Am J Transl Res 2025;17(7):4939-4951
www.ajtr.org /ISSN:1943-8141/AJTR0164740

Original Article

A machine learning model for non-invasive prediction of advanced liver fibrosis in patients with chronic hepatitis B

Jingwei Song[1], Ni Ma[1], Reziwanguli Aini[1], Yuqing Yang[2]

[1]*School of Public Health, Xinjiang Medical University, Urumqi 830017, Xinjiang, China;* [2]*People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi 830001, Xinjiang, China*

Received March 24, 2025; Accepted June 13, 2025; Epub July 15, 2025; Published July 30, 2025

Abstract: Purpose: Chronic Hepatitis B (CHB) is a leading cause of liver fibrosis. Accurate and non-invasive diagnosis of liver fibrosis in CHB patients is of critical clinical importance. This study aimed to develop and validate machine learning (ML)-based models for predicting significant liver fibrosis in CHB patients. Methods: This retrospective cohort study included 328 CHB patients (225 with non-significant liver fibrosis and 103 with significant liver fibrosis) from 2017 to 2022. Four ML models were constructed based on four selected features identified through the least absolute shrinkage and selection operator (LASSO) regression. Model performance was assessed using the receiver operating characteristic (ROC) curve, and the area under the curve (AUC), accuracy, sensitivity, specificity, and SHapley Additive exPlanations (SHAP) analysis. Results: The random forest (RF) model demonstrated the highest predictive performance, with an AUC of 0.874 (95% CI: 0.813-0.934) in the training set and 0.863 (95% CI: 0.772-0.955) in the test set, outperforming extreme gradient boosting (XGBoost), logistic regression (LR), and support vector machine (SVM). Compared with the traditional fibrosis indices such as aspartate aminotransferase to platelet ratio index (APRI) (AUC = 0.585) and fibrosis-4 (FIB-4) (AUC = 0.633), the RF model (AUC = 0.863) demonstrated significantly higher predictive accuracy. SHAP analysis identified platelet count (PLT) as the most influential predictor in the RF model. Conclusion: The ML-based RF model offers a highly accurate, non-invasive interpretable tool for predicting significant liver fibrosis in patients with CHB, offering potential for clinical application in routine fibrosis risk assessment.

Keywords: Chronic hepatitis B, liver fibrosis, machine learning, random forest, non-invasive diagnosis

Introduction

Chronic hepatitis B (CHB) remains a major global public health concern with severe implications for human health [1, 2]. Persistent infection with the hepatitis B Virus (HBV) can lead to chronic hepatic inflammation, gradually progressing to liver fibrosis. Without timely diagnosis and intervention, fibrosis may advance to cirrhosis or hepatocellular carcinoma, severely affecting patient prognosis and quality of life [3, 4]. Although liver biopsy is considered the "gold standard" for diagnosing liver fibrosis, it is an invasive procedure associated with risks such as bleeding and infection. Moreover, the sampling limitations of liver biopsy may lead to inaccurate assessments of the overall fibrosis extent, complicating frequent dynamic monitoring [5, 6].

Currently, numerous studies have focused on developing non-invasive diagnostic models for liver fibrosis, such as the aspartate aminotransferase-to-platelet ratio index (APRI) (AUC = 0.660) and fibrosis-4 (FIB-4) [7]. However, these models have certain limitations in accurately predicting liver fibrosis in CHB patients, particularly in distinguishing different fibrosis stages [8, 9], necessitating further improvement in diagnostic performance. In recent years, machine learning (ML) has emerged as a powerful tool in disease diagnosis and prediction. ML-based models leverage large datasets and complex feature interactions to identify hidden patterns and nonlinear relationships that traditional statistical models might overlook [10, 11]. By integrating clinical, biochemical, imaging, and genetic data, ML algorithms have demonstrated promising potential in im-

https://doi.org/10.62347/KEVQ8263

proving the accuracy, robustness, and generalizability of liver fibrosis prediction [12, 13]. Recent studies have indicated the potential of ML-based models in liver fibrosis grading. For instance, shear wave velocity-based ML algorithm outperformed median shear wave velocity in identifying significant hepatic fibrosis, with performance comparable to MR elastography-based fibrosis staging [14]. Li et al. have revealed that ML-based algorithms effectively stage liver disease by quantifying arterial density and correlating it with collagen proportionate area and METAVIR staging [15]. Despite these advancements, challenges persist in the clinical implementation of ML-based fibrosis models, including dataset heterogeneity, the need for external validation across diverse populations, and the interpretability of complex ML algorithms for clinical decision-making. Therefore, further research is needed to refine ML-based fibrosis models, optimize biomarker selection, and enhance their clinical integration for improved patient management.

In this study, we selected essential predictive variables and developed four ML-based classification models, including random forest (RF), support vector machines (SVM), extreme gradient boosting (XGBoost), and logistic pegression (LR). We evaluated their efficacy in predicting significant liver fibrosis. Our findings suggest that ML models, particularly RF, offer superior diagnostic accuracy and clinical utility compared to traditional scoring systems. By providing a non-invasive and objective fibrosis assessment approach, these models hold potential for facilitating early intervention and individualized patient management in CHB.

Materials and methods

Study population

This retrospective cohort study was conducted at Xinjiang Uygur Autonomous Region People's Hospital. Clinical data were collected from 328 patients diagnosed with CHB between January 2017 and December 2022. The included criteria were: (1) patients who met the diagnostic criteria for CHB; (2) availability of liver biopsy histopathological results; (3) complete clinical data. The exclusion criteria were: (1) presence of other liver diseases, such as non-alcoholic fatty liver disease or autoimmune hepatitis; (2) diagnosis of decompensated cirrhosis, hepatocellular carcinoma, or other severe hepatic dysfunction; (3) presence of other malignancies, or severe heart, kidney, or pulmonary diseases; (4) continued use of medications known to affect liver function for more than three months. This study was approved by the Ethics Committee of Xinjiang Uygur Autonomous Region People's Hospital (No. KY2024052420).

Diagnostic criteria

The diagnostic criteria for CHB were based on the *Guidelines for the Prevention and Treatment of Chronic Hepatitis B* (2022 Edition), which classify CHB into chronic HBV carrier state, HBeAg-positive CHB, inactive HBsAg carrier state, HBeAg-negative CHB, occult HBV infection, and compensated cirrhosis in HBV-related liver cirrhosis. The diagnostic criteria for liver fibrosis followed the *Consensus on the Diagnosis and Treatment of Liver Fibrosis* (2019 Edition), using the Scheuer scoring system to classify fibrosis into five stages. Significant fibrosis was defined as Scheuer stage ≥ S2, characterized by the presence of fibrous septa or bridging fibrosis, whereas S0-S1 was classified as non-significant fibrosis.

Feature selection

Based on previous studies on liver fibrosis and clinical practice at our hospital, a total of 32 variables were collected as candidate features. The selected variables were categorized into three main groups: (1) Baseline characteristics: sex, age, history of hypertension, history of diabetes, smoking history, alcohol consumption history, and body mass index (BMI); (2) Blood biochemical parameters: C-reactive protein (CRP), creatinine (Crea), prealbumin (PA), cystatin C (Cys_C), alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), indirect bilirubin (IBIL), total bilirubin (TBIL), direct bilirubin (DBIL), albumin (ALB), globulin (GLO), total protein (TP), alkaline phosphatase (ALP), alpha-fetoprotein (AFP), cholinesterase (ChE), total bile acid (TBA), Triglycerides (TG), total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), platelet count (PLT), prothrombin time (PT), prothrombin activity (PTA), and fibrinogen (Fbg); (3) Pathological diagnosis results.

BMI classification was based on the Chinese Adult BMI Classification Standard (WS/T 428-

2013): underweight (BMI < 18.5 kg/m^2), normal weight (18.5 kg/m^2 ≤ BMI < 24.0 kg/m^2), overweight (24.0 kg/m^2 ≤ BMI < 28.0 kg/m^2), and obese (BMI ≥ 28.0 kg/m^2).

A comprehensive screening of all 32 variables was performed using univariate statistical tests to identify factors significantly associated with liver fibrosis. To mitigate multicollinearity and reduce the risk of overfitting, the least absolute shrinkage and selection operator (LASSO) regression algorithm was employed, enabling the selection of the most informative variables by penalizing less relevant features and shrinking their coefficients to zero. Key variables identified via LASSO regression were subsequently subjected to multivariate logistic regression to evaluate their independent associations with significant liver fibrosis.

Construction and performances assessment of the ML models

Patients were randomly assigned into two groups, with 70% allocated to the training set and the remaining 30% to the testing set. Using the selected predictive factors, four ML models were developed, including LR, RF, SVM, and XGBoost. The receiver operating characteristic (ROC) curve was plotted, and the area under the curve (AUC) was calculated to examine the overall classification capability of the models. Additionally, the precision-recall (PR) curve was analyzed to evaluate the models' classification performance at various thresholds. Furthermore, the calibration curve was employed to evaluate the concordance between predicted probabilities and actual event occurrences, while the decision curve analysis (DCA) was conducted to quantify the net clinical benefit of each model across different decision thresholds. To comprehensively evaluate model performance, key metrics including the F1-score, accuracy, specificity, and sensitivity were computed and compared. Moreover, to enhance model stability and generalizability, a ten-fold cross-validation approach was employed during both model training and validation, minimizing the impact of data partitioning and improving robustness.

Visualization of the prediction by SHapley Additive exPlanations (SHAP)

SHAP is a machine learning interpretability method that quantifies the contribution of features to model predictions by calculating Shapley values. The principle behind SHAP is to objectively quantify the synergistic and independent effects of various features on a specific prediction outcome by calculating the weighted average effect of all possible feature subset combinations.

In this study, we first calculated the absolute SHAP values of each feature to quantify their contribution to the prediction model, and generated SHAP waterfall and swarm plots to visualize the results. These visualizations were compared with clinical guideline indicators to ensure that the decision-making logic of the model is in line with evidence-based medical knowledge.

Statistical analysis

Data analysis was carried out with SPSS 28.0 and Python 3.11 statistical software. For normally distributed continuous variables, the results were expressed as mean ± standard deviation ($\bar{x}$ ± s), and comparisons between groups were conducted using the t-test. For skewed distribution of continuous variables, the results were presented as median and interquartile range [M (Q1, Q3)], with comparisons conducted using non-parametric tests. Categorical variables were presented as frequency and proportions, with comparisons performed using the χ^2 test. Corrected χ^2 test or Fisher exact test was used when the expectation value did not meet the requirements of the χ^2 test. A *p*-value of < 0.05 was considered statistically significant for all analyses. The λ value corresponding to one standard error above the minimum λ (λ.1se) was chosen as the optimal threshold.

Results

Patient demographics and baseline characteristics

In this study, 328 patients with CHB were enrolled. Based on pathological findings, the patients were categorized into a non-significant liver fibrosis group (225 cases) and a significant liver fibrosis group (≥ S2; 103 cases). Statistical analysis was conducted to compare the clinical characteristics between the two groups. No substantial differences were found between the two groups regarding gender, BMI, history of hypertension, history of diabetes, alcohol

Table 1. Comparison of clinical data between the non-significant liver fibrosis and significant liver fibrosis groups

Variables	Non-significant liver fibrosis group	Significant liver fibrosis group	Statistical values	*P*
Gender (cases)				
Male	135 (60.0%)	66 (64.1%)	χ^2 = 0.495	0.482
Female	90 (40.0%)	37 (35.9%)		
Age (years)				
≤ 40	133 (59.1%)	51 (49.5%)	χ^2 = 10.235	0.006
40-59	84 (37.3%)	39 (37.9%)		
≥ 60	8 (3.6%)	13 (12.6%)		
BMI (kg/m^2)				
Underweight	10 (4.4%)	2 (1.9%)	χ^2 = 6.165	0.104
Normal	102 (45.3%)	37 (35.9%)		
Overweight	84 (37.3%)	42 (40.8%)		
Obese	29 (12.9%)	22 (21.4%)		
History of hypertension				
Yes	17 (7.6%)	7 (6.8%)	χ^2 = 0.060	0.806
No	208 (92.4%)	96 (93.2%)		
History of diabetes mellitus				
Yes	8 (3.6%)	4 (3.9%)	χ^2 = 0.022	0.883
No	217 (96.4%)	99 (96.1%)		
Alcohol consumption history				
Yes	38 (16.9%)	26 (25.2%)	χ^2 = 3.140	0.076
No	187 (83.1%)	77 (74.8%)		
Smoking history				
Yes	44 (19.6%)	27 (26.2%)	χ^2 = 1.847	0.174
No	181 (80.4%)	76 (73.8%)		
PT (s)	11.90 (11.20, 12.78)	12.50 (11.78, 13.39)	Z = -3.920	< 0.001
PTA (%)	96.26 ± 15.19	85.76 ± 16.83	*t* = 5.589	< 0.001
ALB (g/L)	40.47 ± 4.53	38.58 ± 4.78	*t* = 3.433	0.001
LDL-C (mmol/L)	2.58 ± 0.70	2.35 ± 0.62	*t* = 2.855	0.005
HDL-C (mmol/L)	1.09 (0.97, 1.25)	1.04 (0.91, 1.22)	Z = -1.675	0.094
TP (g/L)	68.08 ± 6.04	70.27 ± 6.55	*t* = -2.966	0.003
ChE (U/L)	9.17 ± 2.96	7.91 ± 2.71	*t* = 3.678	< 0.001
CRP (mg/L)	3.23 (1.99, 4.91)	4.10 (2.35, 15.64)	Z = -2.726	0.006
Crea (μmol/L)	65.17 (54.12, 75.00)	64.35 (55.30, 75.58)	*Z* = -0.305	0.760
PA (mg/dL)	22.24 (19.65, 24.35)	19.20 (16.40, 22.10)	*Z* = -5.501	< 0.001
Cys-C (mg/L)	0.85 (0.76, 0.99)	0.94 (0.83, 1.03)	*Z* = -2.972	0.003
Fbg (g/mL)	2.35 (2.01, 2.74)	2.12 (1.83, 2.49)	*Z* = -2.985	0.003
GLO (g/L)	22.80 (18.79, 27.70)	24.34 (19.95, 29.68)	*Z* = -2.360	0.018
PLT (×10^9/L)	217.29 ± 60.16	161.10 ± 56.81	*t* = 7.954	< 0.001
ALT (U/L)	29.00 (20.00, 51.24)	50.00 (32.75, 98.05)	Z = -6.264	< 0.001
AST (U/L)	23.00 (18.00, 33.05)	38.00 (28.00, 67.00)	Z = -7.733	< 0.001
DBIL (μmol/L)	4.50 (2.93, 6.30)	5.70 (4.14, 8.66)	Z = -4.560	< 0.001
IBIL (μmol/L)	7.3 (5.22, 11.00)	8.6 (6.28, 12.80)	Z = -2.365	0.018
TBIL (μmol/L)	12.10 (8.91, 16.94)	15.15 (10.53, 21.08)	Z = -3.530	< 0.001
TC (mmol/L)	4.19 (3.60, 4.67)	4.06 (3.55, 4.49)	Z = -1.987	0.047
TG (mmol/L)	1.10 (0.82, 1.54)	1.03 (0.84, 1.38)	Z = -0.789	0.430

ALP (U/L)	67.21 (55.14, 77.95)	76.50 (60.36, 93.00)	Z = -3.700	< 0.001
TBA (μmol/L)	6.37 (4.06, 9.91)	11.95 (6.83, 19.84)	Z = -6.390	< 0.001
AFP (μg/L)	2.75 (1.85, 4.13)	4.90 (3.31, 10.85)	Z = -7.496	< 0.001
GGT (U/L)	21.00 (15.00, 34.56)	42.00 (24.38, 80.25)	Z = -6.907	< 0.001

BMI: Body Mass Index; PT: Prothrombin Time; PTA: Prothrombin Activity; ALB: Albumin; LDL-C: Low-Density Lipoprotein Cholesterol; HDL-C: High-Density Lipoprotein Cholesterol; TP: Total Protein; ChE: Cholinesterase; CRP: C-Reactive Protein; Crea: Creatinine; PA: Prealbumin; Cys-C: Cystatin C; Fbg: Fibrinogen; GLO: Globulin; PLT: Platelet Count; ALT: Alanine Aminotransferase; AST: Aspartate Aminotransferase; DBIL: Direct Bilirubin; IBIL: Indirect Bilirubin; TBIL: Total Bilirubin; TC: Total Cholesterol; TG: Triglyceride; ALP: Alkaline Phosphatase; TBA: Total Bile Acids; AFP: Alpha-Fetoprotein; GGT: Gamma-Glutamyl Transferase.

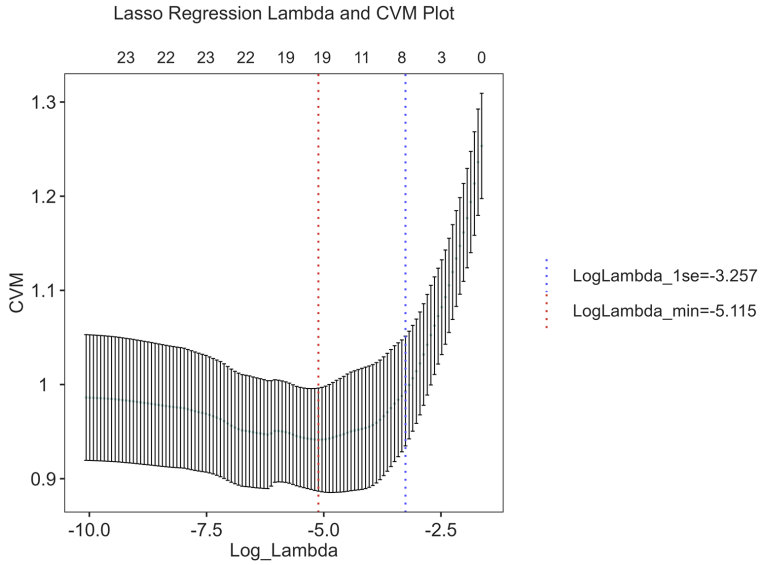

Figure 1. Mean square error path using 10-fold cross validation in the least absolute shrinkage and selection operator (LASSO) regression model. The x-axis represents the log-transformed λ values (Log_Lambda). The y-axis represents the cross-validation error (CVM), with error bars indicating standard deviations.

consumption, smoking history, HDL-C, Crea, and TG levels (P > 0.05). However, significant differences were detected in age, PT, PTA, ALB, LDL-C, TP, ChE, CRP, PA, Cys-C, Fbg, GLO, PLT, ALT, AST, DBIL, IBIL, TBIL, TC, ALP, TBA, AFP, and GGT between the two groups (P < 0.05). Detailed data are presented in **Table 1**.

Identification of key predictors for significant liver fibrosis in CHB using LASSO regression

LASSO regression was employed to identify the most relevant predictors for liver fibrosis diagnosis. Using 10-fold cross-validation, the λ value was selected based on the one-standard-error (1-SE) rule. This rule assumes that the cross-validation error does not exceed the minimum error by one standard error, and the λ value that minimizes model parameters was selected as the optimal point. The LogLambda_1se value identified in this study was -3.257, which was used to finalize the selection of key clinical characteristic variables.

The LASSO regression analysis identified seven key clinical feature variables, including PA, PTA, PLT, ALT, GGT, ALB, and TP (**Figures 1**, **2**). Subsequently, the seven key feature variables were entered into a multivariate logistic regression analysis to assess their independent associations with significant liver fibrosis in CHB patients. The findings indicated that PTA (OR = 0.977, 95% CI: 0.956-0.998), PLT (OR = 0.985, 95% CI: 0.980-0.990), ALB (OR = 0.848, 95% CI: 0.754-0.953), and TP (OR = 1.181, 95% CI: 1.100-1.267) were independent predictors for significant liver fibrosis in CHB patients (P < 0.05; **Table 2**), suggesting their potential clinical utility in fibrosis assessment.

Construction and evaluation of ML-based diagnostic models for significant liver fibrosis in CHB patients

To develop non-invasive diagnostic models for significant liver fibrosis, the dataset was randomly split into a training set (70%) and a test set (30%). Key clinical indicators, including PTA, PLT, ALB, and TP were selected as model input features. Using the training set, four machine learning models (RF, SVM, XGBoost, and LR) were constructed and validated with a 10-fold cross-validation method. The ROC

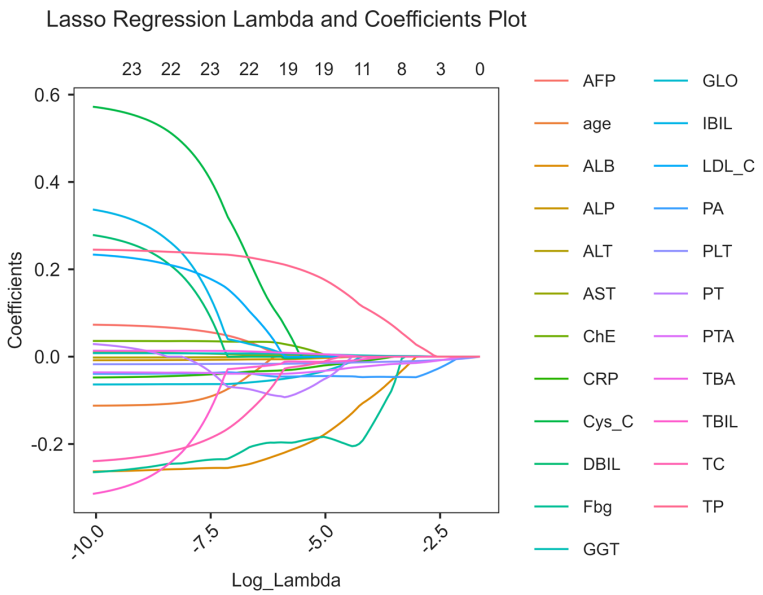

Figure 2. Cross-validation for tuning parameter selection in the LASSO regression model. The x-axis represents the log-transformed λ values (Log_Lambda), where λ is the regularization parameter. The y-axis represents the CVM, with error bars indicating standard deviations.

curve analysis revealed that the RF model exhibited the highest predictive ability, with an AUC of 0.874 (95% CI: 0.813-0.934), outperforming XGBoost (AUC = 0.865, 95% CI: 0.802-0.928), LR (AUC = 0.861, 95% CI: 0.797-0.926), and SVM (AUC = 0.852, 95% CI: 0.785-0.919) (**Figure 3A**). The RF model demonstrated superior performance in both precision and recall, highlighting its strong ability to accurately identify positive cases (**Figure 3B**). To evaluate the clinical applicability of these models, DCA curves were plotted. The results indicated that the RF model yielded the highest net benefit across different probability thresholds (**Figure 3C**). In addition, the RF model showed excellent calibration, with a lower Brier score than the other three models, and the predicted probability was highly consistent with the actual results of significant liver fibrosis (**Table 3**; **Figure 3D**).

Performance evaluation of ML models on the test set

The performance of the four ML models in predicting significant liver fibrosis in CHB patients was then evaluated using the test set. Among the models, RF achieved the highest predictive performance, with an AUC of 0.863 (95% CI: 0.772-0.955), followed by XGBoost (AUC = 0.846, 95% CI: 0.750-0.942), LR (AUC = 0.843, 95% CI: 0.748-0.939), and SVM (AUC = 0.841, 95% CI: 0.746-0.937). Further validation through ROC curves, PR curves, DCA analysis, and calibration plots confirmed the superior predictive performance of RF, particularly in handling imbalanced datasets and demonstrating higher clinical applicability (**Figure 4A-D**). The PR curve analysis indicated that RF had the highest precision-recall AUC (0.768), while DCA showed that RF provided the greatest net benefit across various risk thresholds. Additionally, calibration curve analysis demonstrated that RF exhibited the best agreement between predicted and observed probabilities (**Table 4**), underscoring its potential as a non-invasive and effective diagnostic model for CHB-related significant fibrosis.

Feature importance and contribution to significant liver fibrosis prediction in CHB patients

To further examine the predictive power of the RF model, we compared its performance with traditional fibrosis assessment models, including APRI and FIB-4. FIB-4 contains 4 indexes, ALT, AST, PLT and age, and its calculation formula is $FIB\text{-}4 = \frac{Age \times AST}{PLT \times \sqrt{ALT}}$. FIB-4 score < 1.45 usually indicates no significant liver fibrosis or fibrosis grade ≤ 2, while a score > 3.25 indicates severe fibrosis (grade 3 or 4).

Our results demonstrated that the RF model exhibited superior diagnostic accuracy (AUC = 0.863, 95% CI: 0.772-0.955), significantly surpassing APRI (AUC = 0.585, 95% CI: 0.494-0.677) and FIB-4 (AUC = 0.633, 95% CI: 0.509-0.756) (**Table 5**). Meanwhile, the results of the Delong test showed that the ROC curve performance of the RF model was significantly superior to that of both the APRI and FIB-4 models (**Table 6**). These findings suggest that within the scope of this study, the RF model may be a potentially more accurate and efficient tool for identifying significant liver fibrosis in CHB patients compared to traditional scoring systems.

Table 2. Seven key feature variables were included in the multifactor logistic regression analysis

Variables	*b*	*SE (b)*	*Wald*	*P*	*OR*	95% *CI*	
						Lower limit	Upper limit
PA	-0.051	0.040	1.616	0.204	0.951	0.879	1.028
PTA	-0.024	0.011	4.775	0.029	0.977	0.956	0.998
PLT	-0.015	0.003	30.964	0.000	0.985	0.980	0.990
ALT	0.002	0.002	1.056	0.304	1.002	0.998	1.006
GGT	0.003	0.003	0.937	0.333	1.003	0.997	1.009
ALB	-0.165	0.059	7.710	0.005	0.848	0.754	0.953
TP	0.166	0.036	21.366	0.000	1.181	1.100	1.267

PA: Prealbumin; PTA: Prothrombin Activity; PLT: Platelet Count; ALT: Alanine Aminotransferase; GGT: Gamma-Glutamyl Transferase; ALB: Albumin; TP: Total Protein.

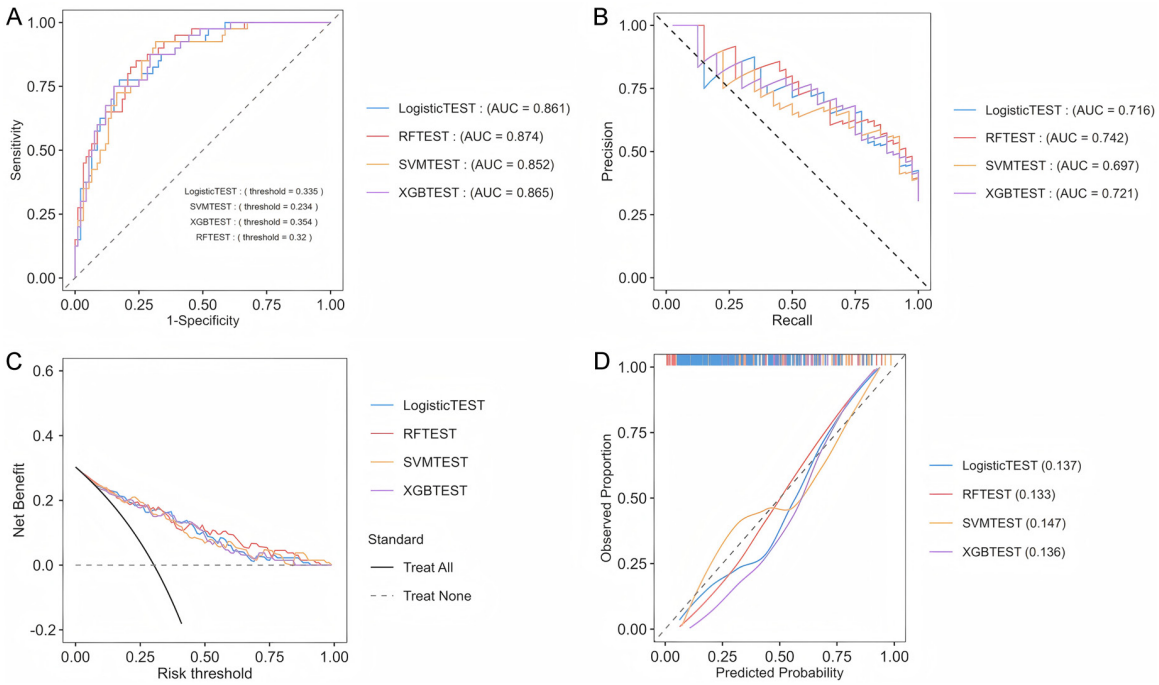

Figure 3. Construction and evaluation of ML-based diagnostic models for significant liver fibrosis in CHB patients using the training set. A. Receiver operating characteristic (ROC) curves for four ML models, including logistic regression (LogisticTEST), random forest (RFTEST), support vector machine (SVMTEST), and extreme gradient boosting (XGBTEST). The x-axis represents 1-specificity (false positive rate), and the y-axis represents sensitivity (true positive rate). The dashed diagonal line represents the random classification baseline (AUC = 0.5). B. The precision-recall (PR) curve provides a comparative analysis of the predictive performance of four ML models. The dashed diagonal line represents the baseline performance of a random classifier. C. Decision curve analysis (DCA) curve evaluates the clinical net benefit of four ML models. The x-axis represents the risk threshold, which determines the probability at which a patient would be classified as high-risk and receive intervention. The y-axis represents the net benefit, calculated by considering both the true positives and false positives. D. The agreement between predicted probabilities and observed outcomes for four ML models was analyzed through the calibration plot. The x-axis represents the predicted probability of the event occurring. The y-axis represents the observed proportion of actual outcomes. The dashed diagonal line represents the ideal calibration line, where predicted probabilities perfectly match observed proportions.

SHAP analysis was employed to determine the contribution of individual features to RF model's predictions. The SHAP feature importance plot indicated that different features had varying impacts on the model's prediction, with PLT showing the highest absolute SHAP value. In

Table 3. Diagnostic efficacy of four machine learning models in predicting significant liver fibrosis in the training set

Models	Accuracy	Precision	F1-Score	AUC	Specificity	Recall	MCC	Brier
RF	0.803	0.733	0.629	0.874	0.913	0.550	0.508	0.133
XGBoost	0.788	0.773	0.548	0.865	0.946	0.425	0.457	0.136
SVM	0.773	0.692	0.545	0.852	0.913	0.450	0.420	0.147
Logistic	0.803	0.769	0.606	0.861	0.935	0.500	0.502	0.137

RF: Random Forest; XGBoost: eXtreme Gradient Boosting; SVM: Support Vector Machine; Logistic: Logistic Regression.

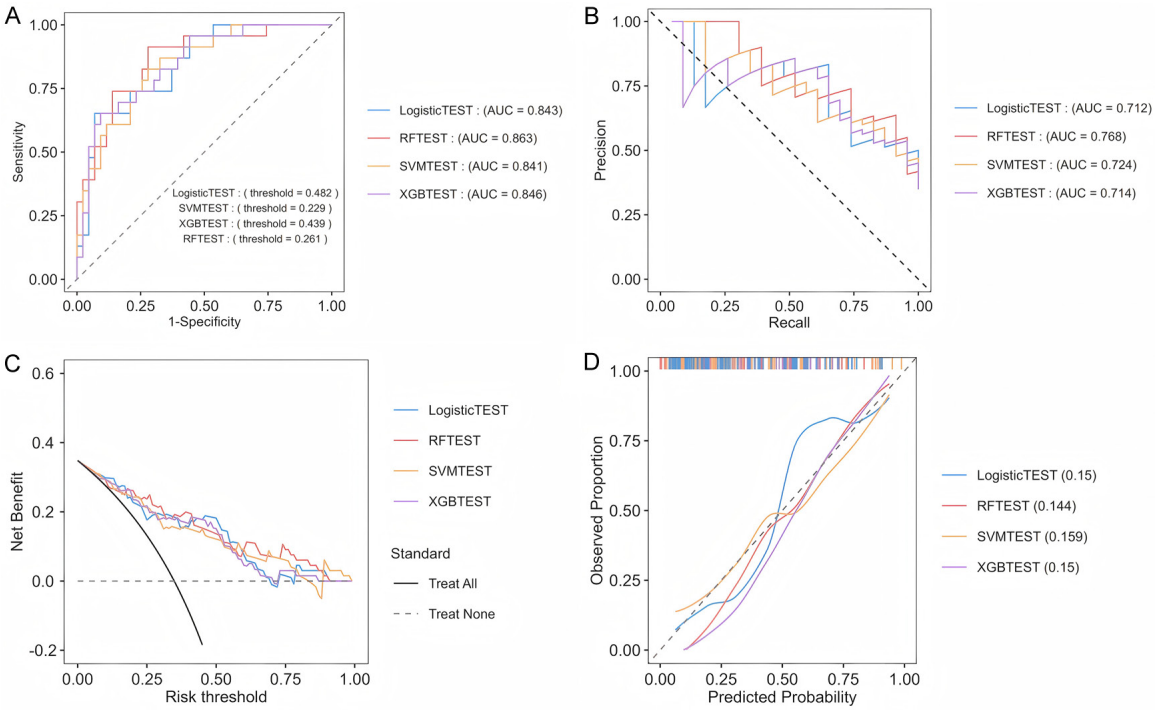

Figure 4. Evaluation of ML-based diagnostic models for significant liver fibrosis in CHB patients using the test set. A. Receiver operating characteristic (ROC) curves of four ML models were compared, with corresponding AUC values and optimal thresholds presented. B. The precision-recall curves for the four ML models and indicated their respective AUC values. C. Decision curve analysis (DCA) curves for four ML models. D. The calibration plot for the four models, comparing the predicted probabilities with the actual observed proportions.

Table 4. Diagnostic efficacy of four machine learning models in predicting significant liver fibrosis in the test set

Models	Accuracy	Precision	F1-Score	AUC	Specificity	Recall	MCC	Brier
RF	0.788	0.714	0.682	0.863	0.860	0.652	0.524	0.144
XGBoost	0.803	0.813	0.667	0.846	0.930	0.565	0.551	0.150
SVM	0.788	0.765	0.650	0.841	0.907	0.565	0.515	0.159
Logistic	0.833	0.833	0.732	0.843	0.930	0.652	0.623	0.150

RF: Random Forest; XGBoost: eXtreme Gradient Boosting; SVM: Support Vector Machine; Logistic: Logistic Regression.

contrast, ALB had the lowest impact on the model's predictions (**Figure 5A**). Individual SHAP values indicated that higher PLT and TP levels positively contributed to fibrosis prediction, whereas elevated PTA and ALB levels were associated with lower fibrosis risk (**Figure 5B**).

Table 5. Diagnostic effect of RF model and traditional models

Models	AUC	*SE (b)*	*P*	95% *CI*	
				Lower limit	Upper limit
RF	0.863	0.047	< 0.001	0.772	0.955
APRI	0.585	0.047	0.068	0.494	0.677
FIB-4	0.633	0.063	0.035	0.509	0.756

RF: Random Forest; APRI: Aspartate Aminotransferase to Platelet Ratio Index; FIB-4: Fibrosis-4 Index.

Table 6. Pairwise comparison of ROC curves (Delong test)

Models	Z	*SE (b)*	*P*	95% *CI*	
				Lower limit	Upper limit
RF-APRI	3.868	0.303	< 0.001	0.173	0.383
RF-FIB-4	5.197	0.327	< 0.001	0.114	0.347
FIB-4-APRI	0.816	0.327	0.414	-0.067	0.162

RF: Random Forest; APRI: Aspartate Aminotransferase to Platelet Ratio Index; FIB-4: Fibrosis-4 Index.

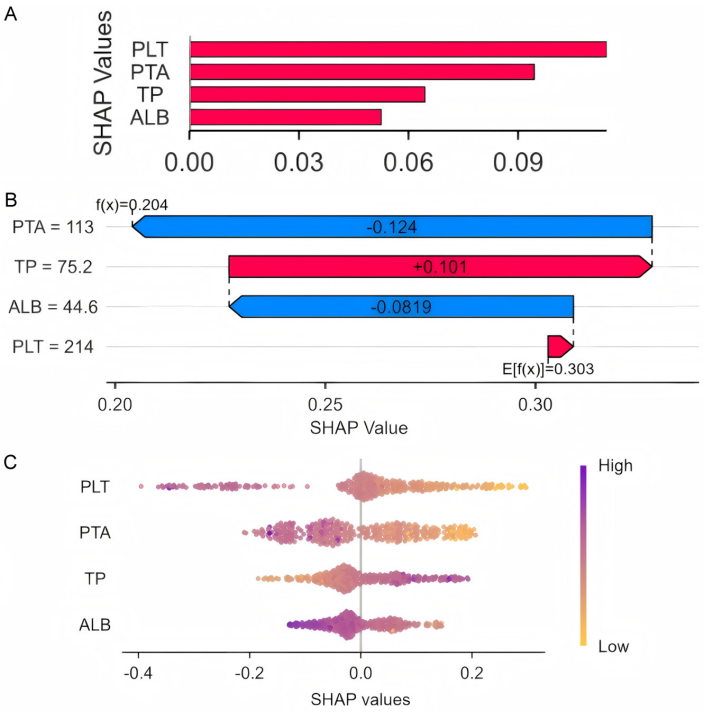

Figure 5. SHapley Additive exPlanations (SHAP) value analysis for feature importance and impact on model predictions. A. SHAP values for feature importance in the predictive model. The y-axis lists the features: PLT (platelet count), PTA (prothrombin activity), TP (total protein), and ALB (albumin). The x-axis represents the SHAP values, indicating the contribution of each feature to the model's predictions. A higher SHAP value suggests a greater impact of the corresponding feature on the model's output. B. SHAP waterfall plot illustrating the impact of features (PLT, PTA, TP, and ALB) on model predictions. The x-axis represents the SHAP value, and the y-axis lists the features along with their corresponding values. Blue bars indicate a negative contribution to the prediction, while red bars indicate a positive contribution. C. SHAP beeswarm plot visualizing the distribution of SHAP values for four features (PLT, PTA, TP, and ALB) and their respective impact on the model's predictions. The color scale indicates the feature values, with "High" in purple and "Low" in orange.

The SHAP beeswarm plot revealed the distribution of feature importance across samples (**Figure 5C**).

Discussion

HBV infection triggers a complex and dynamic cascade of pathological alterations across different disease stages, involving immune activation, hepatocellular injury, and progressive fibrogenesis [16]. Without timely intervention, this pathological progression can cause cirrhosis and hepatocellular carcinoma, substantially increasing morbidity and mortality among patients with CHB [17]. Considering the central role of fibrosis in disease progression, early diagnosis of significant fibrosis (≥ S2) is vital for optimizing clinical management, especially in determining the optimal timing for initiating antiviral therapy [18]. In this research, we identified key predictive biomarkers using a retrospective cohort of 328 CHB patients through LASSO regression and constructed four ML models, including RF, SVM, XGBoost, and LR. Among these models, the RF algorithm demonstrated the highest predictive performance, outperforming traditional fibrosis indices, including APRI and FIB-4.

By LASSO and multivariate logistic regression analyses, we identified PTA, PLT, ALB, and TP as independent predictors. The molecular and cellular mechanisms underlying their association with liver fibrosis warrant further investigation.

There is a complex pathophysiological association between prothrombin activity (PTA) and hepatic fibrosis, with underlying mechanisms potentially involving impaired hepatic synthetic function, activation of inflammatory cascades, and altered intrahepatic hemodynamics [19, 20]. Physiologically, PTA reflects the ability of the liver to synthesize coagulation factors [21]. The liver, as the primary site for coagulation factor synthesis, undergoes structural damage as fibrosis progresses. Hepatic stellate cell activation and excessive extracellular matrix deposition disrupt the liver parenchyma, reducing the number of hepatocytes and impairing coagulation factor production, leading to a decrease in PTA [22]. A rat model of hepatic fibrosis showed that, with disease progression, the expression of relevant genes involved in coagulation factor synthesis was significantly down-regulated in the liver, resulting in decreased coagulation factor production and a gradual reduction in PTA levels [23]. Empirical studies have shown that coagulation indices effectively reflect the level of liver injury and have clinical significance in the diagnosis of liver disease progression [24].

The pathophysiological mechanism underlying the association of PLT with liver fibrosis may involve a combination of portal hypertension and hypersplenism. As liver fibrosis progresses, intrahepatic vascular structure is damaged and the resistance of blood vessels increases, which in turn leads to increased portal pressure [25]. When portal vein pressure becomes excessively high, it causes splenomegaly and hypersplenism. In this condition, a large number of platelets are retained and destroyed in the spleen, resulting in a reduction in circulating PLT levels. Meanwhile, the altered hematopoietic microenvironment in the liver during hepatic fibrosis further impairs PLT production, contributing to decreased PLT levels [26].

The pathophysiological mechanisms underlying the association of ALB with liver fibrosis mainly stem from impaired hepatocyte synthesis and metabolic disturbances driven by the inflammatory microenvironment. The liver is the sole site of ALB synthesis, and this process of ALB synthesis depends on the normal metabolism and function of hepatocytes. As liver fibrosis progresses, the number of hepatic parenchymal cells decreases and their function is impaired, directly impairing ALB synthesis. A study in patients with alcoholic cirrhosis found that plasma ALB levels gradually decreased with the advancement of liver fibrosis, with ALB levels in cirrhotic patients being significantly lower than in healthy controls [27]. The association between ALB and liver fibrosis may also be influenced by nutrient metabolism, as patients with liver fibrosis often experience symptoms such as loss of appetite and digestive and absorption dysfunction, which lead to insufficient protein intake and consequently hinder synthesis of ALB. Furthermore, during fibrosis, the liver experiences dual metabolic abnormalities: hepatocyte damage reduces ALB synthesis, while the chronic inflammatory state may activate the proteolytic pathway, accelerating the degradation of ALB [28].

Total Protein (TP) is the sum of albumin and globulin in the serum, and changes in both can affect TP levels during the course of liver fibrosis. On one hand, as mentioned earlier, liver fibrosis results in hepatocyte damage and decreased albumin synthesis, which lowers TP levels. On the other hand, the body's immune system is activated in response to ongoing inflammatory damage and viral infections, causing an increase in the production of immune cells such as lymphocytes, which in turn promotes a significant increase in globulin synthesis [29]. This compensatory increase in globulin partially offsets the reduction in albumin, and may even lead to elevated TP. In some cases, the rise in globulin can surpass the decline in albumin, ultimately leading to elevated TP levels, which have been associated with an increased risk of liver fibrosis [30].

In this study, the independent predictive value of PTA, PLT, ALB and TP in the CHB population was verified, and an RF model with superior predictive performance was constructed. In this study, by incorporating key indicators reflecting liver synthetic function, coagulation status, and portal hypertension, such as PTA, PLT, ALB and TP, we overcame the limitations of traditional models like APRI and FIB-4, which rely primarily on hepatic enzymes. Additionally,

we utilized the nonlinear modeling and self-help sampling characteristics of the RF model to efficiently capture the complex interactions among the indicators, such as the predictive value of PLT in combination with ALB for identifying advanced fibrosis [31, 32]. Furthermore, SHAP value analysis provided an objective quantification of the contribution of each indicator, ensuring stability and improving diagnostic performance through internal validation. Meanwhile, compared with some existing ML models for predicting liver fibrosis, the RF model constructed in this study achieved a higher predictive performance (AUC = 0.863). In contrast, the model constructed by Rui et al. [33] had a lower AUC of 0.778 in predicting advanced fibrosis, which may be influenced by factors such as study population, sample size, and other variables.

There are limitations to this study, including potential selection bias due to the retrospective single-center design, inclusion of metrics that do not cover emerging fibrosis markers or imaging parameters, and lack of external validation of key machine learning models. Although the RF model demonstrated strong performance in internal training and testing cohorts, its efficacy in CHB patients with different geographic, ethnic, and etiologic backgrounds has not been validated. Variations in laboratory testing protocols and disease heterogeneity may further affect the model's stability, potentially leading to an overestimation of its generalizability. Future studies should prioritize multicenter external validation, incorporate diverse patient cohorts to assess the real-world performance of the model, and integrate multidimensional indicators, such as molecular markers and imaging data, to build a more comprehensive prediction system. Mechanistic studies and the creation of interpretive tools will be crucial for enhancing the clinical reliability and utility of this model, ultimately translating it into a precision diagnostic tool.

Conclusion

The significant liver fibrosis diagnostic model constructed in this study, based on the RF algorithm, demonstrates excellent diagnostic performance and outperforms the traditional APRI and FIB-4 models. Our study demonstrates that ML-based models, particularly RF, provide a highly accurate, non-invasive, and interpretable approach for liver fibrosis assessment in CHB patients. The integration of ML methods with traditional clinical markers represents a significant advancement in fibrosis prediction, offering the potential for earlier diagnosis, improved risk stratification, and personalized treatment strategies for CHB patients.

Acknowledgements

This study was supported by Natural Science Foundation of Xinjiang Uygur Autonomous Region (2024D01C84); and Xinjiang Uygur Autonomous Region Talent Development Fund (TSYC202301A063).

Disclosure of conflict of interest

None.

Address correspondence to: Yuqing Yang, People's Hospital of Xinjiang Uygur Autonomous Region, No. 91, Tianchi Road, Tianshan District, Urumqi 830001, Xinjiang, China. Tel: +86-15899190081; E-mail: 15684582113@163.com

References

[1] Koffas A, Kumar M, Gill US, Jindal A, Kennedy PTF and Sarin SK. Chronic hepatitis B: the demise of the 'inactive carrier' phase. Hepatol Int 2021; 15: 290-300.

[2] Gragnani L, Monti M, De Giorgi I and Zignego AL. The key importance of screening underprivileged people in order to achieve global hepatitis virus elimination targets. Viruses 2025; 17: 265.

[3] Lin MH, Li HQ, Zhu L, Su HY, Peng LS, Wang CY, He CP, Liang XE and Wang Y. Liver fibrosis in the natural course of chronic hepatitis B viral infection: a systematic review with meta-analysis. Dig Dis Sci 2022; 67: 2608-2626.

[4] Lin CL and Kao JH. Development of hepatocellular carcinoma in treated and untreated patients with chronic hepatitis B virus infection. Clin Mol Hepatol 2023; 29: 605-622.

[5] Wang Y, Yang X and Wang S. Progress and prospects of elastography techniques in the evaluation of fibrosis in chronic liver disease. Arch Med Sci 2024; 20: 1784-1792.

[6] Wang L, Li J, Yang K, Zhang H, Wang Q, Lv X and Guan S. Comparison and evaluation of non-invasive models in predicting liver inflammation and fibrosis of chronic hepatitis B virus-infected patients with high hepatitis B virus DNA and normal or mildly elevated alanine

transaminase levels. Medicine (Baltimore) 2020; 99: e20548.
[7] Najafi N, Razavi A, Jafarpour H, Raei M, Azizi Z, Davoodi L, Abdollahi A and Frouzanian M. Evaluation of hepatic injury in chronic hepatitis B and C using APRI and FIB-4 indices compared to fibroscan results. Ann Med Surg (Lond) 2024; 86: 3841-3846.
[8] Khalifa A and Rockey DC. The utility of liver biopsy in 2020. Curr Opin Gastroenterol 2020; 36: 184-191.
[9] Yang L, Zhang Y, Hong X, Zhang K, Liu B, Zhang P, Tang Q, Yu J, Jin XZ, Jin XZ, Zhang N, Targher G, Byrne CD, Zhang Z, Zheng MH and Zhang J. Serum dithiothreitol-oxidizing capacity (DOC) is a promising biomarker for excluding significant liver fibrosis: a proof-of-concept study. BMC Med 2024; 22: 278.
[10] Song J, Gao Z, Lai L, Zhang J, Liu B, Sang Y, Chen S, Qi J, Zhang Y, Kai H and Ye W. Machine learning-based plasma metabolomics for improved cirrhosis risk stratification. BMC Gastroenterol 2025; 25: 61.
[11] Zhu G, Yang N, Yi Q, Xu R, Zheng L, Zhu Y, Li J, Che J, Chen C, Lu Z, Huang L, Xiang Y and Zheng T. Explainable machine learning model for predicting the risk of significant liver fibrosis in patients with diabetic retinopathy. BMC Med Inform Decis Mak 2024; 24: 332.
[12] Xu Y, Zhang B, Zhou F, Yi YP, Yang XL, Ouyang X and Hu H. Development of machine learning-based personalized predictive models for risk evaluation of hepatocellular carcinoma in hepatitis B virus-related cirrhosis patients with low levels of serum alpha-fetoprotein. Ann Hepatol 2024; 29: 101540.
[13] Zhang C, Shu Z, Chen S, Peng J, Zhao Y, Dai X, Li J, Zou X, Hu J and Huang H. A machine learning-based model analysis for serum markers of liver fibrosis in chronic hepatitis B patients. Sci Rep 2024; 14: 12081.
[14] Durot I, Akhbardeh A, Sagreiya H, Loening AM and Rubin DL. A new multimodel machine learning framework to improve hepatic fibrosis grading using ultrasound elastography systems from different vendors. Ultrasound Med Biol 2020; 46: 26-33.
[15] Li Z, Sun X, Zhao Z, Yang Q, Ren Y, Teng X, Tai DCS, Wanless IR, Schattenberg JM and Liu C. A machine learning based algorithm accurately stages liver disease by quantification of arteries. Sci Rep 2025; 15: 3143.
[16] Chang ML and Liaw YF. Hepatitis B flare in hepatitis B e antigen-negative patients: a complicated cascade of innate and adaptive immune responses. Int J Mol Sci 2022; 23: 1552.
[17] Tacke F and Weiskirchen R. An update on the recent advances in antifibrotic therapy. Expert Rev Gastroenterol Hepatol 2018; 12: 1143-1152.
[18] Jiang XY, Huang B, Huang DP, Wei CS, Zhong WC, Peng DT, Huang FR and Tong GD. Long-term follow-up of cumulative incidence of hepatocellular carcinoma in hepatitis B virus patients without antiviral therapy. World J Gastroenterol 2021; 27: 1101-1116.
[19] Croquet V, Vuillemin E, Ternisien C, Pilette C, Oberti F, Gallois Y, Trossaert M, Rousselet MC, Chappard D and Calès P. Prothrombin index is an indirect marker of severe liver fibrosis. Eur J Gastroenterol Hepatol 2002; 14: 1133-1141.
[20] Mussbacher M, Brunnthaler L, Panhuber A, Starlinger P and Assinger A. Till death do us part-the multifaceted role of platelets in liver diseases. Int J Mol Sci 2021; 22: 3113.
[21] Li Y, Wei S, Wang J, Hong L, Cui L and Wang C. Analysis of the factors associated with abnormal coagulation and prognosis in patients with non-small cell lung cancer. Zhongguo Fei Ai Za Zhi 2014; 17: 789-796.
[22] Akkız H, Gieseler RK and Canbay A. Liver fibrosis: from basic science towards clinical progress, focusing on the central role of hepatic stellate cells. Int J Mol Sci 2024; 25: 7873.
[23] Hou W, Hao Y, Yang W, Tian T, Fang P, Du Y, Gao L, Gao Y and Zhang Q. The Jieduan-Niwan (JDNW) formula ameliorates hepatocyte apoptosis: a study of the inhibition of E2F1-mediated apoptosis signaling pathways in Acute-on-Chronic Liver Failure (ACLF) using rats. Drug Des Devel Ther 2021; 15: 3845-3862.
[24] Peng J, He G, Chen H and Kuang X. Study on correlation between coagulation indexes and disease progression in patients with cirrhosis. Am J Transl Res 2021; 13: 4614-4623.
[25] Torres Rojas AM and Lorente S. Liver fibrosis emulation: Impact of the vascular fibrotic alterations on hemodynamics. Comput Biol Med 2023; 166: 107563.
[26] Meng Y, Zhao T, Zhang Z and Zhang D. The role of hepatic microenvironment in hepatic fibrosis development. Ann Med 2022; 54: 2830-2844.
[27] Xu YF, Hao YX, Ma L, Zhang MH, Niu XX, Li Y, Zhang YY, Liu TT, Han M, Yuan XX, Wan G and Xing HC. Difference and clinical value of metabolites in plasma and feces of patients with alcohol-related liver cirrhosis. World J Gastroenterol 2023; 29: 3534-3547.
[28] Bunchorntavakul C, Supanun R and Atsawarungruangkit A. Nutritional status and its impact on clinical outcomes for patients admitted to hospital with cirrhosis. J Med Assoc Thai 2016; 99 Suppl 2: S47-55.
[29] Wang J, Xu Y, Chen Z, Liang J, Lin Z, Liang H, Xu Y, Wu Q, Guo X, Nie J, Lu B, Huang B, Xian H, Wang X, Wu Q, Zeng J, Chai C, Zhang M, Lin Y, Zhang L, Zhao S, Tong Y, Zeng L, Gu X, Chen ZG, Yi S, Zhang T, Delfouneso D, Zhang Y, Nutt

SL, Lew AM, Lu L, Bai F, Xia H, Wen Z and Zhang Y. Liver immune profiling reveals pathogenesis and therapeutics for biliary atresia. Cell 2020; 183: 1867-1883, e1826.

[30] Liu X, Tan S, Liu H, Jiang J, Wang X, Li L and Wu B. Hepatocyte-derived MASP1-enriched small extracellular vesicles activate HSCs to promote liver fibrosis. Hepatology 2023; 77: 1181-1197.

[31] Lee YW, Choi JW and Shin EH. Machine learning model for predicting malaria using clinical information. Comput Biol Med 2021; 129: 104151.

[32] Hu J and Szymczak S. A review on longitudinal data analysis with random forest. Brief Bioinform 2023; 24: bbad002.

[33] Rui F, Xu L, Yeo YH, Xu Y, Ni W, Tan Y, Zheng Q, Tian X, Zeng QL, He Z, Qiu Y, Zhu C, Ding W, Wang J, Huang R, Xue Q, Wang X, Chen Y, Fan J, Fan Z, Ogawa E, Kwak MS, Qi X, Shi J, Wong VW, Wu C and Li J. Machine learning-based models for advanced fibrosis and cirrhosis diagnosis in chronic hepatitis B patients with hepatic steatosis. Clin Gastroenterol Hepatol 2024; 22: 2250-2260, e2212.