



OPEN

A machine learning-based model for predicting the risk of cognitive frailty in elderly patients on maintenance hemodialysis

Meng Cao^{1,2,4}, Bixia Tang^{3,4}, Liwei Yang^{1✉} & Jing Zeng^{1✉}

Elderly patients undergoing maintenance hemodialysis (MHD) face a heightened risk of cognitive frailty (CF), which significantly compromises quality of life. Early identification of at-risk individuals and timely intervention are essential. Nevertheless, current CF risk prediction models fall short in accuracy to adequately fulfill clinical requirements. This study aimed to examine the determinants of CF in elderly patients undergoing MHD and to develop a risk prediction model through machine learning algorithms. The objective is to furnish healthcare professionals with an early prediction tool and to offer insights for personalized CF risk management. A convenience sampling method was employed to select 1,075 elderly MHD patients from various tertiary-level hospitals in Chengdu between October 2023 and March 2024 as the modeling set, and 269 elderly MHD patients from hospitals in Chengdu, Yibin, and Zigong between September 2024 and October 2024 as the external validation set. CF was assessed using the Fried Phenotypic Scale for Frailty (FP) and the Montreal Cognitive Assessment Scale (MOCA). Data on patients' demographics, sleep, nutrition, depression, and social support were collected. Single-factor and multi-factor logistic regression analyses were conducted to identify the factors influencing CF. Five machine learning algorithms—Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Random Forest (RF), Neural Network (NNET), and Logistic Regression (LR)—were employed to develop risk prediction models. These five models served as base classifiers, and 16 ensemble models were constructed using the Stacking method. The optimal ensemble models were identified and compared with the five individual models, followed by the selection and external validation of the most effective predictive models. Finally, the optimal models were deployed on web platforms utilizing the Streamlit library. CF prevalence was 14.2%. Significant CF risk factors included age, mode of residence, medical payment method, exercise, alcohol consumption, dialysis vascular access, serum albumin classification, serum phosphorus classification, total cholesterol classification, blood urea nitrogen classification, malnutrition score and depression score. The Stacking model showed superior performance (AUC = 0.911), with external validation confirming its accuracy (AUC = 0.832). Machine learning models, particularly Stacking, effectively predict CF risk in elderly MHD patients, providing a valuable tool for clinical intervention.

Keywords Maintenance hemodialysis, Elderly patients, Cognitive frailty, Machine learning, Risk prediction models

Maintenance Hemodialysis (MHD) is a widely employed therapeutic approach for individuals with end-stage renal disease. However, MHD itself introduces a range of complications that predispose patients to cognitive impairment. As the treatment duration extends, patients may also experience anemia, malnutrition, and muscle atrophy, culminating in increased frailty¹.

Cognitive Frailty (CF) is a syndrome defined by the coexistence of physical frailty and mild cognitive impairment, while excluding Alzheimer's disease and other forms of dementia². Research has demonstrated a high prevalence of cognitive frailty³, with affected patients being more prone to adverse outcomes such as falls, disability, and mortality^{4,5}, which can significantly diminish quality of life.

¹School of Nursing, Chengdu Medical College, Chengdu, China. ²School of Nursing, Sichuan Vocational College of Health and Rehabilitation, Zigong, China. ³Chengdu Seventh People's Hospital, Chengdu, China. ⁴These authors contributed equally: Meng Cao and Bixia Tang. ✉email: maomao072221@163.com; zengjing@163.com

Risk prediction models can assist healthcare professionals in the early identification of high-risk populations. Currently, CF risk prediction models for elderly patients undergoing MHD are insufficient, with the selection of influencing factors lacking comprehensiveness, and most existing models relying on traditional logistic regression algorithms. In light of this, the present study further investigates the influencing factors of CF in elderly patients with MHD through a cross-sectional survey and it employs five machine learning algorithms (Support Vector Machine, Extreme Gradient Boosting, Random Forest, Neural Networks, Logistic regression) to develop risk prediction models. The above five models are used as base classifiers, and 16 integrated models with different combinations of base classifiers are constructed based on the Stacking method. The optimal integrated models are selected and compared with the five single models, the optimal predictive models are screened and externally validated, and the optimal models are deployed on the web page through the Streamlit library, a move aimed at providing healthcare professionals with a tool for early prediction of CF risk in elderly patients with MHD.

Objects and methods

Subjects

A convenience sampling method was employed to select 1,075 elderly patients undergoing MHD from October 2023 to March 2024 across several tertiary hospitals in Chengdu, as well as 269 elderly patients from September 2024 to October 2024 in various hospitals in Chengdu, Yibin, and Zigong for the study. Inclusion criteria consisted of: ① age ≥ 60 years; ② duration of maintenance hemodialysis treatment ≥ 3 months; ③ demonstrated routine communication abilities and provided informed consent. Exclusion criteria included: ① patients diagnosed with dementia; ② individuals with intellectual disabilities or a history of psychiatric disorders; ③ severe visual or auditory impairments that hindered cooperation in completing the questionnaire and assessments; ④ individuals unable to undergo evaluations of frailty and cognitive impairment due to functional limitations (e.g., arriving at the hospital for dialysis in a wheelchair or being bedridden); and ⑤ patients experiencing acute kidney injury or having undergone renal transplantation within the preceding 3 months.

Methods

CF judging criteria

Referring to the latest international consensus², the patients had undiagnosed clinical dementia; a score of 3 or more on the Frailty Phenotype Scale (FP); and a score of 18 to 25 on the Montreal Cognitive Assessment (MOCA).

Research tools

General information questionnaire Developed by the principal investigator of this study after a review of previous research, this questionnaire encompasses patients' socio-demographic information (age, sex, marital status, education level, mode of residence, medical payment method for medical care), physical indicators (height, weight, BMI), lifestyle and behavioral habits (smoking history, alcohol consumption, and exercise), disease-related information (comorbidities, history of falls within 1 year, months on dialysis, and dialysis vascular access), as well as laboratory indicators (hemoglobin, C-reactive protein, serum albumin, serum calcium, serum phosphorus, parathyroid hormone, total cholesterol, triglycerides, serum sodium, blood urea nitrogen, serum creatinine, and glomerular filtration rate).

Fried's Frailty phenotype scale (FP) The scale, proposed by Fried in 2001⁶, comprises five assessments: decreased body mass, self-reported fatigue, diminished grip strength, reduced gait speed, and low levels of physical activity. The scale has a total score range of 5 points, with the following scoring criteria: 0 indicates non-debilitating status, 1–2 indicates pre-debilitating status, and ≥ 3 indicates debilitating status. It has been extensively utilized within the chronic kidney disease population, demonstrating a content validity coefficient of 0.98 and a Cronbach's alpha coefficient of 0.93.

Montreal Cognitive Assessment Scale (MOCA) The Chinese version of the Montreal Cognitive Assessment (MOCA) was utilized⁷, translated and modified by Wang Wei and colleagues in 2006 to reflect the cultural characteristics of our country. It primarily assesses cognitive function across seven domains, including visuospatial and executive functions, naming, and attention, with a maximum score of 30. Higher scores indicate better cognitive performance, with a score of ≥ 26 on the MOCA considered indicative of normal cognitive function². The MOCA demonstrates high test-retest reliability (0.9 or above), internal consistency (Cronbach's alpha = 0.83), and positive and negative predictive values.

Pittsburgh sleep quality index (PSQI) scale The Pittsburgh Sleep Quality Index (PSQI), compiled by Buysse et al.⁸, was utilized. This scale consists of 19 self-assessed items and five other-assessed items, with the 18 self-assessed items organized into seven components: subjective sleep quality, time taken to fall asleep, sleep duration, sleep efficiency, sleep disturbances, use of hypnotic medication, and daytime dysfunction. Each component is scored from 0 to 3, and the total score is the summation of the 7 component scores, resulting in a range from 0 to 21. A PSQI score greater than 7 indicates the presence of sleep problems. The scale exhibits a Cronbach's alpha coefficient of 0.7962, demonstrating good reliability and validity.

Malnutrition-inflammation score (MIS) The Malnutrition Inflammation Score (MIS)⁹ includes ten evaluation indicators, each scored from 0 (normal) to 3 (malnutrition), resulting in a total score range of 0 to 30 points. A score of 0–8 indicates mild malnutrition, 9–18 indicates moderate malnutrition, and a score more excellent than

18 indicates severe malnutrition. The score increases with the severity of malnutrition and inflammation. The scale demonstrates a Cronbach's alpha coefficient of 0.866.

Geriatric depression scale (GDS-15) The scale, developed by Sheikh¹⁰, comprised 15 items to which respondents replied with “yes” or “no.” Each “yes” was assigned a score of 1 point, while each “no” received a score of 0 points, with a threshold of ≥ 8 points indicating the presence of depressive symptoms. Scores ranged from 0 to 15, reflecting the severity of depressive symptoms, with higher scores correlating to more severe symptoms. The scale primarily assesses respondents' depressive status over the preceding week. The Cronbach's alpha coefficient for the scale was 0.793, and the test-retest reliability was 0.728.

Social Support Rating Scale (SSRS) The scale¹¹, developed by Chinese scholar Xiao Shuiyuan, encompasses three dimensions: objective support, subjective support, and utilization of social support, comprising a total of 10 items. The overall score ranges from 12 to 66 points, with higher scores indicative of greater social support; expressly, scores of ≤ 22 reflect a low level, 23–44 indicate a medium level, and ≥ 45 signify a high level of social support. The internal consistency, measured by Cronbach's alpha coefficient, ranges from 0.825 to 0.896, demonstrating robust reliability and validity, making it suitable for social support research in China.

Survey methods

All investigators received standardized training, with uniform guidelines provided, and were required to use the same brand and model of measurement tools. Surveys were conducted with patients who voluntarily participated in the study, ensuring that they had a thorough understanding of the research. Data collection utilized standardized paper questionnaires, which were distributed and collected on-site. When patients' conditions permitted, they completed the questionnaires independently; however, if patients could not to fill them out due to physical disabilities, reading difficulties, dialysis, or other reasons, the researcher meticulously inquired and accurately recorded responses item by item. Throughout the questionnaire completion process, the researcher clarified any items that the subjects did not understand, while refraining from influencing their selection of answers.

Statistical analysis

Continuous variables were expressed as mean \pm standard deviation, and comparisons between two groups were conducted using an independent samples t-test. Categorical variables were presented as n (%) and comparisons between groups were performed using the chi-square test. Missing values were imputed using the MICE package in R Studio version 4.3.1. Binary logistic regression, including both univariate and multivariate analyses, was employed to identify factors influencing CF in elderly patients undergoing MHD, with factors exhibiting $P < 0.2$ considered for subsequent model construction.

This study included a total of 1,075 modeling samples, with the outcome variables being dichotomous and the predictor variables comprising dichotomous, multichotomous, and continuous variables. Five machine learning algorithms—SVM, XGBoost, RF, NN, and LR—were utilized to construct the model. The `train_test_split` function in Python 3.10 was employed to divide the dataset into a training set and a test set at a 75:25 ratio. Additionally, 269 patients from other hospitals were included as an external validation set.

The predictor variables underwent one-hot encoding and standardization, while 5-fold cross-validation and grid search were utilized for model training and hyperparameter optimization. The five aforementioned models were employed as base classifiers, and an integrated model was constructed using the Stacking method. This method leverages a hierarchical integration strategy, where the prediction outputs of the base classifiers are input into the meta-classifier to generate the final predictions. In this study, XGBoost was selected as the meta-classifier due to its outstanding nonlinear fitting capability and computational efficiency. Each base classifier utilized the optimal hyperparameter settings determined through grid search optimization during its individual model construction. Integration was achieved using the `StackingClassifier`, which combines the prediction outputs of the base classifiers with the original feature inputs for learning by the meta-classifier. Various combinations of three to five base classifiers were generated using a combinatorial approach, and the AUC, accuracy, recall, precision, and F1 score of each combination were evaluated. The optimal integrated model was selected by comparing the AUC of each combination.

The optimal integrated model was compared with five individual models, and the performance of all these models was evaluated using AUC, accuracy, recall, precision, and F1 score. The 95% confidence intervals were calculated using the Bootstrap method, while decision curves and calibration curves were employed to assess the clinical applicability and calibration of the models. The model with the highest AUC was ultimately selected as the best model, and its generalization performance was validated using an external validation set.

Results

Characterization of the study sample

Among the 1,075 elderly patients with MHD, there were 153 patients with CF and 922 without CF, resulting in a CF prevalence of 14.2%. Differences in characteristics between the two groups are detailed in Table 1. Compared to non-CF patients with MHD, those with CF were older (mean age 73.8 ± 7.6 years), more likely to live alone or in a nursing facility, often had residential health insurance or New Rural Cooperative Medical Insurance, did not engage in exercise, regularly consumed alcohol, had long-term central venous cannulas, exhibited lower serum albumin and serum phosphorus levels, presented with abnormal total cholesterol levels, had higher blood urea nitrogen, and experienced more severe malnutrition and depression.

Variable	Description	No (N=922)	Yes (N=153)	OR (95% CI) (univariable)	OR (95% CI) (multivariable)
Age	Mean ± SD	69.1 ± 7.5	73.8 ± 7.6	1.08 (1.06–1.10, <i>p</i> < .001)	1.08 (1.04–1.11, <i>p</i> < .001)
Mode of residence	Living alone	71 (7.7%)	48 (31.4%)	5.58 (3.67–8.49, <i>p</i> < .001)	8.28(4.85–14.14, <i>p</i> < .001)
	Living in a nursing facility	1 (0.1%)	2 (1.3%)	16.50(1.48–183.61, <i>p</i> = .023)	26.59(2.04–346.46, <i>p</i> = .012)
Medical payment method	Resident medical insurance	418 (45.3%)	80 (52.3%)	1.97 (1.34–2.91, <i>p</i> < .001)	2.04 (1.28–3.25, <i>p</i> = .003)
	New Rural Cooperative Medical Insurance	30 (3.3%)	27 (17.6%)	9.28 (5.08–16.97, <i>p</i> < .001)	14.62 (6.77–31.58, <i>p</i> < .001)
Exercise	Occasionally	258 (28%)	25 (16.3%)	0.33 (0.21–0.53, <i>p</i> < .001)	0.41 (0.24–0.72, <i>p</i> = .002)
	Often	306 (33.2%)	24 (15.7%)	0.27 (0.17–0.43, <i>p</i> < .001)	0.57 (0.32–1.03, <i>p</i> = .062)
Alcohol consumption	Often	5 (0.5%)	3 (2%)	3.37 (0.79–14.33, <i>p</i> = .100)	11.24 (1.83–68.97, <i>p</i> = .009)
Dialysis Vascular Access	Long-term central venous catheterization	136 (14.8%)	43 (28.1%)	2.26 (1.52–3.37, <i>p</i> < .001)	1.92 (1.16–3.18, <i>p</i> = .011)
Serum Albumin Classification	Low	489 (53%)	110 (71.9%)	2.32 (1.58–3.39, <i>p</i> < .001)	1.91 (1.17–3.11, <i>p</i> = .010)
Serum Phosphorus Classification	High	514 (55.7%)	69 (45.1%)	0.65 (0.46–0.92, <i>p</i> = .015)	0.72 (0.45–1.14, <i>p</i> = .159)
Total Cholesterol Classification	Abnormal	201 (21.8%)	47 (30.7%)	1.59 (1.09–2.32, <i>p</i> = .016)	1.43 (0.88–2.32, <i>p</i> = .154)
Blood Urea Nitrogen Classification	High	818 (88.7%)	142 (92.8%)	1.72 (0.88–3.37, <i>p</i> = .115)	2.90 (1.21–6.97, <i>p</i> = .017)
Malnutrition Score	Mean ± SD	6.1 ± 3.9	8.8 ± 4.5	1.16 (1.11–1.21, <i>p</i> < .001)	1.04 (0.99–1.10, <i>p</i> = .108)
Depression Score	Mean ± SD	3.7 ± 3.5	6.1 ± 4.0	1.16 (1.12–1.21, <i>p</i> < .001)	1.10 (1.04–1.16, <i>p</i> = .001)

Table 1. Descriptive statistics for variables (*P* < 0.2).

Results of unifactorial and multifactorial analysis

After conducting univariate and multivariate logistic regression analyses, the final results indicated that there were no statistically significant differences between the two groups of patients concerning sex, marital status, education level, smoking history, weight, height, BMI, months on dialysis, comorbidities, history of falls within 1 year, hemoglobin, hemoglobin classification, ultrasensitive c-reactive protein, ultrasensitive c-reactive protein classification, serum albumin, serum calcium, serum calcium classification, serum phosphorus, serum sodium, serum sodium classification, parathyroid hormone, parathyroid hormone classification, total cholesterol, triglycerides, triglyceride classification, blood urea nitrogen, serum creatinine, serum creatinine classification, glomerular filtration rate, glomerular filtration rate classification, PSQI, degree of malnutrition, and social support score (*P* > 0.2) (Supplementary Table 1). However, significant differences (*P* < 0.2) were observed between the two groups in terms of age, mode of residence, medical payment method, exercise, alcohol consumption, dialysis vascular access, serum albumin classification, serum phosphorus classification, total cholesterol classification, blood urea nitrogen classification, malnutrition score, depression score (Table 1).

For comprehensive descriptive analysis, odds ratios (ORs) for each predictor variable were examined through univariate and multivariate logistic regression (refer to Table 1). The study revealed that advanced age, solitary living, residence in nursing facilities, regular alcohol consumption, prolonged central venous catheterization, low serum albumin levels, elevated blood urea nitrogen, abnormal total cholesterol, malnutrition, and depression were identified as risk factors for CF in elderly patients undergoing MHD. (OR 1.08, 95% CI 1.04–1.11, OR 8.28, 95% CI 4.85–14.14, OR 26.59, and 95%CI 2.04–346.46, OR 11.24, 95%CI 1.83–68.97, OR 1.92, 95%CI 1.16–3.18, OR 1.91, 95%CI 1.17–3.11, OR 2.90, 95%CI 1.21–6.97, OR 1.43, 95%CI 0.88–2.32, OR 2.90, 95% CI 1.21–6.97, OR 1.43, 95% CI 0.88–2.32, OR 1.04, 95% CI 0.99–1.10, OR 1.10, 95% CI 1.04–1.16); Elderly patients with MHD enrolled in resident health insurance and the New Rural Cooperative Medical Insurance exhibited a higher risk of CF compared to those enrolled in employee health insurance (OR 2.04, 95% CI 1.28–3.25; OR 14.62, 95% CI 6.77–31.58). Moreover, elderly patients with MHD who engaged in occasional or frequent exercise demonstrated a lower risk of CF compared to those who never exercised (OR 0.41, 95% CI 0.24–0.72; OR 0.57, 95% CI 0.32–1.03). Additionally, elderly patients with MHD exhibiting elevated serum phosphorus levels had a reduced risk of CF in comparison to patients with low or normal serum phosphorus levels (OR 0.72, 95% CI 0.45–1.14).

Comparison of stacking models with varying combinations of base classifiers

By constructing Stacking integrated learning models with different combinations of base classifiers, we compare the performance variations across these models. Table 2 presents the Stacking integrated models with 3, 4, and 5 base classifiers, respectively. The combination “RF + LR + NNET” demonstrates the best performance among all base classifier combinations, with an AUC of 0.911, accuracy of 0.903, precision of 0.912, recall of 0.983, and an F1 score of 0.946. Therefore, the Stacking integration model composed of “RF + LR + NNET” is selected as the optimal integrated model.

Comparison of machine learning algorithm prediction performance

Figure 1 shows the predictive efficacy of the Stacking model and the five single models on the test set in the form of ROC curves. The Stacking model is, the Stacking integrated model with the best predictive performance in Section “Comparison of Stacking models with varying combinations of base classifiers”. The Stacking model (AUC = 0.911) was the best predictor of the occurrence of CF in patients with MHD, followed by XGBoost (AUC = 0.896), SVM (AUC = 0.87), NNET (AUC = 0.866), LR (AUC = 0.858) and RF (AUC = 0.844) models.

Table 3 presents a detailed set of performance metrics for the Stacking model and five individual models in both the training and test sets. In the test set, the Stacking model achieved the highest accuracy and F1 score,

Model Integration	AUC	Accuracy	Precision	Recall	F1
Stacking RF + SVM + LR Test	0.906	0.896	0.915	0.970	0.941
Stacking RF + SVM + XGBoost Test	0.894	0.903	0.915	0.978	0.946
Stacking RF + SVM + NNET Test	0.902	0.903	0.912	0.983	0.946
Stacking RF + LR + XGBoost Test	0.891	0.914	0.916	0.991	0.952
Stacking RF + LR + NNET Test	0.911	0.903	0.912	0.983	0.946
Stacking RF + XGBoost + NNET Test	0.907	0.914	0.916	0.991	0.952
Stacking SVM + LR + XGBoost Test	0.877	0.907	0.912	0.987	0.948
Stacking SVM + LR + NNET Test	0.889	0.903	0.912	0.983	0.946
Stacking SVM + XGBoost + NNET Test	0.884	0.911	0.913	0.991	0.950
Stacking LR + XGBoost + NNET Test	0.884	0.911	0.913	0.991	0.950
Stacking RF + SVM + LR + XGBoost Test	0.887	0.903	0.915	0.978	0.946
Stacking RF + SVM + LR + NNET Test	0.901	0.900	0.912	0.978	0.944
Stacking RF + SVM + XGBoost + NNET Test	0.895	0.903	0.912	0.983	0.946
Stacking RF + LR + XGBoost + NNET Test	0.899	0.914	0.916	0.991	0.952
Stacking SVM + LR + XGBoost + NNET Test	0.883	0.911	0.913	0.991	0.950
Stacking RF + SVM + LR + XGBoost + NNET Test	0.889	0.907	0.916	0.983	0.948

Table 2. Evaluation indicators of different stacking models.

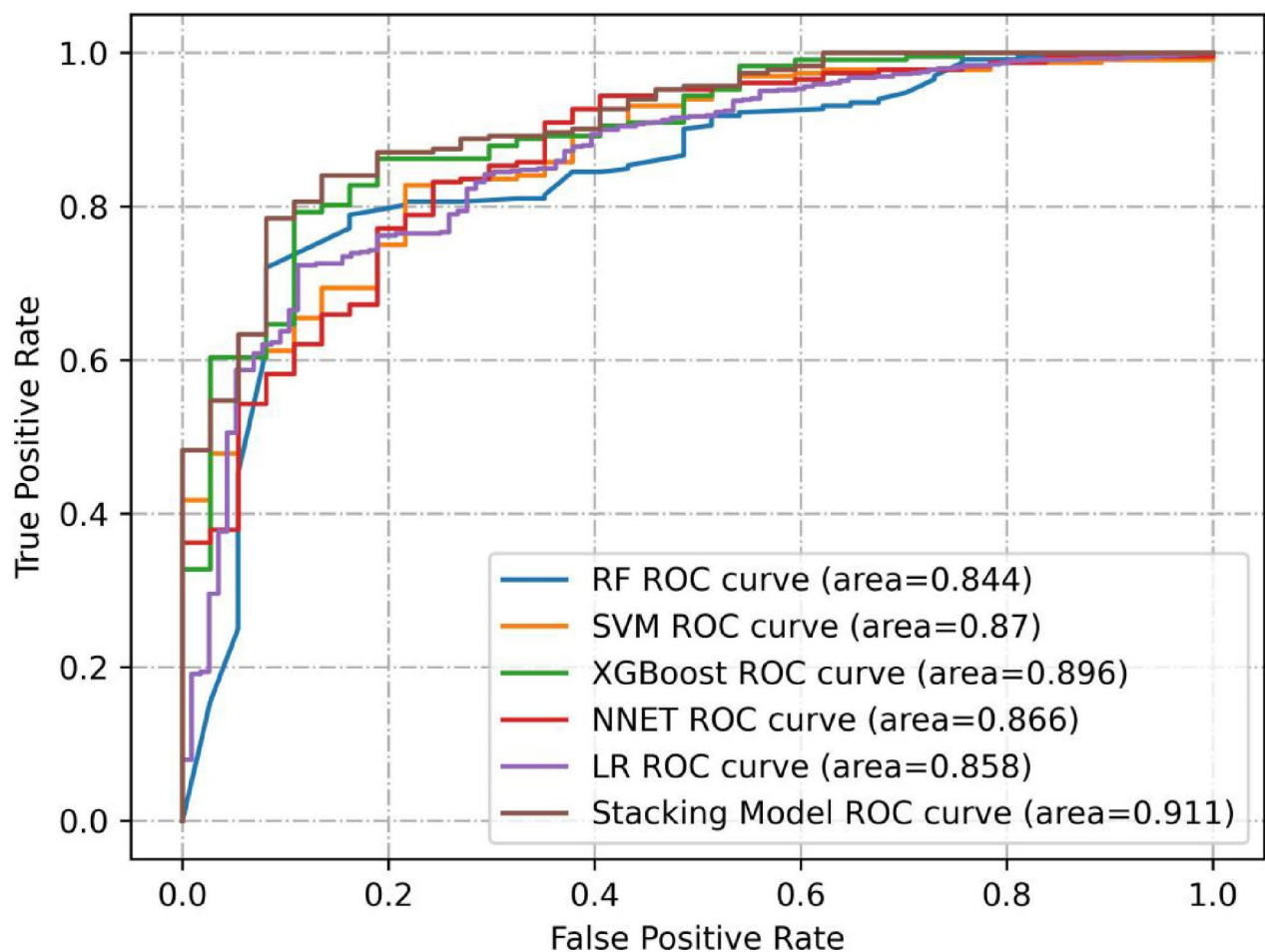


Fig. 1. ROC curves for the test set of six models.

Models	AUC(95% CI)	Accuracy(95% CI)	Precision (95% CI)	Recall (95% CI)	F1 Score(95% CI)
LR Test	0.858 (0.8213, 0.8919)	0.73 (0.6997, 0.7593)	0.974 (0.9603, 0.986)	0.703 (0.6706, 0.738)	0.816 (0.7938, 0.8391)
LR Train	0.874 (0.8165, 0.9234)	0.758 (0.7063, 0.8104)	0.961 (0.9308, 0.9839)	0.75 (0.6936, 0.8034)	0.843 (0.8039, 0.8787)
NNET Test	0.866 (0.8045, 0.9232)	0.881 (0.8401, 0.9182)	0.891 (0.8498, 0.9263)	0.983 (0.9646, 0.9958)	0.934 (0.9098, 0.9558)
NNET Train	0.986 (0.9999, 1.0)	0.893 (0.8722, 0.9143)	0.983 (0.9723, 0.9913)	0.902 (0.8795, 0.9235)	0.94 (0.9275, 0.9526)
RF Test	0.844 (0.7722, 0.9069)	0.862 (0.8215, 0.9033)	0.865 (0.8246, 0.9063)	0.996 (0.9863, 1.0)	0.926 (0.902, 0.949)
RF Train	0.986 (0.95, 0.993)	0.867 (0.8436, 0.8895)	0.998 (0.9956, 1.0)	0.867 (0.8427, 0.8894)	0.928 (0.914, 0.941)
SVM Test	0.87 (0.8124, 0.9188)	0.818 (0.7732, 0.8661)	0.946 (0.9132, 0.9754)	0.836 (0.7885, 0.8844)	0.888 (0.8564, 0.9199)
SVM Train	0.986 (0.9999, 1.0)	0.769 (0.7406, 0.799)	0.765 (0.7353, 0.7973)	0.956 (0.9388, 0.9722)	0.85 (0.8299, 0.8718)
XGBoost Test	0.896 (0.8416, 0.941)	0.896 (0.855, 0.9293)	0.898 (0.8588, 0.9328)	0.991 (0.9781, 1.0)	0.943 (0.9195, 0.9622)
XGBoost Train	0.986 (0.9999, 1.0)	0.909 (0.8895, 0.928)	0.993 (0.9854, 0.9985)	0.91 (0.8888, 0.93)	0.949 (0.9378, 0.9604)
Stacking Test	0.911 (0.8653, 0.9494)	0.903 (0.8661, 0.9368)	0.912 (0.877, 0.9444)	0.983 (0.9646, 0.9957)	0.946 (0.9237, 0.965)
Stacking Train	0.986 (0.9999, 1.0)	0.916 (0.8945, 0.9342)	0.988 (0.9797, 0.9956)	0.919 (0.8981, 0.9388)	0.952 (0.9406, 0.9633)

Table 3. Performance metrics of the models on the training and test set.

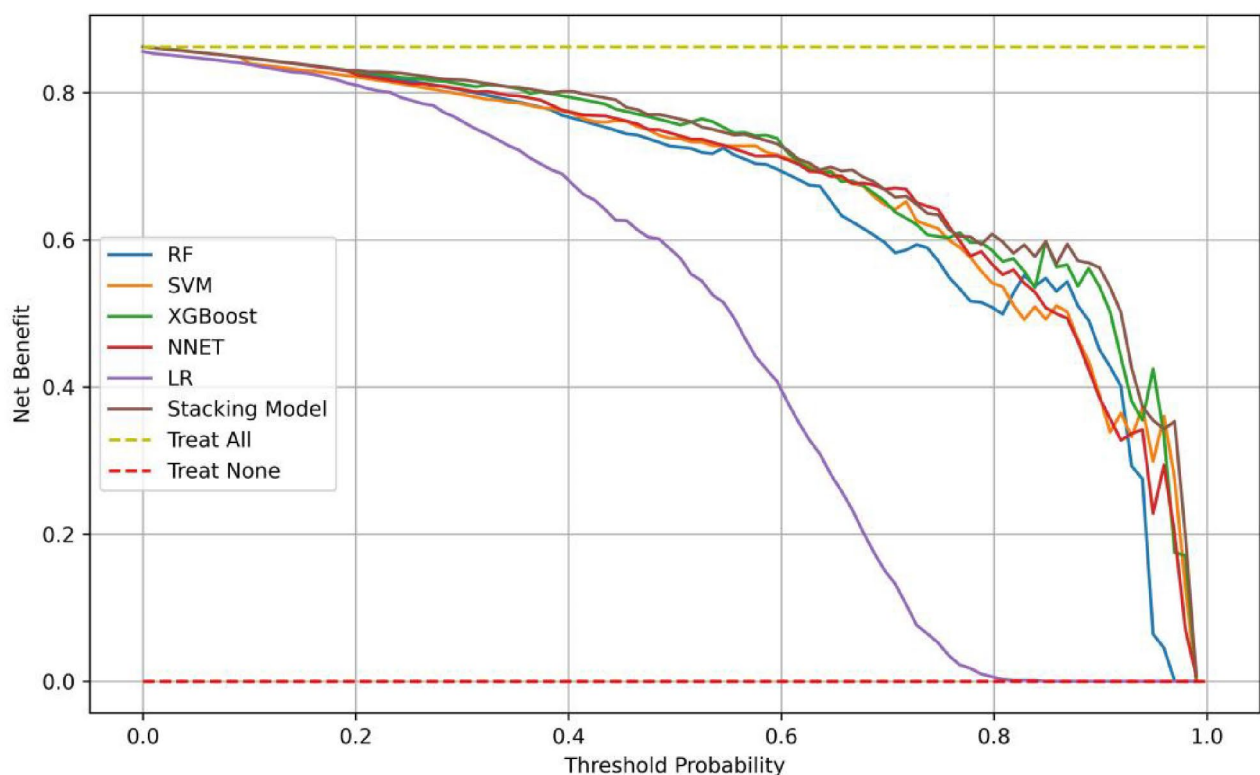


Fig. 2. DCA curves for the test set of six models.

with values of 0.903 and 0.946, respectively. The LR model exhibited the highest precision, with a value of 0.974, while the RF model demonstrated the highest recall, with a value of 0.996.

Figure 2 presents the DCA curves for the test set. The results of the DCA analysis demonstrate that the Stacking model offers the greatest net clinical benefit across the range of threshold probabilities compared to the other models, suggesting that the Stacking model is the optimal model with significant clinical utility.

Figure 3 shows the calibration curves for the Stacking model and the five single models. The calibration curves have a high degree of linear overlap with $y=x$, and the models are in good agreement between the predicted and actual occurrence risk values.

Sensitivity analysis of model parameters

The “max_depth” parameter controls the maximum depth of the tree in the RF model, and various depths ranging from 3 to 27 were tested. As shown in Fig. 4a, the model’s AUC reaches its highest value at a maximum depth of 3. The “hidden_layer_sizes” parameter governs the size of the hidden layers in the NNET model, with sizes ranging from 10 to 100 tested, as shown in Fig. 4b. The model’s AUC reaches its highest value when the

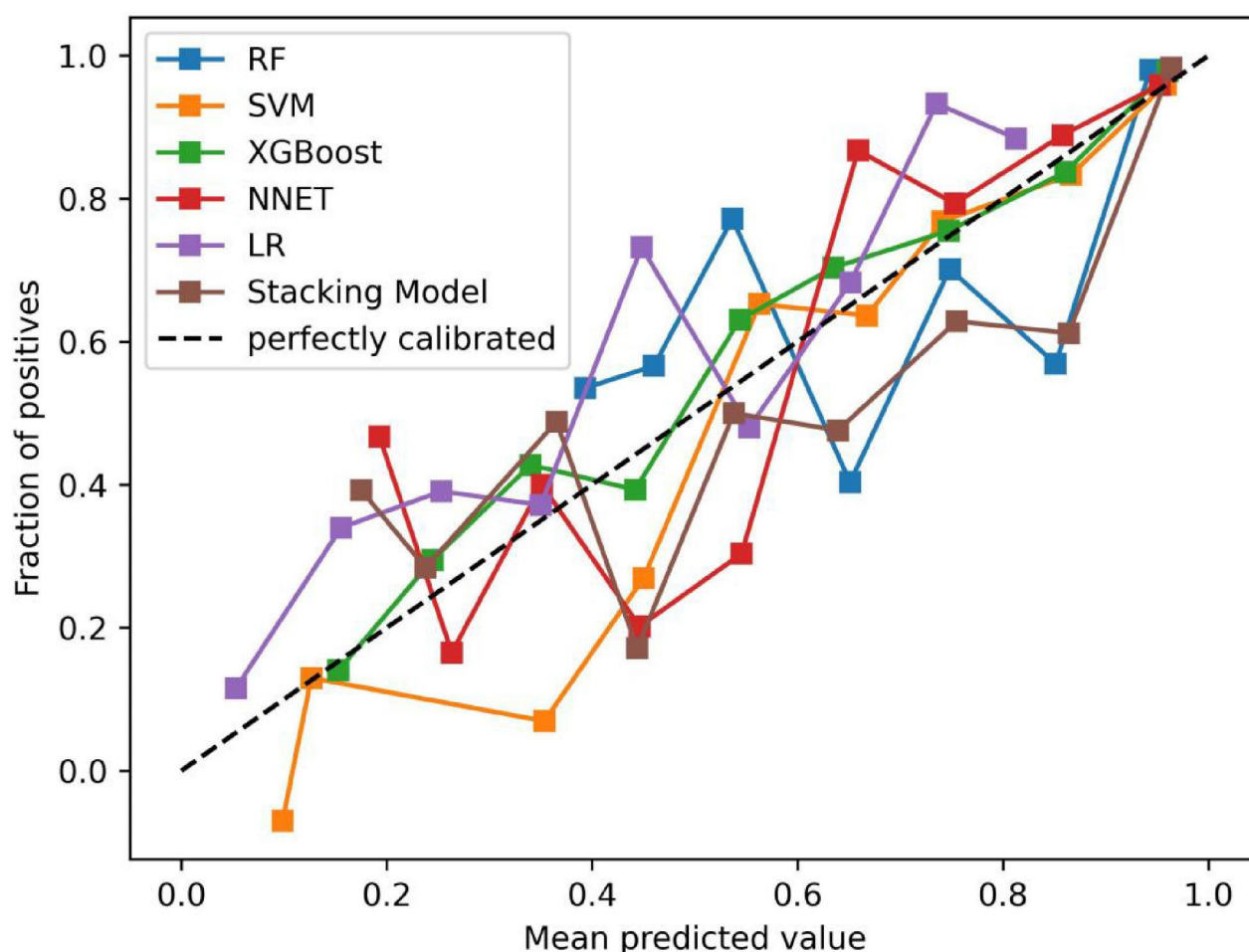


Fig. 3. Calibration curves for the test set of six models.

hidden layer size is 80. The “kernel” parameter controls the kernel function type in the SVM model, with “rbf,” “linear,” “poly,” and “sigmoid” kernel types tested. As depicted in Fig. 4c, the model’s AUC reaches its highest value when the kernel function type is “sigmoid.”

External validation

We validated the Stacking model (AUC=0.832) using an external validation set consisting of 269 patients from other hospitals (Fig. 5a). The DCA curves (Fig. 5b) and calibration curves (Fig. 5c) for the model on the external validation set demonstrated its excellent clinical utility and calibration. Consequently, we concluded that the Stacking model can accurately predict CF in elderly patients with MHD.

Deployment of models on web platforms

The Stacking model’s web page deployment and prediction result visualization were implemented using the Streamlit library. Upon logging into the webpage via the local area network (LAN), users are prompted to input the relevant index data, followed by clicking the prediction button to initiate the forecasting process. The fundamental web interface of the model is depicted in Fig. 6.

Discussion

Analysis of variables related to CF risk in elderly patients with MHD

Age

This study demonstrated that age is a significant risk factor for CF in elderly patients undergoing MHD. Consistent with previous research^{12,13}, age correlates with an increased risk of CF in MHD patients. As patients age, their physiological functions progressively decline, muscle strength in the limbs diminishes, the rate of cerebral white matter atrophy accelerates, and overall organismal weakening intensifies, thereby augmenting the likelihood of developing CF.

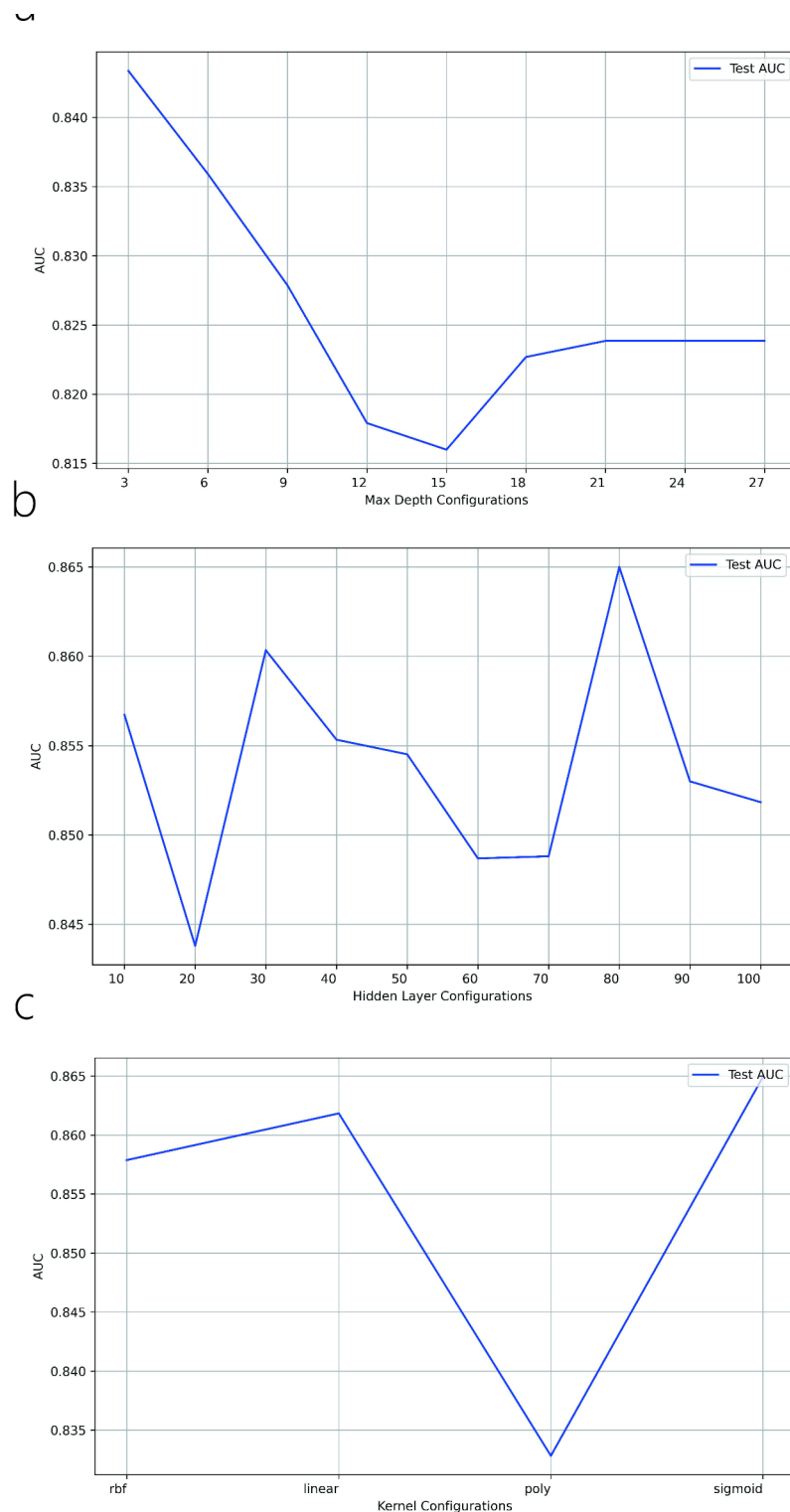


Fig. 4. Model parameter tuning curves. **(a)** Max depth configurations, **(b)** Hidden layer configurations, **(c)** Kernel configurations.

Mode of residence

Compared to living alone, cohabiting with family members not only increases opportunities for older individuals to engage in social activities but also broadens their external social network, exposing them to more impactful information. In a familial setting, older adults can better share experiences and insights with family members, a process less prevalent in healthcare facilities. This interaction fosters cognitive enrichment and supports the maintenance of healthy behaviors. Moreover, family interactions offer emotional support and strengthen social

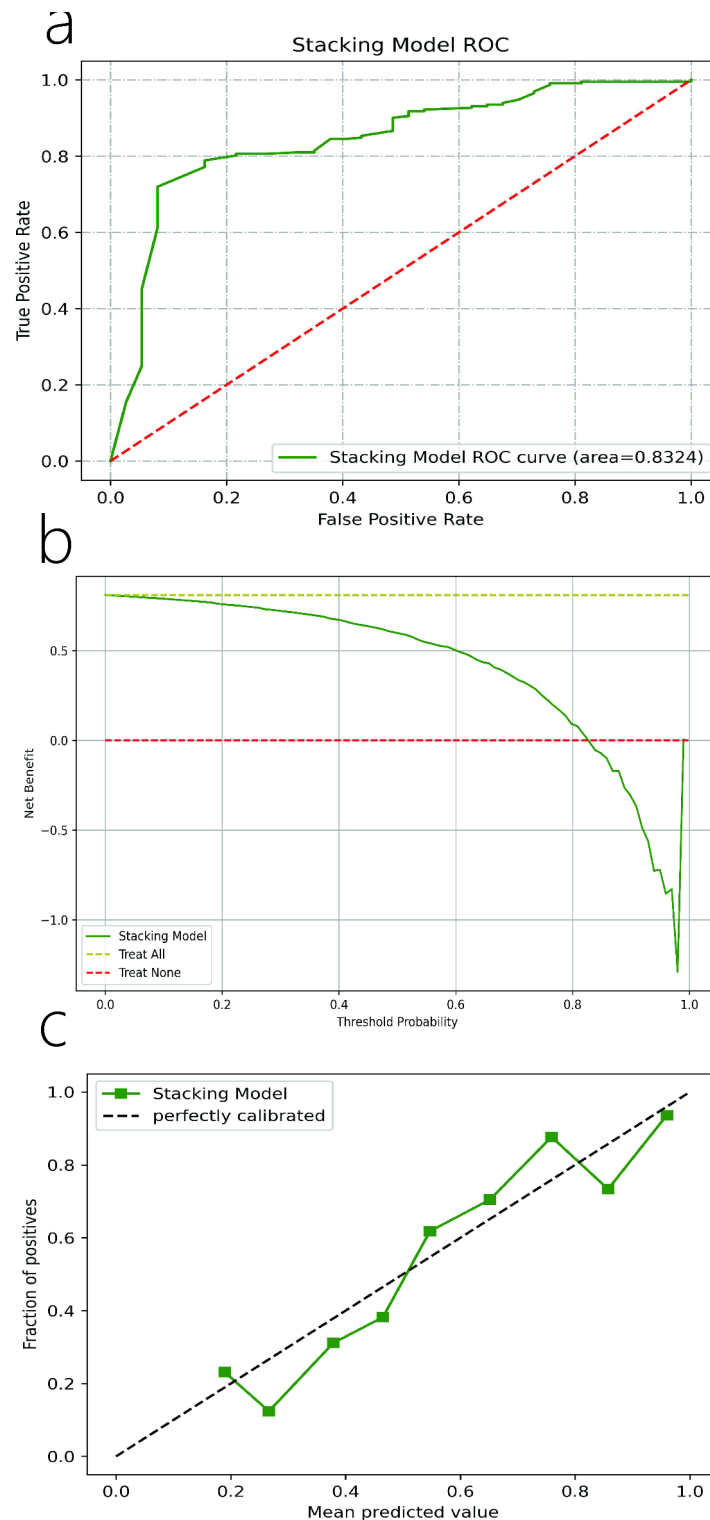


Fig. 5. ROC curve (a), DCA curve (b), and calibration curve (c) of the Stacking model on the external validation set.

connections. Therefore, residing with family members positively influences the overall health and well-being of older adults, thereby mitigating the risk of CF¹⁴.

Medical payment method

Compared to employee health insurance, the risk of CF occurrence was higher in the resident health insurance and new rural medical cooperative health insurance groups, aligning with national studies¹⁵. Older adults

Risk Prediction of Cognitive Frailty in Elderly Patients Undergoing Maintenance Hemodialysis

Age	Serum Albumin Classification
69 - +	3 - +
Mode of residence	Serum Phosphorus Classification
2 - +	1 - +
Medical Payment Method	Total Cholesterol Classification
2 - +	2 - +
Exercise	Blood Urea Nitrogen Classification
2 - +	2 - +
Alcohol consumption	Total Malnutrition Score
1 - +	5 - +
Dialysis Vascular Access	Total Depression Score
1 - +	3 - +
<div>Start to predict</div>	

The risk of cognitive frailty is: 19.94%

Fig. 6. Display of the web prediction interface.

covered by employee health insurance benefit from higher reimbursement rates and more reliable access to healthcare services. In contrast, those under resident health insurance or the new rural medical cooperative health insurance are more likely to face a more significant financial burden in managing disease-related expenses and other healthcare needs.

Exercise

The study by Scholar Tingting Jiang¹⁶ indicated that physical activity was a significant factor influencing the occurrence of CF in MHD patients. Liu et al.¹⁷ observed that exercise could mitigate the risk of CF in these patients. This may be attributed to the positive effects of exercise on cerebral blood flow and oxygen supply, which enhance cerebral circulation and improve cognitive function. Additionally, exercise can increase muscle fiber size or quantity, reduce skeletal muscle cell apoptosis, and improve motor function, thereby lowering the incidence of CF. These findings suggest that healthcare professionals should pay increased attention to elderly MHD patients who engage in limited physical activity and implement targeted exercise interventions.

Alcohol consumption

It has been observed that alcohol consumption affects cognitive function¹⁸. Chronic alcohol intake may result in neuronal damage and cause central nervous system depression, manifesting as reduced sensory perception, prolonged reaction times, memory impairment, diminished motor function, and impaired judgment¹⁹.

Consequently, alcohol consumption is regarded as a critical influencing factor for CF and should be incorporated into the care plan for elderly patients undergoing MHD.

Dialysis vascular access

In this study, we found that elderly MHD patients with long-term central venous cannulation had a higher risk of CF than those using autologous arteriovenous endovascular fistulae or artificially vascularized arteriovenous endovascular fistulae. A foreign survey on the prevalence of debilitation and cognitive dysfunction in MHD patients revealed that complications related to vascular access contributed to an increased prevalence among the elderly²⁰. A domestic study showed that MHD patients with permanent vascular access were less prone to debilitation. This may be attributed to the prevention of frailty associated with repeated vascular punctures and the durable and reliable structure of permanent vascular access, which also facilitates adequate dialysate flow, ensuring effective dialysis and the maintenance of overall health²¹. Therefore, healthcare professionals should closely monitor MHD patients with long-term central venous catheterization, providing meticulous catheter care and vascular protection to minimize repeated punctures, reduce vascular injury, and maintain the proper function of the access.

Serum albumin

Studies have shown^{12,16,22} that low serum albumin levels are a risk factor for CF in elderly MHD patients. As dialysis duration increases, patients typically exhibit signs of malnutrition, which is reflected internally by a decline in serum albumin levels. Serum albumin levels serve as an indicator of nutritional status, and optimal nutritional status can help mitigate the systemic inflammatory response and delay the deterioration of vital organs such as the heart and brain²³. Therefore, healthcare professionals should routinely screen for malnutrition and identify at-risk patients to minimize the occurrence of CF.

Serum phosphorus

National studies have indicated that²² MHD patients with lower blood phosphorus levels are at a higher risk of CF. Disorders in calcium and phosphorus metabolism, coupled with dietary restrictions, lead to diminished protein synthesis; prolonged inactivity further exacerbates the reduction in muscle mass and strength. Consequently, patients are more prone to experiencing decreased grip strength and slowed gait speed²⁴. Healthcare professionals should enhance the assessment of blood phosphorus levels and debilitated conditions in MHD patients, recognizing the potential impact on cognitive function.

Total cholesterol

The risk of CF in elderly MHD patients is closely associated with abnormal total cholesterol levels. A study by Zhang Zanying et al. found that serum total cholesterol levels in the debilitated group were significantly higher than in the non-debilitated group²⁵. Additionally, some studies have highlighted a correlation between elevated total cholesterol levels in the blood of elderly individuals and the onset of cognitive dysfunction²⁶. Therefore, regular monitoring of total cholesterol levels, along with the assessment of related debilitating conditions and cognitive functions, is particularly crucial for elderly MHD patients.

Blood urea nitrogen

Elevated blood urea nitrogen (BUN) levels can place MHD patients in a highly toxic state, often accompanied by severe complications such as acidosis, electrolyte imbalances, and gastrointestinal dysfunction. These complications compromise the patient's nutritional status and mobility, thereby contributing to the onset of frailty. Therefore, it is essential to ensure regular and adequate dialysis. In cases of high BUN levels, increasing the frequency of dialysis sessions or modifying the dialysis modality is necessary to effectively preserve the patients' residual renal function and prevent the development of debilitation²⁷.

Malnutrition

Malnutrition is a significant risk factor for frailty in elderly MHD patients. It acts as a pathophysiological mechanism contributing to both cognitive impairment and physical deterioration in these patients²⁸. Chronic malnutrition can hinder neuronal regeneration in the brain, leading to neurotransmitter imbalances, which in turn damage brain structures and accelerate cognitive decline²⁹. Additionally, malnutrition is closely associated with physical frailty, resulting in decreased body mass and reduced skeletal muscle mass, thereby increasing the risk of fractures, falls, and other accidents³⁰. Therefore, healthcare professionals should collaborate with dietitians to develop personalized nutritional management plans to improve patients' nutritional status to prevent or mitigate frailty.

Depression

Depression is a significant risk factor for frailty in MHD patients. It may be associated with oxidative stress, chronic inflammation, cerebrovascular disease, cerebral white matter lesions, and mitochondrial dysfunction, among other pathological factors³¹. MHD patients with depression often exhibit reduced volitional activity, lack of motivation and initiative, slowed cognition, and decreased social interaction and physical activity, all of which increase the risk of cognitive impairment and physical frailty^{32,33}. Therefore, healthcare professionals should conduct a comprehensive assessment of the patient's psychological state, offer individualized psychological support, and encourage active participation in social activities. These measures are essential for the further prevention of CF²².

In the risk factors screened in this study, in addition to the more widely recognized risk factors for CF in previous studies, such as age, exercise, serum albumin, depression, etc., some specific factors related to the

development of CF in elderly patients with MHD, such as the patient's dialysis vascular access, blood urea nitrogen, etc., were also included. Therefore, the risk factors used in this study are more comprehensive for MHD elderly patients, more targeted and specific, and more suitable to be applied to the screening of CF in MHD elderly patients. Meanwhile, the analysis results of this study showed that there were statistical differences between patients in the CF and non-CF groups in terms of dialysis vascular access and blood urea nitrogen ($P < 0.2$), indicating that the inclusion of MHD-specific indicators in this study is meaningful.

Analysis of CF risk prediction models for MHD patients

Traditional CF risk prediction models are primarily based on logistic regression, which, due to the limitations of the logistic regression algorithm, fails to accurately identify relevant influencing factors and cannot fully leverage the data characteristics. Additionally, selecting factors has not been comprehensive, resulting in limited model accuracy. This study conducted a cross-sectional survey to explore the influencing factors of CF in elderly patients with MHD, enriched the selection of influencing factors, and used five machine learning algorithms to establish a risk prediction model. An optimal integrated model was constructed based on the Stacking method. By comparing the optimal Stacking model with the five single models, the optimal model was identified to provide a tool for clinical healthcare workers to predict the risk of CF in elderly MHD patients at an early stage. The Stacking model and the five single models performed well, with AUC values ≥ 0.844 , showing good differentiation, calibration, and clinical utility. The external validation set confirmed the accuracy of the Stacking model in predicting CF in MHD patients (AUC = 0.832). The web deployment of the model using the Streamlit library enables real-time online prediction, enhancing the model's utility.

Limitations

This study has several limitations. First, participants were not randomly recruited due to resource and time constraints, which may have resulted in the underrepresentation of individuals with CF. Second, using a cross-sectional design raises concerns regarding the causal relationship between predictor variables and CF. Future studies could address this limitation by adopting a prospective cohort study design, which would help establish causality and enhance the predictive accuracy of the model.

Generalizability

The successful deployment of this model through a web-based interface offers the potential for its broader application in real-world clinical settings. With this tool, healthcare professionals can perform real-time risk assessments of patients to guide treatment decisions and improve patient outcomes. In terms of generalizability, the model developed in this study has the potential to be extended to other diseases with similar risk factors to CF. In addition, as machine learning technology continues to evolve, future studies may explore using more advanced algorithms, such as deep learning, to further improve prediction accuracy. In conclusion, the combination of machine learning and web-based deployment for CF risk prediction represents a significant step forward in incorporating artificial intelligence into clinical decision-making. The continued development and refinement of these models are expected to play an essential role in the future of individualized medicine.

Conclusion

Single models constructed with five different algorithms based on machine learning methods and the Stacking model are considered valuable tools for assessing CF risk in MHD patients, with the Stacking model regarded as the best-performing model. It was observed that factors such as age, mode of residence, medical payment method, exercise, alcohol consumption, dialysis vascular access, serum albumin classification, serum phosphorus classification, total cholesterol classification, blood urea nitrogen classification, malnutrition score, and depression score significantly influence the occurrence of CF in MHD patients. Therefore, targeted preventive measures should be implemented to address these factors. These models enable healthcare professionals to objectively assess the risk probability of CF in elderly MHD patients and provide a scientific basis for precise intervention.

Data availability

The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request. Data are located in controlled access data storage at Chengdu Medical College.

Received: 8 July 2024; Accepted: 13 January 2025

Published online: 20 January 2025

References

1. Raja, S. M. & Seyoum, Y. Intradialytic complications among patients on twice-weekly maintenance hemodialysis: an experience from a hemodialysis center in Eritrea. *BMC Nephrol.* **21**(1), 163 (2020).
2. Dartigues, J. F. & Amieva, H. Cognitive frailty: rationale and definition from an (I.A.N.A./I.A.G.G.) international consensus group. *J. Nutr. Health Aging.* **18**(1), 1–6 (2014).
3. Qiu, Y. et al. Prevalence of cognitive frailty among community-dwelling older adults: a systematic review and meta-analysis. *Int. J. Nurs. Stud.* **125**, 104112 (2022).
4. Kim, H. et al. Cognitive frailty in community-dwelling older Japanese people: prevalence and its association with falls. *Geriatr. Gerontol. Int.* **19**(7), 647–653 (2019).
5. Liu, L. K. et al. Cognitive frailty and its association with all-cause mortality among community-dwelling older adults in Taiwan: results from I-Lan Longitudinal Aging Study. *Rejuvenation Res.* **21**(6), 510–517 (2018).
6. Fried, L. P. et al. Frailty in older adults: evidence for a phenotype. *J. Gerontol. Biol. Sci. Med. Sci.* **56**(3), M146–M156 (2001).

7. Wang, W. & Wang, L. N. Montreal Cognitive Assessment Scale in the screening of mild cognitive impairment patients. *Chin. J. Intern. Med.* **46**(5), 414–416 (2007).
8. Buysse, D. J., Reynolds, C. F. 3rd, Monk, T. H., Berman, S. R. & Kupfer, D. J. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* **28**(2), 193–213 (1989).
9. Molnar, M. Z. et al. Association of the malnutrition inflammation score with clinical outcomes in kidney transplant recipients. *Am. J. Kidney Dis.* **58**(1), 101–108 (2011).
10. Sheikh, J. I. & Yesavage, J. A. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin. Gerontol.* **5**, 165–173 (1986).
11. Xiao, S. Y. & Yang, D. S. The impact of social support on physical and mental health. *Chin. J. Mental Health* **4**, 183–187 (1987).
12. Li, M. Establish a frailty risk prediction model for elderly patients with maintenance hemodialysis. *Gen. Pract. Nurs.* **21**(10), 1392–1396 (2023).
13. Qin, T., Li, H. N., Wang, J. Z. & Zhu, X. L. Construction and validation of a frailty risk nomogram model for patients with maintenance hemodialysis. *Qilu Nurs. J.* **29**(7), 13–18 (2023).
14. Sun, H. et al. Study on frailty and cognitive frailty status of elderly people in Zhengzhou community. *South. China J. Prev. Med.* **47**(10), 1262–1266 (2021).
15. Zhou, Q. X. *Study on the Current Status and Influencing Factors of Frailty and Cognitive Frailty in Elderly People in Chongqing Community* (Chongqing Medical University, 2020).
16. Jiang, T. T. et al. Study on the status of cognitive frailty in elderly hemodialysis patients and its correlation with fear of falling. *Chin. Nurs. Manag.* **20**(7), 1005–1009 (2020).
17. Liu, Z. Y. et al. Effect of 24-month physical activity on cognitive frailty and the role of inflammation: the LIFE randomized clinical trial. *BMC Med.* **16**(1), 185–195 (2018).
18. Zhao, C., Liu, S. S., Li, J. B., Wang, S. Y. & Wang, L. D. Exploration of the incidence rate and related risk factors of frailty syndrome in elderly patients with primary hypertension. *Pract. Geriatr.* **37**(2), 147–150 (2023).
19. Yao, F. J., Wang, F. & Li, J. Survey on the current status of frailty and cognitive frailty in patients undergoing maintenance peritoneal dialysis and analysis of influencing factors. *Med. Educ. Mod. Nurs.* **7**(3), 259–263 (2024).
20. Gopinathan, J. C. et al. The prevalence of frailty and its association with cognitive dysfunction among elderly patients on maintenance hemodialysis: a cross-sectional study from South India. *Kidney Med.* **5**(4), 100613 (2023).
21. Wu, S. L., Song, J., Xiao, P., Zang, X. M. & Jang, Q. Q. Systematic review of the incidence and influencing factors of frailty in patients undergoing maintenance hemodialysis. *Chin. Nurs. Educ.* **18**(4), 352–357 (2021).
22. Chen, G. J., Zhang, H. L., Yin, L. X., Du, X. J. & Zhang, Y. P. Current status and influencing factors analysis of cognitive frailty in maintenance hemodialysis patients. *Chin. Nurs. Manag.* **21**(8), 1179–1185 (2021).
23. Kang, S. S., Chang, J. W. & Park, Y. S. Nutritional status predicts 10-year mortality in patients with end-stage renal disease on hemodialysis. *Nutrients* **9**(4), 399–411 (2017).
24. Sabatino, A., Cuppari, L., Stenvinkel, P., Lindholm, B. & Avesan, C. M. Sarcopenia in chronic kidney disease: what have we learned so far? *J. Nephrol.* **33**(3), 365–376 (2020).
25. Zhang, Z. Y. & Gulibannu, L. Y. Q. Study on the correlation between dyslipidemia, high-sensitivity C-reactive protein and frailty in elderly. *J. Cardiovasc. Dis. Integr. Tradit. Chin. Western Med.* **7**(33), 80, 82 (2019).
26. Zhao, B. Y. et al. The gender- and age-dependent relationships between serum lipids and cognitive impairment: a cross-sectional study in a rural area of Xi'an, China. *Lipids Health Dis.* **18**, 4 (2019).
27. Ren, X. M. *Clinical Study on Distribution Characteristics and Influencing Factors of Traditional Chinese Medicine Syndrome in Frailty Patients Undergoing Maintenance Hemodialysis* (Tianjin University of Traditional Chinese Medicine, 2023).
28. Tajbakhsh, R. et al. Effect of hemodialysis on oxidants and antioxidant factors in chronic renal failure. *Saudi J. Kidney Dis. Transpl.* **28**(3), 507–516 (2017).
29. Radić, J. et al. Cognitive psychomotor functions and nutritional status in maintenance hemodialysis patients: are they related? *Ther. Apher. Dial.* **15**(6), 532–539 (2011).
30. Hara, H. et al. Protein energy wasting and sarcopenia in dialysis patients. *Contrib. Nephrol.* **196**, 243–249 (2018).
31. Jung, S., Lee, Y. K., Choi, S. R., Hwang, S. H. & Noh, J. W. Relationship between cognitive impairment and depression in dialysis patients. *Yonsei Med. J.* **54**(6), 1447–1453 (2013).
32. Kwan, R. Y. C. et al. Cognitive frailty and its association with nutrition and depression in community dwelling older people. *J. Nutr. Health Aging* **23**(10), 943–948 (2019).
33. Yuan, H. et al. Exploring psychosocial factors associated with frailty incidence among patients undergoing maintenance hemodialysis. *J. Clin. Nurs.* **29**(9/10), 1695–1703 (2020).

Author contributions

Meng Cao: Conceptualization; Methodology; Literature retrieval; Data acquisition; Data analysis; Writing - Original Draft. Bixia Tang: Conceptualization; Methodology; Literature retrieval; Data acquisition; Data analysis; Writing - Original Draft. Jing Zeng: Supervised the overall project. Liwei Yang: Supervised the overall project.

Funding

This study was supported by Chengdu Seventh People's Hospital(23LHQYZ-02).

Declarations

Competing interests

Author Meng Cao receives research support from the joint project of Chengdu Seventh People's Hospital titled "Construction and Validation of a Risk Prediction Model for Cognitive Decline in Elderly Patients Undergoing Maintenance Hemodialysis Based on Machine Learning". Author Meng Cao reports a relationship with Sichuan Vocational College of Health and Rehabilitation that provides for: employment. Author Bixia Tang reports a relationship with Chengdu Seventh People's Hospital that provides for: employment. Liwei Yang reports a relationship with Chengdu Medical College that provides for: employment. Jing Zeng reports a relationship with Chengdu Medical College that provides for: employment.

Ethical approval

The study was ethically cleared by the Ethics Committee of Chengdu Medical College (Ethics number: 2023NO.116), the study was conducted by the Declaration of Helsinki, and informed consent was obtained from all participants.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-86715-3>.

Correspondence and requests for materials should be addressed to L.Y. or J.Z.

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

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

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

© The Author(s) 2025, corrected publication 2025