BMC Gastroenterology

RESEARCH

Open Access

Leveraging machine learning for precision medicine: a predictive model for cognitive impairment in cholestasis patients

Caixia Fang[1,2], Lina Zhang[3], Lanlan Xu[1], Yongsheng He[2], Xuerong Zhang[2] and Xiaojuan Xing[3*]

Abstract

Background Cholestasis, characterized by impaired bile flow, impacts cognitive function through systemic mechanisms, including inflammation and metabolic dysregulation. Despite its significance, targeted predictive models for cognitive impairment in cholestasis remain underexplored. This study addresses this gap by developing a machine learning-based predictive model tailored to this population.

Methods Clinical and biochemical data from Qingyang People's Hospital (2021–2023) were used to train and validate models for predicting cognitive impairment (MoCA ≤ 17). Recursive feature elimination identified critical predictors, while LightGBM and other machine learning models were evaluated. SHAP analysis enhanced model interpretability, and clinical utility was assessed through decision curve analysis (DCA).

Results LightGBM outperformed other models with an AUC of 0.7955 on the testing dataset. Age, plasma D-dimer, and albumin were key predictors. SHAP analysis revealed non-linear interactions among features, demonstrating the model's clinical alignment. DCA confirmed its utility in improving patient stratification.

Conclusion The developed LightGBM-based model effectively predicts cognitive impairment in cholestasis patients, providing actionable insights for early intervention. Integrating this tool into clinical workflows can enhance precision medicine and improve outcomes in this high-risk population.

Keywords Cholestasis, Cognitive impairment, Machine learning, LightGBM, Predictive modeling

Introduction

Cholestasis is a clinical syndrome characterized by impaired bile flow, which disrupts the enterohepatic circulation and leads to the accumulation of bile components in the bloodstream [1]. Although primarily associated with liver dysfunction, cholestasis exerts systemic effects that extend beyond hepatic damage, impacting cardiovascular, metabolic, and neurocognitive systems [2]. Emerging research highlights the deleterious impact of cholestasis on cognitive function, particularly in domains such as memory, attention, and executive functioning [3]. The underlying mechanisms appear multifaceted, involving chronic systemic inflammation [4], oxidative stress [5], gut-liver-brain axis dysregulation [6], and nutritional deficiencies such as hypoalbuminemia and fat-soluble vitamin depletion [7]. Despite this growing body of evidence, cognitive dysfunction in cholestasis patients often remains underdiagnosed and poorly understood in clinical practice, necessitating robust

*Correspondence:
Xiaojuan Xing
59984456@qq.com
[1] Department of Pharmacy, Clinical Trial Research Center of Qingyang People's Hospital, Qingyang, Gansu, China
[2] Clinical Trial Research Center, Qingyang People's Hospital, Qingyang, Gansu, China
[3] Department of Neurology, Qingyang People's Hospital, Qingyang, Gansu, China

predictive models to facilitate early identification and intervention [8].

The MoCA has been widely recognized as a gold standard for assessing cognitive function due to its sensitivity and multidomain coverage, making it particularly effective in detecting early and moderate-to-severe cognitive impairments [9]. Unlike traditional cognitive screening tools such as the Mini-Mental State Examination (MMSE), MoCA incorporates evaluations of executive function, visuospatial skills, and language, which are often the first domains to decline in systemic diseases like cholestasis [9, 10]. While MoCA has demonstrated utility in patients with neurodegenerative diseases and cerebrovascular disorders, its application in hepatobiliary conditions, particularly in cholestasis, remains underexplored [11]. Previous studies have suggested that cognitive impairment in cholestasis patients shares similarities with cognitive dysfunction seen in other systemic diseases [12]. Although MoCA has not been widely validated in cholestatic patients specifically, its sensitivity in detecting early cognitive deficits in liver-related conditions, such as hepatic encephalopathy, provides a foundation for its potential use in cholestasis [13]. Studies focusing on cholestasis have primarily been descriptive, detailing its association with hepatic encephalopathy or neuroinflammatory states but lacking targeted tools for stratifying cognitive risk. The diagnosis of cholestasis in this study was based on clinical, laboratory, and imaging criteria, as outlined in the guidelines established by the American Association for the Study of Liver Diseases (AASLD) and the European Association for the Study of the Liver (EASL) [14, 15]. Furthermore, existing predictive tools for cognitive impairment often fail to account for the unique biochemical and metabolic profiles of cholestasis patients, which include elevated bilirubin, altered coagulation parameters, and systemic inflammation markers such as D-dimer [16, 17]. This gap underscores the pressing need for disease-specific, data-driven approaches to predict cognitive outcomes.

The current study addresses this unmet need by leveraging machine learning to develop and validate a predictive model for moderate-to-severe cognitive impairment in cholestasis patients, defined by a MoCA score ≤ 17. Using comprehensive clinical data from patients at Qingyang People's Hospital, this study incorporates a wide range of demographic, laboratory, and clinical variables to identify those at highest risk. By applying advanced feature selection techniques and integrating multiple machine learning algorithms, the model is designed to capture complex relationships between variables that traditional statistical methods may overlook. This study not only seeks to provide clinicians with an actionable tool for early risk stratification but also aims to deepen the understanding of the interplay between cholestasis and cognitive dysfunction. Such efforts are vital for advancing precision medicine, where timely, individualized interventions can significantly improve patient outcomes in this high-risk population.

Method

Data collection and preprocessing

Data collection

This study utilized a comprehensive dataset extracted from the electronic medical record (EMR) system of Qingyang People's Hospital. The dataset included records of patients hospitalized with a diagnosis of cholestasis between January 2021 and December 2023. Data collection was carefully planned to ensure completeness, relevance, and consistency, focusing on variables that could potentially contribute to predicting the risk of moderate to severe cognitive impairment, defined as a MoCA score ≤ 17.

The collected data encompassed multiple domains:Demographics: Information such as age, gender, height, weight, and education level was gathered to describe the patient population and serve as potential predictors.

Clinical histories Patients' histories of comorbidities, including hypertension, diabetes, coronary heart disease, cerebral infarction, chronic cholecystitis, acute exacerbation of chronic cholecystitis, and liver dysfunction, were included. Additionally, histories of cardiovascular and cerebrovascular diseases, endocrine disorders, and liver-related conditions such as hepatitis B and fatty liver were captured.

Laboratory data Detailed biochemical and hematological test results were extracted, including liver enzyme levels (ALT, AST, ALP, and GGT), bilirubin profiles (total bilirubin, direct bilirubin, and indirect bilirubin), total bile acid, protein metabolism markers (total protein, globulin, and albumin), lipid profiles (total cholesterol, triglycerides, HDL-C, and LDL-C), fasting blood glucose, uric acid, and coagulation parameters (PT, APTT, TT, fibrinogen, D-dimer, and fibrin degradation products).

Cognitive assessment Cognitive function was assessed using the MoCA. Scores were recorded at admission or during follow-up as part of routine clinical evaluation, providing the primary outcome variable for the study.

Inclusion criteria were defined as follows Patients aged ≥ 18 years with a confirmed diagnosis of cholestasis based on clinical, laboratory, or imaging findings. Complete data for demographic, clinical, and laboratory

variables, along with at least one valid MoCA assessment. Hospitalization within the defined study period (January 2021–December 2023).

Exclusion criteria included Missing or incomplete data for key variables, particularly MoCA scores, laboratory tests, or demographic information. Diagnosed psychiatric or neurodegenerative disorders unrelated to cholestasis, which could independently affect cognitive function.

A standardized data extraction protocol was employed to ensure consistency across the collection process. Data were manually verified for accuracy and completeness by trained medical staff before being anonymized for analysis. These comprehensive and systematically collected data formed the foundation for robust feature selection and predictive modeling.

Data cleaning

The collected dataset underwent rigorous cleaning to ensure analytical quality. Missing data, accounting for less than 5% of the records, were removed without imputation to maintain the integrity of the dataset. Outliers were identified using the interquartile range (IQR) method and excluded to minimize their influence on statistical analyses. Continuous variables were standardized to z-scores, and categorical variables were appropriately encoded for compatibility with predictive algorithms. These cleaning steps ensured that the dataset was suitable for machine learning analysis.

Data splitting

The dataset was divided temporally to align with real-world clinical scenarios. Data from 2021 and 2022 were used as the training set while 2023 data constituted the testing set. The training set was used for feature selection, model training, and hyperparameter tuning, while the testing set served as an independent evaluation cohort for assessing model performance and generalizability.

Ethical considerations

This study adhered to the principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the institutional review board of Qingyang People's Hospital before the initiation of the study(QYRMYY[2021]–002). Patient confidentiality was protected by anonymizing all data before analysis. Only de-identified records were used, ensuring no personal identifiers were included in the analysis or reporting. These measures ensured that the research adhered to ethical and legal standards.

Baseline characteristics analysis

Baseline characteristics were analyzed to provide an overview of the clinical and demographic profiles of the study population and to compare differences between patients with and without moderate to severe cognitive impairment (MoCA $\leq$ 17). Continuous variables were summarized as mean $\pm$ standard deviation (SD) for normally distributed data or as median with interquartile range (IQR) for non-normally distributed data. Categorical variables were expressed as frequencies and percentages. Statistical comparisons were performed to assess differences between groups. For continuous variables, either independent t-tests or Mann–Whitney U tests were used, depending on data distribution. Categorical variables were analyzed using chi-squared tests, with Fisher's exact tests applied for variables with low expected counts. A significance threshold of $p < 0.05$ was used for all analyses. The results provided insight into the differences in baseline characteristics, identifying variables that might be associated with cognitive impairment for further analysis.

Correlation analysis

The association between candidate predictors and moderate to severe cognitive impairment (MoCA $\leq$ 17) was evaluated using a two-step logistic regression approach. First, univariate logistic regression analysis was performed to assess the relationship between each predictor and the outcome variable. Each predictor was modeled independently, and the results were expressed as odds ratios (ORs) with 95% confidence intervals (CIs). This step aimed to explore the potential relationship of each variable with cognitive impairment but did not apply a strict significance threshold for advancing to the next step.

Second, all candidate predictors were included in a backward stepwise selection process to identify the most informative combination of variables. This procedure, based on minimizing the Akaike Information Criterion (AIC), was performed using a multivariate logistic regression model. The backward selection algorithm systematically removed variables contributing the least to the model's explanatory power, iteratively refining the model to achieve the lowest AIC. The final model retained only those variables that minimized AIC while balancing complexity and interpretability.

The results of the final multivariate model were presented as adjusted ORs with 95% CIs. This approach allowed the identification of independently significant predictors of cognitive impairment while controlling for potential confounding effects among variables. By relying on AIC optimization rather than *p*-value thresholds

for univariate analysis, this method ensured a data-driven selection of predictors that contributed meaningfully to the outcome in combination, maximizing model efficiency and robustness.

Recursive Feature Elimination (RFE)

RFE was applied in combination with a random forest algorithm to identify the most relevant features for the predictive model. The process began with the full set of candidate features, iteratively eliminating those with the lowest importance based on the random forest's feature importance scores. After each iteration, the model was retrained and evaluated to determine the optimal feature subset. The selection criterion was based on the model's performance metrics, with the process stopping when further elimination resulted in no significant improvement or a decline in predictive accuracy. The final feature selection was guided by both the feature importance scores provided by the random forest model and the model's ability to achieve the best balance between predictive accuracy and interpretability [18]. This method allowed for the identification of a feature set that optimized model performance while maintaining a manageable number of predictors for downstream analysis.

Model development and optimization

Model construction

To develop robust predictive models for moderate to severe cognitive impairment, 11 machine learning algorithms were implemented, including decision tree, random forest, XGBoost, LightGBM, LASSO regression, Elastic Net regression, ridge regression, k-nearest neighbors (KNN), support vector machine (SVM), multilayer perceptron (MLP), and logistic regression. These models were selected to encompass a range of algorithmic approaches, including linear, non-linear, and tree-based methods, ensuring comprehensive exploration of predictive performance. Each model was trained on the preprocessed training dataset, with hyperparameters initialized to standard values before further tuning.

Model validation

The performance of each model was evaluated using ten-fold cross-validation on the training dataset. This validation strategy divided the dataset into 10 equal parts, using 9 parts for training and 1 part for validation in each iteration. The AUC was used as the primary evaluation metric, quantifying the model's ability to discriminate between patients with and without moderate to severe cognitive impairment. Cross-validation ensured a reliable estimate of each model's generalizability and mitigated overfitting.

To optimize performance, hyperparameter tuning was conducted using Bayesian optimization. This process systematically explored the hyperparameter space to identify configurations that maximized AUC. For tree-based models, parameters such as learning rate, maximum depth, and number of leaves were optimized. For SVM, the kernel type and regularization parameter were adjusted. This systematic optimization ensured that each model was tuned to its maximum potential.

Model integration

To further enhance predictive performance, a stacking ensemble model was constructed. The outputs of the 11 individual models (predicted probabilities) were used as input features for a meta-model, which was trained to produce the final prediction. Logistic regression was selected as the meta-model due to its simplicity and interpretability. This integration strategy aimed to harness the complementary strengths of the individual models, producing a combined prediction with potentially superior accuracy and discrimination.

Performance evaluation on training and testing sets

The predictive performance of all 12 models, including the stacking ensemble, was evaluated on both the training set and the independent testing set. The primary evaluation metric was the AUC, which measures the ability of a model to discriminate between classes. Comparisons focused on identifying the model with the highest AUC on the testing set, reflecting its generalizability and applicability to unseen data.

Evaluation of clinical utility

Decision curve analysis

Decision curve analysis (DCA) was conducted to evaluate the clinical utility of the best-performing model identified in the previous steps. DCA assesses the net benefit of using the model across a range of risk thresholds, comparing its performance to default strategies of treating all patients or treating none. By integrating the trade-offs between true-positive and false-positive predictions at each threshold, DCA provides a quantitative measure of the model's value in guiding clinical decision-making.

Clinical impact curve

The clinical impact curve (CIC) was constructed to further assess the practical implications of the best-performing model. The CIC visualizes the number of patients classified as high-risk and the corresponding number of true positives across different risk thresholds. This analysis complements the DCA by providing a clearer understanding of the model's ability to correctly

identify patients at risk while minimizing unnecessary interventions.

Model interpretability analysis

To enhance the interpretability of the best-performing model, SHapley Additive exPlanations (SHAP) were employed. SHAP values provide a unified framework to explain the contribution of each feature to the model's predictions, offering insights into both global and individual-level model behavior.

Global SHAP feature importance was visualized to identify the overall contribution of each feature to the model's predictive performance. This plot ranks features by their average absolute SHAP values, highlighting the most influential predictors across the dataset. For individual-level explanations, SHAP contribution plots were generated to illustrate how specific features influenced the prediction for a single patient. These plots detail the magnitude and direction of each feature's impact, providing a transparent and actionable interpretation of the model's decision-making process. SHAP interaction plots were created to explore interactions between features, demonstrating how two variables jointly influence the model's predictions.

Result

Baseline characteristics analysis

Table 1 summarizes the baseline characteristics of the study population, comparing individuals with moderate to severe cognitive impairment (MoCA ≤ 17, referred to as the impaired cognition group, \($N=57$ \)) and those without significant impairment (MoCA >17, referred to as the preserved cognition group, \($N=296$ \)). Individuals in the impaired cognition group had a significantly older median age of 72 years compared to 58 years in the preserved cognition group (\($p<0.001$ \)).

The impaired cognition group also exhibited notable differences in nutritional and metabolic markers. Mean height, total protein, albumin, and calcium concentrations were significantly lower in this group ($p=0.029$, $p=0.015$), ($p=0.023$ \), and ($p=0.030$), respectively). Additionally, median uric acid levels were reduced ($p=0.006$)), while plasma D-dimer levels were markedly elevated ($p<0.001$), indicating potential alterations in both metabolic and coagulation pathways.

Differences in chronic health conditions were also observed. The prevalence of chronic cholecystitis was higher among individuals in the impaired cognition group (64.9% vs. 48.3%, ($p=0.022$)), suggesting a possible link between chronic inflammation and cognitive decline. Educational attainment was another key differentiator, with a significantly higher proportion of individuals in the impaired cognition group being illiterate (42.1% vs. 11.5%, ($p<0.001$ \)).

These findings indicate that moderate to severe cognitive impairment is associated with advanced age, metabolic and coagulation imbalances, chronic inflammatory conditions, and lower levels of education, offering critical insights into potential risk factors contributing to cognitive decline.

Correlation analysis

Multivariate analysis of Table 2 showed that MoCA scores were associated with several important factors. The presence of chronic cholecystitis was associated with a higher likelihood of outcome events (adjusted OR, 2.87; 95% CI, 1.39–5.96; $p=0.005$). Conversely, liver dysfunction (adjusted OR, 0.32; 95% CI, 0.11–0.95; $p=0.040$) and a history of cardiovascular or cerebrovascular events (adjusted OR, 0.30; 95% CI, 0.12–0.71; $p=0.007$) were linked to a decreased likelihood. Females were more likely to experience the outcome compared to males (adjusted OR, 2.35; 95% CI, 1.05–5.26; $p=0.038$). Educational attainment showed a trend, with primary school (adjusted OR, 0.26; 95% CI, 0.10–0.66; $p=0.005$) and middle school education (adjusted OR, 0.34; 95% CI, 0.12–1.00; $p=0.050$) being associated with lower odds of the outcome compared to illiteracy. Each unit increase in age was associated with a rise in likelihood (adjusted OR, 1.08; 95% CI, 1.04–1.12; $p<0.001$), whereas increases in total bile acid levels reduced the likelihood (adjusted OR, 0.99; 95% CI, 0.98–1.00; $p=0.019$). Similar effects were observed for total cholesterol levels, where higher levels were associated with increased risk (adjusted OR, 1.37; 95% CI, 1.02–1.85; $p=0.038$). Absolute neutrophil count (adjusted OR, 1.72; 95% CI, 1.03–2.89; $p=0.039$) and neutrophil percentage (adjusted OR, 0.93; 95% CI, 0.87–0.99; $p=0.024$) also showed significant associations, as did lymphocyte percentage (adjusted OR, 0.91; 95% CI, 0.84–0.99; $p=0.027$) and white blood cell count (adjusted OR, 0.60; 95% CI, 0.37–0.98; $p=0.040$). Platelet count was marginally associated with increased risk (adjusted OR, 1.01; 95% CI, 1.00–1.01; $p=0.031$). While the association between triglyceride levels and the outcome was not statistically significant (adjusted OR, 1.22; 95% CI, 0.95–1.55; $p=0.114$), the INR showed a pronounced but non-significant trend towards significance (adjusted OR, 8.57; 95% CI, 0.81–90.38; $p=0.074$).

Feature selection results

Feature selection was conducted using recursive feature elimination (RFE) combined with a random forest algorithm to identify the most relevant predictors for moderate to severe cognitive impairment. The analysis ranked features based on their contribution to the

Table 1 Patient demographics and baseline characteristics

Characteristic	MoCA		*p*-value
	No, $N = 296$[1]	Yes, $N = 57$[1]	
Age	58 (48, 68)	72 (58, 77)	<0.001[2]
Height	170 (161, 175)	165 (160, 173)	0.029[2]
Weight	62 (55, 70)	60 (55, 65)	0.119[2]
Fasting_Blood_Glucose_mmolL	5.45 (4.62, 6.85)	5.93 (4.85, 6.48)	0.845[2]
ALT_U_L	29 (17, 88)	35 (15, 82)	0.726[2]
AST_U_L	28 (20, 63)	31 (21, 108)	0.352[2]
ALP_U_L	99 (77, 170)	103 (81, 152)	0.449[2]
GGT_U_L	38 (18, 219)	42 (18, 247)	0.424[2]
Total_Bilirubin_TBIL_umolL	18 (11, 36)	22 (12, 47)	0.181[2]
Direct_Bilirubin_umolL	5 (3, 15)	8 (4, 23)	0.084[2]
Indirect_Bilirubin_umolL	11 (8, 20)	13 (8, 23)	0.390[2]
Total_Bile_Acid_umolL	5 (3, 13)	6 (3, 13)	0.355[2]
Total_Protein_TP_gL	66 (61, 70)	63 (57, 70)	0.015[2]
Globulin_GLO_gL	26.8 (24.0, 29.5)	25.2 (22.6, 29.3)	0.100[2]
Albumin_ALB_gL	38.9 (35.4, 42.0)	36.4 (33.3, 40.6)	0.023[2]
Urea_mmolL	5.28 (4.27, 6.57)	6.08 (4.73, 7.54)	0.013[2]
Calcium_mmolL	2.24 (2.14, 2.32)	2.19 (2.09, 2.30)	0.030[2]
Uric_Acid_umolL	255 (191, 314)	221 (153, 279)	0.006[2]
Total_Cholesterol_TCHO_mmolL	3.66 (2.99, 4.25)	3.56 (3.07, 4.33)	0.876[2]
Triglycerides_TG_mmolL	1.20 (0.83, 1.68)	0.99 (0.65, 1.46)	0.037[2]
HDL_Cholesterol_HDL_C_mmolL	1.00 (0.85, 1.19)	1.07 (0.87, 1.24)	0.339[2]
LDL_Cholesterol_LDL_C_mmolL	2.15 (1.62, 2.79)	2.10 (1.69, 2.83)	0.926[2]
Neutrophils_Absolute_NEUT_10.9L	4.5 (3.2, 7.2)	4.7 (3.5, 7.0)	0.640[2]
Neutrophils_Percentage_NEUT_Percent	71 (61, 85)	73 (63, 85)	0.531[2]
Lymphocytes_Absolute_10.9L	1.13 (0.77, 1.65)	1.13 (0.73, 1.43)	0.414[2]
Lymphocytes_Percentage_Percent	21 (16, 29)	20 (12, 26)	0.165[2]
White_Blood_Cell_WBC	6.4 (5.3, 9.1)	6.3 (5.1, 8.2)	0.816[2]
Hemoglobin_HGB_gL	134 ± 19	130 ± 20	0.140[3]
Platelets_PLT_10.9L	175 (123, 220)	165 (134, 236)	0.671[2]
APTT_Coagulation_Time	28.0 (25.8, 32.0)	28.0 (26.0, 31.0)	0.849[2]
Prothrombin_Time_PT	12.70 (11.90, 13.60)	12.90 (12.30, 14.00)	0.062[2]
Thrombin_Time_TT	15.76 (14.90, 16.66)	16.10 (15.10, 16.80)	0.219[2]
INR_International_Normalized_Ratio	1.06 (1.01, 1.13)	1.10 (1.04, 1.18)	0.009[2]
Fibrinogen_FIB_gL	2.68 (2.24, 3.56)	3.05 (2.39, 3.71)	0.115[2]
Plasma_D_Dimer_ugml	0.24 (0.11, 0.48)	0.43 (0.25, 0.99)	<0.001[2]
Fibrin_Degradation_Products_ugml	1.9 (1.2, 3.6)	3.0 (1.6, 6.4)	0.006[2]
Common_Bile_Duct_Stone			0.076[4]
No	131 (44.3%)	18 (31.6%)	
Yes	165 (55.7%)	39 (68.4%)	
Hypertension			0.939[4]
No	201 (67.9%)	39 (68.4%)	
Yes	95 (32.1%)	18 (31.6%)	
Diabetes			0.199[4]
No	264 (89.2%)	54 (94.7%)	
Yes	32 (10.8%)	3 (5.3%)	
Coronary_Heart_Disease			0.438[4]
No	251 (84.8%)	46 (80.7%)	
Yes	45 (15.2%)	11 (19.3%)	

Table 1 (continued)

Characteristic	MoCA		*p*-value
	No, $N = 296$[1]	Yes, $N = 57$[1]	
Tumor			0.759[5]
No	279 (94.3%)	53 (93.0%)	
Yes	17 (5.7%)	4 (7.0%)	
Cerebral_Infarction			0.542[4]
No	228 (77.0%)	46 (80.7%)	
Yes	68 (23.0%)	11 (19.3%)	
Chronic_Cholecystitis			0.022[4]
No	153 (51.7%)	20 (35.1%)	
Yes	143 (48.3%)	37 (64.9%)	
Acute_Exacerbation_Chronic_Cholecystitis			0.166[4]
No	228 (77.0%)	39 (68.4%)	
Yes	68 (23.0%)	18 (31.6%)	
Liver_Dysfunction			0.719[4]
No	238 (80.4%)	47 (82.5%)	
Yes	58 (19.6%)	10 (17.5%)	
Hepatitis_B			0.395[5]
No	288 (97.3%)	54 (94.7%)	
Yes	8 (2.7%)	3 (5.3%)	
Fatty_Liver			0.208[4]
No	245 (82.8%)	51 (89.5%)	
Yes	51 (17.2%)	6 (10.5%)	
Cardiovascular_Cerebrovascular_History			0.047[4]
No	194 (65.5%)	45 (78.9%)	
Yes	102 (34.5%)	12 (21.1%)	
Endocrine_Disease_History			0.485[5]
No	281 (94.9%)	56 (98.2%)	
Yes	15 (5.1%)	1 (1.8%)	
Gender			0.143[4]
male	201 (67.9%)	33 (57.9%)	
female	95 (32.1%)	24 (42.1%)	
Education_Level			< 0.001[4]
illiteracy	34 (11.5%)	24 (42.1%)	
Primary school	117 (39.5%)	18 (31.6%)	
Middle school	116 (39.2%)	12 (21.1%)	
University and above	29 (9.8%)	3 (5.3%)	

[1] Median (IQR); Mean ± SD; n (%)

[2] Wilcoxon rank sum test

[3] Welch Two Sample t-test

[4] Pearson's Chi-squared test

[5] Fisher's exact test

model's predictive performance, as measured by importance scores. In particular, the inclusion of variables like age and plasma D-dimer was driven by their strong association with cognitive decline observed in previous clinical research, which indicated their potential relevance in predicting cognitive impairment in cholestasis patients. Figure 1A shows the ranking of variable importance, with fasting blood glucose, total protein, plasma D-dimer, albumin, and calcium identified as the most influential predictors. Additional important features included HDL cholesterol and other markers of metabolic and coagulation status. These variables demonstrated strong contributions to the model's ability to discriminate between individuals with and

Table 2 Univariate and multivariate analysis of influencing factors (Logistic regression)

Characteristic	Univariable			Multivariable		
	OR[1]	95% CI[1]	*p*-value[2]	OR[1]	95% CI[1]	*p*-value[2]
Common_Bile_Duct_Stone						
No	—	—				
Yes	1.72	0.94, 3.15	0.078			
Hypertension						
No	—	—				
Yes	0.98	0.53, 1.80	0.939			
Diabetes						
No	—	—				
Yes	0.46	0.14, 1.55	0.210			
Coronary_Heart_Disease						
No	—	—				
Yes	1.33	0.64, 2.77	0.440			
Tumor						
No	—	—		—	—	
Yes	1.24	0.40, 3.83	0.710	0.27	0.06, 1.10	0.067
Cerebral_Infarction						
No	—	—				
Yes	0.80	0.39, 1.63	0.543			
Chronic_Cholecystitis						
No	—	—		—	—	
Yes	1.98	1.10, 3.57	0.023*	2.87	1.39, 5.96	0.005**
Acute_Exacerbation_Chronic_Cholecystitis						
No	—	—				
Yes	1.55	0.83, 2.88	0.168			
Liver_Dysfunction						
No	—	—		—	—	
Yes	0.87	0.42, 1.83	0.719	0.32	0.11, 0.95	0.040*
Hepatitis_B						
No	—	—				
Yes	2.00	0.51, 7.78	0.317			
Fatty_Liver						
No	—	—				
Yes	0.57	0.23, 1.39	0.213			
Cardiovascular_Cerebrovascular_History						
No	—	—		—	—	
Yes	0.51	0.26, 1.00	0.051	0.30	0.12, 0.71	0.007**
Endocrine_Disease_History						
No	—	—				
Yes	0.33	0.04, 2.58	0.294			
Gender						
male	—	—		—	—	
female	1.54	0.86, 2.75	0.145	2.35	1.05, 5.26	0.038*
Education_Level						
illiteracy	—	—		—	—	
Primary school	0.22	0.11, 0.45	< 0.001***	0.26	0.10, 0.66	0.005**
Middle school	0.15	0.07, 0.32	< 0.001***	0.34	0.12, 1.00	0.050*
University and above	0.15	0.04, 0.54	0.004**	0.41	0.08, 2.12	0.285
Age	1.07	1.04, 1.09	< 0.001***	1.08	1.04, 1.12	< 0.001***

Table 2 (continued)

Characteristic	Univariable			Multivariable		
	OR[1]	95% CI[1]	*p*-value[2]	OR[1]	95% CI[1]	*p*-value[2]
Height	0.97	0.95, 1.00	0.073			
Weight	0.98	0.95, 1.01	0.112			
Fasting_Blood_Glucose_mmolL	0.95	0.84, 1.08	0.423			
ALT_U_L	1.00	1.00, 1.00	0.302			
AST_U_L	1.00	1.00, 1.00	0.922			
ALP_U_L	1.00	1.00, 1.00	0.589			
GGT_U_L	1.00	1.00, 1.00	0.785			
Total_Bilirubin_TBIL_umolL	1.00	1.00, 1.01	0.082			
Direct_Bilirubin_umolL	1.00	1.00, 1.01	0.070			
Indirect_Bilirubin_umolL	1.01	1.00, 1.02	0.111			
Total_Bile_Acid_umolL	1.00	0.99, 1.00	0.675	0.99	0.98, 1.00	0.019*
Total_Protein_TP_gL	0.97	0.94, 1.00	0.070			
Globulin_GLO_gL	0.98	0.92, 1.04	0.452			
Albumin_ALB_gL	0.97	0.94, 1.01	0.107			
Urea_mmolL	1.07	0.97, 1.19	0.156			
Calcium_mmolL	0.14	0.02, 0.93	0.042*			
Uric_Acid_umolL	1.00	0.99, 1.00	0.003**	1.00	0.99, 1.00	0.066
Total_Cholesterol_TCHO_mmolL	1.18	0.96, 1.45	0.109	1.37	1.02, 1.85	0.038*
Triglycerides_TG_mmolL	1.09	0.91, 1.31	0.352	1.22	0.95, 1.55	0.114
HDL_Cholesterol_HDL_C_mmolL	1.17	0.73, 1.88	0.515			
LDL_Cholesterol_LDL_C_mmolL	1.14	0.86, 1.50	0.363			
Neutrophils_Absolute_NEUT_10.9L	1.03	0.95, 1.11	0.487	1.72	1.03, 2.89	0.039*
Neutrophils_Percentage_NEUT_Percent	1.01	0.99, 1.03	0.537	0.93	0.87, 0.99	0.024*
Lymphocytes_Absolute_10.9L	0.77	0.50, 1.18	0.230			
Lymphocytes_Percentage_Percent	0.98	0.96, 1.01	0.165	0.91	0.84, 0.99	0.027*
White_Blood_Cell_WBC	1.00	0.92, 1.08	0.994	0.60	0.37, 0.98	0.040*
Hemoglobin_HGB_gL	0.99	0.97, 1.00	0.124			
Platelets_PLT_10.9L	1.00	1.00, 1.00	0.591	1.01	1.00, 1.01	0.031*
APTT_Coagulation_Time	1.00	0.95, 1.05	0.967			
Prothrombin_Time_PT	1.11	0.97, 1.28	0.132			
Thrombin_Time_TT	1.04	0.88, 1.22	0.671			
INR_International_Normalized_Ratio	6.17	1.00, 37.99	0.050*	8.57	0.81, 90.38	0.074
Fibrinogen_FIB_gL	1.09	0.92, 1.28	0.330			
Plasma_D_Dimer_ugml	1.12	0.94, 1.33	0.193			
Fibrin_Degradation_Products_ugml	1.03	1.00, 1.07	0.061			

[1] *OR* Odds Ratio, *CI* Confidence Interval

[2] *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Null deviance = 312; Null df = 352; Log-likelihood = −110; AIC = 259; BIC = 336; Deviance = 219; Residual df = 333; No. Obs. = 353

without cognitive impairment. As shown in Fig. 1B, the RFE process determined that the optimal model performance was achieved using 13 features. This is indicated by the peak accuracy marked by the red dashed line, where further addition of features did not result in improved performance. These 13 features provided the best trade-off between predictive accuracy and model complexity, serving as the final subset of predictors used in subsequent modeling and validation steps.

Model performance results

The performance of the models was evaluated using ROC curves and AUC values for both the training and testing datasets, as shown in Fig. 2A and B.In the training

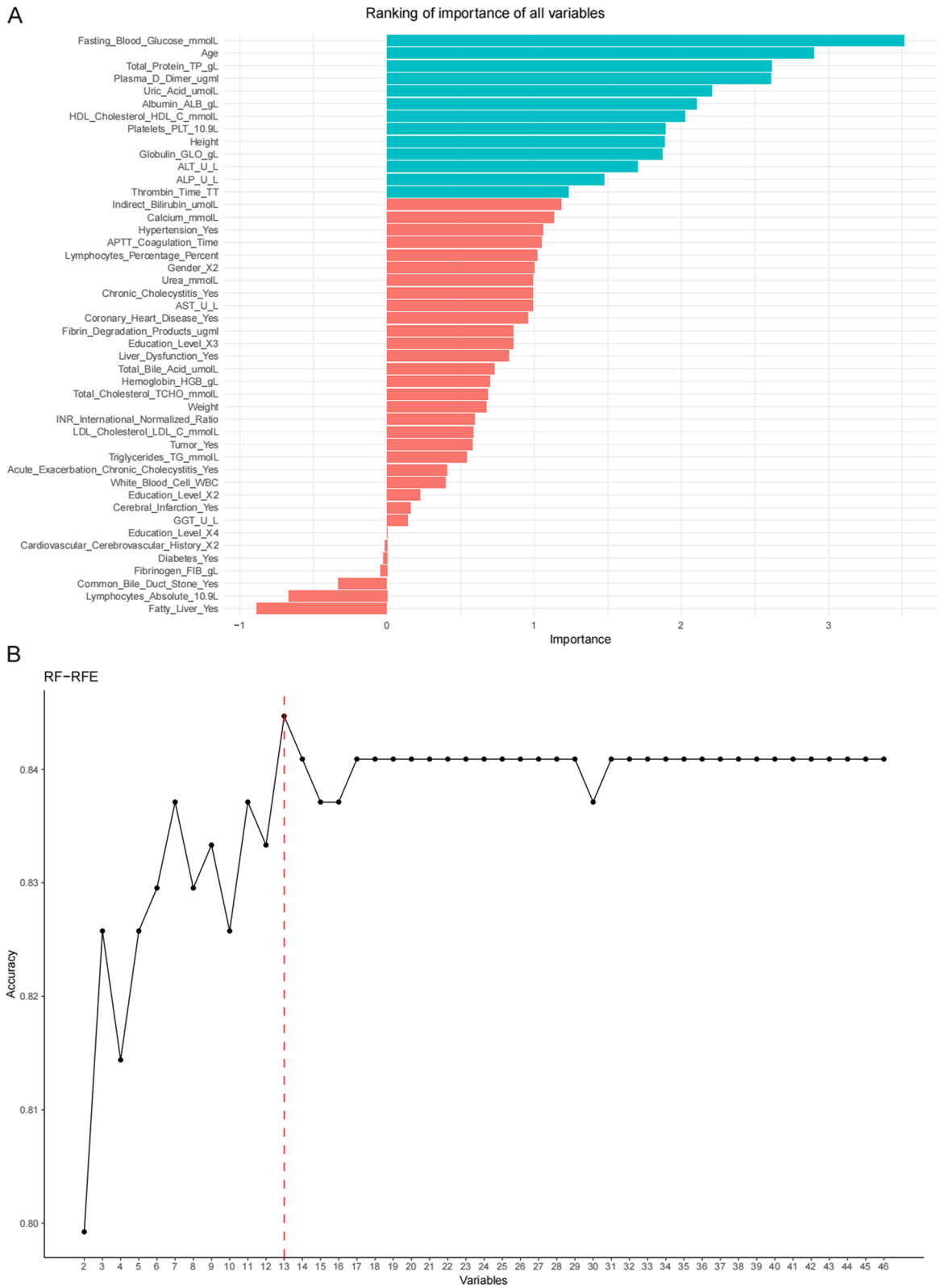

Fig. 1 Baseline characteristics and feature selection results. **A** Clinical and laboratory characteristics of the study population, stratified by cognitive impairment status (MoCA ≤ 17). Significant group differences are marked. **B** Feature importance ranking using recursive feature elimination (RFE) with a random forest algorithm. (C) Final selection of 13 features contributing most significantly to the predictive model

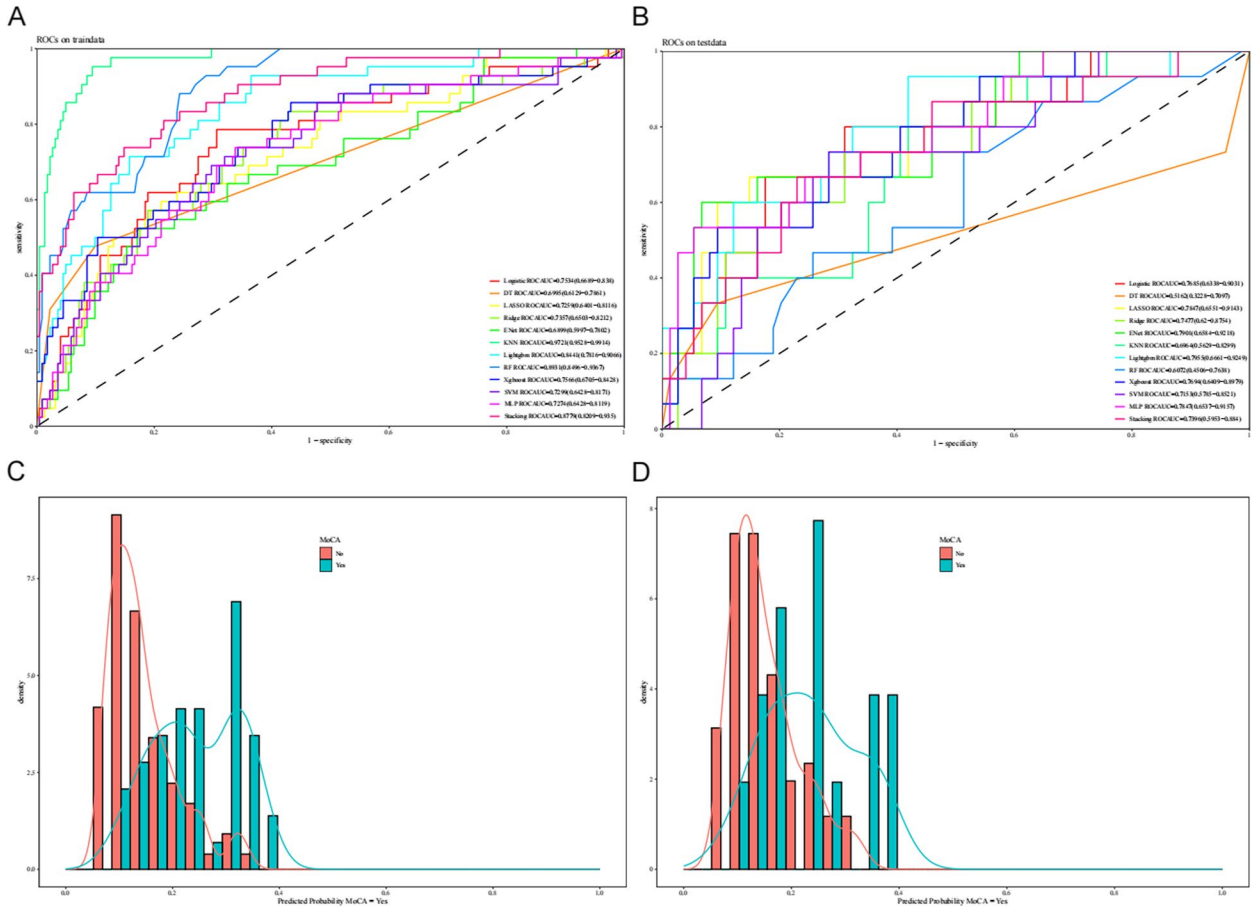

Fig. 2 Performance of machine learning models. **A**, **B** ROC curves for training (**A**) and testing (**B**) datasets comparing all models. LightGBM shows the best balance of performance and generalizability. **C**, **D** Predicted probability distributions for training (**C**) and testing (**D**) datasets, showing the separation between cognitive impairment and non-impairment groups

dataset (Fig. 2A), the random forest model achieved the highest AUC (0.8931), followed by the stacking ensemble (0.8779) and LightGBM (0.8441). However, in the testing dataset (Fig. 2B), LightGBM demonstrated the highest AUC (0.7955), outperforming all other models, including the stacking ensemble and random forest, which experienced more pronounced drops in performance. Considering both training and testing results, LightGBM strikes the best balance between high discriminatory ability and robust generalization, making it the most suitable model for predicting moderate to severe cognitive impairment while minimizing the risk of overfitting.

Figure 2C and D show the predicted probability distributions of the LightGBM model for moderate to severe cognitive impairment (MoCA ≤ 17) in the training and testing datasets, respectively. In the training dataset (Fig. 2C), the positive class (teal) and negative class (red) are well-separated, indicating that the LightGBM model effectively differentiates between the two groups based on learned patterns. In the testing dataset (Fig. 2D), the separation remains evident but with a slightly larger overlap, reflecting the real-world variability in unseen data. These results demonstrate the LightGBM model's strong discriminatory ability and reliable generalization across both datasets.

Clinical utility evaluation

Figure 3A and C present the decision curve analysis (DCA) and clinical impact curve (CIC) for the LightGBM model in the training dataset, while Fig. 3B and D display the corresponding results for the testing dataset. These analyses assess the clinical utility and practical effectiveness of the model in identifying individuals with moderate to severe cognitive impairment (MoCA ≤ 17).

In the training dataset (Fig. 3A), the DCA shows that the LightGBM model provides a higher net benefit across a wide range of threshold probabilities compared to treating all or none. This suggests that the model effectively balances sensitivity and specificity, making it valuable for guiding clinical decisions. The CIC

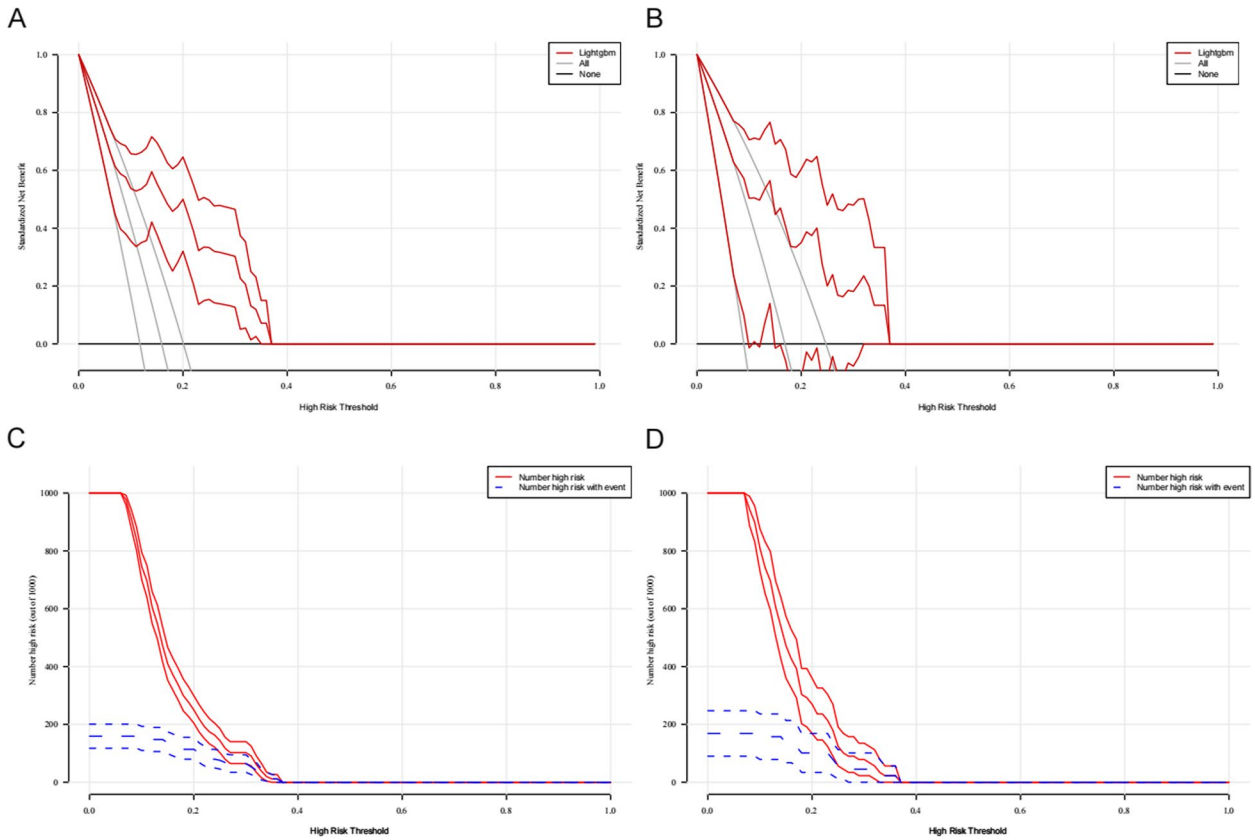

Fig. 3 Clinical utility of the LightGBM model. **A**, **B** Decision curve analysis (DCA) for training (**A**) and testing (**B**) datasets, showing net benefit across risk thresholds. **C**, **D** Clinical impact curve (CIC) for training (**C**) and testing (**D**) datasets, illustrating the alignment of predicted high-risk cases with true positives

in Fig. 3C indicates that the model accurately identifies high-risk individuals, as shown by the alignment of true positives (blue line) with the total number of predicted high-risk cases (red line) at various thresholds.

In the testing dataset (Fig. 3B), the DCA confirms the model's generalizability, maintaining a higher net benefit across clinically relevant thresholds. Similarly, the CIC in Fig. 3D demonstrates consistent performance in predicting high-risk individuals, with a clear distinction between true positives and overestimated cases.

SHAP-based model interpretability

The SHAP analysis provides a detailed explanation of how the LightGBM model predicts moderate to severe cognitive impairment (MoCA $\leq$ 17) by assessing the contribution of features at both global and individual levels.

Figures 4A and B present the overall importance of features in the LightGBM model. Age and plasma D-dimer were identified as the most influential predictors, with higher age and elevated plasma D-dimer levels strongly increasing the risk of cognitive impairment. Albumin and platelet counts were also significant, where low albumin levels and high platelet counts were associated with

(See figure on next page.)

Fig. 4 Model Interpretability Using SHAP Analysis (**A**) SHAP summary plot showing the overall importance of features in the LightGBM model, with age, plasma D-dimer, and albumin identified as the most influential predictors. The color gradient represents the feature values, where higher or lower values impact predictions in specific directions. **B** Mean absolute SHAP values ranking feature importance, highlighting the significant contribution of age, plasma D-dimer, and metabolic markers. **C** SHAP dependence plots illustrating the relationships between feature values and their contributions to the model, revealing nonlinear effects for age, plasma D-dimer, and albumin. **D** SHAP interaction plot showing the interplay between age and plasma D-dimer, demonstrating how their combined effects influence predictions. **E** SHAP waterfall plot providing an individual prediction explanation, detailing how specific features cumulatively contribute to the predicted probability of cognitive impairment

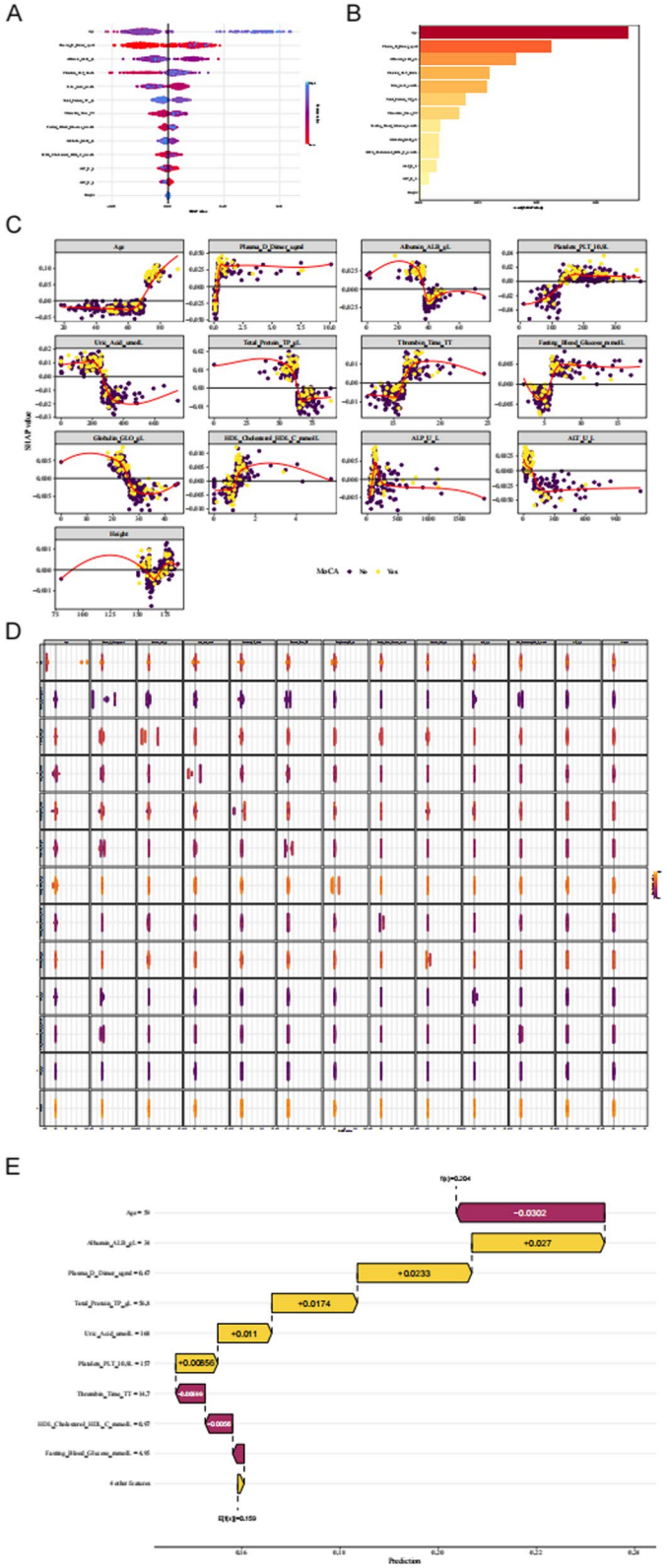

Fig. 4 (See legend on previous page.)

elevated risk. Figure 4B ranks these features by their average SHAP contributions, highlighting their central roles in the model's predictions.

Figure 4C explores the relationship between feature values and their SHAP contributions. The likelihood of cognitive impairment increases sharply with age beyond 65 years and with plasma D-dimer levels exceeding 0.5 μg/mL. Nonlinear effects are evident for features such as uric acid and thrombin time, where both high and low values contribute to predictions in different ways. These plots illustrate the nuanced and context-specific influence of features, providing deeper insights into the clinical risk factors associated with cognitive impairment.

Figure 4D examines interactions between features. Notably, the combined effect of age and plasma D-dimer levels significantly increases the predicted risk of cognitive impairment, demonstrating how these features interact synergistically. Other interactions, such as those between albumin and total protein, further highlight the importance of evaluating predictors in combination rather than in isolation.

Figure 4E provides a detailed decomposition of the prediction for a single individual. Starting from the baseline prediction ($E[f(x)] = 0.159$), each feature contributes incrementally to the final probability ($f(x) = 0.204$). For this individual, elevated plasma D-dimer levels (+0.027) and low albumin levels (+0.0233) were the main contributors to the higher risk, while age (−0.0302) slightly reduced the probability. This personalized explanation clarifies how the model arrives at its predictions, enhancing its interpretability and trustworthiness in clinical contexts.

Discussion

This study identified several key predictors associated with the risk of moderate to severe cognitive impairment (MoCA ≤ 17) in patients with cholestasis, including age, plasma D-dimer, albumin, and platelet count [19]. Age emerged as a critical factor, with older patients demonstrating a significantly higher risk, while elevated plasma D-dimer and low albumin levels indicated potential contributions from coagulation abnormalities and nutritional deficiencies, respectively. These findings underscore the multifactorial nature of cognitive impairment in cholestasis, highlighting the interplay between metabolic, inflammatory, and hemostatic pathways. Compared to other diseases, such as neurodegenerative disorders or cerebrovascular diseases, cholestasis presents unique cognitive impairments associated with distinct metabolic and coagulation abnormalities, such as elevated bilirubin and D-dimer levels [20–22]. The LightGBM model demonstrated superior performance among the 12 machine learning algorithms tested, achieving the highest AUC on the independent testing dataset. While it did not show the highest AUC on the training dataset, its consistent performance on unseen data underscores its robust generalizability and resistance to overfitting. Compared to other models, such as random forest and stacking ensembles, LightGBM demonstrated the least overfitting, as evidenced by its lower AUC difference between the training dataset (AUC 0.8441) and the testing dataset (AUC 0.7955). The AUC gap between training and testing datasets is understandable and acceptable in machine learning models, as some variability between datasets is typical. Moreover, LightGBM's performance on the testing dataset still remains strong, suggesting its robust generalization and clinical applicability.

In addition to LightGBM, other classification models such as random forest and stacking ensembles were evaluated. Random forest, for instance, showed competitive performance with an AUC of 0.7905 on the testing dataset, while the stacking ensemble achieved an AUC of 0.7872.This balance between discrimination and generalization positions LightGBM as the optimal model for predicting cognitive impairment in this population [23]. SHAP analysis provided critical insights into the clinical interpretability of the model. By identifying feature importance and visualizing their contributions to individual predictions, SHAP reinforced the model's alignment with clinical reasoning. For instance, the significant impact of plasma D-dimer and age on individual predictions aligns with established clinical evidence, enabling clinicians to understand not only the "what" but also the "why" behind the model's outputs, enhancing trust and applicability in practice.

To further assess the clinical utility of the model, decision curve analysis (DCA) was performed. DCA demonstrated that the LightGBM model provides a higher net benefit across a wide range of threshold probabilities compared to treating all or none. Figures 3A and B present the DCA results for both the training and testing datasets, clearly showing that the model provides actionable insights for clinical decision-making, where using the model as a risk assessment tool would improve patient outcomes. Moreover, the clinical impact curve (CIC), shown in Fig. 3C and D, further supports the model's clinical utility by demonstrating the alignment of predicted high-risk cases with actual true positives, offering additional evidence of the model's value in real-world clinical scenarios.

The developed LightGBM-based predictive model holds significant promise for assisting clinicians in early identification of high-risk patients with cholestasis [24]. By stratifying risk based on key predictors such as plasma D-dimer and albumin, the model enables targeted interventions to mitigate or delay the progression of cognitive

impairment [25]. This proactive approach could significantly improve patient outcomes, particularly in a vulnerable population where cognitive decline often remains underrecognized [26]. Integrating the model into electronic medical record (EMR) systems offers the potential for real-time risk assessment during routine clinical care. With automated data processing, clinicians can receive immediate risk estimates, supporting timely decision-making and resource allocation [27]. This integration would enhance the efficiency and precision of patient management, aligning predictive analytics with day-to-day clinical workflows. The use of machine learning in this context highlights its transformative potential to complement traditional clinical methods. Unlike conventional models, machine learning algorithms can uncover complex, nonlinear relationships between predictors, providing a more nuanced understanding of disease risk [28]. This capability to identify subtle interactions and dependencies positions machine learning as a valuable tool in advancing personalized medicine, particularly in diseases with multifactorial etiologies like cholestasis [29].

This study advances the current understanding of cognitive impairment in cholestasis patients by introducing a machine learning-based predictive model, a novel approach not previously explored in this population. Prior studies, such as those focusing on hepatic encephalopathy, have identified cognitive impairment as a secondary outcome of cholestasis, often attributing it to systemic inflammation and neurotoxic metabolite accumulation [30], these works primarily relied on descriptive or correlational analyses without integrating predictive methodologies [25]. In contrast, this study highlights the importance of specific metabolic and coagulation factors, such as D-dimer and albumin levels, as key predictors of cognitive decline, distinguishing it from studies on neurodegenerative diseases or cerebrovascular conditions, which typically emphasize structural or neurochemical markers [31, 32]. The use of machine learning in this study, particularly LightGBM, demonstrates the ability to handle complex, high-dimensional datasets and identify non-linear interactions among predictors, such as age, plasma D-dimer, and albumin [26]. Unlike traditional statistical approaches, such as logistic regression, which are limited in capturing intricate feature relationships, the model used in this study achieved superior performance and interpretability through SHAP analysis [33, 34]. This interpretability allows for direct clinical insights, such as the identification of specific metabolic markers, consistent with earlier literature that emphasizes the role of coagulation and inflammation markers in cognitive outcomes [35]. By buundational findings, this study bridges the gap between observational research and actionable predictive tools, offering a transformative approach to precision medicine for cholestasis-related cognitive impairment [19, 36]. One limitation of this study is the single-center dataset from Qingyang People's Hospital, which may limit the generalizability of the findings to other populations. The cohort was largely composed of patients from a specific geographical region, potentially introducing regional biases that may not fully reflect the diversity of the broader cholestasis patient population.

Additionally, the lack of external validation with data from other centers or diverse populations limits the assessment of the model's robustness. The use of static clinical and laboratory data, without longitudinal measures such as follow-up MoCA scores, may also constrain the model's ability to capture dynamic changes over time. Future work should focus on expanding datasets to include multi-center and diverse populations to improve generalizability. Incorporating additional biomarkers, such as inflammatory markers or neuroimaging, could enhance predictive accuracy. Prospective validation studies and the development of adaptive models capable of real-time updates are crucial to ensure clinical applicability and evaluate the model's impact on patient outcomes.

Conclusion

This study developed and validated multiple machine learning models based on clinical data from patients with cholestasis at Qingyang People's Hospital, ultimately identifying LightGBM as the optimal predictive model. The LightGBM model demonstrated strong predictive performance and clinical utility, providing a robust tool for early identification of patients at risk of moderate to severe cognitive impairment. By integrating key predictors such as age, plasma D-dimer, and albumin, the model offers a scientific basis for risk stratification and supports targeted clinical interventions. This work highlights the potential of machine learning in advancing precision medicine for cholestasis-related complications and improving patient outcomes.

Authors' contributions

CXF and XJX designed the study. CXF and LNZ performed the study and analyzed the data. LLX and CXF wrote the manuscript. YSH and XRZ provided the expert consultations and clinical suggestions.CXF and XJX conceived of the study, XRZ participated in its design and coordination, ALL authors helped to draft the manuscript.

Funding

This study was supported by grants from Science and Technology Program of Qingyang City(QY2021A-S031), Science and Technology Innovation Platform and Talents Program of Qingyang City (QCSJ-[2022]-42–33).

Data availability
The data used in this study can be obtained from the corresponding author on reasonable grounds.

Declarations

Ethics approval and consent to participate
This study was approved by the Qingyang Municipal People's Hospital Ethics Committee (QYRMYY[2021]-002). Written informed consent to participate was obtained from all participants prior to their inclusion in the study.

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Received: 22 November 2024 Accepted: 18 February 2025
Published online: 18 March 2025

References

1. Niknahad H, Nadgaran A, Alidaee S, Arjmand A, Abdoli N, Mazloomi SM, Akhlagh A, Nikoozadeh A, Kashani SMA, Mehrabani PS, et al. Thiol-reducing agents abate cholestasis-induced lung inflammation, oxidative stress, and histopathological alterations. Trends Pharm Sci. 2023;9(1):55–70.
2. EslimiEsfahani D, Zarrindast MR. Cholestasis and behavioral disorders. Gastroenterol Hepatol Bed Bench. 2021;14(2):95–107.
3. Kronsten VT, Tranah TH, Pariante C, Shawcross DL. Gut-derived systemic inflammation as a driver of depression in chronic liver disease. J Hepatol. 2022;76(3):665–80.
4. Huang F, Pariante CM, Borsini A. From dried bear bile to molecular investigation: a systematic review of the effect of bile acids on cell apoptosis, oxidative stress and inflammation in the brain, across preclinical models of neurological, neurodegenerative and neuropsychiatric disorders. Brain Behav Immun. 2022;99:132–46.
5. Pierzchala K, Simicic D, Sienkiewicz A, Sessa D, Mitrea S, Braissant O, McLin VA, Gruetter R, Cudalbu C. Central nervous system and systemic oxidative stress interplay with inflammation in a bile duct ligation rat model of type C hepatic encephalopathy. Free Radical Biol Med. 2022;178:295–307.
6. Jena PK, Sheng L, Di Lucente J, Jin LW, Maezawa I, Wan YY. Dysregulated bile acid synthesis and dysbiosis are implicated in Western diet-induced systemic inflammation, microglial activation, and reduced neuroplasticity. Faseb J. 2018;32(5):2866–77.
7. Huang L-T, Tiao M-M, Tain Y-L, Chen C-C, Hsieh C-S. Melatonin ameliorates bile duct ligation-induced systemic oxidative stress and spatial memory deficits in developing rats. Pediatr Res. 2009;65(2):176–80.
8. HertišPetek T, Petek T, Močnik M, MarčunVarda N. Systemic inflammation, oxidative stress and cardiovascular health in children and adolescents: a systematic review. Antioxidants. 2022;11(5):894.
9. Adhyaru BB, Jacobson TA. Safety and efficacy of statin therapy. Nat Rev Cardiol. 2018;15(12):757–69.
10. Giuffrè M, Merli N, Pugliatti M, Moretti R. The metabolic impact of non-alcoholic fatty liver disease on cognitive dysfunction: a comprehensive clinical and pathophysiological review. Int J Mol Sci. 2024;25(6):3337.
11. Niemeyer-Guimaraes M, Parsons H, Carvalho R. Elderly cancer patients in the Intensive Care Unit (ICU): a case for the need of Palliative Care (PC) assessment. 2014.
12. Berry K, Ruck JM, Barry F, Shui AM, Cortella A, Kent D, Seetharaman S, Wong R, VandeVrede L. Lai JC. Prevalence of cognitive impairment in liver transplant recipients. Clin Transplant. 2024;38(1):e15229.
13. Giuffrè M, Merli N, Pugliatti M. Moretti R. The metabolic impact of non-alcoholic fatty liver disease on cognitive dysfunction: a comprehensive clinical and pathophysiological review. Int J Mol Sci. 2024;25(6):3337.
14. Bowlus CL, Arrivé L, Bergquist A, Deneau M, Forman L, Ilyas SI, Lunsford KE, Martinez M, Sapisochin G, Shroff R, Tabibian JH. AASLD practice guidance on primary sclerosing cholangitis and cholangiocarcinoma. Hepatology. 2023;77(2):659–702.
15. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines on the Management of Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) [published correction appears in Obes Facts. 2024;17(6):658. https://doi.org/10.1159/000541386. Obes Facts. 2024;17(4):374-444. https://doi.org/10.1159/000539371.
16. Yu L, Liu Y, Wang S, Zhang Q, Zhao J, Zhang H, Narbad A, Tian F, Zhai Q, Chen W. Cholestasis: exploring the triangular relationship of gut microbiota-bile acid-cholestasis and the potential probiotic strategies. Gut Microbes. 2023;15(1):2181930.
17. Pieters A, Gijbels E, Cogliati B, Annaert P, Devisscher L, Vinken M. Biomarkers of cholestasis. Biomark Med. 2021;15(6):437–54.
18. Senan EM, Al-Adhaileh MH, Alsaade FW, Aldhyani TH, Alqarni AA, Alsharif N, Uddin MI, Alahmadi AH, Jadhav ME. Alzahrani MY. Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques. J Healthc Eng. 2021;2021(1):1004767.
19. Ashley EA. Towards precision medicine. Nat Rev Genet. 2016;17(9):507–22.
20. Pieruccini-Faria F, Black SE, Masellis M, Smith EE, Almeida QJ, Li KZ, Bherer L, Camicioli R, Montero-Odasso M. Dementia: gait variability across neurodegenerative and cognitive disorders: results from the Canadian Consortium of Neurodegeneration in Aging (CCNA) and the Gait and Brain Study. Alzheimer's Dementia. 2021;17(8):1317–28.
21. Khaw J, Subramaniam P, Abd Aziz NA, Ali Raymond A, Wan Zaidi WA, Ghazali SE. Current update on the clinical utility of MMSE and MoCA for stroke patients in Asia: a systematic review. Int J Environ Res Public Health. 2021;18(17):8962.
22. Hosoya M, Toi S, Seki M, Saito M, Hoshino T, Yoshizawa H, Kitagawa K. Association between total cerebral small vessel disease score and cognitive function in patients with vascular risk factors. Hypertens Res. 2023;46(5):1326–34.
23. Yan J, Xu Y, Cheng Q, et al. LightGBM: accelerated genomically designed crop breeding through ensemble learning. Genome Biol. 2021;22(1):271. https://doi.org/10.1186/s13059-021-02492-y.
24. GuolinKe QM, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY. Lightgbm: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst. 2017;30:52.
25. Mursil M, Rashwan HA, Cavallé-Busquets P, Santos-Calderón LA, Murphy MM, Puig D. Maternal nutritional factors enhance birthweight prediction: a super learner ensemble approach. Information. 2024;15(11):714.
26. Meng L, Treem W, Heap GA, Chen J. A stacking ensemble machine learning model to predict alpha-1 antitrypsin deficiency-associated liver disease clinical outcomes based on UK Biobank data. Sci Rep. 2022;12(1):17001.
27. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. J Clin Oncol. 2016;34(21):2534–40.
28. Ye J, Wang S, Yang X, Tang X. Gene prediction of aging-related diseases based on DNN and Mashup. BMC Bioinformatics. 2021;22(1):597.
29. Clarke KSP, Kingdon CC, Hughes MP, et al. The search for a blood-based biomarker for Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome (ME/CFS): from biochemistry to electrophysiology. J Transl Med. 2025;23(1):149. https://doi.org/10.1186/s12967-025-06146-6.
30. Terlapu P, Sadi RP, Pondreti R, Tippana C. Intelligent identification of liver diseases based on incremental hidden layer neurons ann model. Int J Comput Digit Syst. 2022;11(1):1027–50.
31. Riva N, Attard LM, Vella K, Squizzato A, Gatt A, Calleja-Agius J. Diagnostic accuracy of D-dimer in patients at high-risk for splanchnic vein thrombosis: a systematic review and meta-analysis. Thromb Res. 2021;207:102–12.
32. Wu T, Simonetto DA, Halamka JD, Shah VH. The digital transformation of hepatology: The patient is logged in. Hepatology. 2022;75(3):724-39. https://doi.org/10.1002/hep.32329.
33. Van den Broeck G, Lykov A, Schleich M, Suciu D. On the tractability of SHAP explanations. J Artif Intell Res. 2022;74:851–86.

34. Founta K, Dafou D, Kanata E, et al. Gene targeting in amyotrophic lateral sclerosis using causality-based feature selection and machine learning. Mol Med. 2023;29(1):12. https://doi.org/10.1186/s10020-023-00603-y.
35. Wang K, Tian J, Zheng C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med. 2021;137:104813. https://doi.org/10.1016/j.compbiomed.2021.104813.
36. Kosorok MR, Laber EB. Precision medicine. Annu Rev Stat Appl. 2019;6(1):263–86.

Publisher's Note