

Construction of a Preoperative Prediction Model for TACE Resistance in Primary Hepatocellular Carcinoma Based on Machine Learning Algorithms

Huyu Jiao*, Zhengang Zhang*

Department of Gastroenterology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, 430030, People's Republic of China

*These authors contributed equally to this work

Correspondence: Zhengang Zhang, Department of Gastroenterology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095 Jiefang Avenue, Wuhan, Hubei, 430030, People's Republic of China, Email zgzhang@tjh.tjmu.edu.cn

Purpose: Transcatheter arterial chemoembolization (TACE) resistance compromises prognosis in unresectable hepatocellular carcinoma (HCC). This study aimed to develop an interpretable prediction model using machine learning (ML) and Shapley Additive Explanations (SHAP) for preoperative assessment of TACE resistance.

Patients and Methods: A single-center retrospective analysis included 562 HCC patients who received ≥ 3 TACE sessions (2013–2024). Multi-modal features (blood routine, coagulation, biochemistry, imaging) were integrated. Seven ML models (LR, RF, DT, XGBoost, LightGBM, SVM, ANN) were constructed. Feature selection used univariate Logistic regression and Lasso regression. Model performance was evaluated via AUC, F1 score, and accuracy; SHAP analyzed feature importance.

Results: Data were split into training ($n=394$, 70%) and validation ($n=168$, 30%) sets. Seven core predictors (NLR, tumor capsule integrity, AFP, etc.) were identified. XGBoost outperformed other models, with AUCs of 0.942 (95% CI: 0.919–0.966) and 0.898 (95% CI: 0.853–0.944) in training and validation sets, respectively, and an F1 score of 0.741. SHAP revealed NLR (mean Shapley value=0.13) and tumor capsule absence (0.08) as the strongest predictors.

Conclusion: This interpretable ML model efficiently predicts TACE resistance using multi-modal data, with $AUC > 0.8$. It offers a preoperative tool to identify high-risk patients, optimize treatment strategies, and holds significant clinical translational value.

Keywords: hepatocellular carcinoma, TACE resistance, machine learning, prediction model, SHAP analysis

Introduction

Primary hepatocellular carcinoma (HCC) is the sixth most common malignant tumor globally and the third leading cause of cancer-related death, with a 5-year survival rate of less than 20%.¹ Epidemiological data show that more than 70% of patients are diagnosed at an advanced stage (Barcelona Clinic Liver Cancer [BCLC] stage B/C) for the first time, losing the opportunity for radical surgery, and their prognosis is significantly worse than that of early-stage patients (the difference in 5-year survival rate exceeds 40%).² For unresectable HCC (uHCC), transcatheter arterial chemoembolization (TACE) remains the first-line treatment recommended by international guidelines.³ However, clinical observations have shown that approximately 65–70% of uHCC patients experience tumor progression within 1 year after the first TACE treatment. This drug resistance phenomenon is closely related to the remodeling of the tumor microenvironment, manifested by the upregulation of angiogenic factors (such as VEGF), increased infiltration of immunosuppressive cells, and activation of hypoxia-inducible factor, ultimately leading to treatment resistance.⁴

Machine learning (ML) technology in artificial intelligence is revolutionizing the mode of tumor prognosis prediction due to its excellent multi-dimensional data analysis capabilities. Compared with traditional statistical methods, ML has the advantage of being able to model both linear correlations and non-linear interactions simultaneously, making it particularly suitable for processing heterogeneous data commonly seen in clinical practice (such as continuous laboratory indicators, categorical imaging features, and time-series treatment records).^{5–7} By integrating multi-modal data (clinical indicators + radiomics + pathological features), ML models can construct high-precision prediction frameworks, with a false positive rate 30–40% lower than that of traditional models. More importantly, combined with interpretable technologies such as Shapley Additive Explanations (SHAP), clinicians can intuitively understand the model's decision-making logic, such as identifying key predictive factors and their contribution degrees, thereby enhancing the clinical trust and practicality of the model. Therefore, this study aimed to develop and validate an interpretable prediction model based on ML, which can realize preoperative individualized prediction of TACE resistance by integrating preoperative blood routine, coagulation function, imaging features, and clinical indicators.

Methods

Data Collection

This study strictly followed the ethical guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of Tongji Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology (Ethics Approval No.: TJ-IRB-2024-032). Due to the retrospective design of the study, written informed consent from patients was exempted in accordance with the Measures for the Ethical Review of Biomedical Research Involving Humans. We pledge to maintain strict confidentiality of patient data and ensure that the data will be used solely for the purposes of this study. The data source was limited to HCC patients who received standardized TACE treatment in our hospital from January 2013 to October 2024. The inclusion criteria were strictly formulated with reference to the Guidelines for the Diagnosis and Treatment of Primary Liver Cancer in China (2024 Edition).

The specific inclusion criteria were as follows: (1) Age ≥ 18 years, diagnosed with HCC by pathology or imaging (enhanced CT/MRI); (2) Preoperative liver function classified as Child-Pugh grade A or B; (3) Received ≥ 3 consecutive and standardized TACE treatments, without surgery, ablation, targeted therapy, or immunotherapy during the treatment period; (4) Eastern Cooperative Oncology Group Performance Status (ECOG PS) score ≤ 1 ; (5) Completed imaging re-examination (enhanced CT/MRI) within 1–3 months after the last TACE, with measurable target lesions (according to mRECIST criteria).

The exclusion criteria included: (1) Received other anti-tumor treatments (such as radiotherapy, systemic therapy) before TACE; (2) Complicated with severe cardiopulmonary and renal dysfunction (eGFR < 30 mL/min/1.73m², New York Heart Association [NYHA] cardiac function grade III or above); (3) Complete embolization of the main portal vein with tumor thrombus and no collateral circulation established; (4) Complicated with active infection or hemorrhagic diseases.

Additional Clarification: It is important to note that among the included patients, those classified as BCLC stage C had no evidence of distant metastasis. This was confirmed through comprehensive imaging studies prior to enrollment, ensuring that the patients selected for this study were suitable for TACE treatment alone without the need for additional systemic therapy.

Treatment Process

TACE Treatment Procedure

All patients received superselective TACE. Under local anesthesia, a 5F catheter was inserted through the femoral artery to the celiac trunk/superior mesenteric artery for digital subtraction angiography (DSA) to clarify the tumor-feeding artery. Then, a 2.7F microcatheter was introduced for superselective catheterization to the target vessel. The embolization protocol adopted the “chemotherapeutic drug + lipiodol + gelatin sponge” sandwich method: first, an emulsion of epirubicin (20 mg) + lipiodol (10–20 mL) was injected, followed by embolization with gelatin sponge particles (350–500 μ m) until blood flow stagnated. Routine symptomatic supportive treatments such as liver protection and anti-emesis were given after surgery.

Follow-up and Efficacy Evaluation

All patients underwent computed tomography (CT) and/or magnetic resonance imaging (MRI), liver function, blood routine, and tumor marker follow-up within 1–3 months after each TACE treatment. According to the mRECIST criteria, the short-term efficacy was divided into Complete Response (CR), Partial Response (PR), Stable Disease (SD), and Progressive Disease (PD). PD was defined as an increase of $\geq 20\%$ in the sum of the diameters of target lesions or the appearance of new lesions. Patients with PR and SD needed to undergo TACE again, while the decision to continue TACE for patients with PD was made based on the disease assessment. Within 1–3 months after 3 consecutive standardized and refined TACE treatments, enhanced CT/MRI was used to determine whether the intrahepatic target lesions were still in the state of PD compared with those before the first TACE treatment. The evaluation criteria for this outcome referred to the definition of TACE resistance in China.⁸

Data Preprocessing

Thirty-eight candidate variables were extracted from the hospital electronic medical record system, covering demographic characteristics, serological indicators, imaging features, and clinical staging. For variables with random missing values (missing rate $< 15\%$) such as INR and APTT, multiple imputation was used to generate 5 complete datasets for combined analysis. Continuous variables were standardized by Z-score, and categorical variables were processed by one-hot encoding. The dataset was randomly divided into a training set and a validation set at a ratio of 7:3, and stratified sampling was used to ensure a balanced incidence of TACE resistance between the two groups.

Data used to construct ML models must meet the requirements of integrity, no missing values, and scale standardization. In this study, the predictive variables with missing values included INR, APTT, platelets, etc. The main reasons for missing values were that some patients were transferred from the emergency department or other hospitals, and relevant items were not carried out to avoid duplication during the second visit to our hospital, and the examination results from other hospitals were not included in our hospital's system. The above data missing were random missing, which was not suitable for mean imputation or univariate imputation. Multiple imputation is one of the mainstream methods widely used at present, which realizes missing value imputation through multi-chain equations.⁹ In this study, the MICE package in R 4.4.2 software was used to impute the original data, and a complete analysis dataset was finally generated.

Model Construction

In this study, stratified random sampling was used to divide the dataset into a training set ($n=394$) and a validation set ($n=168$) at a ratio of 7:3, ensuring a balanced incidence of TACE resistance between the two groups (42.7% in the training set vs 41.3% in the validation set). The feature selection process combined traditional statistical methods and machine learning techniques: first, univariate Logistic regression was used to screen variables significantly associated with TACE resistance ($P < 0.05$), then LASSO regression was used for further dimensionality reduction, and the optimal regularization parameter was determined by 10-fold cross-validation (λ .1SE criterion). Finally, 7 core predictive factors that met both univariate analysis ($P < 0.05$) and non-zero coefficients of LASSO regression were retained, including Neutrophil-to-Lymphocyte Ratio (NLR), tumor capsule integrity, Alpha-Fetoprotein (AFP) level, bilateral liver lobe involvement of tumor, platelet count, primary tumor size, and fibrinogen level. Based on the selected features, 7 machine learning models were constructed: Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), XGBoost, LightGBM, Support Vector Machine (SVM), and Artificial Neural Network (ANN). All models were subjected to hyperparameter tuning through 5-fold cross-validation.

Model Evaluation Indicators

This study adopted a multi-dimensional indicator system to comprehensively evaluate model performance: first, the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) were used to quantify the discriminative ability of the model, and the DeLong test was used to compare the AUC differences between different models; second, accuracy, F1 score, and Brier score were combined to comprehensively evaluate the prediction accuracy, where accuracy reflects the classification correctness rate, F1 score balances precision and recall, and Brier score measures the calibration degree of probability prediction; further, calibration curves were drawn to verify the consistency between predicted probabilities and actual observed probabilities, and Decision Curve Analysis (DCA) was performed to quantify the net clinical benefit under different threshold probabilities to

evaluate clinical practicality; finally, SHAP values were generated for the optimal model (XGBoost) to analyze feature contribution and directional correlation, enhancing the interpretability of model decisions. This multi-modal evaluation framework systematically integrates discriminative power, calibration degree, and clinical decision-making value, ensuring the robustness and practicality of the model.

Statistical Analysis

Dual platforms of R 4.4.2 and Python 3.10 were used for data analysis in this study. For continuous variables, median (interquartile range) was used to describe the data distribution characteristics, and the Kolmogorov–Smirnov test was used to evaluate normality. Corresponding statistical test methods were selected according to the data distribution type: the Mann–Whitney *U*-test was used for comparison between two groups, and the Kruskal–Wallis *H*-test was used for comparison among multiple groups. Categorical variables were presented as frequency, and inter-group differences were judged by the Chi-square test or Fisher’s exact test. All statistical analyses followed the two-tailed test principle, with $P < 0.05$ as the threshold for statistical significance, ensuring the scientificity and reliability of the results.

Results

Baseline Characteristics

The flow diagram of this study is shown in [Figure 1](#). Based on the aforementioned inclusion and exclusion criteria, a total of 562 HCC patients who received TACE treatment were finally included in this study, including 382 cases in the TACE-responsive group (67.97%) and 180 cases in the TACE-resistant group (32.03%).

Comparison of baseline characteristics ([Table 1](#)) showed that the TACE-resistant group had significant abnormalities in tumor burden, liver function status, and inflammatory indicators: the median tumor diameter was 97.50 mm, which was significantly larger than 75.00 mm in the responsive group ($P < 0.001$); the proportion of multiple tumors was 77.78%, the proportion of bilateral liver lobe involvement was 55.00%, the rate of tumor capsule absence was as high as 59.44%, and the rate of tumor thrombus formation was 44.44%, all of which were significantly higher than those in the responsive group ($P < 0.001$). In terms of laboratory tests, the proportion of $\text{AFP} \geq 400$ ng/mL in the resistant group was 63.33%, the median NLR was 4.04, the median fibrinogen was 3.66 g/L, and the median platelet count was 169.00×10^9 /L, all of which were significantly higher than those in the responsive group; while the lymphocyte count was significantly lower. Liver function evaluation showed that the proportion of Child–Pugh grade B in the resistant group was 12.78%, the proportion of ECOG PS score 1 was 19.44%, and the proportion of BCLC stage C was 42.78%, all of which were significantly worse than those in the responsive group. There were no statistically significant differences in age, albumin, coagulation function (except fibrinogen) and other indicators between the two groups. The above features were verified by the Mann–Whitney *U*-test or Chi-square test, and all *P* values were < 0.05 , providing a clinical biological basis for the prediction of TACE resistance.

Comparison of Model Performance

The performance of the 7 machine learning models in the validation set is shown in [Table 2](#): the XGBoost model showed the best comprehensive performance with an AUC value of 0.898 (95% CI: 0.853–0.944), and its accuracy (0.833), precision (0.741), sensitivity (0.741), and F1 score (0.741) were all the best. At the same time, its specificity (0.877) was comparable to that of Random Forest (0.860) and SVM (0.860), reflecting the balanced advantage of high specificity and sensitivity. Random Forest (AUC=0.894) and ANN (AUC=0.896) had similar performance, but the sensitivity of ANN (0.759) was slightly higher than that of XGBoost. Support Vector Machine (SVM) showed stable overall performance (AUC=0.893), while the Decision Tree model had relatively low indicators (AUC=0.808). Through the Receiver Operating Characteristic (ROC) curve ([Figure 2A](#)), Decision Curve Analysis (DCA) curve ([Figure 2B](#)), and comparison of specific indicators in [Table 2](#) and [Table 3](#), XGBoost was finally selected as the optimal prediction model. It achieves the best balance between sensitivity and precision while ensuring high specificity, providing a reliable tool for preoperative prediction of TACE resistance in clinical practice.

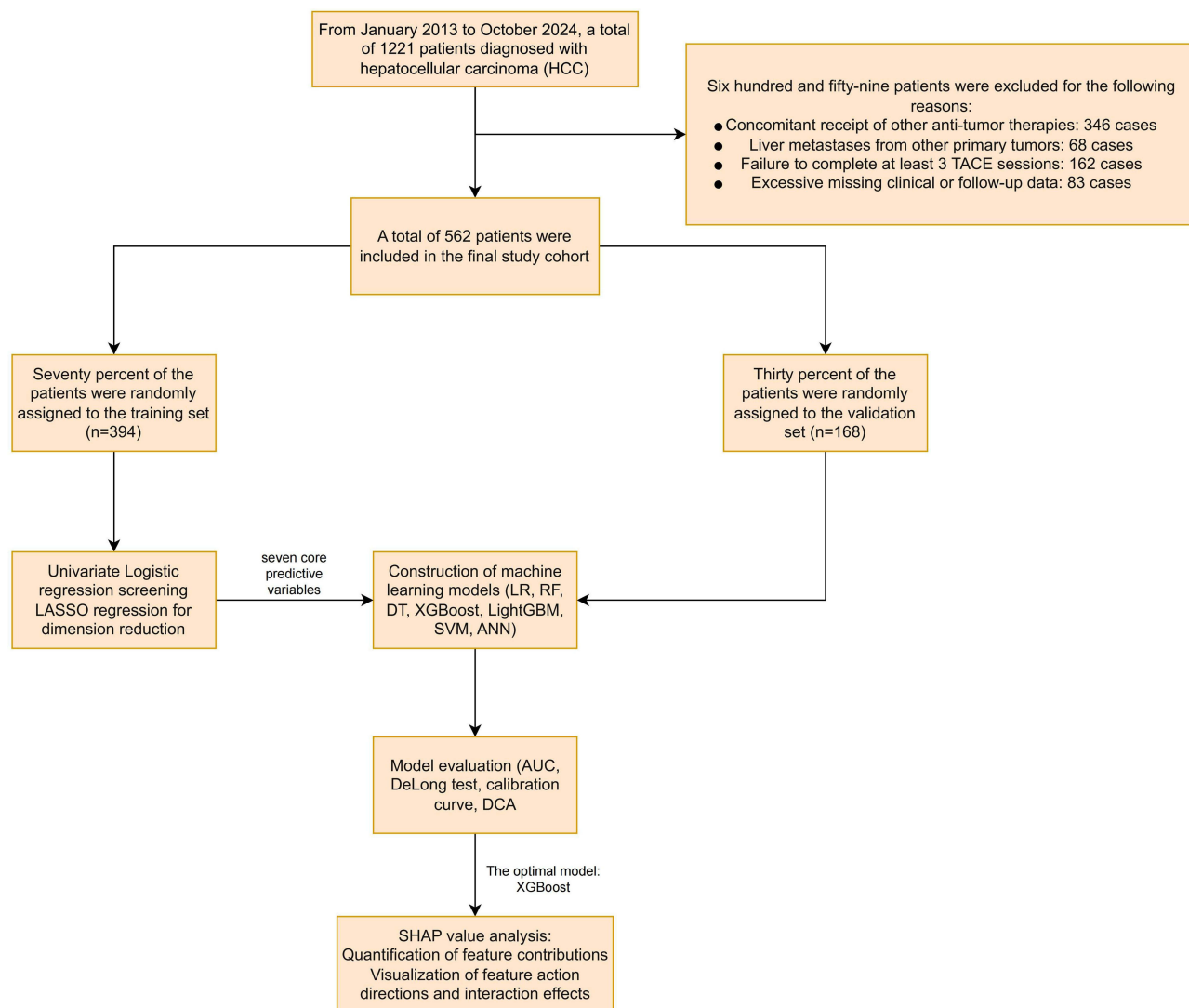


Figure 1 Flow Diagram of This Study.

Abbreviations: HCC, hepatocellular carcinoma; TACE, transarterial chemoembolization; LR, logistic regression; RF, random forest; DT, decision tree; XGBoost, extreme gradient boosting; LightGBM, light gradient boosting machine; SVM, support vector machine; ANN, artificial neural network; AUC, area under the roc curve; DCA, decision curve analysis.

Evaluation of Model Calibration

This study used calibration curves to evaluate the goodness-of-fit of each machine learning model. [Figure 3](#) shows the calibration curves and Brier scores of each model. In the validation cohort, the calibration curves of all 7 machine learning models performed well, and the Brier scores were all less than 0.2, indicating that the predicted probabilities of the models were highly consistent with the actual incidence of TACE resistance in HCC patients. Among them, the XGBoost model performed the best, with an AUC (95% CI) of 0.898.

Model Comparison

To assess the predictive performance of our developed model in accordance with existing clinical criteria, we selected the best - performing XGBoost model and compared it with the up - to - seven score, a widely - used predictive scoring system in clinical practice. As illustrated in [Figure 4](#), in terms of the AUC, the XGBoost model demonstrated significantly superior performance compared to the up - to - seven score. In the training set, the XGBoost model achieved an AUC of 0.942 (95% confidence interval: 0.919–0.966), while the up - to - seven score had an AUC of 0.712

Table 1 Baseline Characteristics of Patients

Characteristics	ALL N=562	Effective N=382	Resist N=180	p.Overall
Gender (%)				1
Female	63 (11.21)	43 (11.26)	20 (11.11)	
Male	499 (88.79)	339 (88.74)	160 (88.89)	
Schistosomiasis (%)				0.327
No	529 (94.13)	363 (95.03)	166 (92.22)	
Yes	33 (5.87)	19 (4.97)	14 (7.78)	
Comorbid diabetes (%)				0.379
No	508 (90.39)	342 (89.53)	166 (92.22)	
Yes	54 (9.61)	40 (10.47)	14 (7.78)	
Comorbid hypertension (%)				1
No	470 (83.63)	321 (84.03)	149 (82.78)	
Yes	92 (16.37)	61 (15.97)	31 (17.22)	
Comorbid hepatitis_B (%)				0.549
No	82 (14.59)	52 (13.61)	30 (16.67)	
Yes	480 (85.41)	330 (86.39)	150 (83.33)	
AFP (%)				<0.001
<400	348 (61.92)	282 (73.82)	66 (36.67)	
≥ 400	214 (38.08)	100 (26.18)	114 (63.33)	
Number of tumors (%)				0.001
Single	182 (32.38)	142 (37.17)	40 (22.22)	
Multiple	380 (67.62)	240 (62.83)	140 (77.78)	
Tumor capsule (%)				<0.001
No	173 (30.78)	66 (17.28)	107 (59.44)	
Yes	389 (69.22)	316 (82.72)	73 (40.56)	
Tumor thrombus (%)				<0.001
No	424 (75.44)	324 (84.82)	100 (55.56)	
Yes	138 (24.56)	58 (15.18)	80 (44.44)	
Bilobar distribution of tumor				<0.001
No	335 (59.61)	254 (66.49)	81 (45.00)	
Yes	227 (40.39)	128 (33.51)	99 (55.00)	
Child-Pugh grade (%)				0.005
A	518 (92.17)	361 (94.50)	157 (87.22)	
B	44 (7.83)	21 (5.50)	23 (12.78)	
ECOG-PS (%)				0.028
0	480 (85.41)	335 (87.70)	145 (80.56)	
1	82 (14.59)	47 (12.30)	35 (19.44)	
BCLC (%)				<0.001
0	30 (5.34)	24 (6.28)	6 (3.33)	
A	193 (34.34)	162 (42.41)	31 (17.22)	
B	206 (36.65)	140 (36.65)	66 (36.67)	
C	133 (23.67)	56 (14.66)	77 (42.78)	
Albumin_Bilirubin_Score (%)				0.232
Level 1	252 (44.84)	164 (42.93)	88 (48.89)	
Level 2	293 (52.14)	207 (54.19)	86 (47.78)	
Level 3	17 (3.02)	11 (2.88)	6 (3.33)	
Age (mean (SD))	55.06 (10.62)	55.64 (10.48)	53.84 (10.84)	0.136
White blood cells (median [IQR])	5.38 [3.92, 6.66]	4.93 [3.68, 6.19]	6.11 [4.87, 7.73]	<0.001
Lymphocytes (median [IQR])	1.23 [0.92, 1.60]	1.32 [0.94, 1.71]	1.12 [0.87, 1.34]	<0.001
Hemoglobin (median [IQR])	133.00 [120.00, 147.00]	132.50 [120.00, 146.00]	134.50 [119.75, 147.00]	0.353
Platelets (median [IQR])	129.00 [86.12, 187.00]	112.50 [76.00, 164.00]	169.00 [121.75, 220.25]	<0.001
Neutrophils (median [IQR])	3.40 [2.32, 4.53]	2.90 [2.07, 3.97]	4.48 [3.39, 5.62]	<0.001

(Continued)

Table 1 (Continued).

Characteristics	ALL N=562	Effective N=382	Resist N=180	p.Overall
NLR (median [IQR])	2.70 [1.89, 3.93]	2.29 [1.69, 3.10]	4.04 [2.95, 5.44]	<0.001
A/G Ratio (median [IQR])	1.25 [1.04, 1.44]	1.25 [1.03, 1.44]	1.24 [1.06, 1.44]	0.76
Albumin (median [IQR])	38.60 [35.00, 42.20]	37.90 [34.60, 42.10]	39.50 [35.98, 42.50]	0.059
ALT (median [IQR])	33.00 [23.00, 53.00]	33.00 [23.00, 52.00]	35.00 [24.00, 55.00]	0.222
Indirect bilirubin (median [IQR])	8.40 [5.62, 12.47]	8.60 [6.00, 12.30]	8.10 [5.07, 12.70]	0.536
Globulin (median [IQR])	31.70 [28.20, 35.27]	31.60 [27.92, 35.48]	31.70 [28.98, 34.70]	0.922
AST (median [IQR])	44.00 [31.00, 69.00]	41.00 [30.25, 65.00]	48.50 [35.00, 79.25]	0.005
Direct bilirubin (median [IQR])	5.80 [3.90, 9.17]	5.85 [4.00, 9.45]	5.75 [3.70, 8.72]	0.458
Total bilirubin (median [IQR])	14.90 [9.90, 21.78]	15.35 [10.10, 22.17]	13.65 [9.28, 21.05]	0.391
Total protein (mean (SD))	70.50 [66.00, 75.00]	70.45 [65.80, 75.10]	70.70 [67.45, 74.32]	0.321
GT (median [IQR])	104.00 [53.25, 202.75]	88.00 [49.00, 173.50]	142.00 [70.00, 223.00]	<0.001
INR (median [IQR])	1.10 [1.03, 1.21]	1.11 [1.04, 1.22]	1.09 [1.03, 1.17]	0.08
APTT (median [IQR])	38.30 [35.60, 41.30]	38.60 [35.90, 41.27]	37.90 [35.38, 41.32]	0.376
TT (median [IQR])	17.50 [16.70, 18.60]	17.60 [16.70, 18.70]	17.35 [16.70, 18.30]	0.055
Prothrombin activity (mean (SD))	85.00 [75.00, 95.00]	85.00 [74.00, 95.00]	86.50 [78.00, 98.00]	0.154
PT (median [IQR])	14.10 [13.40, 15.10]	14.20 [13.40, 15.28]	14.00 [13.30, 14.80]	0.058
Fibrinogen (median [IQR])	3.09 [2.50, 3.99]	2.83 [2.35, 3.50]	3.66 [3.04, 4.68]	<0.001
Primary tumor size (mm)(median [IQR])	82.00 [50.25, 102.00]	75.00 [53.00, 86.75]	97.50 [76.00, 119.00]	<0.001

Abbreviations: AFP, alpha-fetoprotein; NLR, neutrophil-to-lymphocyte ratio; A/G Ratio, albumin-to-globulin ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GT, gamma-glutamyl transferase; INR, international normalized ratio; APTT, activated partial thromboplastin time; TT, thrombin time; PT, prothrombin time; ECOG-PS, eastern cooperative oncology group performance status; BCLC, barcelona clinic liver cancer staging system; SD, standard deviation; IQR, interquartile range.

Table 2 Performance of Each Model in the Validation Set

Model	AUC	95% CI Lower	95% CI Upper	Accuracy	Precision	Sensitivity	Specificity	F1 Score
Logistic	0.891	0.844	0.938	0.821	0.722	0.722	0.868	0.722
Decision Tree	0.808	0.749	0.868	0.750	0.588	0.741	0.754	0.656
Random Forest	0.894	0.847	0.940	0.810	0.704	0.704	0.860	0.704
XGBoost	0.898	0.853	0.944	0.833	0.741	0.741	0.877	0.741
LightGBM	0.885	0.837	0.933	0.786	0.661	0.685	0.833	0.673
SVM	0.893	0.846	0.940	0.815	0.709	0.722	0.860	0.716
ANN	0.896	0.849	0.942	0.827	0.719	0.759	0.860	0.739

Abbreviations: AUC, area under the ROC curve; SVM, support vector machine; ANN, artificial neural network.

(95% confidence interval: 0.682–0.751). Similarly, in the validation set, the XGBoost model attained an AUC of 0.898 (95% confidence interval: 0.853–0.944), whereas the up-to-seven score had an AUC of 0.675 (95% confidence interval: 0.642–0.734). These results clearly indicate that our XGBoost model outperforms the traditional up-to-seven score in predicting the occurrence of TACE resistance in hepatocellular carcinoma patients. This performance advantage may be attributed to the ability of the XGBoost algorithm to capture complex non-linear relationships and interactions among multiple features, which the up-to-seven score, based on a limited number of clinical parameters, may fail to fully consider.

Results of SHAP Analysis

Figure 5 is a SHAP value bar chart, which is used to quantify and visualize the contribution of each feature to the model prediction. The SHAP value represents the influence degree of each feature on the final prediction result under different feature combinations, and a higher SHAP value indicates a greater influence of the feature on the model prediction.

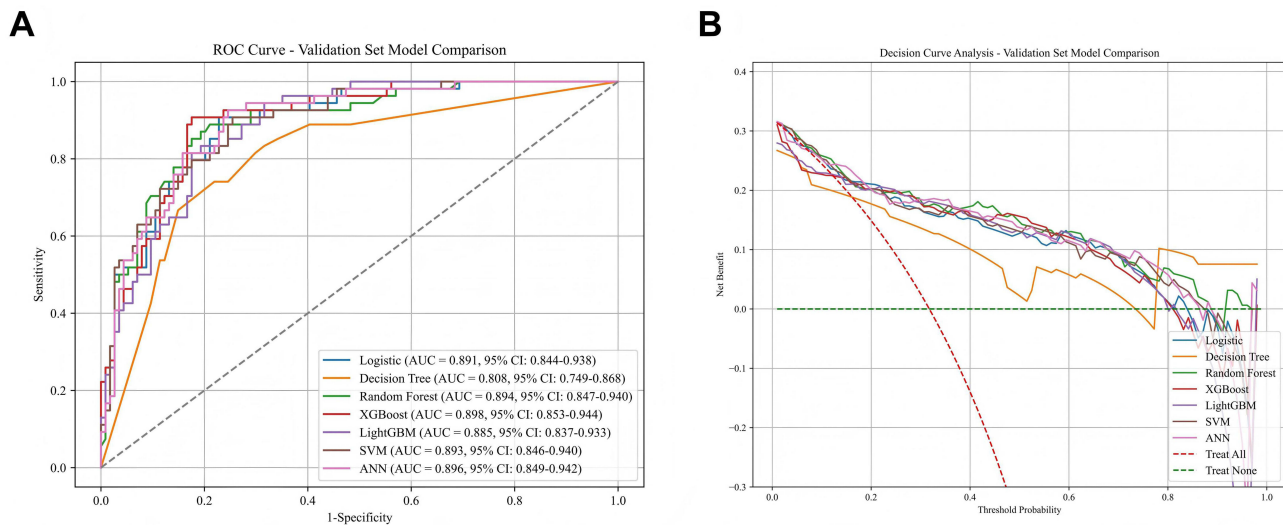


Figure 2 Comparative Performance Analysis of Machine Learning Models (A) ROC curves of each mo (B) DCA curves of each model.

Figure 6 is a SHAP beeswarm plot, which shows the distribution of SHAP values of each feature and intuitively presents the influence of each feature on the model prediction result in different value ranges. The SHAP values are color-coded (blue indicates lower feature values, and purple indicates higher feature values) to reveal the relationship between feature values and prediction results. The chart is sorted in descending order of feature influence, showing that NLR is the strongest predictive factor (SHAP value range: -2.5 to 3.0), and its high value is significantly positively associated with resistance risk; the absence of tumor capsule (purple dots concentrated in the positive interval) and AFP \geq 400 ng/mL (sharp increase in contribution of high values) further verify their core predictive status; fibrinogen and primary tumor size show a non-linear dose-response effect.

Figure 7 is a SHAP dependence plot, presenting the contribution of core clinical features to TACE resistance prediction. Each subplot corresponds to a predictive factor, the X-axis is the range of feature values, the Y-axis is the SHAP value (reflecting the positive/negative contribution of the feature to the prediction result), and the color gradient indicates the size of the feature value (yellow→high, purple→low). Key findings include: when the tumor capsule is absent (value=1), the SHAP value is significantly higher than 0, indicating that it is a strong positive indicator for resistance prediction; when AFP \geq 400 ng/mL (yellow area), the SHAP value increases significantly, verifying its positive predictive value; indicators such as fibrinogen and platelet count show a non-linear positive correlation trend between feature values and SHAP values. The orange dashed line (SHAP=0) serves as the contribution threshold, and the red trend line reveals the XGBoost model’s ability to capture complex clinical relationships.

Figure 8 is an individual-level SHAP explanation plot: A and C correspond to TACE-responsive patients. Low NLR, intact tumor capsule, and normal AFP are the main factors leading to reduced risk, which significantly reduce the

Table 3 Performance of Each Model in the Training Set

Model	AUC	AUC 95% CI Lower	AUC 95% CI Upper	Accuracy	Precision	Sensitivity	Specificity	F1 Score
Logistic	0.926	0.9	0.953	0.863	0.807	0.76	0.913	0.783
Decision Tree	0.896	0.865	0.927	0.834	0.725	0.785	0.857	0.754
Random Forest	0.94	0.916	0.964	0.941	0.909	0.909	0.956	0.909
XGBoost	0.942	0.919	0.966	0.933	0.9	0.893	0.952	0.896
LightGBM	0.943	0.92	0.967	0.941	0.909	0.909	0.956	0.909
SVM	0.881	0.849	0.914	0.836	0.768	0.711	0.897	0.738
ANN	0.901	0.871	0.931	0.858	0.793	0.76	0.905	0.776

Abbreviations: AUC, area under the ROC curve; SVM, support vector machine; ANN, artificial neural network.

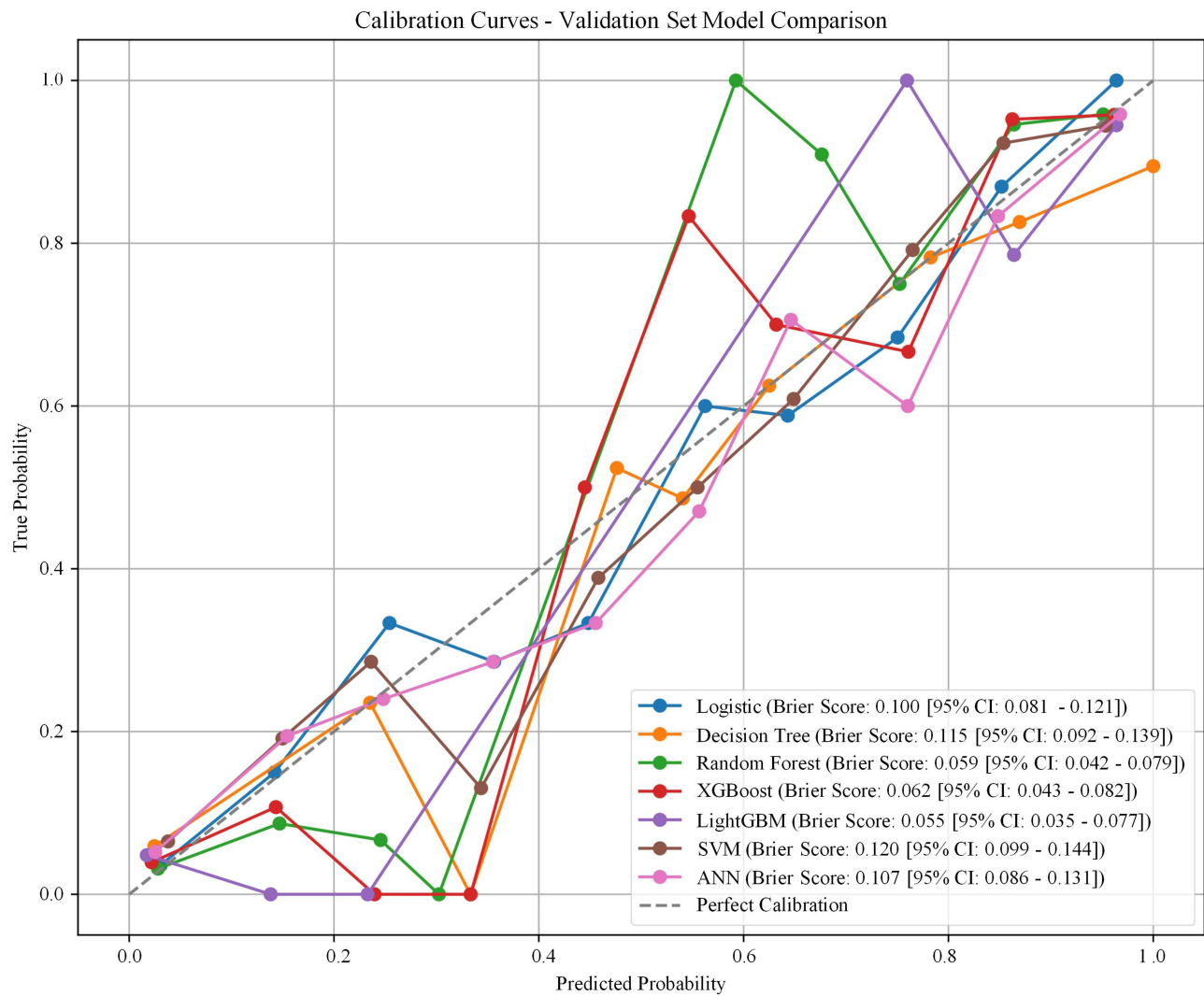


Figure 3 Calibration Curves Comparison for Different Machine Learning Models on Validation Set.

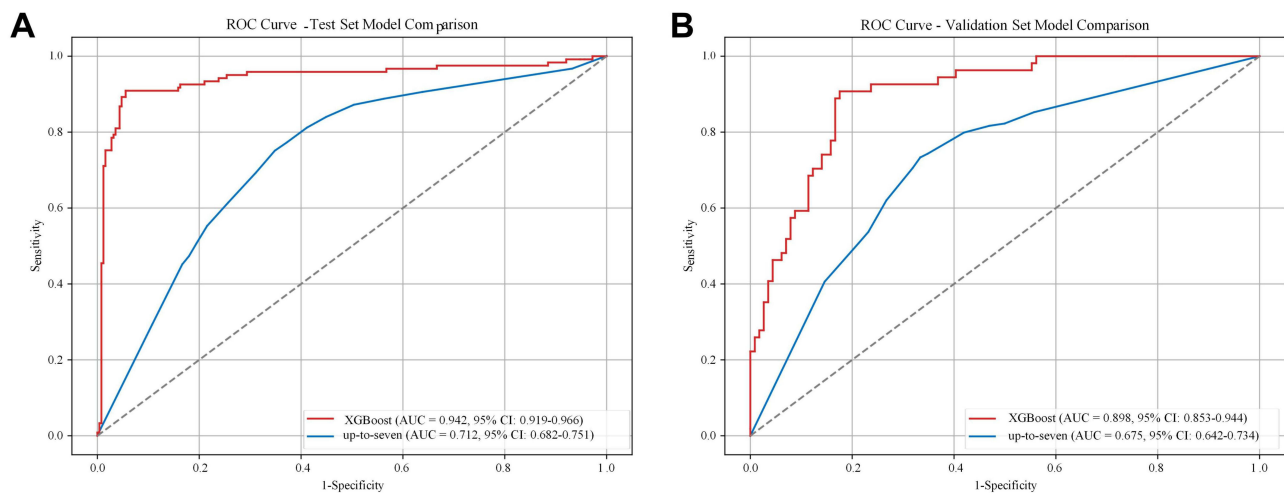


Figure 4 ROC Curves Comparing ML Model and Existing Predictive Scores. **(A)** Comparison in the test set; **(B)** Comparison in the validation set.

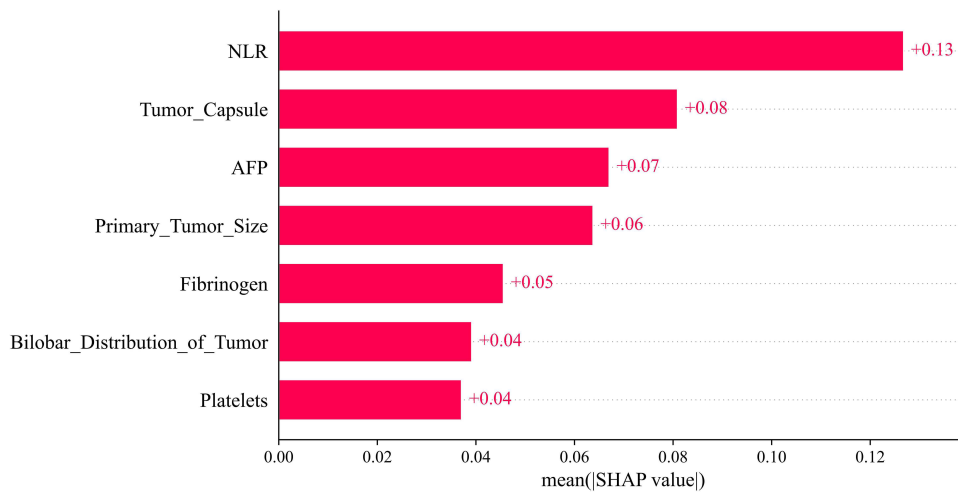


Figure 5 Feature Importance Analysis: Mean SHAP Values for Key Clinical Variables.
Abbreviations: AFP, alpha-fetoprotein; NLR, neutrophil-to-lymphocyte ratio.

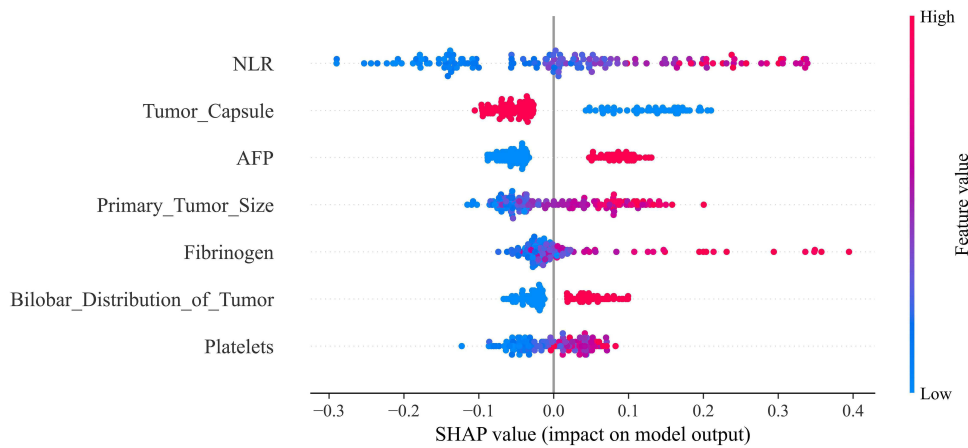


Figure 6 SHAP Value Distribution for Clinical Features: Impact on Model Output.
Abbreviations: AFP, alpha-fetoprotein; NLR, neutrophil-to-lymphocyte ratio.

predicted value of the model. The final predicted value is 0.02, indicating that the patient has a low risk of TACE resistance; at the same time, slightly larger tumor size and higher fibrinogen level are risk-increasing factors, but their contributions are relatively small. B and D correspond to TACE-resistant patients. The significantly increased NLR level leads to a substantial increase in the predicted value of the model. The final predicted value is 0.997, indicating that the patient has an extremely high risk of TACE resistance.

Discussion

Hepatocellular carcinoma is a serious global health challenge, especially in developing countries.¹⁰ Despite continuous advances in medical technology, the cure rate of HCC remains low, so more research is urgently needed to improve treatment outcomes.¹¹

Machine learning has shown significant advantages in the field of disease prediction.^{12,13} In this study, we used 7 different machine learning models to predict TACE-resistant patients. The experimental results showed that the XGBoost model performed excellently, especially in processing multi-feature and high-complexity data, and its generalization ability and predictive performance were far superior to traditional models.¹⁴ The performance of XGBoost was significantly better than other models, indicating that this model can not only capture the contribution of individual biomarkers to the risk of TACE

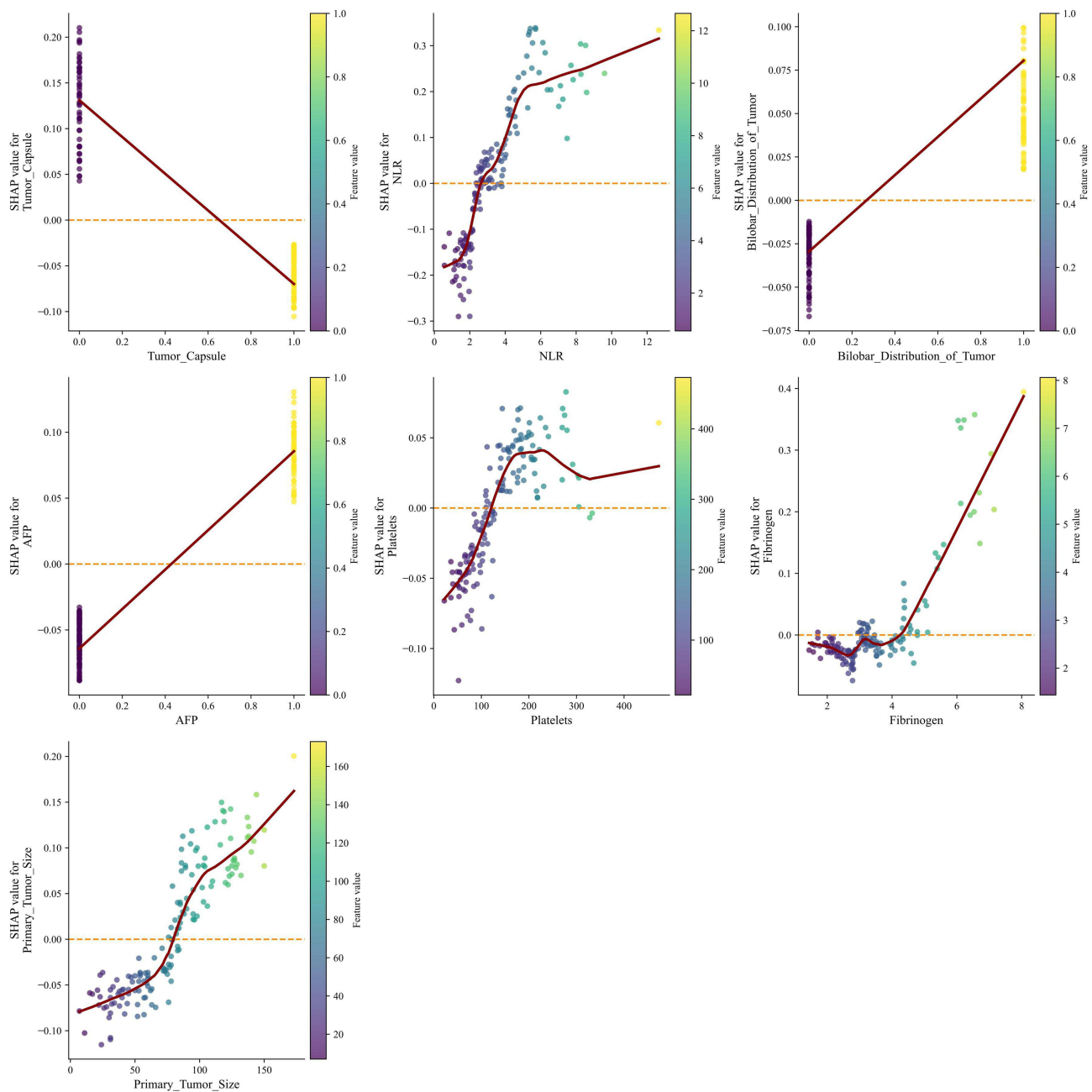


Figure 7 SHAP Dependency Plots: Relationship Between Feature Values and Their Impact on Model Output.

resistance but also identify the interactions between different features, thereby more accurately predicting the occurrence of TACE resistance in HCC patients. As an efficient gradient boosting algorithm, XGBoost is good at processing non-linear and high-dimensional sparse data. Its built-in regularization mechanism can effectively control model complexity and reduce the risk of overfitting; while the parallel computing function can accelerate the training process, ensuring high computational efficiency even for large datasets.¹⁵ In addition, the automated characteristics and scalability of this model give it significant advantages in clinical applications, being able to effectively process incomplete or inconsistent data, thereby reducing prediction bias caused by data missing or heterogeneity.

The 7 independent risk factors identified in this study (tumor capsule integrity, bilateral liver lobe involvement of tumor, platelet count, NLR, fibrinogen, AFP, and primary tumor size) are all closely related to TACE efficacy, and their mechanisms of action can be explained by tumor biological characteristics and microenvironment regulation.

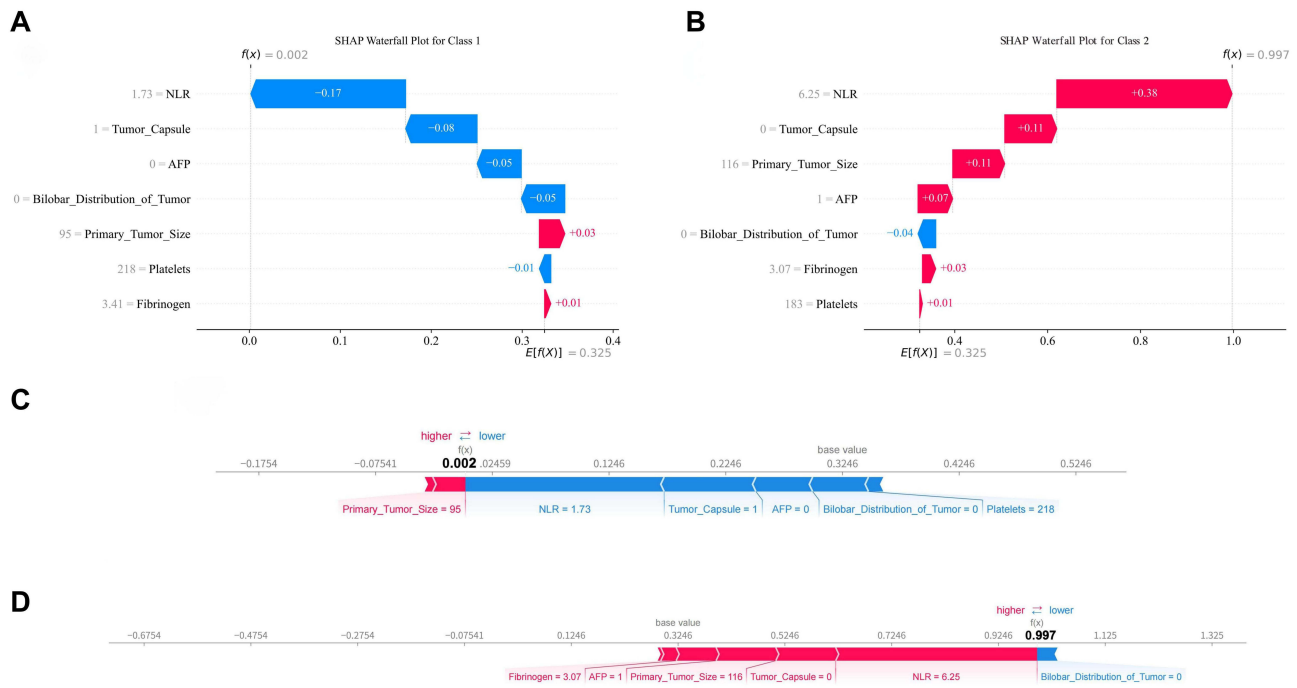


Figure 8 Individual-level SHAP explanation plot: (A and C) are TACE-responsive patients; (B and D) are TACE-resistant patients. **Note:** the values 0.002 and 0.997 are presented in bold. These values represent the final predicted probabilities output by the machine learning model for patient.

The integrity of the tumor capsule is a key morphological feature affecting TACE efficacy. An intact tumor capsule can limit the invasive growth of tumors, reduce the disordered proliferation of tumor blood vessels, and make chemotherapeutic drugs more likely to accumulate in the tumor during TACE treatment; while the absence of the capsule often indicates active proliferation of tumor cells and blurred boundaries, which may lead to uneven drug penetration and increase the risk of resistance. Bilateral liver lobe involvement of the tumor reflects the extensiveness of tumor burden. Bilateral lobe involvement means a more complex blood supply network and higher tumor heterogeneity, which may reduce the control efficiency of a single TACE treatment on all lesions. The positive correlation between primary tumor size and resistance risk is also consistent with clinical cognition: larger tumors are often accompanied by central necrosis, insufficient blood supply, or enrichment of drug-resistant clones, leading to difficulty in chemotherapeutic drugs exerting their effects fully.¹⁶

As a classic tumor marker for HCC, the association between elevated AFP and TACE resistance has a clear biological basis. A high AFP level usually indicates poor differentiation and high proliferative activity of tumor cells, and may be accompanied by an epithelial-mesenchymal transition phenotype. Such cells have low sensitivity to chemotherapeutic drugs. In addition, AFP can construct an immunosuppressive microenvironment by inhibiting the immune system (such as reducing the maturation of dendritic cells), indirectly weakening the anti-tumor immune response after TACE treatment.

Among hematological indicators, NLR was identified as the most significant predictive factor by SHAP analysis, which is consistent with the role of NLR as a “barometer” of the tumor immune microenvironment in previous studies.¹⁷ A high NLR reflects an imbalance between enhanced neutrophil-mediated pro-inflammatory responses and weakened lymphocyte-mediated anti-tumor immunity: neutrophils promote chronic liver inflammation and tumor angiogenesis by releasing Neutrophil Extracellular Traps (NETs),¹⁸ while lymphocyte reduction directly impairs the killing function of tumor-infiltrating T cells. Both together constitute an “immunosuppressive microenvironment”, which may reduce the sensitivity of tumor cells to chemotherapeutic drugs after TACE treatment.¹⁹ Increased platelet count may promote tumor stromal fibrosis and abnormal vascular remodeling by releasing cytokines such as Platelet-Derived Growth Factor and Transforming Growth Factor- β ,^{20,21} thereby inhibiting the effective penetration of chemotherapeutic drugs into the tumor. Fibrinogen activates the coagulation system to form a dense fibrin network around the tumor, which not only physically blocks drug diffusion but also recruits immunosuppressive cells such as macrophages, further exacerbating the

occurrence of TACE resistance.^{22,23} In addition, platelets and fibrinogen can also exert a synergistic effect, interfering with the clearance of intravascular tumor cells by NK cells, making it easier for tumor cells to survive and spread in the blood vessels, and ultimately enhancing the metastatic potential of the tumor.²⁴

It is worth noting that the conclusion of this study that increased platelet count is associated with TACE resistance is different from the study by Christopher Schrecker et al, who reported thrombocytopenia as a risk factor for HCC recurrence.²⁵ This apparent contradiction may stem from the bidirectional regulatory effect of platelet count on tumor biological behavior: high platelet count promotes resistance through pro-fibrotic and immune escape mechanisms, while low platelet count may be related to tumor nutritional competition or treatment-related myelosuppression. The average tumor diameter of patients in the resistant group in this study was 97.5 mm, which was significantly larger than the 50 mm reported in Schrecker's study,²² suggesting that differences in tumor burden may affect the direction of the association between platelets and prognosis. Existing evidence shows that in most solid tumors, increased platelet count is positively correlated with poor prognosis and increased thrombus risk,^{26,27} and its mechanism of action may show tissue specificity due to differences in the tumor microenvironment and treatment methods.

This study integrated 7 risk factors through the XGBoost model for the first time, and its predictive efficacy was significantly better than traditional statistical methods and other machine learning models. This advantage stems from the ability of XGBoost to process high-dimensional and non-linear data: the model can not only quantify the independent role of individual factors (such as the high weight of NLR) but also capture the interaction between features (such as the synergistic enhancement of "large tumor + high AFP" on resistance risk), which is highly consistent with the clinical reality that TACE efficacy is jointly regulated by multiple factors.

Compared with previous prediction studies based on a single indicator, the advantages of this model are reflected in: ① Integrating multi-dimensional information of tumor morphology, imaging, and hematology, improving the comprehensiveness of prediction; ② Automated feature selection and generalization ability enable it to be directly applied to clinical data input, providing convenience for real-time risk assessment; ③ The regularization mechanism effectively reduces the risk of overfitting, ensuring stability in different clinical scenarios. These characteristics make the XGBoost model expected to become an auxiliary tool for clinical decision-making, such as combining targeted drugs or immunotherapy in advance for high-risk patients to reduce the incidence of resistance.

This study has certain limitations: first, as a single-center retrospective study, there may be selection bias in the sample, and there is a lack of an external validation cohort, so the generalizability of the model needs further verification; second, molecular biological indicators (such as gene mutations and epigenetic modifications) were not included, which may miss potential predictive factors; in addition, the dynamic predictive ability of the model for TACE resistance (such as efficacy monitoring during treatment) has not been explored. Future research can be advanced from three aspects: ① Conduct multi-center, prospective studies to expand the sample size and perform external validation to improve the generalizability of the model; ② Integrate genomics and proteomics data to deeply explore the molecular mechanism of TACE resistance and enrich the biological connotation of the prediction model; ③ Develop a convenient mobile-based prediction tool to promote the transformation of the model to clinical practice and realize the closed-loop management of "individualized risk stratification - dynamic adjustment of treatment plans".

In conclusion, this study revealed the key risk factors of TACE resistance and their predictive value through machine learning technology, providing a new idea for precision treatment of HCC. With further optimization and validation of the model, it is expected to become an important clinical tool for improving TACE efficacy and prolonging patient survival.

Ethical Considerations

This study strictly followed the ethical guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of Tongji Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology (Ethics Approval No.: TJ-IRB-2024-032). Due to the retrospective design of the study, written informed consent from patients was exempted in accordance with the Measures for the Ethical Review of Biomedical Research Involving Humans. We pledge to maintain strict confidentiality of patient data and ensure that the data will be used solely for the purposes of this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research did not receive any external funding support.

Disclosure

The authors report no conflicts of interest in this work.

References

- Li C, He W-Q. Comparison of primary liver cancer mortality estimates from World Health Organization, global burden disease and global cancer observatory. *Liver Int.* 2022;42(10):2299–2316. doi:10.1111/liv.15357
- Wang CY, Li S. Clinical characteristics and prognosis of 2887 patients with hepatocellular carcinoma: a single center 14 years experience from China. *Medicine.* 2019;98(4):e14070. doi:10.1097/MD.00000000000014070
- Reig M, Bruix J. Reply to: “Correspondence on the <BCLC>”. *J Hepatol.* 2022;76(5):1240–1241. doi:10.1016/j.jhep.2022.02.026
- Wang L, Cao J, Liu Z, et al. Enhanced interactions within microenvironment accelerates dismal prognosis in HBV-related HCC after TACE. *Hepatol Commun.* 2024;8(10):e0548. doi:10.1097/H9C.0000000000000548
- Zhang K, Jiao B, Sun J, et al. Predicting high-risk factors for postoperative inadequate analgesia and adverse reactions in cesarean delivery surgery: a prospective study. *Int J Surg.* 2025;111(6):3859. doi:10.1097/JS9.0000000000002354
- Ma Y, Luo M, Guan G, Liu X, Cui X, Luo F. An explainable predictive machine learning model of gangrenous cholecystitis based on clinical data: a retrospective single center study. *World J Emerg Surg.* 2025;20(1):1. doi:10.1186/s13017-024-00571-6
- Dong W, Jiang H, Li Y, et al. Interpretable machine learning analysis of immunoinflammatory biomarkers for predicting CHD among NAFLD patients. *Cardiovasc Diabetol.* 2025;24(1):263. doi:10.1186/s12933-025-02818-1
- Clinical Guidelines Committee of the Interventional Physicians Branch of the Chinese Medical Doctor Association. Chinese clinical practice guidelines for transarterial chemoembolization of hepatocellular carcinoma (2023 edition). *National Medical Journal of China.* 2023;103(34):2674–2694. doi:10.3760/cma.j.cn112137-20230630-01114
- Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: a data analyst’s perspective. *Multivariate Behav Res.* 1998;33(4):545–571. doi:10.1207/s15327906mbr3304_5
- Llovet JM, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers.* 2021;7(1):6. doi:10.1038/s41572-020-00240-3
- Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. *Hepatology.* 2018;68(2):723–750. doi:10.1002/hep.29913
- Zhang Y, Tong S, Yang J, et al. Explainable machine learning model for predicting the transarterial chemoembolization response and subtypes of hepatocellular carcinoma patients. *BMC Gastroenterol.* 2025;25(1):503. doi:10.1186/s12876-025-04105-5
- Ahn JC, Connell A, Simonetto DA, Hughes C, Shah VH. Application of artificial intelligence for the diagnosis and treatment of liver diseases. *Hepatology.* 2021;73(6):2546–2563. doi:10.1002/hep.31603
- Chen R, Zhang S, Li J, et al. A study on predicting the length of hospital stay for Chinese patients with ischemic stroke based on the XGBoost algorithm. *BMC Med Inform Decis Mak.* 2023;23(1):49. doi:10.1186/s12911-023-02140-4
- Ogunleye A, Wang QG. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinform.* 2020;17(6):2131–2140. doi:10.1109/TCBB.2019.2911071
- Moustafa AS, Abdel Aal AK, Ertel N, Saad N, DuBay D, Saddekni S. Chemoembolization of hepatocellular carcinoma with extrahepatic collateral blood supply: anatomic and technical considerations. *Radio Graphics.* 2017;37(3):963–977. doi:10.1148/rg.2017160122
- Asiri M, BinAbdu A, Al-Ammar A, et al. Survival outcomes and prognostic factors of systemic therapy for advanced hepatocellular carcinoma: a multidisciplinary clinic experience from Saudi Arabia. *JHC.* 2025;12:1623–1632. doi:10.2147/JHC.S525984
- Margetts J, Ogle LF, Chan SL, et al. Neutrophils: driving progression and poor prognosis in hepatocellular carcinoma? *Br J Cancer.* 2018;118(2):248–257. doi:10.1038/bjc.2017.386
- Wang C, Wang M, Zhang X, et al. The neutrophil-to-lymphocyte ratio is a predictive factor for the survival of patients with hepatocellular carcinoma undergoing transarterial chemoembolization. *Ann Transl Med.* 2020;8(8):541. doi:10.21037/atm.2020.02.113
- Lucatelli P, Burrell M, Guiu B, de Rubeis G, van Delden O, Helmberger T. CIRSE standards of practice on hepatic transarterial chemoembolization. *Cardiovasc Intervent Radiol.* 2021;44(12):1851–1867. doi:10.1007/s00270-021-02968-1
- Razavi AS, Mohtashami M, Razi S, Rezaei N. TGF- β signaling and the interaction between platelets and T-cells in tumor microenvironment: foes or friends? *Cytokine.* 2022;150:155772. doi:10.1016/j.cyto.2021.155772
- Kluz N, Grabowska H, Chmiel P, et al. Platelets in hepatocellular carcinoma—from pathogenesis to targeted therapy. *Cancers.* 2025;17(14):2391. doi:10.3390/cancers17142391
- Hua N, Chen A, Yang C, et al. The correlation of fibrinogen-like protein-1 expression with the progression and prognosis of hepatocellular carcinoma. *Mol Biol Rep.* 2022;49(8):7911–7919. doi:10.1007/s11033-022-07624-6
- Palumbo JS, Talmage KE, Massari JV, et al. Platelets and fibrin(ogen) increase metastatic potential by impeding natural killer cell-mediated elimination of tumor cells. *Blood.* 2005;105(1):178–185. doi:10.1182/blood-2004-06-2272

25. Schrecker C, Waidmann O, El Youzouri H, et al. Low platelet count predicts reduced survival in potentially resectable hepatocellular carcinoma. *Current Oncol.* 2022;29(3):1475–1487. doi:10.3390/curroncol29030124
26. Lu Z, Huang Y, Huang J, et al. High platelet count is a potential prognostic factor of the early recurrence of hepatocellular carcinoma in the presence of circulating tumor cells. *J Hepatocell Carcinoma.* 2023;10:57–68. doi:10.2147/JHC.S398591
27. Belluco C, Forlin M, Delrio P, et al. Elevated platelet count is a negative predictive and prognostic marker in locally advanced rectal cancer undergoing neoadjuvant chemoradiation: a retrospective multi-institutional study on 965 patients. *BMC Cancer.* 2018;18(1):1094. doi:10.1186/s12885-018-5022-1

Journal of Hepatocellular Carcinoma

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal>

Dovepress
Taylor & Francis Group