHAP values represent the combination of feature importance and feature effects. Figure 7B shows the order of importance of the features: age > waist circumference > urban green coverage > educational levels > hours of annual sunshine > area whole-day average noise. Current results indicate that age is the most important feature in predicting Shapley value. Aging is positively associated with Shapley value, and older age was more likely to be predicted to have cognitive impairment. Secondary to age is waist circumference. Having a larger waist circumference, higher area whole-day average noise (colored pink) was associated with

Shapley values and was a positive predictor of early cognitive impairment in hypertension. Having lower urban green coverage and educational attainment, fewer hours of annual sunshine (blue color) were associated with Shapley values and were negative predictors of early cognitive impairment in hypertension. In summary, age, waist circumference, and area whole-day average noise were positive predictors of early cognitive impairment in hypertension, while urban green coverage, educational levels, and annual sunshine hours were negative predictors of cognitive impairment in hypertension. It suggests the threat of personal factors (older age, larger waist circumference, lower education level) and environmental factors (lower urban green coverage, fewer hours of annual sunshine, and higher area whole-day average noise value) to the risk of early cognitive impairment in hypertension.

4 Discussion

Our findings suggested that age, waist circumference, urban green coverage, educational levels, annual sunshine hours, and area whole-day average noise were significant predictors of early cognitive impairment in hypertension. Five ML models of XGBoost, LR, AdaBoost, GNB and SVM based on multiple predictors were developed and validated. The XGBoost model performed significantly better than other classifier models, with
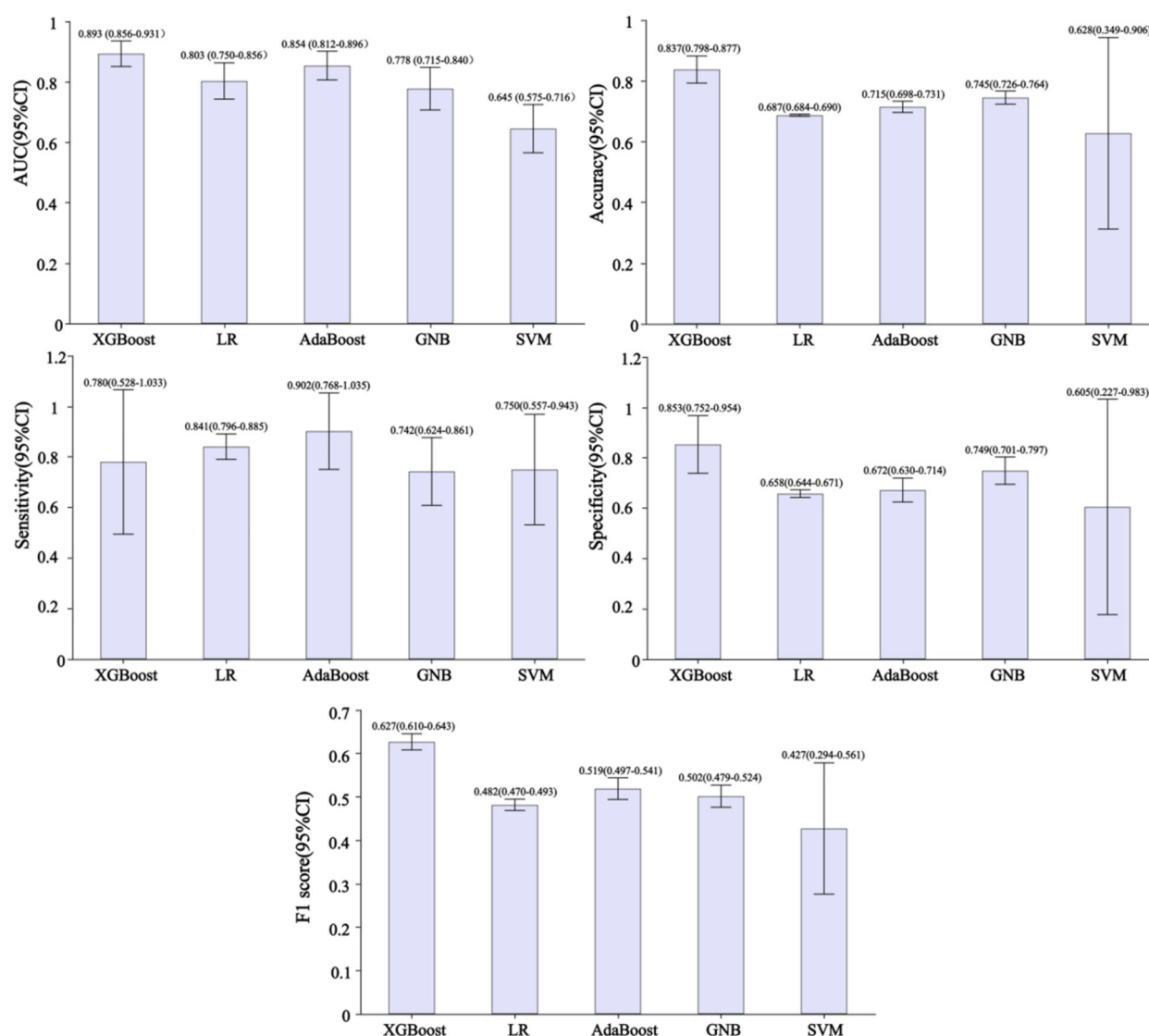
FIGURE 3
Performance of developed predictive models based on different classifiers. AUC, area under the curve; XGBoost, extreme gradient boosting; LR, logistic regression; AdaBoost, adaptive boosting; GNB, Gaussian naive Bayes; SVM, support vector machines.

AUC (0.893), sensitivity (0.780), specificity (0.853), accuracy (0.837), and F1 score (0.627). DCA analysis shows that the current predictive model has excellent clinical practicability. These findings may provide evidence for diagnosis and risk prediction of early cognitive impairment in hypertension.

Evidence for the use of predictors in clinical practice remains highly controversial and evidence-based data is lacking to recommend a diagnostic assessment of an individual risk for preclinical stages of cognitive impairment. The predictive model of cognitive impairment based on risk factors is still in the exploratory stage, but promising. To date, most of the development of predictive models for cognitive impairment has focused on a single personal factor, while designs for environmental factors have not been reported. Recently, there has been growing recognition that cognitive decline is often attributed to a combination of risk factors (5), in which both personal and natural environmental factors are

considered to be significant influencing factors on cognitive impairment. In this study, with the advantages of multicollinearity and overfitting among effective control variables (36), LASSO regression helped us screen for six predictors of early cognitive impairment in hypertension, including age, education level, waist circumference, annual sunshine duration, area whole-day average noise, and urban green coverage. Further SHAP analysis revealed the feature importance of the predictive model.

At present, exploring the relationship between personal factors and cognitive impairment is still mainstream. In this study, we observed that aging is the most important risk factor. Aging is a common risk factor for hypertension and cognitive decline. Multidimensional homeostasis dysfunction and impaired cellular stress recovery caused by aging can accelerate the destruction of the cerebral blood supply and blood-brain barrier in patients with chronic hypertension, leading to cognitive decline (3).
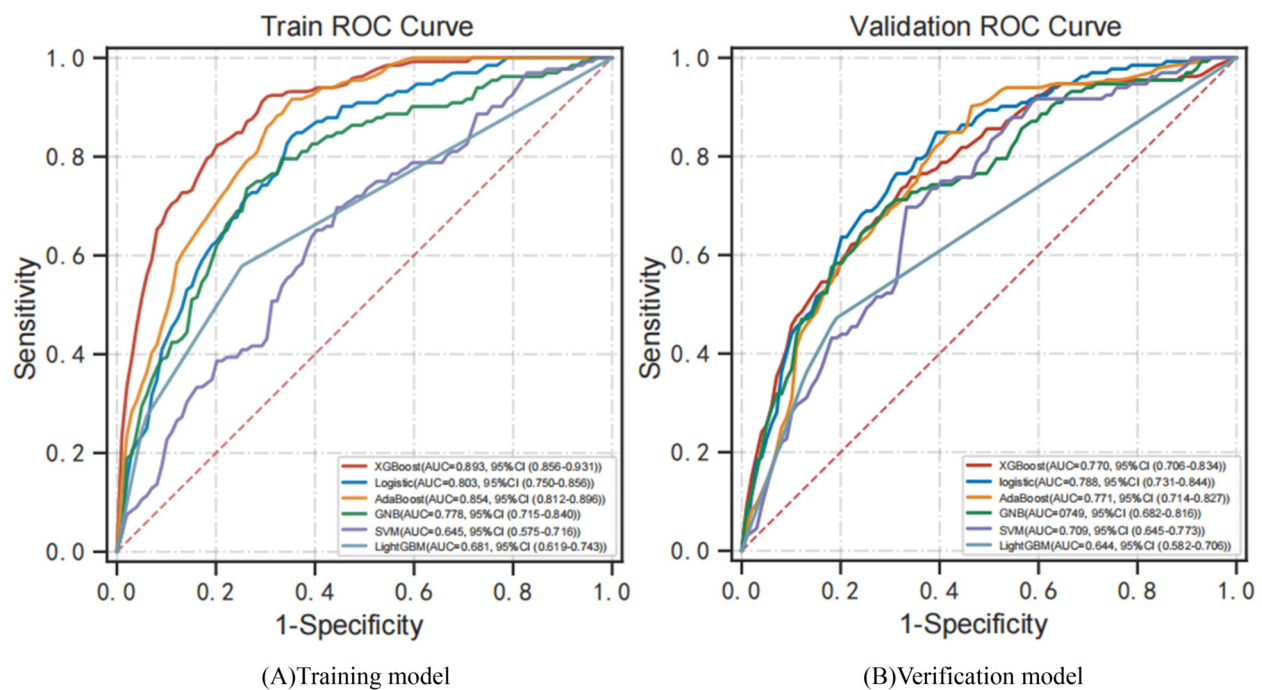
FIGURE 4
ROC curves for models developed with six predictors. (A) Training ROC analysis results of five classifiers. (B) Verification ROC analysis results for five classifiers. ROC curve, receiver operating characteristic curve; AUC, area under the curve; XGBoost, extreme gradient boosting; AdaBoost, adaptive boosting; GNB, Gaussian naive Bayes; SVM, support vector machines.
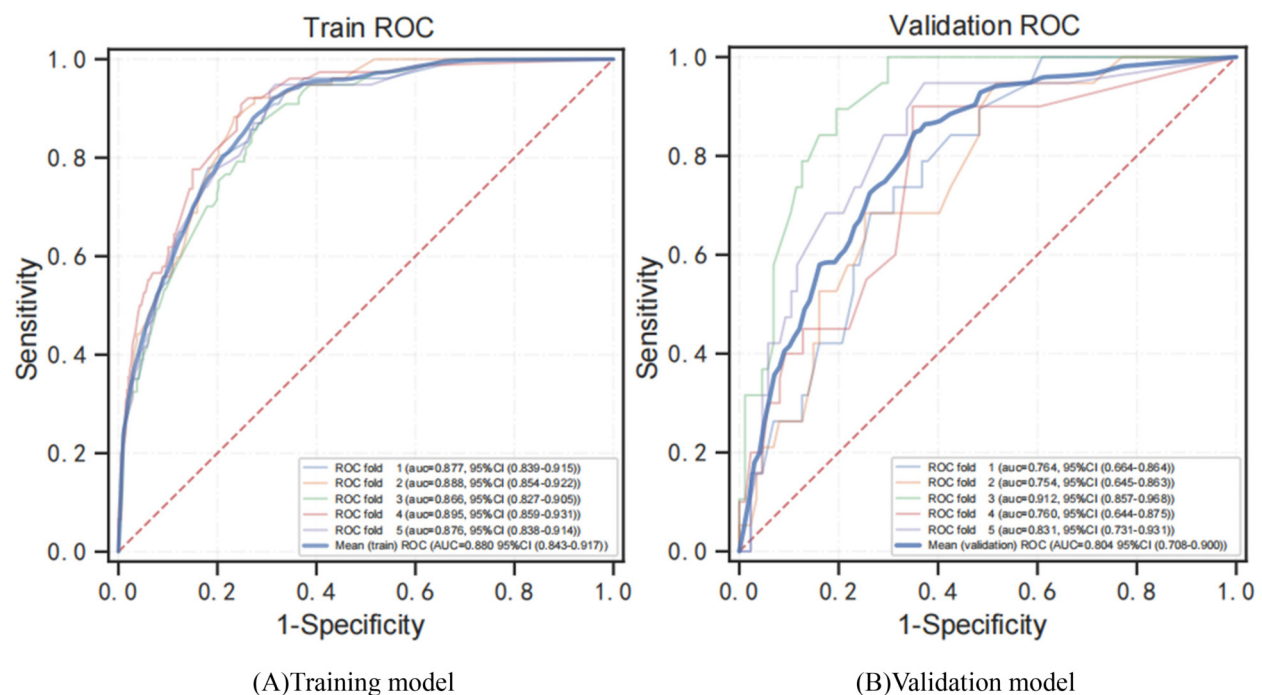

FIGURE 5
ROC curves for training and verification of the XGBoost model by 5-fold CV. (A) ROC analysis results of the XGBoost training model. (B) ROC analysis results of the XGB verification model. ROC curve, receiver operating characteristic curve; AUC, area under curve; XGBoost, adaptive boosting.
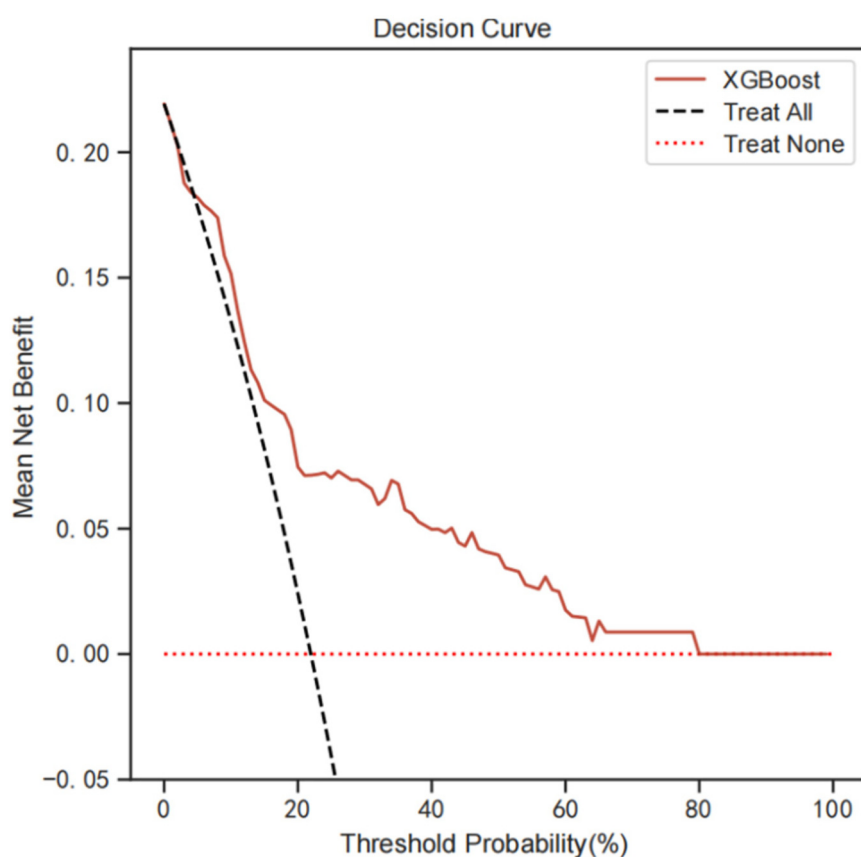
FIGURE 6
DCA analysis was performed to evaluate the clinical usefulness of the XGB model. The *y*-axis indicated the net benefit; the *x*-axis indicated the threshold probability. The solid red line shows the net benefit rate of the XGB forecast model. Within a certain threshold range, the XGB model has a higher net benefit. DCA, decision curve analysis.
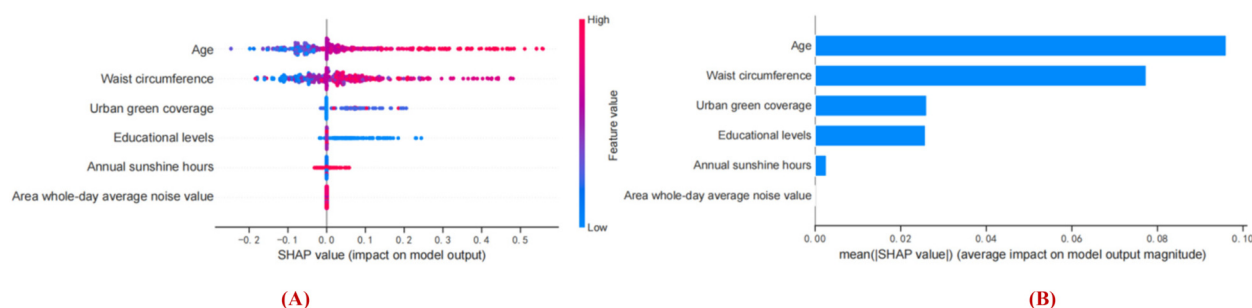

FIGURE 7
Interpretation of the obtained model based on SHAP analysis. **(A)** Prediction direction of model features. **(B)** Ranking the importance of model features. The vertical axis of the graph represents features and the horizontal axis represents SHAP observations. Each point represents a feature and a Shapley value, which represents the contribution of each feature to the model output. Feature values are represented by color, and feature importance is arranged from top to bottom along the *Y*-axis. Pink shows a positive association with early cognitive impairment in hypertension, blue shows a negative association with early cognitive impairment in hypertension.

Epidemiological evidence strongly support that hypertension is associated with cognitive decline, especially in midlife (37). A Finnish study reported that middle-aged people with SBP ≥ 160 mmHg were independently associated with a >2-fold increased risk of Alzheimer's disease (AD) (38). In older adults,

however, the link between hypertension and cognitive function remains controversial (39). Waist circumference, an important measure of central obesity, was observed in this study as a secondary risk factor for cognitive impairment in hypertension. Several recent studies have also reported a strong correlation

between waist circumference and cognitive decline (40–42). As waist circumference increases, chronic low inflammation associated with obesity can lead to multiple chronic diseases and neuroinflammation, which is an important pathway for cognitive decline (43). Previous studies have shown that that educational attainment can have a positive impact on cognitive function (44, 45). A similar conclusion was observed in our study, which showed an independent positive association between low educational attainment and cognitive impairment in hypertension. Emerging evidence suggests that the effect of education level on cognitive function in late-life is associated with promoting individual differences in cognitive ability that occur in early adulthood but persist into old age (44). Interestingly, we observed that patients with early cognitive impairment simultaneously exhibited larger waist circumference and a higher proportion of engaging in physical labor; however, this finding is not contradictory. Many of these individuals are manual laborers, often shift workers, who typically have higher alcohol consumption, larger waistlines, and increased energy intake (46). Evidence strongly links shift work to a higher risk of abdominal obesity (47), likely due to metabolic dysfunction from circadian rhythm disruption. Thus, these findings indicate shared pathophysiological mechanisms, such as metabolic disturbances from circadian dysrhythms, rather than contradictory outcomes.

Recently, increasing attention has been paid to the impact of environmental factors on cognitive function. In this study, we observed less annual sunlight hours as a risk factor for cognitive impairment in hypertension. Sunlight exposure is an important factor affecting circadian rhythm and sleep-wake time (48), and disruption of circadian rhythm and lack of sunlight are related to mood fluctuations (49), all of which may have an impact on cognitive function. Both Zhu et al. (50) and Chantranupong et al. (51) have reported that sunlight can alter mood, behavior, and cognition, while enhancing learning and memory. Moreover, previous studies consistently indicated that long-term residential natural sunlight exposure is associated with a lower risk of cognitive decline (49, 52–54). In general, less annual sunshine means that patients live in rainy and hazy weather, high humidity, low oxygen content in the air, low sympathetic excitability, and easy to produce mood swings. In addition, the relative decline in outdoor activities is not conducive to interpersonal interaction, and may induce insomnia, anxiety, and depression, which can lead to cognitive decline. Area whole-day average noise is another important environmental factor in current models for predicting cognitive impairment in hypertension. Similar to our study, a large number of scholars have also demonstrated that chronic noise exposure (CNE) is an independent risk factor for cognitive impairment (55–58). Animal research has shown that CNE can cause neuroendocrine disorders, overactivation of sympathetic regions of the autonomic nervous system, and an increase in stress hormones that affect the brain and behavior (59), meanwhile, noise exposure can also trigger endothelial and neuronal dysfunction, which activates inflammatory and oxidative stress pathways (60), which may explain the adverse cognitive effects of noise. In addition, anxiety and depression caused by environmental noise exposure have also become longitudinal

predictors of impaired cognitive function (61, 62). Green spaces are conducive to better cognitive function. In this study, urban green coverage was also identified as a significant factor in early cognitive impairment in hypertension. Urban greening benefits residents by regulating climate, air quality and water resources, and a comfortable experience can bring a good mood to residents. Hu et al. (63) reported that participants with the most green space had a 20% lower risk of cognitive impairment [hazard ratio (HR): 0.80, 95% CI: 0.73–0.89]. Recent systematic review and meta-analysis (57) suggested that the protective effect of more greenness (OR = 0.97, 95% CI: 0.95–0.995) on cognitive function. In addition, a population-based cohort study (9) also highlighted the protective effects of greenness on cognitive function. Taken together, our findings may shed light on future research to better understand the role of natural environmental factors in the association between cognitive impairment.

Today, ML has provided a new paradigm for monitoring cognitive impairment, but there is little evidence of its application to hypertensive cognitive impairment. Hu et al. (21) established a logistic regression model for Chinese community-dwelling elderly with normal cognition using age, instrumental activities of daily living, marital status, and baseline cognitive function, with a consistency index of 0.814 (95% CI 0.781–0.846). Yadgir et al. (22) developed an ML model to assist in screening for cognitive impairment in the emergency department, the best performing algorithm was the XGBoost model with an AUC of 0.72, sensitivity of 0.73, and specificity of 0.64. Tan et al. (34) developed an ensemble model based on three classifiers using age, race, highest education, and neuroimaging markers with an AUC of 0.80, accuracy of 0.83, F1 score of 0.87, sensitivity of 0.86, and specificity of 0.74. Liu et al. (24) noted that random forest models had high accuracy for all outcomes at Year 2 (AUC = 0.81), Year 4 (AUC = 0.79), and cross-sectional Year 4 (AUC = 0.80). Ciarmiello et al. (64) developed a deep learning model using radiomic and amyloid PET load, with an AUC of 0.71, accuracy of 0.57, and F1 score of 0.48. Na et al. (65) developed a gradient enhancer (GBM) model to predict cognitive impairment after 2 years, which showed good performance with sensitivity of 0.967, specificity of 0.825, and AUC of 0.921. Similar ML predictive models for cognitive impairment have also been reported in other studies as well (66–69). According to the latest systematic review (70), SVM is the most common machine model developed to predict dementia, with an average accuracy of 75.4%, while convolutional neural network (CNN) has a higher average accuracy of 78.5%. In this study, XGBoost showed the best predictive performance in our prediction models based on five classifiers, with AUC (0.893), sensitivity (0.780), specificity (0.853), accuracy (0.837), and F1 score (0.627). The several recent studies conducted by Du et al. (71) have consistently demonstrated the superior performance of the XGBoost model in disease prediction and diagnosis. Our developed model has some advantages. Our previous study revealed that hip circumference, age, education level, and physical activity are significant predictors of early cognitive impairment in individuals with hypertension (72). The XGB model, which is based on hip circumference, age, education level, and physical activity, demonstrated a stronger predictive effect on the risk of

cognitive impairment in hypertensive patients. It achieved an AUC of 0.88, F1 score of 0.59, accuracy of 0.81, sensitivity of 0.84, and specificity of 0.80. However, it should be noted that environmental exposures were not included in this analysis. In this study, we investigated a multi-center population of hypertensive patients and integrated natural environmental factors into the model development process. Subsequently, we identified the most superior predictive model for further analysis.

Inevitably, the limitations of the current study must be considered. First, we only analyzed patient data from four hospitals in Shandong Province, with a limited sample size, and this data may not cover all regions and residents. Second, we selected only a few important individual and natural environmental factors as predictors for model development and may have overlooked some other important indicators, such as lifestyle (73), genetics (74), air pollution (75), and drinking water quality (76). Third, the design of the cross-sectional study makes it difficult to determine causality between hypertension and cognitive impairment. Fourth, although the MMSE questionnaire assessment has a better diagnostic value for early cognitive impairment, the Montreal Cognitive Assessment (MoCA) may be more appropriate for screening in this study (77). Finally, given the variability of natural environmental factors at different latitudes, the generalizability of current results to populations at other latitudes remains limited. Nevertheless, we conducted a comprehensive assessment of cognitive function in hypertensive patients from multiple regions, and used detailed residence information to link physical environment data to each patient. To our knowledge, this is the first reliable ML predictive model of early cognitive impairment in hypertension developed and validated based on simple personal and natural environmental factors, which indeed provides a new perspective for diagnostic decision-making and prevention of cognitive impairment in hypertension in a clinical setting. The preliminary findings warrant further validation through longitudinal cohort studies to establish temporal causality and robustness of observed patterns.

5 Conclusions

In conclusion, the XGBoost model based on age, waist circumference, urban green coverage, educational levels, annual sunshine hours, and area whole-day average noise has superior predictive performance and clinical practicability, which provides a reliable and economical tool for the diagnosis and risk prediction of early cognitive impairment in hypertension. Further investigations, including the design of large-sample multicenter research, multimodal and multidimensional data, predictor thresholds, and model optimization, should be considered in future research.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Affiliated Hospital of Shandong University of Chinese Medicine with the informed consent of all participants. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *Lancet Neurol.* (2016) 15(5):455–532. doi: 10.1016/S1474-4422(16)00062-4

2. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the lancet commission. *Lancet.* (2020) 396(10248):413–46. doi: 10.1016/S0140-6736(20)30367-6

3. Ungvari Z, Toth P, Tarantini S, Prodan CI, Sorond F, Merkely B, et al. Hypertension-induced cognitive impairment: from pathophysiology to public health. *Nat Rev Nephrol.* (2021) 17:639–54. doi: 10.1038/s41581-021-00430-6

4. Jia L, Du Y, Chu L, Zhang Z, Li F, Lyu D, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. *Lancet Public Health.* (2020) 5(12):e661–71. doi: 10.1016/S2468-2667(20)30185-7

5. Dominguez LJ, Veronese N, Vernuccio L, Catanese G, Inzerillo F, Salemi G, et al. Nutrition, physical activity, and other lifestyle factors in the prevention of cognitive decline and dementia. *Nutrients.* (2021) 13(11):4080. doi: 10.3390/nu13114080

6. Wen C, Hu H, Ou YN, Bi YL, Ma YH, Tan L, et al. Risk factors for subjective cognitive decline: the CABLE study. *Transl Psychiatry.* (2021) 11(1):576. doi: 10.1038/s41398-021-01711-1

7. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA.* (2014) 312(23):2551–61. doi: 10.1001/jama.2014.13806

8. Wei Y, Wang Y, Lin CK, Yin K, Yang J, Shi L, et al. Associations between seasonal temperature and dementia-associated hospitalizations in new England. *Environ Int.* (2019) 126:228–33. doi: 10.1016/j.envint.2018.12.054

9. Yuchi W, Sbihi H, Davies H, Tamburic L, Brauer M. Road proximity, air pollution, noise, green space and neurologic disease incidence: a population-based cohort study. *Environ Health.* (2020) 19(1):8. doi: 10.1186/s12940-020-0565-4

10. Chandra M, Rai CB, Kumari N, Sandhu VK, Chandra K, Krishna M, et al. Air pollution and cognitive impairment across the life course in humans: a systematic review with specific focus on income level of study area. *Int J Environ Res Public Health.* (2022) 19(3):1405. doi: 10.3390/ijerph19031405

11. Huang L, Zhang Y, Wang Y, Lan Y. Relationship between chronic noise exposure, cognitive impairment, and degenerative dementia: update on the experimental and epidemiological evidence and prospects for further research. *J Alzheimers Dis.* (2021) 79(4):1409–27. doi: 10.3233/JAD-201037

12. Zhang T. *Discussion on Cognitive Dysfunction and Risk Factors in Plateau Hypertension Population.* Ningxia: Ningxia Medical University (2020).

13. Li H, Zhao C, Lin ZQ, Wang L. Clinical characteristics and risk factors of cognitive dysfunction in elderly patients with essential hypertension. *J Cardio Cerebrovasc Dis Integr Tradit Chinese West Med.* (2022) 20:565–9. Available online at: https://kns.cnki.net/kcms2/article/abstract?v=9IId9Ku_yBaD0akQ_aqZ_RXsdy0Vsv8BM-d_E_kUr6eb_-szqp9TEHy99H9QmyNcV3psfmr1PwA60Fh7dejCeK9lbh1mdGd7Swl5VYlgOfupsrG4GD7UrEWxN84XkP5797hd6k0L-y5qo7p8TqVvtIMWYsK82MGdHLOweLX3RaEhjHrgPZUUfQ==&uniplatform=NZKPT&language=CHS

14. Ma L, Feng M, Qian Y, Yang W, Liu J, Han R, et al. Insulin resistance is an important risk factor for cognitive impairment in elderly patients with primary hypertension. *Yonsei Med J.* (2015) 56(1):89–94. doi: 10.3349/ymj.2015.56.1.89

15. Lu JY. *Building of Risk Prediction Model for Mild Cognitive Impairment in Elderly Hypertension Patients in Community.* Guangxi: Guangxi University of Chinese Medicine (2021).

16. Bratić B, Kurbalija V, Ivanović M, Oder I, Bosnić Z. Machine learning for predicting cognitive diseases: methods, data sources and risk factors. *J Med Syst.* (2018) 42(12):243. doi: 10.1007/s10916-018-1071-x

17. Reel PS, Reel S, van Kralingen JC, Langton K, Lang K, Erlic Z, et al. Machine learning for classification of hypertension subtypes using multi-omics: a multi-centre, retrospective, data-driven study. *EBioMedicine.* (2022) 84:104276. doi: 10.1016/j.ebiom.2022.104276

18. Poss AM, Maschek JA, Cox JE, Hauner BJ, Hopkins PN, Hunt SC, et al. Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease. *J Clin Invest.* (2020) 130(3):1363–76. doi: 10.1172/JCI131838

19. Zhou J, Lee S, Liu Y, Chan JSK, Li G, Wong WT, et al. Predicting stroke and mortality in mitral regurgitation: a machine learning approach. *Curr Probl Cardiol.* (2023) 48(2):101464. doi: 10.1016/j.cpcardiol.2022.101464

20. Reinke C, Doblhammer G, Schmid M, Welchowski T. Dementia risk predictions from German claims data using methods of machine learning. *Alzheimers Dement.* (2023) 19(2):477–86. doi: 10.1002/alz.12663

21. Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among Chinese community-dwelling elderly people with normal cognition: development and validation study. *J Med Internet Res.* (2021) 23(2):e20298. doi: 10.2196/20298

22. Yadgir SR, Engstrom C, Jacobsohn GC, Green RK, Jones CMC, Cushman JT, et al. Machine learning-assisted screening for cognitive impairment in the emergency department. *J Am Geriatr Soc.* (2022) 70(3):831–7. doi: 10.1111/jgs.17491

23. Liu Z, Zhang L, Wu J, Zheng Z, Gao J, Lin Y, et al. Machine learning-based classification of circadian rhythm characteristics for mild cognitive impairment in the elderly. *Front Public Health.* (2022) 10:1036886. doi: 10.3389/fpubh.2022.1036886

24. Liu H, Zhang X, Liu H, ST C. Using machine learning to predict cognitive impairment among middle-aged and older Chinese: a longitudinal study. *Int J Public Health.* (2023) 68:1605322. doi: 10.3389/ijph.2023.1605322

25. James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the eighth joint national committee (JNC 8). *JAMA.* (2014) 311(5):507–20. doi: 10.1001/jama.2013.284427

26. Ren Z, Li Y, Li X, Shi H, Zhao H, He M, et al. Associations of body mass index, waist circumference and waist-to-height ratio with cognitive impairment among Chinese older adults: based on the CLHLS. *J Affect Disord.* (2021) 295:463–70. doi: 10.1016/j.jad.2021.08.093

27. You S, Wang X, Lindley RI, Robinson T, Anderson CS, Cao Y, et al. Early cognitive impairment after intracerebral hemorrhage in the INTERACT1 study. *Cerebrovasc Dis.* (2017) 44(5-6):320–4. doi: 10.1159/000481443

28. Han F, Luo C, Lv D, Tian L, Qu C. Risk factors affecting cognitive impairment of the elderly aged 65 and over: a cross-sectional study. *Front Aging Neurosci.* (2022) 14:903794. doi: 10.3389/fnagi.2022.903794

29. Sun JY, Hua Y, Zou HY, Qu Q, Yuan Y, Sun GZ, et al. Association between waist circumference and the prevalence of (Pre) hypertension among 27,894 US adults. *Front Cardiovasc Med.* (2021) 8:717257. doi: 10.3389/fcvm.2021.717257

30. Adza WK, Hursthouse AS, Miller J, Boakye D. Exploring the combined association between road traffic noise and air quality using QGIS. *Int J Environ Res Public Health.* (2022) 19(24):17057. doi: 10.3390/ijerph192417057

31. Feng L, Liu Z, Li C, Li Z, Lou X, Shao L, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicentre observational study. *Lancet Digit Health.* (2022) 4(1):e8–17. doi: 10.1016/S2589-7500(21)00215-6

32. Rasool A, Bunterngchit C, Tiejian L, Islam MR, Qu Q, Jiang Q. Improved machine learning-based predictive models for breast cancer diagnosis. *Int J Environ Res Public Health.* (2022) 19(6):3211. doi: 10.3390/ijerph19063211

33. Bunterngchit C, Bunterngchit Y. A comparative study of machine learning models for Parkinson's disease detection. *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Chiangrai, Thailand; (2022). p. 465–9.

34. Tan WY, Hargreaves C, Chen C, Hilal S. A machine learning approach for early diagnosis of cognitive impairment using population-based data. *J Alzheimers Dis.* (2023) 91:449–61. doi: 10.3233/JAD-220776

35. Shi J, He Q, Pan Y, Zhang X, Li M, Chen S. Estimation of appendicular skeletal muscle mass for women aged 60–70 years using a machine learning approach. *J Am Med Dir Assoc.* (2022) 23(12):1985.e1–.e7. doi: 10.1016/j.jamda.2022.09.002

36. Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect Psychol Sci.* (2017) 12:1100–22. doi: 10.1177/1745691617693393

37. Iadecola C, Gottesman RF. Neurovascular and cognitive dysfunction in hypertension. *Circ Res.* (2019) 124(7):1025–44. doi: 10.1161/CIRCRESAHA.118.313260

38. Kivipelto M, Helkala EL, Laakso MP, Hänninen T, Hallikainen M, Alhainen K, et al. Midlife vascular risk factors and Alzheimer's disease in later life: longitudinal, population based study. *Br Med J.* (2001) 322(7300):1447–51. doi: 10.1136/bmj.322.7300.1447

39. Walker KA, Power MC, Gottesman RF. Defining the relationship between hypertension, cognitive decline, and dementia: a review. *Curr Hypertens Rep.* (2017) 19(3):24. doi: 10.1007/s11906-017-0724-3

40. Tang X, Zhao W, Lu M, Zhang X, Zhang P, Xin Z, et al. Relationship between central obesity and the incidence of cognitive impairment and dementia from cohort studies involving 5,060,687 participants. *Neurosci Biobehav Rev.* (2021) 130:301–13. doi: 10.1016/j.neubiorev.2021.08.028

41. Lin WY. Associations of five obesity indicators with cognitive performance in 30,697 Taiwan biobank participants. *BMC Geriatr.* (2022) 22(1):839. doi: 10.1186/s12877-022-03457-x

42. Chuang SY, Liu WL, Chang HY, Hsu CC, Pan WH. Central obesity and elevated blood pressure in middle life are associated with physical and cognitive impairment in later life: a retrospective design with repeated measures. *Exp Gerontol.* (2023) 173:112093. doi: 10.1016/j.exger.2023.112093

43. Flores-Cordero JA, Pérez-Pérez A, Jiménez-Cortegana C, Alba G, Flores-Barragán A, Sánchez-Margalet V. Obesity as a risk factor for dementia and Alzheimer's disease: the role of leptin. *Int J Mol Sci.* (2022) 23(9):5202. doi: 10.3390/ijms23095202

44. Lövdén M, Fratiglioni L, Glymour MM, Lindenberger U, Tucker-Drob EM. Education and cognitive functioning across the life span. *Psychol Sci Public Interest.* (2020) 21:6–41. doi: 10.1177/1529100620920576

45. Zhang Q, Wu Y, Han T, Liu E. Changes in cognitive function and risk factors for cognitive impairment of the elderly in China: 2005–2014. *Int J Environ Res Public Health.* (2019) 16(16):2847. doi: 10.3390/ijerph16162847

46. Ahn CW, Shin S, Lee S, Park HS, Hong N, Rhee Y. Association of shift work with normal-weight obesity in community-dwelling adults. *Endocrinol Metab (Seoul).* (2022) 37(5):781–90. doi: 10.3803/EnM.2022.1532

47. Albishri MA, Alsubaie DM, Abugad HA, Abdel Wahab MM. Association between glycemic control and shift working among healthcare workers with diabetes, Dammam, Saudi Arabia: five years' experience. *Saudi Med J.* (2021) 42(12):1296–301. doi: 10.15537/smj.2021.42.12.20210416

48. Elovainio M, Komulainen K, Lipsanen J, Partonen T, Pesonen AK, Pulkki-Råback L, et al. Long-term cumulative light exposure from the natural environment and sleep: a cohort study. *J Sleep Res.* (2022) 31(3):e13511. doi: 10.1111/jsr.13511

49. Komulainen K, Hakulinen C, Lipsanen J, Partonen T, Pulkki-Råback L, Kähönen M, et al. Long-term residential sunlight exposure associated with cognitive function among adults residing in Finland. *Sci Rep.* (2022) 12(1):20818. doi: 10.1038/s41598-022-25336-6

50. Zhu H, Wang N, Yao L, Chen Q, Zhang R, Qian J, et al. Moderate UV exposure enhances learning and memory by promoting a novel glutamate biosynthetic pathway in the brain. *Cell.* (2018) 173(7):1716–1727.e17. doi: 10.1016/j.cell.2018.04.014

51. Chantranupong L, Sabatini BL. Sunlight brightens learning and memory. *Cell.* (2018) 173(7):1570–2. doi: 10.1016/j.cell.2018.05.044

52. Gao Q, Luan D, Wang X, Xin S, Liu Y, Li J. Effect of sun exposure on cognitive function among elderly individuals in northeast China. *Clin Interv Aging.* (2018) 13:2075–82. Published 2018 October 18. doi: 10.2147/CIA.S179355

53. Kent ST, Kabagambe EK, Wadley VG, Howard VJ, Crosson WL, Al-Hamdan MZ, et al. The relationship between long-term sunlight radiation and cognitive decline in the REGARDS cohort study. *Int J Biometeorol.* (2014) 58(3):361–70. doi: 10.1007/s00484-013-0631-5

54. Kent ST, McClure LA, Crosson WL, Arnett DK, Wadley VG, Sathiakumar N. Effect of sunlight exposure on cognitive function among depressed and non-depressed participants: a REGARDS cross-sectional study. *Environ Health.* (2009) 8:34. doi: 10.1186/1476-069X-8-34

55. Paul KC, Haan M, Mayeda ER, Ritz BR. Ambient air pollution, noise, and late-life cognitive decline and dementia risk. *Annu Rev Public Health.* (2019) 40:203–20. doi: 10.1146/annurev-publhealth-040218-044058

56. Oudin A. Short review: air pollution, noise and lack of greenness as risk factors for Alzheimer's disease epidemiologic and experimental evidence. *Neurochem Int.* (2020) 134:104646. doi: 10.1016/j.neuint.2019.104646

57. Zhao YL, Qu Y, Ou YN, Zhang YR, Tan L, Yu JT. Environmental factors and risks of cognitive impairment and dementia: a systematic review and meta-analysis. *Ageing Res Rev.* (2021) 72:101504. doi: 10.1016/j.arr.2021.101504

58. Thompson R, Smith RB, Bou Karim Y, Shen C, Drummond K, Teng C, et al. Noise pollution and human cognition: an updated systematic review and meta-analysis of recent evidence. *Environ Int.* (2022) 158:106905. doi: 10.1016/j.envint.2021.106905

59. Jafari Z, Kolb BE, Mohajerani MH. Noise exposure accelerates the risk of cognitive impairment and Alzheimer's disease: adulthood, gestational, and prenatal mechanistic evidence from animal studies. *Neurosci Biobehav Rev.* (2020) 117:110–28. doi: 10.1016/j.neubiorev.2019.04.001

60. Hahad O, Prochaska JH, Daiber A, Muenzel T. Environmental noise-induced effects on stress hormones, oxidative stress, and vascular dysfunction: key factors in the relationship between cerebrocardiovascular and psychological disorders. *Oxid Med Cell Longev.* (2019) 2019:4623109. doi: 10.1155/2019/4623109

61. Jigeer G, Tao W, Zhu Q, Xu X, Zhao Y, Kan H, et al. Association of residential noise exposure with maternal anxiety and depression in late pregnancy. *Environ Int.* (2022) 168:107473. doi: 10.1016/j.envint.2022.107473

62. Freire ACC, Pondé MP, Liu A, Caron J. Anxiety and depression as longitudinal predictors of mild cognitive impairment in older adults. *Can J Psychiatry.* (2017) 62(5):343–50. doi: 10.1177/0706743717699175

63. Hu W, Zhang H, Ni R, Cao Y, Fang W, Chen Y, et al. Interaction between the animal-based dietary pattern and green space on cognitive function among Chinese older adults: a prospective cohort study. *Int J Hyg Environ Health.* (2023) 250:114147. doi: 10.1016/j.ijheh.2023.114147

64. Ciarmiello A, Giovannini E, Pastorino S, Ferrando O, Foppiano F, Mannironi A, et al. Machine learning model to predict diagnosis of mild cognitive impairment by using radiomic and amyloid brain PET. *Clin Nucl Med.* (2023) 48(1):1–7. doi: 10.1097/RLU.0000000000004433

65. Na KS. Prediction of future cognitive impairment among the community elderly: a machine-learning based approach. *Sci Rep.* (2019) 9(1):3335. doi: 10.1038/s41598-019-39478-7

66. Tunvirachaisakul C, Supasitthumrong T, Tangwongchai S, Hemrunroj S, Chuchuen P, Tawankanjanachot I, et al. Characteristics of mild cognitive impairment using the Thai version of the consortium to establish a registry for Alzheimer's disease tests: a multivariate and machine learning study. *Dement Geriatr Cogn Disord.* (2018) 45(1–2):38–48. doi: 10.1159/000487232

67. Lin H, Himali JJ, Satizabal CL, Beiser AS, Levy D, Benjamin EJ, et al. Identifying blood biomarkers for dementia using machine learning methods in the Framingham heart study. *Cells.* (2022) 11(9):1506. doi: 10.3390/cells11091506

68. Lin RH, Wang CC, Tung CW. A machine learning classifier for predicting stable MCI patients using gene biomarkers. *Int J Environ Res Public Health.* (2022) 19(8):4839. doi: 10.3390/ijerph19084839

69. Kang SH, Cheon BK, Kim JS, Jang H, Kim HJ, Park KW, et al. Machine learning for the prediction of amyloid positivity in amnestic mild cognitive impairment. *J Alzheimers Dis.* (2021) 80(1):143–57. doi: 10.3233/JAD-201092

70. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. *Alzheimers Res Ther.* (2021) 13(1):162. doi: 10.1186/s13195-021-00900-w

71. Du J, Tao X, Zhu L, Qi W, Min X, Deng H, et al. A risk prediction system for depression in middle-aged and older adults grounded in machine learning and visualization technology: a cohort study. *Front Public Health.* (2025) 13:1606316. doi: 10.3389/fpubh.2025.1606316

72. Zhong X, Yu J, Jiang F, Chen H, Wang Z, Teng J, et al. A risk prediction model based on machine learning for early cognitive impairment in hypertension: development and validation study. *Front Public Health.* (2023) 11:1143019. doi: 10.3389/fpubh.2023.1143019

73. Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nat Rev Neurol.* (2018) 14(11):653–66. doi: 10.1038/s41582-018-0070-3

74. Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* (2022) 54(4):412–36. doi: 10.1038/s41588-022-01024-z

75. Schikowski T, Altuğ H. The role of air pollution in cognitive impairment and decline. *Neurochem Int.* (2020) 136:104708. doi: 10.1016/j.neuint.2020.104708

76. Pan X, Luo Y, Zhao D, Zhang L. Associations among drinking water quality, dyslipidemia, and cognitive function for older adults in China: evidence from CHARLS. *BMC Geriatr.* (2022) 22(1):683. doi: 10.1186/s12877-022-03375-y

77. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* (2005) 53(4):695–9. doi: 10.1111/j.1532-5415.2005.53221.x