



# Prediction model for hepatocellular carcinoma recurrence after hepatectomy: Machine learning-based development and interpretation study

Rongqiang Liu<sup>a,b,1</sup>, Shinan Wu<sup>a,c,1</sup>, Hao yuan Yu<sup>a,1</sup>, Kaining Zeng<sup>a,1</sup>, Zhixing Liang<sup>a</sup>, Siqi Li<sup>a</sup>, Yongwei Hu<sup>a</sup>, Yang Yang<sup>a,\*\*</sup>, Linsen Ye<sup>a,\*</sup>

<sup>a</sup> Department of Hepatic Surgery and Liver Transplantation Center, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

<sup>b</sup> Department of Hepatobiliary Surgery, Renmin Hospital of Wuhan University, Wuhan, Hubei Province, China

<sup>c</sup> Institute of Xiamen University, School of Medicine, Xiamen University, Xiamen, Fujian, China

## ARTICLE INFO

### Keywords:

Hepatocellular carcinoma  
Machine learning  
Multilayer perceptron  
Shapley additive explanations  
Risk factor

## ABSTRACT

**Background:** Identifying patients with hepatocellular carcinoma (HCC) at high risk of recurrence after hepatectomy can help to implement timely interventional treatment. This study aimed to develop a machine learning (ML) model to predict the recurrence risk of HCC patients after hepatectomy.

**Methods:** We retrospectively collected 315 HCC patients who underwent radical hepatectomy at the Third Affiliated Hospital of Sun Yat-sen University from April 2013 to October 2017, and randomly divided them into the training and validation sets at a ratio of 7:3. According to the postoperative recurrence of HCC patients, the patients were divided into recurrence group and non-recurrence group, and univariate and multivariate logistic regression were performed for the two groups. We applied six machine learning algorithms to construct the prediction models and performed internal validation by 10-fold cross-validation. Shapley additive explanations (SHAP) method was applied to interpret the machine learning model. We also built a web calculator based on the best machine learning model to personalize the assessment of the recurrence risk of HCC patients after hepatectomy.

**Results:** A total of 13 variables were included in the machine learning models. The multilayer perceptron (MLP) machine learning model was proved to achieve optimal predictive value in test set (AUC = 0.680). The SHAP method displayed that  $\gamma$ -glutamyl transpeptidase (GGT), fibrinogen, neutrophil, aspartate aminotransferase (AST) and total bilirubin (TB) were the top 5 important factors for recurrence risk of HCC patients after hepatectomy. In addition, we further demonstrated the reliability of the model by analyzing two patients. Finally, we successfully constructed an online web prediction calculator based on the MLP machine learning model.

**Conclusion:** MLP was an optimal machine learning model for predicting the recurrence risk of HCC patients after hepatectomy. This predictive model can help identify HCC patients at high recurrence risk after hepatectomy to provide early and personalized treatment.

\* Corresponding author. Department of Hepatic Surgery and Liver transplantation Center, The Third Affiliated Hospital of Sun Yat-sen University, No 600, TianHe Street, TianHe District, Guangzhou, 510220, Guangdong, China.

\*\* Corresponding author. Department of Hepatic Surgery and Liver transplantation Center, The Third Affiliated Hospital of Sun Yat-sen University, China.

E-mail addresses: [yysysu@163.com](mailto:yysysu@163.com) (Y. Yang), [ye\\_linsen@163.com](mailto:ye_linsen@163.com) (L. Ye).

<sup>1</sup> These authors contributed equally to this work.

<https://doi.org/10.1016/j.heliyon.2023.e22458>

Received 24 May 2023; Received in revised form 10 September 2023; Accepted 13 November 2023

Available online 19 November 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Hepatocellular carcinoma is the most common tumor and the second most common cause of cancer-related death worldwide [1]. The major risk factors for liver cancer include hepatitis B virus, hepatitis C virus, and alcoholic and nonalcoholic liver disease [2]. The annual death rate of liver cancer in China accounts for more than 50 % of the global death rate of liver cancer each year [3]. At present, surgical resection is still the main method for liver cancer patients in clinical practice. Due to the high malignancy and aggressive nature of HCC, patients still have a high recurrence rate even after radical hepatectomy [4]. Research has demonstrated that the majority of patients with hepatocellular carcinoma (HCC) who undergo hepatectomy can experience recurrence within one year, with a median survival time from the point of recurrence to death being 21 months [5]. Effective evaluation of postoperative recurrence of HCC patients can timely help doctors to implement individualized treatment for patients and improve postoperative survival time.

Machine learning (ML) is a mathematical method that generalizes and analyzes data to achieve artificial intelligence. It has been widely used in many different sciences. Research has shown that ML is more accurate than traditional statistical methods, and it can detect interactions among variables [6]. In the medical field, ML is widely used to build models to increase the understanding of the diagnosis or prognosis of diseases [7,8]. Currently, ML has been widely used in a variety of tumors due to its good predictive performance. For example, Liu et al. used ML to successfully develop a network predictor to predict the risk of bone metastases in patients with prostate cancer [9]. Li et al. successfully constructed a model to assess the risk of lung metastases in patients with Ewing's sarcoma [10]. In a novel prognostic prediction model constructed by Hasnain et al. they found that the sensitivity and specificity of predicting tumor recurrence after bladder cancer resection exceeded 70 % [11]. Liang et al. constructed the support vector machine model to predict patient prognosis by analyzing 83 patients with liver cancer after radiofrequency ablation [12]. We recently developed a risk stratification model for the prognosis of acute ischemic stroke using machine learning methods and achieved excellent discrimination, diagnosis, and prediction performance [13]. However, studies about the application of ML to predict the recurrence risk of HCC patients after hepatectomy were still limited [14,15]. In the past, there have been prediction models for post-hepatectomy recurrence in liver cancer using CT-based radiomics deep learning techniques [16]. However, research on predicting its recurrence using various serum markers and tumor-related indicators and providing visual explanations for black-box models like machine learning is still lacking [17]. ML has not been fully explored in predicting the recurrence risk of HCC patients after hepatectomy.

In the study, we identified potential risk factors for the recurrence risk of liver cancer patients after hepatic resection by using preoperative indicators, including general demographic data, serological indicators, and tumor pathological characteristics, and constructed prediction models based on a machine learning algorithm. By comparing different models, we selected the optimal ML model. Finally, a web calculator was developed to achieve convenient and personalized prediction for liver cancer patients.

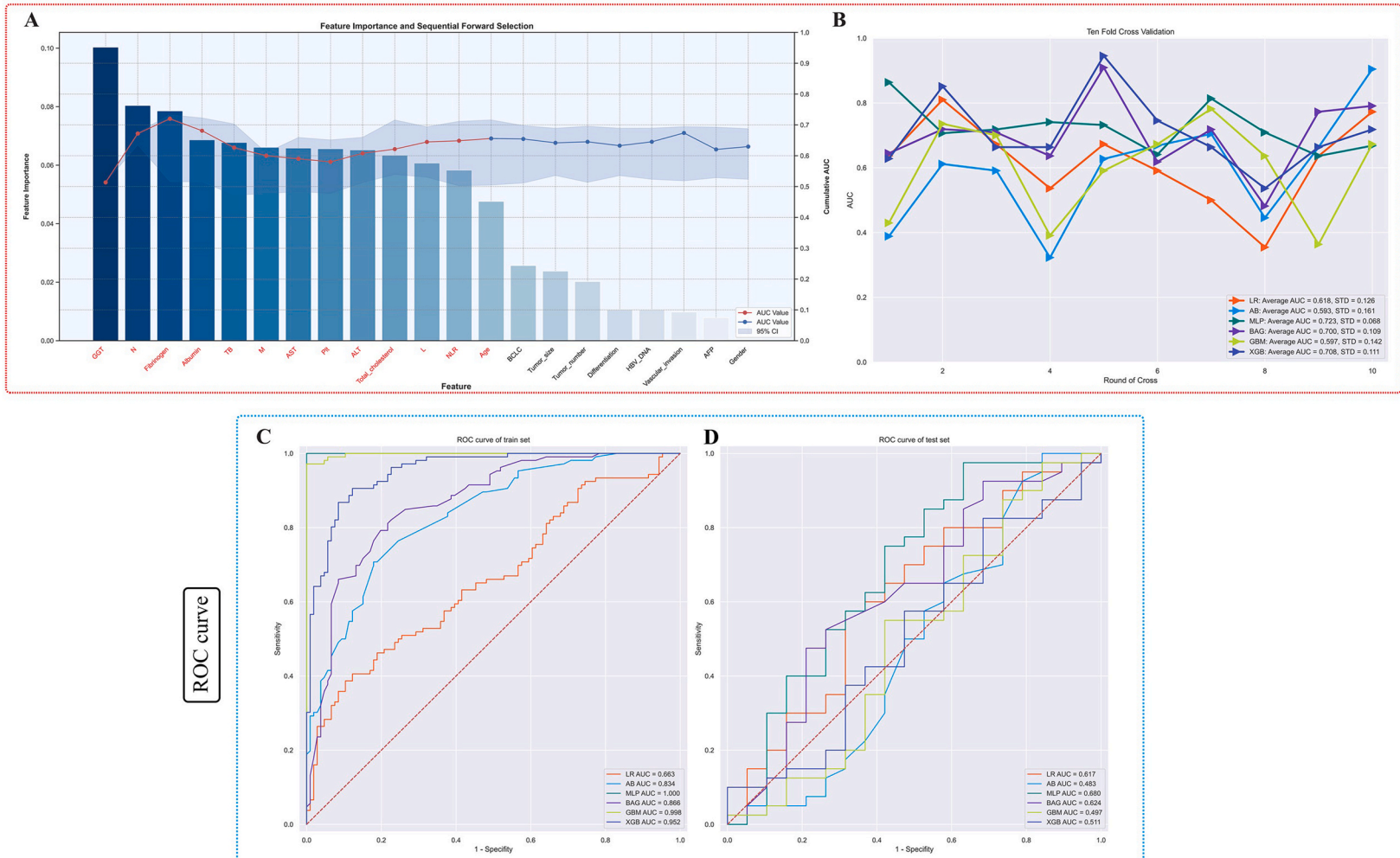
## 2. Methods

### 2.1. Subjects

We recruited a total of 315 patients undergoing radical hepatectomy at the Third Affiliated Hospital of Sun Yat-sen University from April 2013 to October 2017. All patients were randomly divided into a training set and a validation set at a ratio of 7:3. To mitigate the potential impact of imbalanced data on model training, we first applied the Synthetic Minority Over-sampling Technique (SMOTE) method to oversample the training set after splitting [18]. Additionally, we removed the samples that contained missing value.; As a result, the training set included 212 cases, and the testing set included 59 cases. The inclusion criteria were: (1) histopathologically confirmed HCC after hepatectomy, (2) no other tumors or extrahepatic metastases, (3) no preoperative anticancer treatment. (4) no other underlying diseases and chronic diseases. All clinicopathological data were retrieved electronically from the hospital or handwritten medical records, including patient factors, laboratory parameters, and histological features of the tumor. During the follow-up period, the patients were divided into recurrence and non-recurrence groups. All patients signed the informed consent form. The Ethics Committee of the Third Affiliated Hospital of Sun Yat-sen University approved the study.

### 2.2. Selection of machine learning feature variable

This study collected clinical variables at the baseline data collection for liver cancer hepatectomy patients. These included gender, tumor size, number of tumors, presence of vascular invasion by the tumor, BCLC staging (Barcelona Clinic Liver Cancer staging classification), AFP (alpha-fetoprotein), HBV DNA, tumor differentiation degree, neutrophils (N), macrophages (M), fibrinogen, age, platelets (Plt), lymphocytes (L), albumin, the ratio of neutrophils to lymphocytes (NLR), aspartate aminotransferase (AST), total bilirubin (TB),  $\gamma$ -glutamyl transpeptidase (GGT), alanine aminotransferase (ALT), and Total cholesterol. The criteria for selecting feature variables for subsequent machine learning algorithms were: feature importance ranking based on the random forest machine learning algorithm and Sequential Forward Selection [19,20]. On this basis, all the aforementioned feature variables were included for feature importance ranking and screening. Subsequently, these variables were inputted into a hierarchical clustering algorithm to remove feature variables with multicollinearity. Then, we re-ordered the pre-selected variables. Finally, the optimal machine learning feature variables were determined based on when the area under the curve (AUC) of the receiver operating characteristic reached a stable value. We chose the top 13 feature variables for subsequent machine learning algorithm development, as there was no observed significant improvement in the AUC after the 14th iteration (Fig. 1A). The feature variables included in the subsequent machine learning model construction were GGT, N, Fibrinogen, Albumin, TB, M, AST, Plt, ALT, Total Cholesterol, L, NLR, and Age.



**Fig. 1.** Machine learning feature variable extraction and internal ten-fold cross-validation were conducted, resulting in ROC curves for both the training and testing sets.

**Notes:** Fig. 1A displays feature importance ranking and forward sequential feature selection based on random forests, with features selected highlighted in red font. Fig. 1B shows the results of internal ten-fold cross-validation on the training set, with the best machine learning model being MLP, achieving an AUC of  $0.723 \pm 0.068$ . Fig. 1C and D respectively depict the ROC curve results for the training and testing sets under different machine learning algorithms. Notably, the MLP machine learning model achieves an AUC value of 1 in the training set and 0.680 in the testing set.

**Abbreviation:** ML, machine learning; AUC, the area under the curve; AB, adaptive boosting; MLP, multilayer perceptron; BAG, bootstrapped aggregating; LR, logistic regression; GBM, gradient boosting machine; XGB, extreme gradient boost.

### 2.3. Model building

Univariate analysis was performed to screen recurrence risk factors in HCC patients after hepatectomy. In this study, we used six different machine learning algorithms including multilayer perceptron (MLP) [21], adaptive boosting (AB), bootstrapped aggregating (BAG) [22], logistic regression (LR) [23], gradient boosting machine (GBM) [22], and extreme gradient boost (XGB) [24]. A 10-fold cross-validation was performed to analyze the prediction performance of different ML models in the training set. The receiver operating characteristic curve (ROC) of the six ML models was plotted using the validation cohort. The model with the best prediction performance was selected as the final model. The prediction performance of the final model was evaluated using the confusion matrix. Finally, an online risk calculator was constructed based on the final model.

### 2.4. Model interpretation

Shapley additive explanations (SHAP) is a method for interpreting the results of predictive models based on cooperative game theory [25]. This method can quantify the SHAP value of each characteristic variable, representing the contribution of different factors to postoperative recurrence in HCC patients. The SHAP method demonstrates each factor's positive or negative effect on the predicted outcome. When the SHAP value is positive, it indicates that the corresponding feature leads to a higher probability of recurrence risk. When the SHAP value is negative, it indicates that the corresponding feature leads to a lower recurrence risk. We used the SHAP method to establish the importance ranking of recurrence risk factors in HCC patients undergoing hepatectomy and made individual interpretations of the model.

**Table 1**  
Demographic and clinical characteristics.

Variables	Total (n = 315)	No recurrence (n = 137)	Recurrence (n = 178)	P value
Gender, n (%)				0.127
male	288 (91)	121 (88)	167 (94)	
female	27 (9)	16 (12)	11 (6)	
Tumor size n (%)				<0.001
≤3 cm	85 (27)	52 (38)	33 (19)	
> 3 cm	230 (73)	85 (62)	145 (81)	
Tumor number, n (%)				<0.001
1	229 (73)	115 (84)	114 (64)	
> 1	86 (27)	22 (16)	64 (36)	
Vascular invasion, n (%)				0.133
No	198 (63)	93 (68)	105 (59)	
Yes	117 (37)	44 (32)	73 (41)	
BCLC, n (%)				<0.001
0	55 (17)	38 (28)	17 (10)	
1	115 (37)	48 (35)	67 (38)	
2	34 (11)	10 (7)	24 (13)	
3	111 (35)	41 (30)	70 (39)	
AFP, n (%)				0.195
≤400	219 (70)	101 (74)	118 (66)	
> 400	96 (30)	36 (26)	60 (34)	
HBV DNA, n (%)				0.005
≤1000	97 (31)	54 (39)	43 (24)	
> 1000	218 (69)	83 (61)	135 (76)	
Differentiation, n (%)				0.907
poor	39 (12)	18 (13)	21 (12)	
moderate	236 (75)	101 (74)	135 (76)	
well	40 (13)	18 (13)	22 (12)	
Age, Mean ± SD	50.09 ± 11.45	50.35 ± 11.72	49.88 ± 11.27	0.721
Plt, Median (Q1,Q3)	171 (128, 216.5)	172 (123, 209)	169.5 (130, 222.5)	0.711
N, Median (Q1,Q3)	3.25 (2.46, 4.28)	3.12 (2.41, 3.98)	3.37 (2.49, 4.69)	0.087
L, Median (Q1,Q3)	1.62 (1.25, 2.08)	1.64 (1.3, 2.03)	1.61 (1.22, 2.1)	0.662
M, Median (Q1,Q3)	0.42 (0.32, 0.56)	0.38 (0.3, 0.52)	0.44 (0.33, 0.6)	0.011
NLR, Median (Q1,Q3)	1.94 (1.54, 2.69)	1.91 (1.52, 2.55)	1.98 (1.58, 3.01)	0.205
TB, Median (Q1,Q3)	14.1 (10.45, 18.1)	14.2 (10.6, 18.7)	14.05 (10.4, 17.95)	0.766
Albumin, Mean ± SD	40.09 ± 4.24	40.47 ± 4.48	39.79 ± 4.04	0.16
ALT, Median (Q1,Q3)	38 (28, 52)	36 (24, 48)	40 (30, 56.75)	0.042
AST, Median (Q1,Q3)	38 (29, 50)	34 (27.5, 47.5)	40 (30, 50.75)	0.041
Total cholesterol, Median (Q1,Q3)	4.45 (3.76, 5.26)	4.44 (3.72, 5.18)	4.49 (3.8, 5.32)	0.379
GGT, Median (Q1,Q3)	61 (38, 107)	57 (32, 85.5)	65 (42.25, 118.5)	0.004
Fibrinogen, Median (Q1,Q3)	2.98 (2.47, 3.65)	2.88 (2.39, 3.37)	3.12 (2.54, 3.89)	0.031

**Abbreviation:** Plt: platelet; N: neutrophils; L: lymphocyte; M: macrophages; NLR: neutrophils to lymphocytes ratio; TB: total bilirubin; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT:  $\gamma$ -glutamyl transpeptidase; BCLC: Barcelona Clinic Liver Cancer staging classification.

## 2.5. Statistical analysis

The data were analyzed by Python (version 3.8, Python Software Foundation) and R software (version 4.0.2). Models were constructed using the training set, and internal test sets were used for model validation and evaluation. The ratio of the training set to the internal test set is 7:3. Continuous numerical variables were assessed for normality using the Kolmogorov-Smirnov test [26]. Data that follows a normal distribution is represented as mean  $\pm$  standard deviation (mean  $\pm$  SD), while non-normally distributed data is represented using the median (Q1, Q3) quartiles. Categorical variables are represented using frequency distribution. Two independent samples t-tests were used for continuous variables. For categorical variables, the Chi-square test was used. Logistic regression was used to analyze the recurrence risk factors in patients with liver cancer after hepatectomy. Python SHAP package was used to perform the SHAP. Python programming language (version 3.8) was applied to build and evaluate ML models and design the online risk calculator. P-Values < 0.05 were considered statistically significant with 95 % confidence intervals (CIs) applied for all logistic regression analyses.

## 3. Result

### 3.1. Clinical characteristics

315 patients with HCC after hepatectomy were enrolled. 178 patients experienced recurrence and 137 patients did not experience recurrence during the follow-up time. There were significant differences in tumor size, tumor number, BCLC stage, HBV DNA, macrophages, alanine aminotransferase (ALT), aspartate aminotransferase (AST),  $\gamma$ -glutamyl transpeptidase (GGT) and fibrinogen between the recurrence and non-recurrence groups (P < 0.05). The demographic and clinicopathological characteristics of the patients were detailed in Table 1.

### 3.2. Univariate and multivariate logistic regression

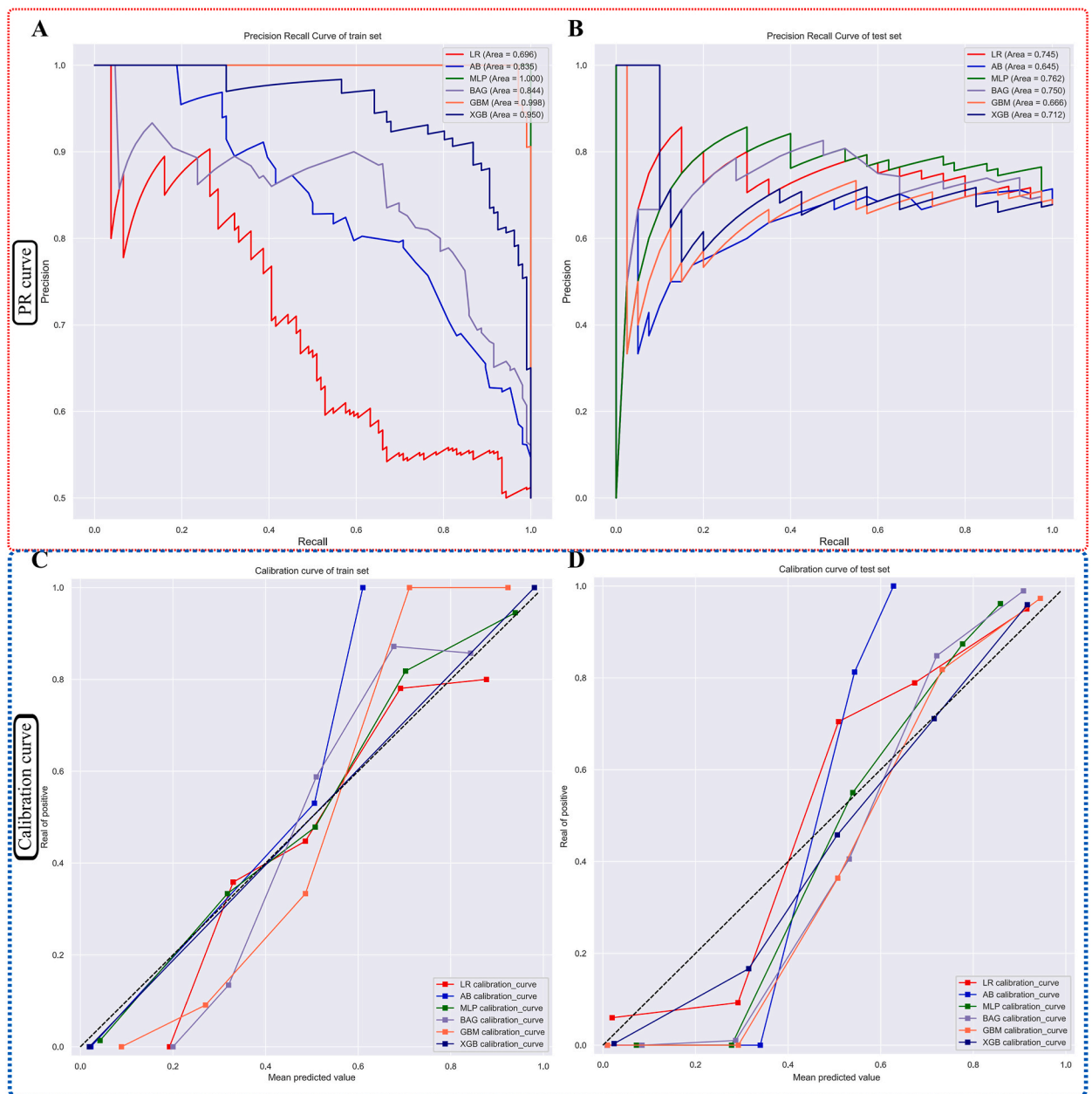
Variables with P value < 0.1 in univariate analysis were screened for multivariate logistic regression analysis to determine the risk

**Table 2**  
Univariate and multivariate logistic regression.

Characteristics	Category	Univariate analysis		Multivariate analysis	
		OR (95 % CI)	P value	OR (95 % CI)	P value
Gender	male	Ref	Ref	Ref	Ref
	female	0.498 (0.223–1.111)	0.089	0.894 (0.3–2.66)	0.84
Tumor size	$\leq$ 3 cm	Ref	Ref	Ref	Ref
	> 3 cm	2.688 (1.611–4.485)	<0.001	1.417 (0.488–4.111)	0.521
Tumor number	1	Ref	Ref	Ref	Ref
	> 1	2.935 (1.694–5.083)	<0.001	4.723 (1.924–11.598)	0.001
Vascular invasion	No	Ref	Ref	Ref	Ref
	Yes	1.469 (0.921–2.343)	0.106	\	\
BCLC	0	Ref	Ref	Ref	Ref
	1	3.12 (1.578–6.168)	0.001	1.841 (0.497–6.822)	0.361
	2	5.365 (2.109–13.644)	<0.001	0.721 (0.124–4.18)	0.715
	3	3.816 (1.915–7.606)	<0.001	1.34 (0.335–5.367)	0.679
AFP	$\leq$ 400	Ref	Ref	Ref	Ref
	> 400	1.427 (0.873–2.332)	0.156	\	\
HBV DNA	$\leq$ 1000	Ref	Ref	Ref	Ref
	> 1000	2.043 (1.258–3.318)	0.004	1.747 (0.861–3.547)	0.123
Differentiation	poor	Ref	Ref	Ref	Ref
	moderate	1.146 (0.58–2.262)	0.695	\	\
	well	1.048 (0.432–2.54)	0.918	\	\
N	\	1.178 (1.012–1.37)	0.034	0.931 (0.684–1.266)	0.647
M	\	4.84 (1.472–15.909)	0.009	2.367 (0.461–12.151)	0.302
Fibrinogen	\	1.34 (1.027–1.749)	0.031	1.178 (0.855–1.623)	0.317
Age	\	0.996 (0.977–1.016)	0.719	\	\
Plt	\	1.001 (0.998–1.004)	0.663	\	\
L	\	0.946 (0.674–1.327)	0.746	\	\
Albumin	\	0.962 (0.912–1.015)	0.155	\	\
NLR	\	1.185 (1–1.405)	0.05	1.219 (0.909–1.633)	0.185
AST	\	1.01 (0.998–1.023)	0.1	\	\
TB	\	0.994 (0.979–1.009)	0.403	\	\
GGT	\	1.003 (0.999–1.006)	0.118	\	\
ALT	\	1.004 (0.995–1.012)	0.375	\	\
Total cholesterol	\	1.098 (0.876–1.375)	0.417	\	\

**Abbreviation:** Plt: platelet; N: neutral granulocyte; L: lymphocyte; M: macrophages; NLR: neutrophils to lymphocyte ratio; TB: total bilirubin; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT:  $\gamma$ -glutamyl transpeptidase; BCLC: Barcelona Clinic Liver Cancer staging classification.

factors for postoperative recurrence in HCC patients after hepatectomy (Table 2) [27,28]. In univariate analysis, gender, tumor size, tumor number, BCLC, HBV-DNA, neutrophil, macrophages, fibrinogen, and neutrophils to lymphocytes ratio (NLR) were risk factors for postoperative recurrence. Multivariate logistic regression showed that tumor number was an independent risk factor for HCC recurrence (P value < 0.05).



**Fig. 2.** Precision-Recall Curves and Calibration Curves for six different machine learning algorithms on both the training and testing sets. **Notes:** Fig. 2A displays the Precision-Recall (PR) curve for the training set, while Fig. 2B shows the PR curve for the testing set. MLP performs the best in both the training and testing sets, with an Area under the Curve (AUC) of 1 in the training set and 0.762 in the testing set. Fig. 2C and D represent the calibration curves for the training and testing sets, respectively. In these curves, the predictions from the MLP machine learning algorithm are closest to the actual values, indicating superior calibration performance. **Abbreviation:** AB, adaptive boosting; MLP, multilayer perceptron; BAG, bootstrapped aggregating; LR, logistic regression; GBM, gradient boosting machine; XGB, extreme gradient boost.

### 3.3. Model performance

We evaluated the recurrence risk of HCC patients after hepatectomy by six different machine learning models. A 10-fold cross-validation was performed to evaluate the predictive performance of all models in both the training and internal validation sets. Fig. 1B showed the average AUC of the six models, with the MLP model achieving the best performance (AUC =  $0.723 \pm 0.068$ ) (Fig. 1B). The AUC value of the MLP machine learning algorithm, which predicts the recurrence after hepatectomy in liver cancer patients in the training set, is 1, and its performance in the internal test set reached 0.680 (Fig. 1C). According to the Precision-Recall Curve results for the training and testing sets, the area under the curve (AUC) for the MLP machine learning model is 1 in the training set (Fig. 2A) and 0.762 in the testing set (Fig. 2B). The calibration curve results demonstrate that the predictive performance of the MLP machine learning model is closer to the actual real-world results compared to other machine learning models (Fig. 2C and D). The detailed results of the validation set are presented in Table 3. Furthermore, we plotted the confusion matrix of the model, in the training set after SMOTE processing, all predictions were correct (Fig. 3A). In the original distribution of the testing set, there were 37 correct predictions and 22 incorrect predictions, resulting in an accuracy of 0.627 (Fig. 3B). Using the best machine learning model, MLP, and performing internal five-fold cross-validation, the average AUC value is 0.66 with a standard deviation of 0.08 (Fig. 3C). For the above six ML models, we also plotted radar plots to evaluate the performance of different models. Compared with other ML models, MLP has the best value in the evaluation of various indicators including sensitivity, F1 score, AUC and accuracy (Fig. 3D).

### 3.4. Model interpretability

We used SHAP to establish the importance ranking of recurrence risk factors based on MLP (Fig. 4.). The variable importance ranking results show that GGT, fibrinogen, neutrophils, AST, TB, and lymphocytes are important factors influencing postoperative recurrence in liver cancer patients. In Fig. 4A, the factors highlighted in red are significant risk factors for recurrence, while those in blue are protective factors against recurrence. Corresponding to this, in the samples used in this study, there are violin plots depicting the feature importance rankings for each sample in Fig. 4B. In addition, according to SHAP values, we selected two subjects respectively, including members of the recurrence group and non-recurrence group. In the recurrent samples, N = 4.11, AST = 46.0, Albumin = 43.1, TB = 15, Total cholesterol = 5.24, and L = 0.79 are important risk factors for recurrence (Fig. 4C). Conversely, factors like Fibrinogen = 2.21 and NLR = 5.2 are important protective factors against recurrence. In the non-recurrent samples, Albumin = 34.7, Fibrinogen = 3.42, GGT = 202.0, and NLR = 0.86 are important risk factors for recurrence, while factors like TB = 13.1, AST = 71.0, N = 2.82, L = 3.28, and Age = 67 are important protective factors against recurrence (Fig. 4D).

### 3.5. Web calculator

Based on the MLP model, we developed a web predictor to predict the risk of postoperative recurrence in HCC patients after hepatectomy. We can predict the postoperative recurrence risk in HCC patients after hepatectomy by setting variables in the sidebar of the website, and provide the importance contributions of each indicator for the occurrence of postoperative recurrence in the respective participants (<https://livercancerrecurrence-jpbrmbtbsfn4jojso4anta.streamlit.app/>) (Fig. 5). In addition, readers can also access the code used in our research, as well as replicate the figures in the article, by visiting ([https://github.com/Wu-Shi-Nan/Liver\\_cancer/blob/main/code.py](https://github.com/Wu-Shi-Nan/Liver_cancer/blob/main/code.py)). The parameters for each machine learning model and the versions of the libraries used are also available in the GitHub repository.

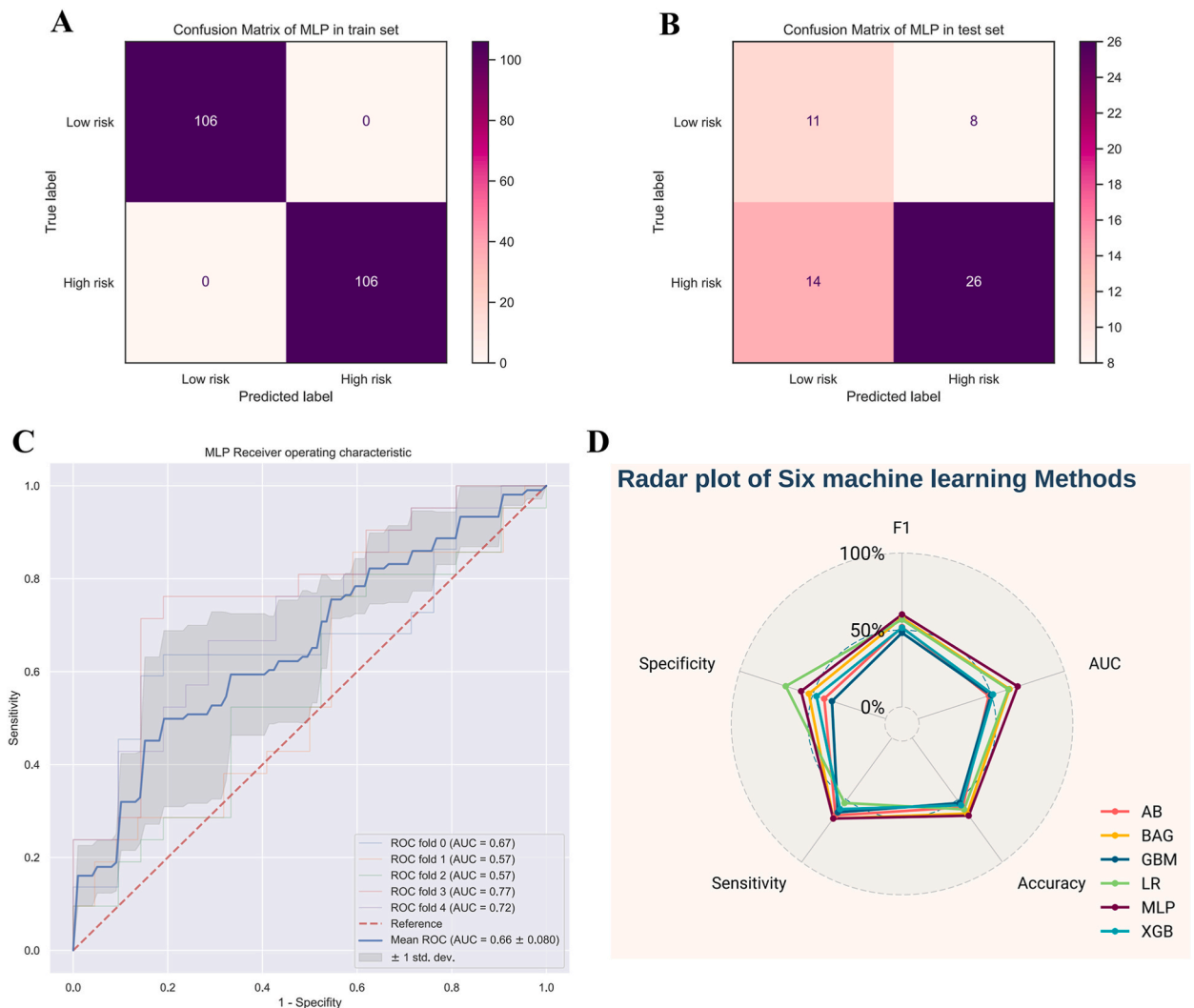
## 4. Discussion

Studies revealed that the recurrence risk within 5 years remained as high as 60 % even in patients with early-stage hepatocellular carcinoma after hepatectomy [29]. The high recurrence risk of HCC greatly reduced the survival rate of patients. Multiple preoperative indicators influenced cancer patient recurrence after hepatectomy. Zhang et al. found that GGT can serve as an independent factor for assessing postoperative survival prognosis in liver cancer patients, and it is closely associated with postoperative overall survival (OS) and disease-free survival (DFS) [17]. In our current study, the results also indicate that GGT is a positive factor influencing the risk of postoperative recurrence in liver cancer patients. Dai et al. demonstrated that preoperative high fibrinogen was associated with rapid recurrence of HCC [30]. Similarly, based on our feature variable importance ranking using MLP machine learning, it is evident that

**Table 3**  
Performance comparison of six ML models.

Model	F1	AUC	Accuracy	Sensitivity	Specificity
AB	0.519	0.483	0.559	0.625	0.421
LR	0.568	0.617	0.576	0.525	0.684
BAG	0.579	0.624	0.610	0.650	0.526
MLP	0.601	0.680	0.627	0.650	0.579
GBM	0.482	0.497	0.525	0.600	0.368
XGB	0.515	0.511	0.542	0.575	0.474

**Abbreviation:** ML, machine learning; AUC, area under the curve; AB, adaptive boosting; MLP, multilayer perceptron; BAG, bootstrapped aggregating; LR, logistic regression; GBM, gradient boosting machine; XGB, extreme gradient boost.



**Fig. 3.** The confusion matrices for the training and testing sets, the five-fold cross-validated ROC curve for the best machine learning model MLP, and a comparative radar chart of different machine learning methods under various evaluation metrics.

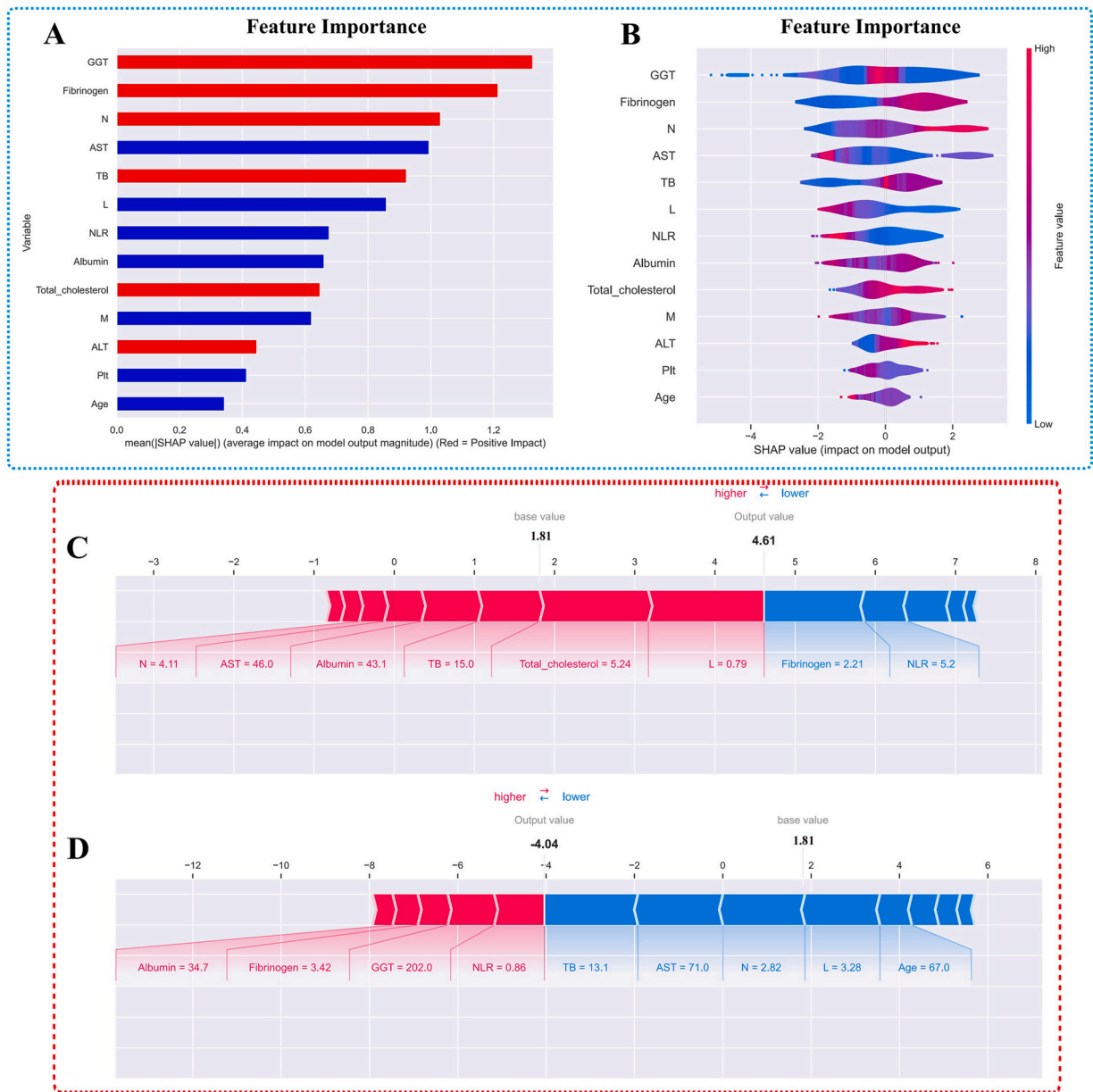
**Notes:** Fig. 3A and B displays the confusion matrices for the training and testing sets. In the testing set’s confusion matrix, the accuracy is observed to be 0.63. Fig. 3C shows the five-fold cross-validated ROC curve results for the best machine learning model, MLP, with an AUC of  $0.66 \pm 0.08$ . Fig. 3D presents a radar chart visualization of different machine learning algorithms under various evaluation metrics. In this chart, the MLP machine learning algorithm exhibits the best performance in terms of F1-score, sensitivity, accuracy, and AUC value.

**Abbreviation:** MLP, multilayer perceptron; ROC, receiver operating characteristic; AUC, the area under the ROC curve.

fibrinogen significantly contributes as a positive factor influencing its recurrence. Recent studies have shown that the immune microenvironment plays a crucial role in the invasive and malignant tendencies of tumors. Neutrophils, as important stress cells in cellular inflammation, have been previously shown to be an important indicator for assessing tumor staging and invasive characteristics in liver cancer when considering the neutrophil-to-lymphocyte ratio [31]. In our current study, the abnormally elevated levels of neutrophils were also significantly correlated with postoperative recurrence in liver cancer patients. We employed the MLP machine learning algorithm to predict postoperative recurrence in liver cancer patients and analyzed the feature variables that significantly impact recurrence. Some indicators emerged as significant risk factors, such as GGT, fibrinogen, neutrophils, and TB, while others served as significant protective factors, including AST and lymphocytes. Through visual explanations of the black-box machine learning model, we were able to provide insights into the effects of abnormal serum value elevations or reductions on predicting postoperative recurrence in liver cancer patients.

Researchers have proposed several methods to assess the prognosis of patients with liver cancer, such as tumor-node-metastasis, cancer of the liver Italian program (CLIP) score, the Singapore liver cancer recurrence (SLICER) score, and china liver cancer (CNLC) staging [32–34]. However, due to inherent defects, the prognostic value of these methods was still not widely recognized. In addition, many prognostic models based on cox regression have also been developed, such as ERASL and ALBI grade [35,36]. However,





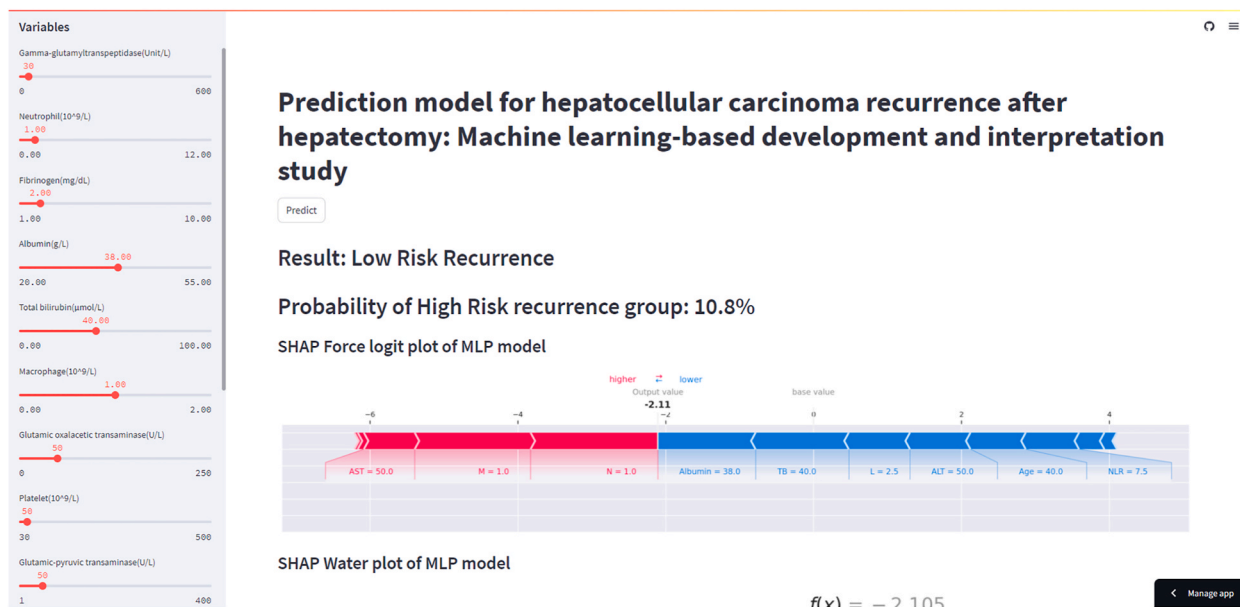
**Fig. 4.** SHAP summary plot and SHAP explanation of two patients.

**Notes:** Fig. 4A and B displays the importance ranking of feature variables based on the MLP machine learning model. Red represents variables that are risk factors for postoperative recurrence in hepatocellular carcinoma, while blue represents variables that act as protective factors. Fig. 4C presents a high-risk SHAP interpretation model for postoperative recurrence in patients with hepatocellular carcinoma after hepatectomy, and Fig. 4D shows a low-risk SHAP interpretation model for postoperative recurrence in the same patient group.

**Abbreviation:** SHAP, Shapley additive explanations; MLP, multilayer perceptron; N: neutrophil; M: macrophages; NLR: neutrophils to lymphocytes ratio; GGT:  $\gamma$ -glutamyl transpeptidase; Plt: platelets, L: lymphocytes, AST: aspartate aminotransferase, TB: total bilirubin, ALT: alanine aminotransferase.

the prediction methods based on cox regression are not very accurate due to their statistical drawbacks. ML can be adapted to more complex nonlinear relationships because it can assess the nonlinear impact of risk factors on patient survival [37]. Therefore, more and more predictive models based on ML have been developed to predict the prognosis of tumor patients and show a bright application prospect.

In the present study, we compared six different machine learning models to assess the recurrence risk of HCC patients after hepatectomy and finally determined that the MLP model was the best predictive model. We further developed a web calculator based on the MLP model. Clinicians can assess the patient's recurrence risk by entering the patient's indicators on the web page.



**Fig. 5.** Web calculator for predicting the recurrence of patients with hepatocellular carcinoma after hepatectomy.

Furthermore, we conducted a comprehensive assessment of the various indicators for each participant. If an indicator was identified as a risk factor for recurrence, it was displayed in red on the web calculator. Conversely, if an indicator was deemed a protective factor for recurrence, it was represented in blue. MLP model is a machine learning algorithm based on an artificial neural network, which can effectively analyze linear and nonlinear characteristic variables to make effective and accurate predictions [38]. Previous studies used MLP to preprocess input images and achieve an accurate diagnosis of bladder cancer [39]. However, although machine learning models are more powerful and accurate than traditional statistical models, the interpretation of the models is correspondingly more complex. They like a black box, which limits their further clinical application. In this study, we further explained the MLP model through the SHAP method. SHAP is an independent machine learning model interpretation technique, which can explain the global and individual sample black box models and help understand the relationship between prediction indicators and results in the MLP model [40]. We applied SHAP to enhance the global interpretation of the MLP model applied to the recurrence risk prediction of HCC patients after liver resection and help to increase clinicians' trust in the clinical application of the machine learning model.

There were still some limitations in the current study. Firstly, this study was a single-center retrospective study, and the performance of the ML model may vary depending on the characteristics of patients in different regions and the datasets of different institutions. Moreover, as this is a retrospective study, it introduces the possibility of selection bias and limits the establishment of causal relationships. Secondly, because our sample size is relatively small, there is a significant margin of error in the AUC values during the ten-fold cross-validation process. However, multiple cross-validation results show that the optimal machine learning model is still MLP. We will further increase the sample size in our subsequent studies to obtain stable cross-validation results. Lastly, the specificity of our best machine learning model, MLP, in this study is relatively low. One potential reason might be the limited number of feature variables included. In subsequent research, we will collect more pre-operative, intra-operative, and post-operative indicators to establish a comprehensive and standardized decision-making system.

## 5. Conclusion

We successfully established an MLP model to predict the recurrence risk of HCC patients after hepatectomy. This predictive model can help identify patients at high recurrence risk and can provide patients with early and personalized treatment to further improve the prognosis and quality of life.

## Supportive foundations

None.

## Data availability statement

All data was in the manuscript and can be obtained from the corresponding author.

## Authors' contributions

Yang Yang and Linsen Ye contributed to the study's inception and design. Rongqiang Liu, Shinan Wu, Hao Yuan Yu and Kaining Zeng equally contributed to the literature search, analysis and writing of the manuscript. Other authors contributed to the study design and study supervision. All authors approved the final version of the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Abbreviation

HCC	hepatocellular carcinoma
ML	machine learning
MLP	multilayer perceptron
SHAP	Shapley additive explanations
AB	adaptive boosting
BAG	bootstrapped aggregating;
LR	logistic regression
GBM	gradient boosting machine;
XGB	extreme gradient boost
ROC	receiver operating characteristic curve
ALT	aminotransferase
AST	aspartate aminotransferase
GGT	$\gamma$ -glutamyl transpeptidase
NLR	neutrophils to lymphocytes ratio

## References

- [1] A. Vogel, T. Meyer, G. Sapichochin, R. Salem, A. Saborowski, Hepatocellular carcinoma, *Lancet* 400 (10360) (2022) 1345–1362.
- [2] H.B. El-Serag, Hepatocellular carcinoma, *N. Engl. J. Med.* 365 (12) (2011) 1118–1127.
- [3] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *Ca - Cancer J. Clin.* 71 (3) (2021) 209–249.
- [4] H. Maki, K. Hasegawa, Advances in the surgical treatment of liver cancer, *Biosci. Trends* 16 (3) (2022) 178–188.
- [5] P. Tabrizian, G. Jibara, B. Shrager, M. Schwartz, S. Roayaie, Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis, *Ann. Surg.* 261 (5) (2015) 947–955.
- [6] R.C. Deo, Machine learning in medicine, *Circulation* 132 (20) (2015) 1920–1930.
- [7] L.M. Fleuren, T.L.T. Klausch, C.L. Zwager, L.J. Schoonmade, T. Guo, L.F. Roggeveen, E.L. Swart, A.R.J. Girbes, P. Thoral, A. Ercole, et al., Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy, *Intensive Care Med.* 46 (3) (2020) 383–400.
- [8] M. Montazeri, R. ZahediNasab, A. Farahani, H. Mohseni, F. Ghasemian, Machine learning models for image-based diagnosis and prognosis of COVID-19: systematic review, *JMIR Med. Inform* 9 (4) (2021), e25181.
- [9] W.C. Liu, M.X. Li, W.X. Qian, Z.W. Luo, W.J. Liao, Z.L. Liu, J.M. Liu, Application of machine learning techniques to predict bone metastasis in patients with prostate cancer, *Cancer Manag. Res.* 13 (2021) 8723–8736.
- [10] W. Li, Q. Zhou, W. Liu, C. Xu, Z.R. Tang, S. Dong, H. Wang, W. Li, K. Zhang, R. Li, et al., A machine learning-based predictive model for predicting lymph node metastasis in patients with Ewing's sarcoma, *Front. Med.* 9 (2022), 832108.
- [11] Z. Hasnain, J. Mason, K. Gill, G. Miranda, I.S. Gill, P. Kuhn, P.K. Newton, Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients, *PLoS One* 14 (2) (2019), e0210976.
- [12] J.D. Liang, X.O. Ping, Y.J. Tseng, G.T. Huang, F. Lai, P.M. Yang, Recurrence predictive models for patients with hepatocellular carcinoma after radiofrequency ablation using support vector machines with feature selection methods, *Comput. Methods Progr. Biomed.* 117 (3) (2014) 425–434.
- [13] K. Wang, T. Hong, W. Liu, C. Xu, C. Yin, H. Liu, X. Wei, S.N. Wu, W. Li, L. Rong, Development and validation of a machine learning-based prognostic risk stratification model for acute ischemic stroke, *Sci. Rep.* 13 (1) (2023), 13782.
- [14] Y. Huang, H. Chen, Y. Zeng, Z. Liu, H. Ma, J. Liu, Development and validation of a machine learning prognostic model for hepatocellular carcinoma recurrence after surgical resection, *Front. Oncol.* 10 (2020), 593741.
- [15] J. Zeng, J. Zeng, K. Lin, H. Lin, Q. Wu, P. Guo, W. Zhou, J. Liu, Development of a machine learning model to predict early recurrence for hepatocellular carcinoma after curative resection, *Hepatobiliary Surg. Nutr.* 11 (2) (2022) 176–187.
- [16] G.W. Ji, F.P. Zhu, Q. Xu, K. Wang, M.Y. Wu, W.W. Tang, X.C. Li, X.H. Wang, Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: a multi-institutional study, *EBioMedicine* 50 (2019) 156–165.
- [17] I. Ekanayake, D. Meddage, U. Rathnayake, A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP), *Case Stud. Constr. Mater.* 16 (2022), e01059.
- [18] A. Fernández, S. García, F. Herrera, N.V. Chawla, SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary, *J. Artif. Intell. Res.* 61 (2018) 863–905.
- [19] Y. Liu, Y. Wang, J. Zhang, New machine learning algorithm: random forest, in: *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14–16, 2012 Proceedings* 3: 2012, Springer, 2012, pp. 246–252.
- [20] A. Marciano-Cedeño, J. Quintanilla-Domínguez, M. Cortina-Januchs, D. Andina, Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network, in: *IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society: 2010, IEEE, 2010*, pp. 2845–2850.
- [21] S.K. Pal, S. Mitra, Multilayer Perceptron, *Fuzzy Sets, Classification*, 1992.

- [22] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, *Front. Neurobot.* 7 (2013).
- [23] J.C. Stoltzfus, Logistic regression: a brief primer, *Acad. Emerg. Med.* 18 (10) (2011) 1099–1104.
- [24] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, Xgboost: extreme gradient boosting, R package version 04-2 1 (4) (2015) 1–4.
- [25] R. Rodríguez-Pérez, J. Bajorath, Interpretation of compound activity predictions from complex machine learning models using local approximations and Shapley values, *J. Med. Chem.* 63 (16) (2020) 8761–8777.
- [26] F.J. Massey Jr., The Kolmogorov-Smirnov test for goodness of fit, *J. Am. Stat. Assoc.* 46 (253) (1951) 68–78.
- [27] M.A. López-Espejo, M.H. Chávez, I. Huete, Short-term outcomes after a neonatal arterial ischemic stroke, *Childs Nerv. Syst.* 37 (4) (2021) 1249–1254.
- [28] A. Alawneh, H. Anshasi, Place of death for patients treated at a tertiary cancer center in Jordan, *Support. Care Cancer* 29 (4) (2021) 1837–1842.
- [29] S. Roayaie, K. Obeidat, C. Sposito, L. Mariani, S. Bhoori, A. Pellegrinelli, D. Labow, J.M. Llovet, M. Schwartz, V. Mazzaferro, Resection of hepatocellular cancer  $\leq 2$  cm: results from two Western centers, *Hepatology* 57 (4) (2013) 1426–1435.
- [30] T. Dai, L. Peng, G. Lin, Y. Li, J. Yao, Y. Deng, H. Li, G. Wang, W. Liu, Y. Yang, et al., Preoperative elevated plasma fibrinogen level predicts tumor recurrence and poor prognosis in patients with hepatocellular carcinoma, *J. Gastrointest. Oncol.* 10 (6) (2019) 1049–1063.
- [31] L. Wong, K. Bozhilov, B. Hernandez, S. Kwee, O. Chan, L. Ellis, L. LeMarchand, Underlying liver disease and advanced stage liver cancer are associated with elevated neutrophil-lymphocyte ratio, *Clin. Mol. Hepatol.* 25 (3) (2019) 305–316.
- [32] J.A. Marrero, R.J. Fontana, A. Barrat, F. Askari, H.S. Conjeevaram, G.L. Su, A.S. Lok, Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort, *Hepatology* 41 (4) (2005) 707–716.
- [33] S.F. Ang, E.S. Ng, H. Li, Y.H. Ong, S.P. Choo, J. Ngeow, H.C. Toh, K.H. Lim, H.Y. Yap, C.K. Tan, et al., The Singapore Liver Cancer Recurrence (SLICER) Score for relapse prediction in patients with surgically resected hepatocellular carcinoma, *PLoS One* 10 (4) (2015), e0118658.
- [34] D.Y. Xie, Z.G. Ren, J. Zhou, J. Fan, Q. Gao, Chinese clinical guidelines for the management of hepatocellular carcinoma: updates and insights, *Hepatobiliary Surg. Nutr.* 2020 9 (4) (2019) 452–463.
- [35] A.W.H. Chan, J. Zhong, S. Berhane, H. Toyoda, A. Cucchetti, K. Shi, T. Tada, C.C.N. Chong, B.D. Xiang, L.Q. Li, et al., Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection, *J. Hepatol.* 69 (6) (2018) 1284–1293.
- [36] C.Y. Lin, C.C. Lin, C.C. Wang, C.L. Chen, T.H. Hu, C.H. Hung, P.Y. Huang, M.C. Tsai, The ALBI grade is a good predictive model for very late recurrence in patients with hepatocellular carcinoma undergoing primary resection, *World J. Surg.* 44 (1) (2020) 247–257.
- [37] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction, *Comput. Struct. Biotechnol. J.* 13 (2015) 8–17.
- [38] R.U. Khan, S. Almakdi, M. Alshehri, R. Kumar, I. Ali, S.M. Hussain, A.U. Haq, I. Khan, A. Ullah, M.I. Uddin, Probabilistic approach to COVID-19 data analysis and forecasting future outbreaks using a multi-layer perceptron neural network, *Diagnostics* 12 (10) (2022).
- [39] I. Lorencin, N. Andelić, J. Španjol, Z. Car, Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis, *Artif. Intell. Med.* 102 (2020), 101746.
- [40] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.I. Lee, From local explanations to global understanding with explainable AI for trees, *Nat. Mach. Intell.* 2 (1) (2020) 56–67.