Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Predicting Resistance and Survival of HCC Patients Post-HAIC: Based on Shapley Additive exPlanations and Machine Learning

Fan Yao[1,2,*], Jianliang Miao[3,*], Bing Quan[1,2], Jinghuan Li[1,2], Bei Tang[1,2], Shenxin Lu[1,2], Xin Yin[1,2]

[1]Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai, People's Republic of China; [2]National Clinical Research Center for Interventional Medicine, Shanghai, People's Republic of China; [3]First Affiliated Hospital of Dalian Medical University, Dalian Medical University, Dalian, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xin Yin, Liver Cancer Institute, Zhongshan Hospital, Fudan University, 136 Yi Xue Yuan Road, Shanghai, People's Republic of China, Email yin.xin@zs-hospital.sh.cn

Purpose: To establish prediction models using Shapley Additive exPlanations (SHAP) and multiple machine learning (ML) algorithms to identify clinical features influencing hepatic arterial infusion chemotherapy (HAIC) resistance and survival in patients with hepatocellular carcinoma (HCC).

Patients and Methods: We recruited 286 patients with unresectable HCC who underwent HAIC. Patients were divided into training and validation datasets (7:3 ratio). eXtreme Gradient Boosting (XGBoost) was used to build the preliminary resistance prediction model. The SHAP values explained the importance of the clinical features. Recursive Feature Elimination with Cross-Validation (RFECV) was used to select the optimum number of features. Seven ML methods were used to construct further resistance prediction models, and ten ML algorithms were employed to establish the survival prognosis models.

Results: The areas under the curve (AUC) of the XGBoost model were 1.000 and 0.812 for the training and validation groups, respectively. SHAP identified 27 of the 38 clinical features affecting resistance, with pre-HAIC treatment being the main factor. RFECV showed the best model performance with six features (pre-HAIC treatment, tumor size, HBV DNA, alkaline phosphatase (AKP), prothrombin time (PT), and portal vein tumor thrombosis (PVTT)). Random Forest had the best performance among the seven ML algorithms (AUC=0.935 for training, AUC=0.876 for validation). The combination of Stepcox [forward] and Gradient Boosting Machine was the best for predicting survival (AUC=0.98 in training, AUC=0.83 in validation). Based on the above clinical characteristics, patients were categorized into high-risk and low-risk groups based on the median risk score, and it was found that these characteristics also performed well in the prognostic model for predicting the survival of patients with HCC.

Conclusion: Pre-HAIC treatment, tumor size, HBV DNA, AKP, PT, and PVTT are effective predictors of post-HAIC resistance and survival in patients with unresectable advanced HCC.

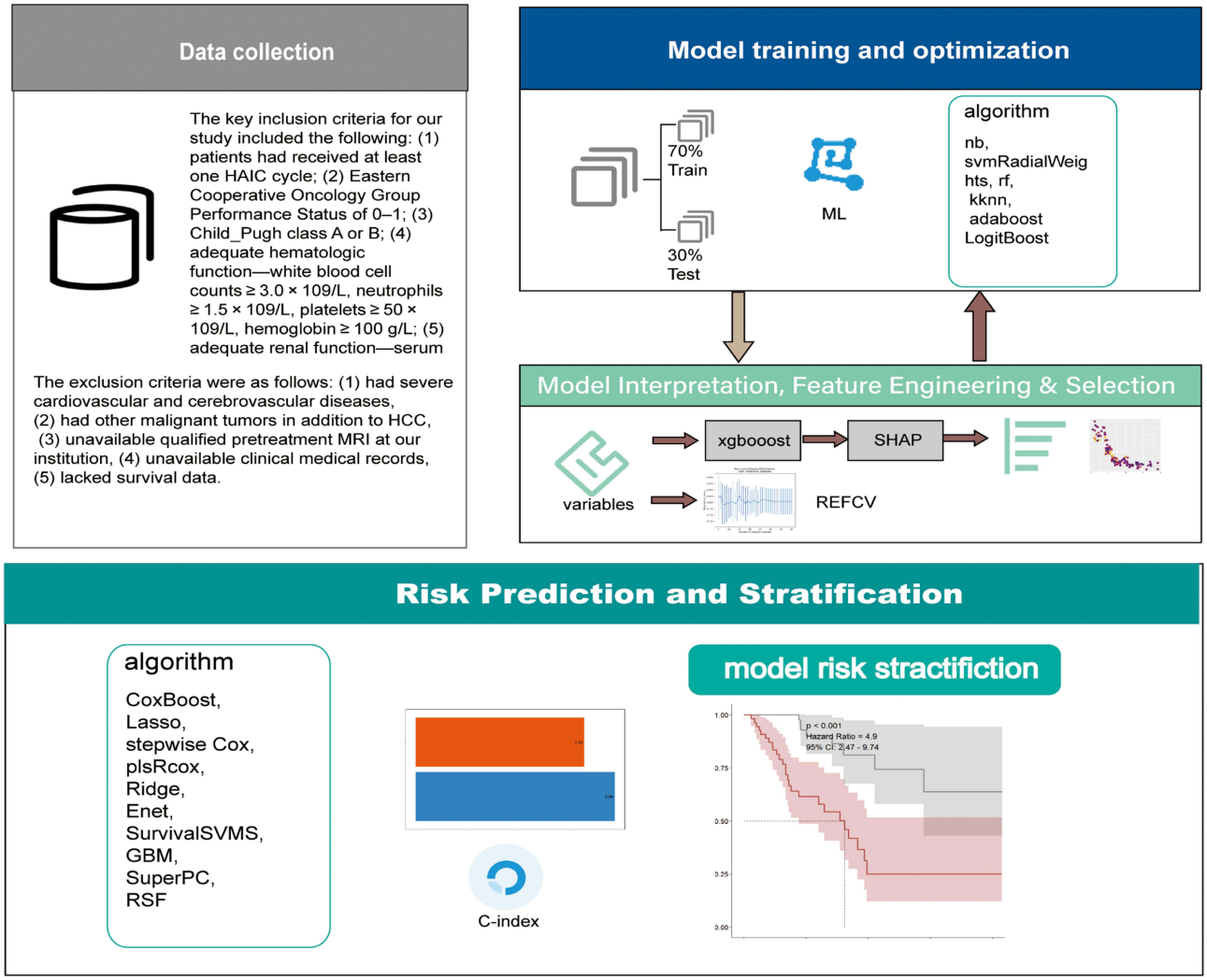
Keywords: interpretable AI, treatment response prediction, prognostic modeling, hepatocellular carcinoma outcomes

Introduction

Hepatocellular carcinoma (HCC), the most prevalent form of primary liver cancer, is the third leading cause of cancer-related death globally. It is estimated that approximately 850,000 new cases emerge annually.[1] While liver transplantation is theoretically an optimal treatment approach, its implementation is hindered by several factors, including the shortage of liver donors and the physical condition of the patients.[2] Consequently, liver resection has become the standard treatment for HCC. However, a substantial number of patients do not qualify for surgery due to conditions such as multiple tumors or liver dysfunction.[3] For those with unresectable or advanced HCC, multitargeted tyrosine kinase inhibitors (TKIs) such as sorafenib and lenvatinib have been regarded as the standard first-line systematic treatment.[4] In recent years, immune checkpoint inhibitors (ICIs) have revolutionized HCC treatment. The combination of atezolizumab

Received: 5 March 2025
Accepted: 17 May 2025
Published: 31 May 2025

Graphical Abstract



(anti-programmed death-ligand 1, anti-PDL1) and bevacizumab (anti-vascular endothelial growth factor) has demonstrated superior efficacy in improving overall survival (OS) in patients with HCC, leading to its approval by the Food and Drug Administration (FDA).[5,6] Additionally, the combination of sintilimab (anti-PD1) and IBI305 (a bevacizumab biosimilar) has also shown remarkable therapeutic effects in clinical trials.[7] Despite significant progress in ICIs therapy across multiple cancers and efforts to enhance the clinical benefits of HCC immunotherapy, the immune response remains limited, and the majority of patients still face drug resistance and disease progression.[8,9]

Hepatic arterial infusion chemotherapy (HAIC) serves as an alternative treatment for intermediate and advanced HCC, utilizing catheter technology to deliver chemotherapeutic drugs directly to tumors via the hepatic artery.[9,10] Clinical trials have shown that the combination of HAIC and FOLFOX (fluorouracil, leucovorin, and oxaliplatin) significantly enhanced survival rates in patients with unresectable HCC when compared to transarterial chemoembolization (TACE).[11,12] Furthermore, research conducted by Zhao et al revealed that OS in the HAIC-FOLFOX group was nearly twice as favorable as that in the sorafenib group, with a median OS of 10.8 months versus 5.7 months (hazard ratio [HR] 0.343; 95% confidence interval [CI], 0.219 to 0.538; $P < 0.001$).[13] In addition to these findings, HAIC has also achieved promising results when used in conjunction with targeted immunotherapy.[14,15] In summary, HAIC is an

effective therapeutic option for patients with advanced unresectable HCC. However, owing to disease progression such as metastasis and tumor enlargement, there remains a certain probability of resistance in patients following HAIC treatment.

In the past decade, machine learning (ML) has experienced a significant surge in the medical field, particularly in oncology, where it has been utilized to discern data patterns and structures for disease diagnosis, prognosis, and treatment recommendations.[16] Chen et al employed ML methods (LASSO, random forest (RF), SVM-RF, and eXtreme Gradient Boosting (XGBoost)) to differentiate various stemness classifications of HCC patients, thereby predicting prognosis and aiding in guiding clinical immunotherapy and targeted therapy strategies.[17] XGBoost, a novel ML algorithm, is an ensemble learning algorithm based on gradient lifting decision trees. It outperforms other algorithms by continuously iterating and optimizing the loss function to progressively reduce the residual error.[18] Other ML methods, such as support vector classification (SVC), multilayer perceptron (MLP), linear discriminant analysis (LDA), have also been applied in breast cancer, esophageal cancer and gastric cancer.[19,20] However, the "black box" nature of ML often makes predictions ambiguous, as the importance of the features included is non-transparent, limiting the interpretability of ML methods.[21] Consequently, SHapley Additive exPlanations (SHAP) has been introduced to interpret the results of ML models. SHAP elucidates the importance of each feature in an ML model through visualization and quantification, thereby enhance the credibility of the model.[22] The core concept of SHAP is to treat model predictions as a cooperative game, with each feature acting as a participant. The SHAP can quantify the significance of each feature by calculating its marginal contribution to the final prediction.[23]

Most current clinical trials have focused on the response to HAIC in combination with other treatments, but there is a paucity of reports predicting the probability of resistance and survival of patients with unresectable HCC following HAIC surgery. In this study, we applied the SHAP values to interpret the XGBoost model and predict the efficacy and survival prognosis of patients with HCC following HAIC using multiple ML algorithms.

Materials and Methods

Patients

This retrospective study was conducted at Zhongshan Hospital of Fudan University between May 2019 and March 2022, enrolling patients with unresectable HCC who had undergone HAIC treatment. The key inclusion criteria were as follows: 1) all patients had received at least one cycle of HAIC; 2) Child-Pugh grade A or B; 3) Eastern Cooperative Oncology Group Performance Status of 0–1; 4) adequate renal function, defined as serum creatinine ≤ 2 mg/dL; and 5) normal blood routine, including white blood cell count $\geq 3.0 \times 10^9$/L, platelet count $\geq 50 \times 10^9$/L, and hemoglobin level ≥ 100 g/L. Additionally, patients were excluded if they 1) had severe cardiovascular and cerebrovascular disease, 2) had other malignancies besides HCC, 3) had unavailable clinical medical records. For construction of survival models, those who lacked survival data were excluded. Patients who were followed-up for less than 3 months or who had visited other hospitals during follow-up were excluded. This study was conducted in accordance with the institutional ethics committee of Zhongshan Hospital of Fudan University and the Helsinki Declaration (ethics review number KY2022078).

Data Collection

We collected comprehensive clinical data including age, sex, pre-HAIC treatment, perfusion chemotherapy regimen, tumor number, maximum tumor size, thrombus presence, extrahepatic metastasis, tumor stage, liver function, and other relevant parameters. For a valid and prognostic assessment of all patients, we used disease progression (resistance/ineffective) and OS as key outcome indicators. In this study, resistance to HAIC therapy was defined by disease progression (PD) following treatment, assessed through the following criteria: 1) RECIST criteria: Confirmed PD on two consecutive imaging assessments, excluding confounding factors (eg, post-procedural inflammation or necrosis). PD was defined as a $\geq$ 20% increase in the diameters of target lesion or the appearance of new lesions; 2) Serological progression: In patients with elevated baseline alpha-fetoprotein (AFP), a persistent rise in AFP levels after 2–3 treatment cycles; 3) Clinical deterioration: An increase in Child-Pugh score or development of jaundice, ascites, or hepatic encephalopathy. Resistance was considered if any one of the above criteria was met.

Application of XGBoost, Logistic Regression and SHAP

We initially divided the data into a training set and a validation set at a 7:3 ratio, integrating all clinical baseline features into the XGBoost model for predictive model training. Subsequently, we employed SHAP to interpret the XGBoost-trained model and leveraged it to elucidate the importance of each feature within the model and its impact on clinical outcomes. SHAP values provide an approach for interpreting predictive outcomes by quantifying the contribution of each feature to the model's predictions, thereby enabling a more intuitive comprehension of the model's specific predictions. Logistic regression was implemented via the R package "survival", XGBoost was implemented using the R package "XGBoost", and SHAP was executed via the "shapviz" package in R, version 4.3.2.

Feature Selection

Recursive Feature Elimination with Cross-Validation (RFECV) is a robust feature-selection tool that integrates RFE with CV to ascertain the optimal number of features. The RFE iteratively selects features by considering progressively smaller sets and systematically eliminating the least important sets. RFECV enhances this process by performing RFE within a CV loop, ensuring that the reduction in features does not compromise model performance. The process stops when further feature reduction leads to a decline in the performance. The dataset was randomly split into 10 equal subsets. Initially, all features were included in the model. Then, based on feature importance, the least important features were iteratively removed. In each iteration, one-fold was used for validation and the remaining nine for training. This was repeated 10 times, with each fold serving once as the validation set. By assessing model performance across iterations, the optimal feature subset was identified. RFECV was implemented using the "sklearn" library in Python version 3.7, which provides a systematic approach for feature selection that balances model complexity with predictive accuracy.

Construction of Resistance Prediction Models

After partitioning the data into training and validation sets at a 7:3 ratio, we constructed resistance prediction models for patients undergoing HAIC treatment. We used a total of 7 ML methods: "Naïve Bayes (NB)", "svmRadialWeights", "RF", "kk-nearest neighbor (kknn)", "Adaptive Boosting (Adaboost)", "LogitBoost" and "cancerclass". These models are based on the features selected through the REFCV. Following the data partitioning, we employed 10-fold cross-validation via the R package "caret" to train and evaluate these models, identifying the optimal one. Subsequently, the sensitivity of the models was assessed using a Receiver Operating Characteristic (ROC) curve. This process was implemented using the R package "mime" in R, version 4.3.2.

Construction of Survival Prediction Models

Using the same set of features, we evaluated their performance in survival prediction models. We employed an ensemble of 10 ML algorithms (CoxBoost, Lasso, Stepwise Cox, plsRcox, Ridge, Enet, SurvivalSVMS, Gradient Boosting Machine (GBM), SuperPC, and RSF) to construct the survival models. Subsequently, all patients in the training set were stratified into high- and low-risk groups based on the median risk score. Kaplan-Meier (K-M) survival analysis was conducted to compare the OS between the high-risk and low-risk groups. This analysis was performed using the R package "mime" in R, version 4.3.2. In addition, survival risks were calculated using CoxPH function from the R package "survival", and the forest plot for visualization were created with the R package "forestploter".

Statistical Analysis

Continuous variables that followed a normal distribution were reported as mean ± standard deviation (SD). For continuous variables that did not follow a normal distribution, the data were reported as medians, and normality was verified through histograms. Parametric continuous variables were compared using Student's *t*-test, while non-parametric variables were analyzed using the Mann–Whitney *U*-test or Kruskal–Wallis test. Categorical variables are represented as counts and percentages and were analyzed using χ^2-tests or Fisher's exact tests when necessary. Survival curves were plotted using the K-M method and compared using the multivariate Log rank test. The HR for the events was calculated

using the Cox proportional hazards model. All statistical tests were two-tailed, and statistical significance was set at $P < 0.05$. All analyses and calculations were performed using Python version 3.7 and R version 4.3.2.

Results

Clinical Characteristics

A total of 461 patients with unresectable HCC who underwent HAIC were included in the study. We excluded those who were followed up for less than three months or who had visited other hospitals during follow-up. Patients with incomplete clinical data were also excluded. A total of 286 participants were enrolled in this study. The clinical characteristics of the study participants are summarized in Table 1.

Table 1 Clinical characteristics of the study population

Characteristic	Effective (N=208)[a]	Resistance (N=78)[a]	p-value[b]
Age (years)	56 (49, 65)	58 (52, 63)	0.6
Gender			0.01
Male	186 (89%)	77 (99%)	
Female	22 (11%)	1 (1.3%)	
Pre_HAIC treatment			<0.001
Ablation	26 (13%)	12 (15%)	
Hepatectomy	5 (2.4%)	33 (42%)	
No	82 (39%)	5 (6.4%)	
TACE	95 (46%)	28 (36%)	
HAIC times			>0.9
1	82 (39%)	34 (44%)	
2	59 (28%)	25 (32%)	
3	39 (19%)	11 (14%)	
4	15 (7.2%)	5 (6.4%)	
5	4 (1.9%)	1 (1.3%)	
>5	9 (4.4%)	2 (2.6%)	
Initial chemotherapy			0.4
Full dose	53 (25%)	24 (31%)	
Reduced dose	155 (75%)	54 (69%)	
Combine targeted therapy	107 (51%)	35 (45%)	0.3
Combine immunotherapy	74 (36%)	21 (27%)	0.2
After_HAIC treatment			0.005
Ablation	2 (1.0%)	1 (1.3%)	
Comprehensive cancer therapy	26 (13%)	6 (7.7%)	

(*Continued*)

Table 1 (Continued).

Characteristic	Effective (N=208)[a]	Resistance (N=78)[a]	p-value[b]
Hepatectomy	1 (0.5%)	0 (0%)	
No	151 (73%)	46 (59%)	
TACE	28 (13%)	25 (32%)	
Lesion number			0.083
1	25 (12%)	3 (3.8%)	
2	5 (2.4%)	2 (2.6%)	
3	178 (86%)	73 (94%)	
Tumor size (cm)	8.6 (6.0, 12.7)	4.7 (3.0, 6.8)	<0.001
Branchor/hepatic vein tumor thrombus (+)	135 (65%)	27 (35%)	<0.001
Portal vein tumor thrombus (+)	89 (43%)	14 (18%)	<0.001
Superior mesenteric vein tumor thrombus (+)	2 (1.0%)	1 (1.3%)	>0.9
Lymph node metastasis (+)	76 (37%)	23 (29%)	0.3
Pulmonary metastasis (+)	27 (13%)	12 (15%)	0.6
Bone metastasis (+)	6 (2.9%)	4 (5.1%)	0.5
Adrenal metastasis (+)	11 (5.3%)	2 (2.6%)	0.5
HBsAg (+)	170 (82%)	65 (83%)	0.8
HBeAb (+)	160 (77%)	60 (77%)	>0.9
HBcAb (+)	197 (95%)	75 (96%)	0.8
HBV DNA (10^2)	78 (0, 1,056)	0 (0, 102)	<0.001
Anti_virus treatment	160 (77%)	66 (85%)	0.2
HCV (+)	7 (3.4%)	0 (0%)	0.2
AFP (μg/L)	1,009 (48, 18,350)	134 (7, 6,680)	0.004
APT (mAU/mL)	11,600 (1,210, 60,750)	2,565 (120, 30,900)	0.001
PT (s)	13.10 (12.30, 13.80)	12.40 (11.90,13.00)	<0.001
PLT (10^9/L)	149 (101, 216)	123 (98, 165)	0.007
Neutrophil (10^9/L)	3.30 (2.40, 4.60)	3.05 (2.10, 4.00)	0.082
Lymphocyte (10^9/L)	1.05 (0.70, 1.50)	1.10 (0.90, 1.50)	0.094
Monocyte (10^9/L)	0.49 (0.34, 0.67)	0.43 (0.33, 0.58)	0.12
TB (μmol/L)	16 (11, 24)	13 (10, 20)	0.006
DB (μmol/L)	6.2 (3.9, 10.0)	3.9 (2.5, 6.5)	<0.001
ALT (U/L)	40 (28, 58)	39 (26, 53)	0.5
AST (U/L)	63 (43, 109)	47 (31, 79)	0.003

(*Continued*)

Table 1 (Continued).

Characteristic	Effective (N=208)[a]	Resistance (N=78)[a]	p-value[b]
AKP (U/L)	161 (125, 237)	115 (93, 184)	<0.001
r-GT/GGT (U/L)	178 (106, 312)	110 (43, 174)	<0.001
ALB (g/L)	39.0 (35.0, 42.5)	40.0 (37.0, 44.0)	0.078
CRP (mg/L)	13 (5, 35)	6 (3, 17)	0.001

Notes: [a]Median (Q1, Q3); n (%), [b]Wilcoxon rank sum test; Pearson's chi-square test; Fisher's exact test.
Abbreviations: HBsAg, Hepatitis B surface antigen; HBeAg, Hepatitis B e antigen; HBcAg, hepatitis B core antigen; HCV, Hepatitis C virus; AFP, Alpha-fetoprotein; APT, Abnormal prothrombin; PT, Prothrombin time; PLT, Platelet; TB, Total bilirubin; DB, Direct bilirubin; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; AKP, Alkaline Phosphatase; r-GT/GGT, gamma-glutamyl transferase; ALB, albumin; CRP, C Reactive protein.

Clinical Feature Explanation Based on SHAP

XGBoost and Logistic Regression Prediction Model

We divided the entire patient cohort into training and validation datasets in a 7:3 ratio. XGBoost was trained using 38 candidate variables, without feature elimination or extraction. Figure 1A shows the performance of the model, with an area under curve (AUC) of 1.000 for the training set and an AUC of 0.812 for the validation set. This indicates that the XGBoost-trained prediction model exhibits excellent sensitivity and accuracy when all the features are incorporated. In addition, the AUC of the prediction model constructed by logistic regression was 0.731 in the training group and 0.727 in the validation group (Figure 1B). This means that compared with the traditional prediction model, the XGBoost model has better performance.

Feature Explanation Based on SHAP

Following training of the prediction model using XGBoost, we applied SHAP values to interpret the significance of each feature within the model. This allowed us to visually assess the impact of each feature on patient resistance outcomes and to quantify the importance of these features. The SHAP summary plot (Figure 2A) ranks the absolute SHAP values from highest to lowest for each factor, presenting 27 influential features. Pre-HAlC treatment, major tumor size, alkaline phosphatase (AKP), direct bilirubin (DB), alanine aminotransferase (ALT), neutrophil, total bilirubin (TB), platelet (PLT), and C-reactive protein (CRP) levels emerged as the primary factors influencing the model's output results. Figure 2B shows the SHAP values for each sample, with each dot representing an individual sample. Yellow indicates a higher SHAP value, whereas purple indicates a lower value. The SHAP force plot (Figure 2C) revealed how various baseline clinical features collectively affected the prediction outcomes. It not only indicates whether clinical features promote or inhibit resistance but also provides cut-off values for each variable. For continuous variables, a positive SHAP direction (indicated by yellow in Figure 2C, such as albumin (ALB), HBV DNA, neutrophil, and lymphocyte) implies that if the feature value exceeds the threshold, the model is more likely to predict HAIC resistance. Conversely, a negative direction (indicated by purple in Figure 2C, such as CRP, major size, AFP, ALT, AKP, TB, prothrombin time (PT), AST, r_GT/GGT, and monocyte) suggests that if the variable surpasses the threshold, the model is more likely to predict therapeutic efficacy. Regarding categorical variables, including pre-HAIC treatment, initial chemotherapy dose, portal vein tumor thrombosis (PVTT), branch or hepatic vein tumor thrombus after_HAIC treatment, targeted therapy, and lymph node metastasis, Figure 2C highlights the conditions that have the most significant impact on the outcome. Specifically, HAIC was more likely to be less effective in patients who had undergone hepatectomy (pre_HAIC treatment=4) prior to HAIC. The probability of effectiveness is highest when the initial chemotherapy dose is full dose (initial chemotherapy dose = "0"), and there are no portal vein and other tumor thrombi present.

The SHAP dependence plot (Figure 3) illustrates the relationship between the SHAP values and numerical values of the four continuous variables and two categorical variables. Pre_HAIC treatment, PVTT, tumor size, HBV DNA load, PT and AKP levels significantly impact HAIC treatment effectiveness. The analysis reveals that smaller tumors exhibit high positive SHAP values, suggesting potential treatment resistance. In contrast, when tumor size surpasses 7.1 cm, the SHAP values become negative, indicating improved treatment efficacy. Lower HBV DNA load means high positive

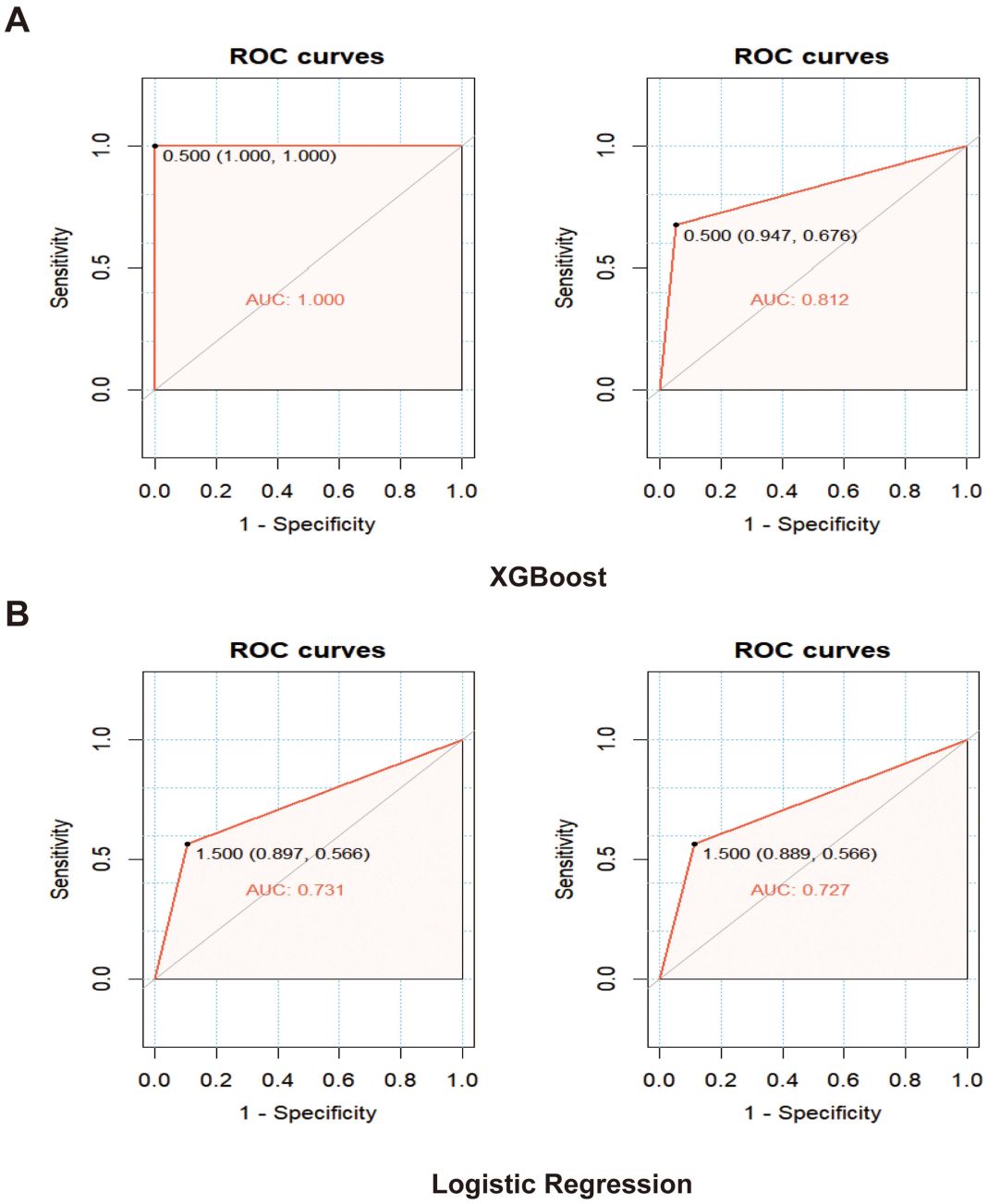


Figure 1 The ROC curves of the prediction model based on XGBoost and logistic regression (without feature selection).
Notes: (**A**) The ROC curves of the model based on XGBoost. (**B**) The ROC curves of the model based on logistic regression.
Abbreviations: ROC, Receiver Operating Characteristic; XGBoost, eXtreme Gradient Boosting.

SHAP values and likely resistance, while higher HBV DNA level leads to negative SHAP values, implying effective. Meanwhile, samples with higher PLT levels and elevated HBV DNA load show lower SHAP values, suggesting better HAIC response. Low PT values associated with positive SHAP values, indicating reduced efficacy. However, when PT exceeds 14s, SHAP values shift to negative, reflecting improved outcomes. In addition, larger tumor sizes and elevated PT values predict negative SHAP values, closely linked to HAIC effectiveness. Low AKP aligns with low positive SHAP values (resistance), whereas levels exceeding 150 U/L switch SHAP values to negative (effective response). Notably, combined high PLT and high AKP levels further reinforce negative SHAP values, underscoring HAIC sensitivity. All pre_HAIC therapies exhibit positive SHAP values, with hepatectomy (treatment 4) showing the highest values, suggesting strong resistance signals. Absence of PVTT corresponds to low negative SHAP values (favorable response),

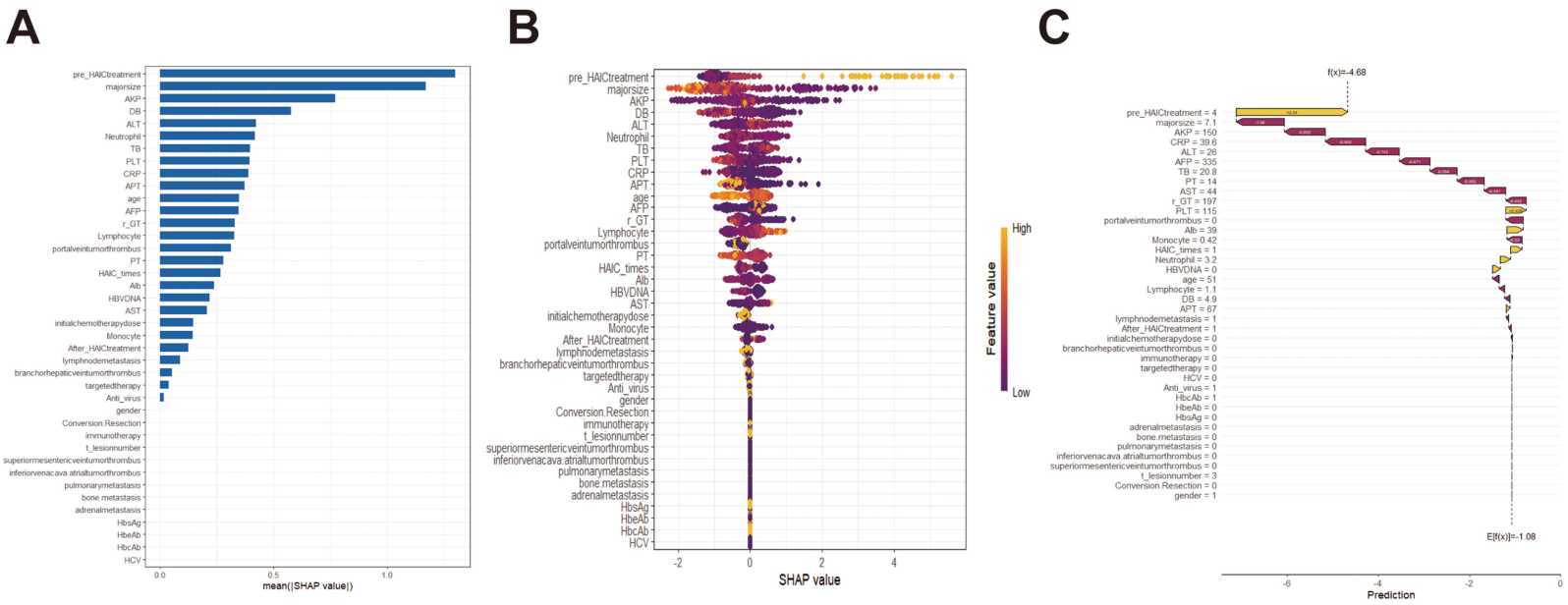


Figure 2 SHAP explanation based on XGBoost.

Notes: (**A**) The SHAP summary plot displays the clinical baseline features that most significantly influence the recurrence outcome of HAIC post-surgery using the XGBoost model, ranked from highest to lowest. (**B**) The position of each feature is arranged in descending order of importance based on the model predictions, with each dot representing a patient sample, where purple indicates a lower SHAP value, and yellow indicates a higher SHAP value. (**C**) The SHAP force plot reveals how various clinical baseline features collectively affect the prediction outcome, allowing a clear understanding of the specific contribution of each feature to the prediction results.

Abbreviations: SHAP, SHapley Additive explanations; XGBoost, extreme Gradient Boosting; REFCV, Recursive Feature Elimination with Cross-Validation; HBsAg, Hepatitis B surface Antigen; HBeAg, Hepatitis B e antigen; HBcAg, Hepatitis B core antigen; HCV, Hepatitis C virus; AFP, Alpha-fetoprotein; APT, Abnormal prothrombin; PT, Prothrombin time; PLT, Platelet; TB, Total bilirubin; DB, Direct bilirubin; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; AKP, Alkaline phosphatase; r-GT/GGT, gamma-glutamyl transferase; ALB, albumin; CRP, C-reactive protein.

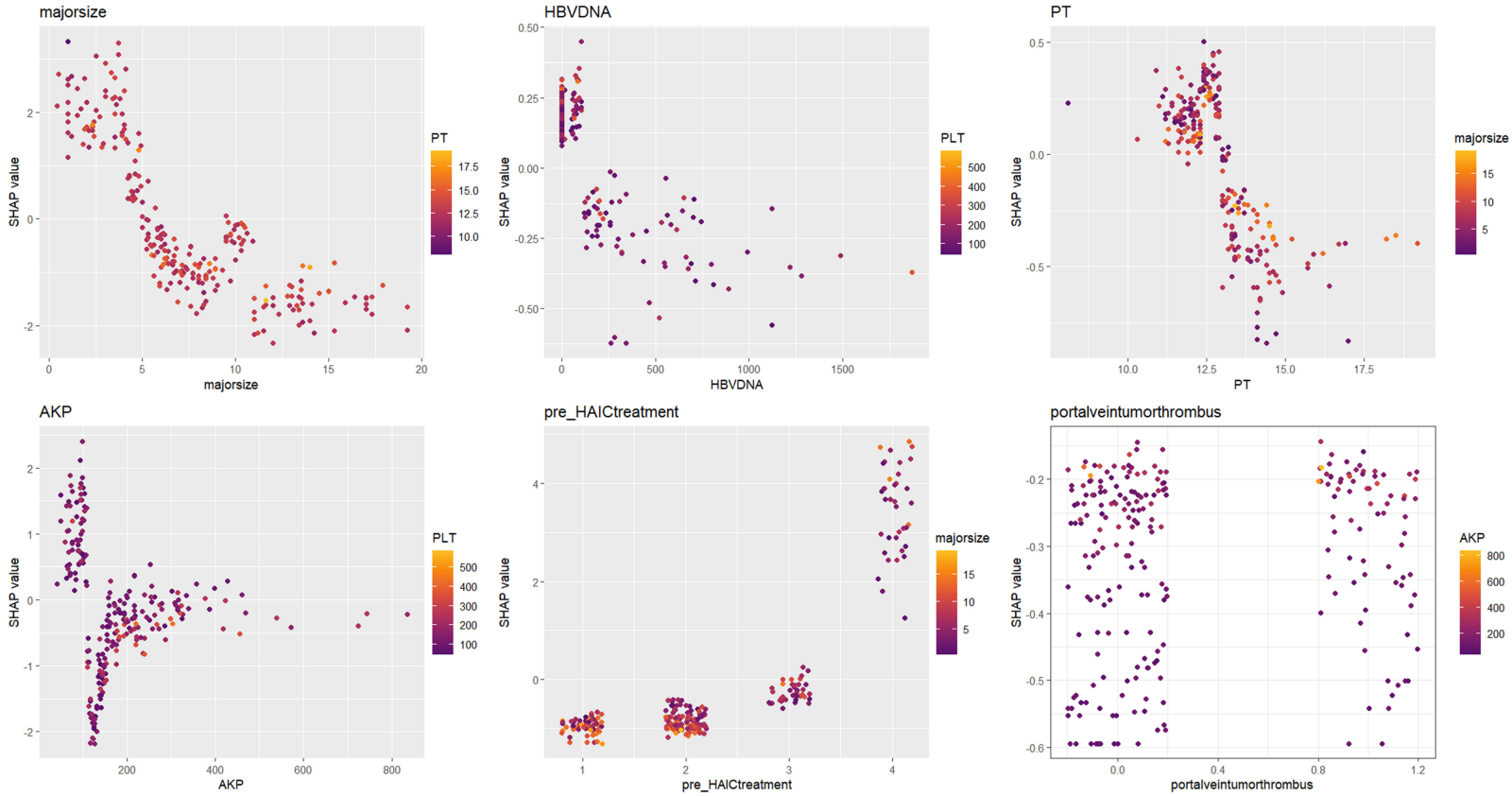


Figure 3 The dependence plots of six variables based on SHAP values.
Abbreviations: SHAP, SHapley Additive explanations; PT, Prothrombin time; AKP, Alkaline phosphatase.

https://doi.org/10.2147/JHC.S523806

while PVTT presence elevates SHAP values to positive (resistance). AKP levels are particularly sensitive to PVTT status, serving as a potential biomarker for resistance risk. These insights underscore the importance of considering these clinical features and their interactions with other clinical indicators when predicting HAIC therapeutic efficacy.

Feature Selection and Construction of Prediction Models

Feature Selection Based on RFECV

Figures 2 and 3 present a preliminary analysis of all included variables. However, the accuracy of the algorithm is questionable because no feature selection is performed. Subsequently, we applied the RFECV method to determine the optimal number of clinical features for predicting the outcomes. As shown in Figure 4A, the accuracy of the model

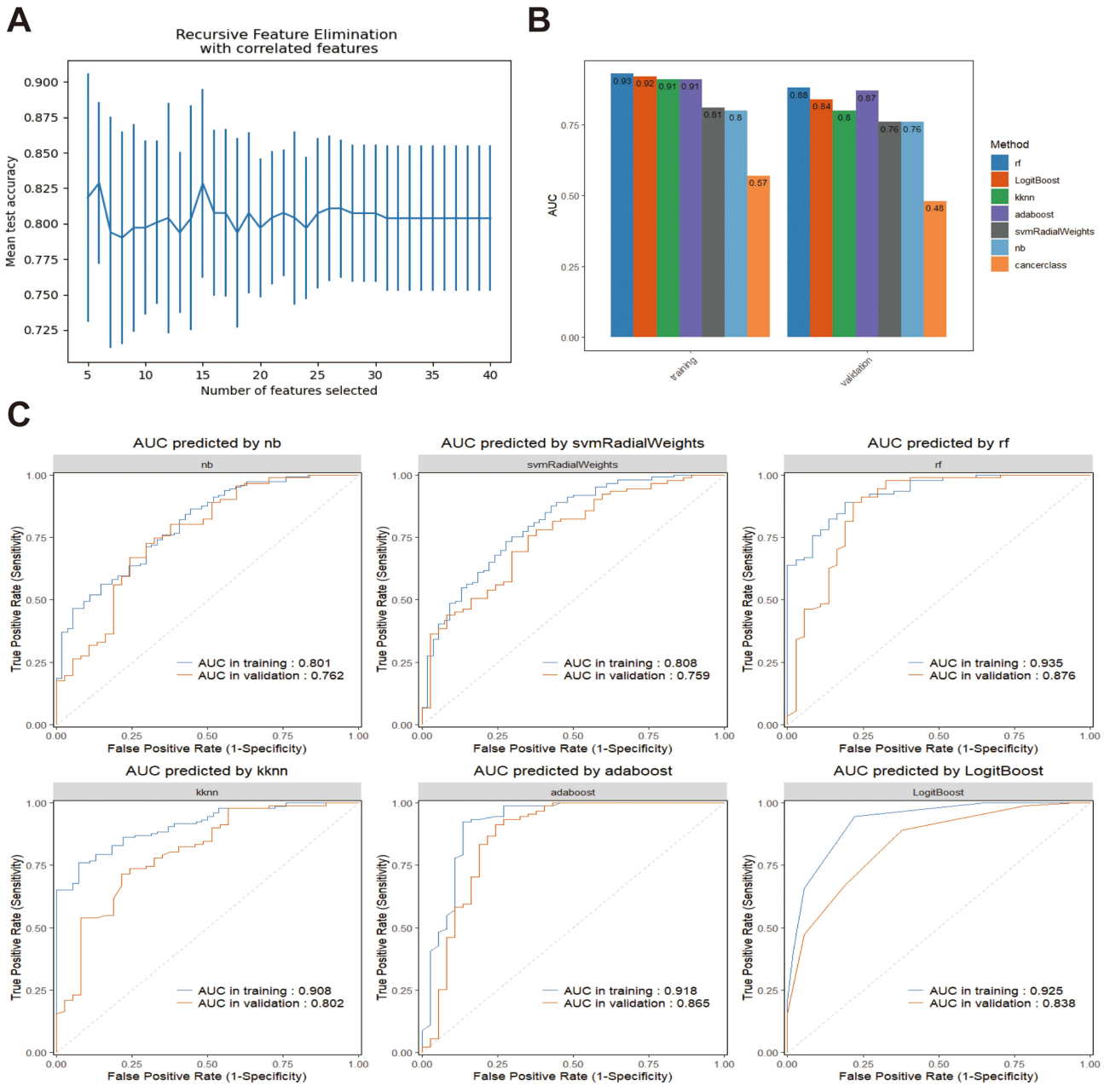


Figure 4 Clinical baseline feature selection and prediction model construction.

Notes: (**A**) Feature selection based on RFECV. (**B**) AUC of the prediction models built based on seven ML methods. (**C**) The ROC curve of prediction models built with "nb", "svmRadialWeights", "RF", "kknn", "adaboost", "LogitBoost".

Abbreviations: AUC, Area Under the Curve; ML, Machine learning; RFECV, Recursive Feature Elimination with Cross-Validation; ROC, Receiver Operating Characteristic; NB, Naïve Bayes; RF, Random Forest; kknn, k-nearest neighbor; Adaboost, adaptive boosting.

https://doi.org/10.2147/JHC.S523806

peaked when six features were included (HBV DNA, tumor size, AKP, PT, PVTT, and pre-HAIC treatment), with a mean test accuracy of 0.83.

Construction of Prediction Models via Different ML Methods

Next, by leveraging the features identified by RFECV, we applied seven ML algorithms to predict the resistance probability of patients with HCC undergoing HAIC. As depicted in Figure 4B and C, the AUC of different models are as follows: RF has a training AUC of 0.935 and a testing AUC of 0.876; Adaboost has a training AUC of 0.918 and a testing AUC of 0.865; LogitBoost has a training AUC of 0.925 and a testing AUC of 0.838; kknn has a training AUC of 0.908 and a testing AUC of 0.802; svmRadialWeights has a training AUC of 0.808 and a testing AUC of 0.759; and NB has a training AUC of 0.801 and a testing AUC of 0.762. This indicates that the inclusion of these characteristics for predicting resistance probability is both accurate and sensitive. The RF algorithm clearly outperformed the other algorithms in predicting the outcome (AUC=0.935 in the training group and 0.876 in the validation group).

Construction of Survival Models Based on Screening Results

The six clinical features selected by RFECV showed remarkable performance in predicting the risk of resistance following HAIC. Since the response of patients to HAIC treatment is closely related to survival, we integrated 10 distinct ML methods, namely CoxBoost, Lasso, stepwise Cox, plsRcox, Ridge, Enet, SurvivalSVM-GBM, SuperPC, and RSF, to train a survival model by utilizing these features. After excluding patients lacking survival information, a total of 260 patients were included in the survival models. Figure 5 illustrates the C-index for predicting patient survival in the training, validation, and entire cohorts after combining the two ML methods. We observed that the combination of Stepcox [forward] and GBM yielded the highest C-index (training set, 0.98; validation set, 0.83; whole cohort, 0.905), indicating the best accuracy (Figure 6A). In addition, we obtained the risk score for each patient using these two algorithms. Subsequently, we stratified the training set into high- and low-risk groups based on the median risk score. Thereafter, we performed K-M survival analysis to compare the OS time between the two groups. The results demonstrated that Stepcox [forward] and GBM enabled the six clinical factors to predict the prognosis of patients following HAIC treatment (Figure 6B). These findings indicate that pre-HAIC treatment, tumor size, HBV DNA, AKP, PT, and PVTT can not only predict the likelihood of resistance after HAIC but also predict the survival prognosis of these participants.

To quantify the survival risk of HCC patients, we utilized the CoxPH function and adopted a forest plot to visualize the risk. As shown in Figure 7, we incorporated six characteristics into the model and found that patients who had undergone hepatectomy before HAIC and whose tumor size≥7.1 cm had a better survival prognosis, while those with PVTT, HBV in the replication stage, PT≥14 s, and AKP≥150 U/L had a worse survival prognosis.

Discussion

In this study, we harnessed the SHAP value and several ML algorithms to predict the resistance and survival prognosis of patients with unresectable HCC following HAIC. Additionally, combining SHAP with RFECV improves the accuracy of the model and makes the feature selection process more transparent by revealing how features function in the model. Our findings highlight the importance of these analytical techniques in identifying the pivotal features that significantly affect patient outcomes.

HAIC has emerged as a promising treatment option for HCC, particularly in patients with advanced or unresectable disease. This strategy allows a higher concentration of drugs to be delivered to the tumor site, thereby increasing the local cytotoxic effect while minimizing systemic toxicity.[9] Despite the efficacy of HAIC in combination therapies, post-operative disease progression remains a concern.[12] At this time, predicting the probability of therapeutic resistance through patient clinical indicators become crucial, facilitating the clinical development of prevention measures.

ML models are increasingly applied across various domains, including HCC.[24] For instance, a recent study by Bo et al also utilized ML to predict the response to lenvatinib monotherapy for unresectable HCC, emphasizing the importance of integrating multiple clinical and radiomic features.[25] However, the interpretability of these models has

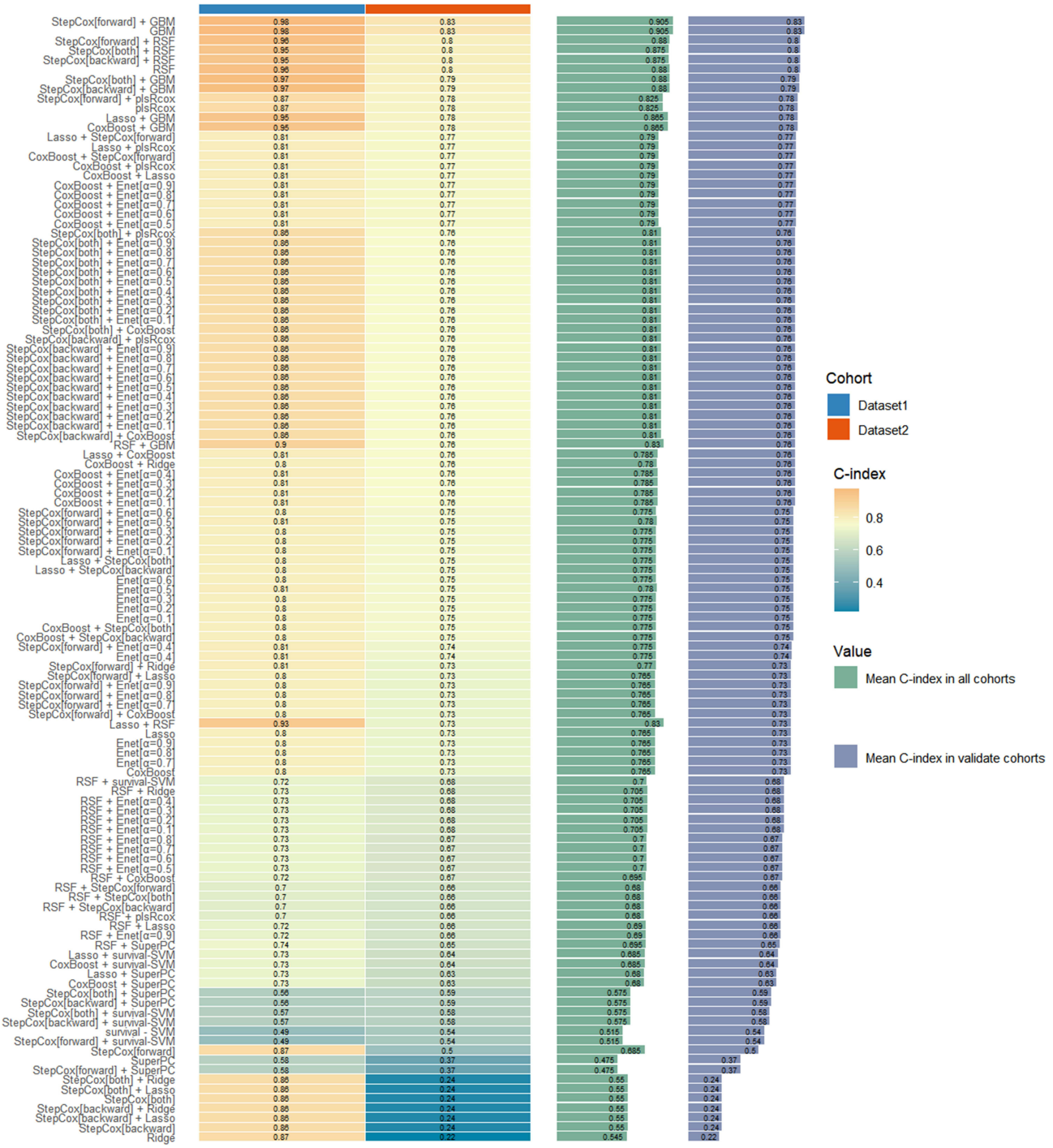


Figure 5 The C-index of 10 ML algorithms in training set (Dataset1), the validation set (Dataset2) and the whole cohort after cross-combining.
Abbreviation: ML, Machine learning.

become a significant research focus. SHAP, a model interpretation method grounded in game theory, offers a unified and effective framework to explain model predictions.

Initially, we used XGBoost to train a preliminary predictive model for the probability of resistance post-HAIC in enrolled participants. We then used SHAP to conduct an interpretability analysis based on XGBoost. Through SHAP, we gained a deeper understanding of the importance of the features included in the model and their functional dynamics. Subsequently, we deployed the RFECV method for feature selection by employing CV to retain the feature set with an

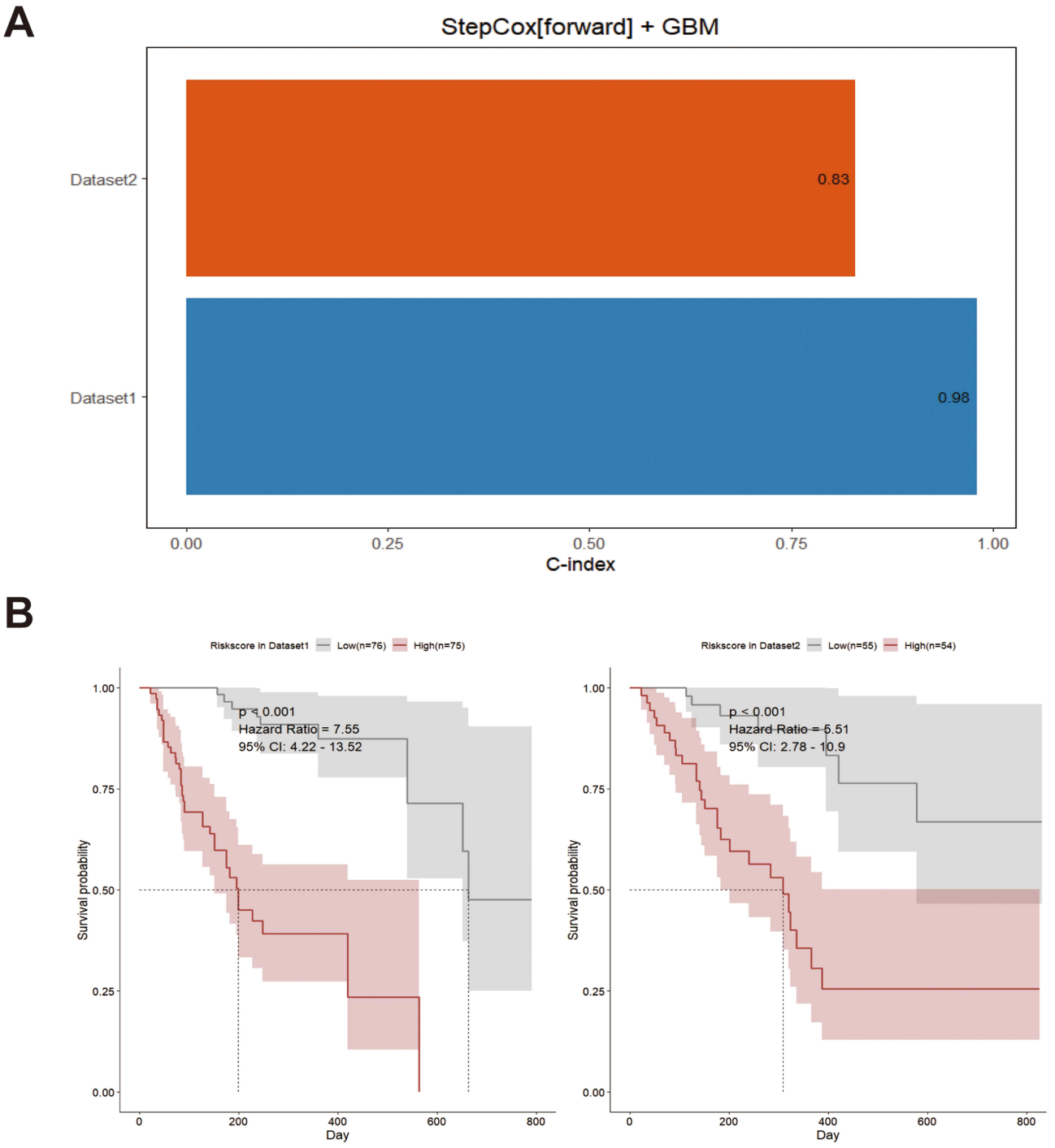


Figure 6 The C-index and K-M survival curve of Stepcox [forward]+GBM.
Notes: (**A**) C-index of the StepCox[forward]+GBM model on both the training set (Dataset1) and validation set (Dataset2). (**B**) The training and validation groups were divided into high- and low-risk groups based on the median risk score and plotting the K-M curves of both groups.
Abbreviations: GBM, gradient boosting machine; K-M, Kaplan–Meier method.

optimal performance. By calculating the sum of their decision coefficients, we ascertained the importance of different features to the score and retained the best feature combination. The results (pre-HAIC treatment approach, tumor size, HBV DNA, AKP, PT, and PVTT) corroborate the analysis of clinical feature importance conducted by SHAP. Moreover, SHAP enables a better understanding of the functionality of these features. For categorical features, we observed that the

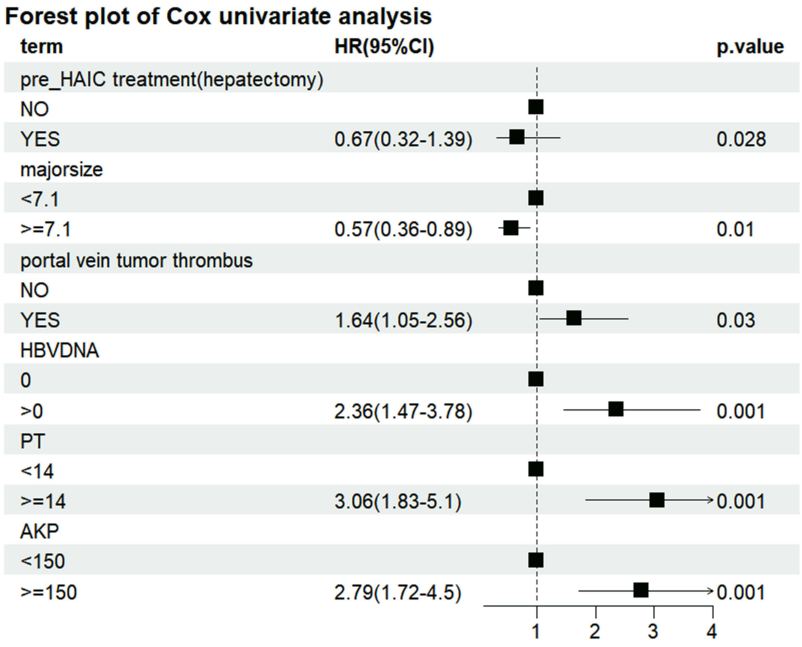


Figure 7 The forest plot of Cox univariate analysis.
Abbreviations: PT, Prothrombin time; AKP, Alkaline phosphatase.

conditions under which they exerted a greater impact on predicting resistance. For continuous features, we clarified their thresholds and their correlation with the SHAP values.

Finally, predictive models were constructed based on the multiple ML models. We also assessed the reliability of these six features in predicting the survival of patients with HCC after HAIC. We utilized pairwise combinations of ten different ML algorithms to predict survival. When the combination of StepCox [forward] and GBM ML algorithms was employed, the six features, led by the pre-HAIC treatment approach, were the most sensitive in predicting patient survival. These features continue to perform admirably in predict the survival of patients with HCC following HAIC treatment. Overall, our findings demonstrated that the pre-HAIC treatment approach, tumor size, HBV DNA, AKP, PT, and PVTT can effectively predict the probability of postoperative resistance and survival in patients with HCC.

For pre-HAIC treatment and PVTT, a poor HAIC effect or resistance is evident when patients have hepatectomy before HAIC and when there exists PVTT. Pre_HAIC treatment method is the main factor affecting the therapeutic effect of HAIC. Consistent with the previous research results, patients who received other treatments before HAIC, such as TKI, TACE and radiotherapy, had influence on the therapeutic effect of HAIC.[26,27] Table 1 shows that most patients who underwent ablation, TACE or no other treatment before HAIC are sensitive to the treatment, while those who have undergone hepatectomy are prone to resistance. Some studies have shown that HAIC after TACE can be used as a rescue plan for HCC refractory to TACE, that is, a sequential strategy of TACE and HAIC, to improve the prognosis of patients.[28] In addition, TACE and ablation before HAIC can be used as a combined treatment to improve the effect of HAIC, which may be because it can further control the residual lesions or micrometastases.[29,30] For those HCC patients who did not receive other treatments before HAIC, the tumor cells maintained a high sensitivity to the chemotherapy drugs of HAIC, and the vascular anatomy was not damaged by treatments such as TACE, thus ensuring that the chemotherapy drugs fully reached the tumor site. However, hepatectomy may change the anatomical structure or blood flow distribution of the hepatic artery, affecting the drug-targeted delivery of HAIC. Previous studies have also indicated that Chinese patients with HCC often present with larger tumors, with the average or median tumor diameter being greater than approximately 10 cm.[31] HAIC may be more suitable for patients with larger tumors (> 5 cm) and abundant blood supply.[32] This aligns with the results we obtained based on SHAP. As for the negative relationship between HBV DNA and SHAP values, studies have shown that high levels of HBV DNA may make tumor cells more vulnerable to chemotherapy drugs in the immune microenvironment by suppressing the immune system, thus making

https://doi.org/10.2147/JHC.S523806

HAIC treatment more effective.[33] Prolonged PT may imply abnormal coagulation leading to the formation of tumor thrombosis, and increased AKP may represent increased tumor load.[34,35] Overall, an increase in PT and AKP may represent greater tumor burden, resulting in better treatment response. Furthermore, in the forest plot of our survival analysis, we found that when the above six clinical features were included, the effects of pre-HAIC, HBV DNA, PT, and AKP on HAIC resistance prediction and survival prognostic prediction results were different. This indicates that although these six clinical features are both influencing factors for the efficacy of HAIC and for survival outcomes, however, when evaluating the prognosis of HCC patients, the therapeutic outcomes of HAIC should be considered alongside other critical prognostic factors. Various factors should also be comprehensively considered to make the prediction more convincing. In conclusion, the application of ML methods and SHAP values helps us to better understand how patients' clinical baseline characteristics play a role in predicting resistance and survival, which is more conducive to our application in practical clinical treatment.

Limitations

Although our results show that the six clinical features can predict resistance and survival probability after HAIC surgery, there are still some limitations to this study. First, the retrospective nature of data collection may introduce biases, and the sample size, although substantial, may limit the generalizability of our findings to other populations. Moreover, since HAIC is generally applicable to patients with a larger tumor burden, the overall condition of the included participants is poor, resulting in a lack of representativeness of the samples. In addition, reliance on a single-center dataset could affect the performance of the model when applied to diverse clinical settings. Future studies should consider larger, multicenter datasets to validate the robustness of our findings.

Conclusions

In patients with unresectable advanced HCC, pre-HAIC treatment, tumor size, HBV DNA, AKP, PT, and PVTT effectively predicted post-HAIC resistance and survival.

Abbreviations

HCC, hepatocellular carcinoma; TKIs, tyrosine kinase inhibitors; ICIs, immune checkpoint inhibitors; OS, overall survival; FDA, the Food and Drug Administration; HAIC, hepatic arterial infusion chemotherapy; TACE, transarterial chemoembolization; HR, hazard ratio; CI, confidence interval; ML, machine learning; XGboost, eXtreme Gradient Boosting; SVC, support vector classification; MLP, multilayer perceptron; LDA, linear discriminant analysis; SHAP, SHapley Additive exPlanations; RFECV, Recursive Feature Elimination with Cross-Validation; NB, Naïve Bayes; RF, random forest; kknn, kk-nearest neighbor; Adaboost, adaptive boosting; ROC, Receiver Operating Characteristic; GBM, Gradient Boosting Machine; K-M, Kaplan-Meier; SD, standard deviation; PD, disease progression; AUC, Area Under Curve; AKP, alkaline phosphatase; DB, direct bilirubin; ALT, alanine aminotransferase; CRP, C reactive protein; TB, total bilirubin; PVTT, portal vein tumor thrombosis; PT, prothrombin time; PLT, platelet.

Ethics Approval and Informed Consent

Informed consent was obtained from the participants or their guardians, and the study was approved by the Ethics Committee of Zhongshan Hospital in accordance with the ethical guidelines of the Declaration of Helsinki (Approval No. KY2022078).

Consent for Publication

All the authors confirm that the details and the article contents can be published.

Acknowledgments

We acknowledge the clinicians and patients who participated in this study. We thank R and Python for allowing us to analyze the data more easily. We are also grateful to the National Nature Science Foundation of China for their financial support. Finally, we acknowledge Springer Nature for polishing the manuscript.

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 study was supported by the National Natural Science Foundation of China (No. 81972889; Dr. Yin).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*. 2019;16(10):589–604. doi:10.1038/s41575-019-0186-y
2. Brown ZJ, Tsilimigras DI, Ruff SM, et al. Management of hepatocellular carcinoma: a review. *JAMA Surg*. 2023;158(4):410–420. doi:10.1001/jamasurg.2022.7989
3. Fan ST. Hepatocellular carcinoma--resection or transplant?. *Nat Rev Gastroenterol Hepatol*. 2012;9(12):732–737. doi:10.1038/nrgastro.2012.158
4. Qin S, Chan SL, Gu S, et al. Camrelizumab plus rivoceranib versus sorafenib as first-line therapy for unresectable hepatocellular carcinoma (CARES-310): a randomised, open-label, international phase 3 study. *Lancet*. 2023;402(10408):1133–1146. doi:10.1016/S0140-6736(23)00961-3
5. Llovet JM, Castet F, Heikenwalder M, et al. Immunotherapies for hepatocellular carcinoma. *Nat Rev Clin Oncol*. 2022;19(3):151–172. doi:10.1038/s41571-021-00573-2
6. Shen KY, Zhu Y, Xie SZ, Qin LX. Immunosuppressive tumor microenvironment and immunotherapy of hepatocellular carcinoma: current status and prospectives. *J Hematol Oncol*. 2024;17(1):25. doi:10.1186/s13045-024-01549-2
7. Yang X, Yang C, Zhang S, et al. Precision treatment in advanced hepatocellular carcinoma. *Cancer Cell*. 2024;42(2):180–197. doi:10.1016/j.ccell.2024.01.007
8. Liu Y, Xun Z, Ma K, et al. Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. *J Hepatol*. 2023;78(4):770–782. doi:10.1016/j.jhep.2023.01.011
9. Iwamoto H, Shimose S, Shirono T, Niizeki T, Kawaguchi T. Hepatic arterial infusion chemotherapy for advanced hepatocellular carcinoma in the era of chemo-diversity. *Clin Mol Hepatol*. 2023;29(3):593–604. doi:10.3350/cmh.2022.0391
10. Yuan Y, He W, Yang Z, et al. TACE-HAIC combined with targeted therapy and immunotherapy versus TACE alone for hepatocellular carcinoma with portal vein tumour thrombus: a propensity score matching study. *Int J Surg*. 2023;109(5):1222–1230. doi:10.1097/JS9.0000000000000256
11. Li QJ, He MK, Chen HW, et al. Hepatic arterial infusion of oxaliplatin, fluorouracil, and leucovorin versus transarterial chemoembolization for large hepatocellular carcinoma: a randomized phase III trial. *J Clin Oncol*. 2022;40(2):150–160. doi:10.1200/JCO.21.00608
12. Li SH, Mei J, Cheng Y, et al. Postoperative adjuvant hepatic arterial infusion chemotherapy with FOLFOX in hepatocellular carcinoma with microvascular invasion: a multicenter, phase III, randomized study. *J Clin Oncol*. 2023;41(10):1898–1908. doi:10.1200/JCO.22.01142
13. Lyu N, Wang X, Li JB, et al. Arterial chemotherapy of oxaliplatin plus fluorouracil versus sorafenib in advanced hepatocellular carcinoma: a biomolecular exploratory, randomized, phase III trial (FOHAIC-1). *J Clin Oncol*. 2022;40(5):468–480. doi:10.1200/JCO.21.01963
14. Zhang TQ, Geng ZJ, Zuo MX, et al. Camrelizumab (a PD-1 inhibitor) plus apatinib (an VEGFR-2 inhibitor) and hepatic artery infusion chemotherapy for hepatocellular carcinoma in Barcelona clinic liver cancer stage C (TRIPLET): a phase II study. *Signal Transduct Target Ther*. 2023;8(1):413. doi:10.1038/s41392-023-01663-6
15. Lai Z, He M, Bu X, et al. Lenvatinib, toripalimab plus hepatic arterial infusion chemotherapy in patients with high-risk advanced hepatocellular carcinoma: a biomolecular exploratory, phase II trial. *Eur J Cancer*. 2022;174:68–77. doi:10.1016/j.ejca.2022.07.005
16. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med*. 2021;13(1):152. doi:10.1186/s13073-021-00968-x
17. Chen D, Liu J, Zang L, et al. Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. *Int J Biol Sci*. 2022;18(1):360–373. doi:10.7150/ijbs.66913
18. Yi F, Yang H, Chen D, et al. XGBoost-SHAP-based interpretable diagnostic framework for Alzheimer's disease. *BMC Med Inform Decis Mak*. 2023;23(1):137. doi:10.1186/s12911-023-02238-9
19. Zuo D, Yang L, Jin Y, Qi H, Liu Y, Ren L. Machine learning-based models for the prediction of breast cancer recurrence risk. *BMC Med Inform Decis Mak*. 2023;23(1):276. doi:10.1186/s12911-023-02377-z
20. Zhou CM, Wang Y, Ye HT, et al. Machine learning predicts lymph node metastasis of poorly differentiated-type intramucosal gastric cancer. *Sci Rep*. 2021;11(1):1300. doi:10.1038/s41598-020-80582-w
21. Yang X, Qiu H, Wang L, Wang X. Predicting colorectal cancer survival using time-to-event machine learning: retrospective cohort study. *J Med Internet Res*. 2023;25:e44417. doi:10.2196/44417
22. Ali S, Akhlaq F, Imran AS, Kastrati Z, Daudpota SM, Moosa M. The enlightening role of explainable artificial intelligence in medical & healthcare domains: a systematic literature review. *Comput Biol Med*. 2023;166:107555. doi:10.1016/j.compbiomed.2023.107555
23. Wang Y, Lang J, Zuo JZ, et al. The radiomic-clinical model using the SHAP method for assessing the treatment response of whole-brain radiotherapy: a multicentric study. *Eur Radiol*. 2022;32(12):8737–8747. doi:10.1007/s00330-022-08887-0

24. Ho CT, Tan EC, Lee PC, et al. Conventional and machine learning-based risk scores for patients with early-stage hepatocellular carcinoma. *Clin Mol Hepatol*. 2024;30(3):406–420. doi:10.3350/cmh.2024.0103
25. Bo Z, Chen B, Zhao Z, et al. Prediction of response to lenvatinib monotherapy for unresectable hepatocellular carcinoma by machine learning radiomics: a multicenter cohort study. *Clin Cancer Res*. 2023;29(9):1730–1740. doi:10.1158/1078-0432.CCR-22-2784
26. Kawaoka T, Aikata H, Hyogo H, et al. Comparison of hepatic arterial infusion chemotherapy versus sorafenib monotherapy in patients with advanced hepatocellular carcinoma. *J Dig Dis*. 2015;16(9):505–512. doi:10.1111/1751-2980.12267
27. Lyu N, Kong Y, Mu L, et al. Hepatic arterial infusion of oxaliplatin plus fluorouracil/leucovorin vs. sorafenib for advanced hepatocellular carcinoma. *J Hepatol*. 2018;69(1):60–69. doi:10.1016/j.jhep.2018.02.008
28. Hien PN, Chun HJ, Oh JS, Kim SH, Choi BG. Hepatic arterial infusion chemotherapy for hepatocellular carcinoma refractory to transarterial chemoembolization: exploring the influence of prior transarterial chemoembolization and additional transarterial chemoembolization on survival outcomes. *J Gastrointest Oncol*. 2024;15(2):721–729. doi:10.21037/jgo-23-1006
29. Liu Y, Qiao Y, Zhou M, et al. Efficacy and safety of hepatic arterial infusion chemotherapy combined with lenvatinib and sequential ablation in the treatment of advanced hepatocellular carcinoma. *Cancer Med*. 2023;12(5):5436–5449. doi:10.1002/cam4.5366
30. Zhou SA, Zhou QM, Wu L, et al. Efficacy of hepatic arterial infusion chemotherapy and its combination strategies for advanced hepatocellular carcinoma: a network meta-analysis. *World J Gastrointest Oncol*. 2024;16(8):3672–3686. doi:10.4251/wjgo.v16.i8.3672
31. He MK, Le Y, Li QJ, et al. Hepatic artery infusion chemotherapy using mFOLFOX versus transarterial chemoembolization for massive unresectable hepatocellular carcinoma: a prospective non-randomized study. *Chin J Cancer*. 2017;36(1):83. doi:10.1186/s40880-017-0251-2
32. He M, Liu S, Lai Z, et al. Hepatic arterial infusion chemotherapy for patients with hepatocellular carcinoma: applicability in Western countries. *Curr Opin Pharmacol*. 2023;70:102362. doi:10.1016/j.coph.2023.102362
33. Chen QJ, Lin KY, Lin ZW, et al. Association of hepatitis B virus DNA levels with efficacy and safety outcomes in patients with hepatitis B virus-associated advanced hepatocellular carcinoma receiving tyrosine kinase inhibitor plus anti-PD-1 antibody: a multicenter propensity-matched study. *Int Immunopharmacol*. 2023;125(Pt A):111098. doi:10.1016/j.intimp.2023.111098
34. Nickel KF, Ronquist G, Langer F, et al. The polyphosphate-factor XII pathway drives coagulation in prostate cancer-associated thrombosis. *Blood*. 2015;126(11):1379–1389. doi:10.1182/blood-2015-01-622811
35. Carr BI, Guerra V. Hepatocellular carcinoma size: platelets, γ-glutamyl transpeptidase, and alkaline phosphatase. *Oncology*. 2013;85(3):153–159. doi:10.1159/000354416