Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and validation of an interpretable machine model for predicting recurrence after radiofrequency ablation in hepatocellular carcinoma

Chunhui Cui
The Second Affiliated Hospital of Nanchang University
Jianwei Yi
The Second Affiliated Hospital of Nanchang University
Tao Tan
Affiliated Hospital of Jiangxi University of Traditional Chinese Medicine
Huajun Wu
The Second Affiliated Hospital of Nanchang University
Jiafu Guan
The Second Affiliated Hospital of Nanchang University
Jian Wang
Pingxiang People's Hospital
Zhimeng Chen
The Second Affiliated Hospital of Nanchang University
Xuliang Hu
The Second Affiliated Hospital of Nanchang University
Binghai Zhou
zhoubinghainc@163.com

The Second Affiliated Hospital of Nanchang University

Research Article

Keywords:

Posted Date: September 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10853617/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Hepatocellular carcinoma (HCC) remains a leading cause of cancer-related mortality worldwide, and radiofrequency ablation (RFA) is widely used for early-stage disease, yet postoperative recurrence hampers long-term survival. Accurate risk stratification is therefore essential for individualized management. This study developed and validated an explainable machine learning model to predict recurrence after RFA in a large, long-term single-center cohort.

Methods

Between January 2018 and December 2022, 1,481 HCC patients were screened; after applying exclusion criteria, 1,115 patients (training set: 905; validation set: 210) were included. Demographic, clinical, tumor, and biochemical data were collected. Both Cox proportional hazards and XGBoost models were constructed; performance was assessed by ROC analysis, SHapley Additive exPlanations (SHAP), and decision curve analysis (DCA).

Results

XGBoost markedly outperformed Cox (training AUC = 0.827 vs 0.742; validation AUC = 0.738 vs 0.629). SHAP revealed postoperative treatment, number of tumors, maximum tumor diameter, and AFP as top predictors. Waterfall plots illustrated distinct feature contributions across risk strata, clarifying drivers of high, intermediate, and low recurrence probability. DCA confirmed superior standardized net benefit over “treat-all” and “treat-none” strategies in both cohorts.

Conclusions

Our findings demonstrate that XGBoost provides accurate, interpretable recurrence prediction, capturing complex risk patterns and enabling transparent clinical decision-making. Integration of this model may optimize surveillance and adjuvant therapy, improving outcomes for HCC patients undergoing RFA.

Background

Hepatocellular carcinoma (HCC) is the most common primary malignancy of the liver and a leading cause of cancer-related death worldwide, with hepatitis B virus infection accounting for the majority of cases in endemic regions[1]. In recent years, the incidence of HCC has continued to rise, posing a growing public health burden[2]. Although radiofrequency ablation (RFA) has become a widely adopted minimally invasive treatment for early-stage HCC, offering favorable local control and acceptable survival

outcomes, postoperative recurrence remains a major obstacle to long-term cure[3]. Once recurrence develops, prognosis deteriorates substantially, and repeated interventions may be required, increasing both morbidity and healthcare costs[4]. Therefore, timely identification of patients at high risk of recurrence is essential for optimizing surveillance strategies and tailoring adjuvant therapies.

Although several prognostic models for HCC recurrence after RFA have been proposed in previous studies[5, 6], these efforts have notable limitations. Many models rely on traditional statistical methods with linear assumptions[6], which may fail to capture complex, nonlinear interactions among diverse clinical, laboratory, and treatment variables. In addition, a substantial proportion of existing studies are based on single-center or modest-size cohorts, limiting statistical power and generalizability[7]. Some investigations focus solely on risk factor exploration without constructing implementable prediction tools for clinical use[8]. Furthermore, although machine learning (ML) techniques have shown promise in improving predictive performance, many reported models lack interpretability, offering little insight into how individual variables influence predictions, and thus limiting clinicians' trust and adoption in routine practice[9]. Even when ML models are developed, they are sometimes trained on incomplete datasets lacking key variables such as detailed treatment history or comprehensive laboratory panels, constraining their real-world applicability[10]. Importantly, few studies have leveraged long-term follow-up data from large samples to build and validate recurrence prediction models, which reduces their reliability for guiding extended postoperative management. Consequently, there remains a clear need for a robust, interpretable model based on a large, long-duration cohort to accurately stratify recurrence risk in HCC patients undergoing RFA.

In this study, we aimed to develop and validate an explainable ML model for the early and accurate prediction of recurrence risk in HCC patients treated with RFA, using a large, long-term single-center cohort with extensive clinical and biochemical data spanning from 2018 to 2022 and prolonged follow-up thereafter. We employed the SHapley Additive exPlanation (SHAP) method to elucidate feature importance and directional contributions, thereby bridging the gap between high predictive accuracy and clinical transparency, and providing a practical tool for individualized postoperative management in routine practice.

Methods

Study Design and Participants

This retrospective cohort study reviewed the medical records of patients with HCC who underwent RFA at our institution between January 2018 and December 2022. A total of 1,481 patients were initially identified. Those meeting predefined exclusion criteria—including presence of vascular invasion or extrahepatic metastasis prior to ablation, intractable ascites or hepatic encephalopathy, severe hepatic or renal dysfunction or other major organ disorders, active infections before or after ablation, history of other malignancies, and loss to follow-up—were excluded. Ultimately, 1,115 patients were eligible for

analysis and randomly divided into a training cohort (n = 905) and a validation cohort (n = 210). The study protocol was approved by the institutional ethics committee, and patient information was handled in accordance with the Declaration of Helsinki.

Inclusion criteria:

1. (a) Pathologically confirmed HCC;
2. (b) Received first-time RFA as primary or salvage treatment;
3. (c) Availability of complete baseline demographic, clinical, laboratory, and imaging data;
4. (d) Complete follow-up records for recurrence status.

Exclusion criteria:

1. (a) Pre-existing distant metastasis or macrovascular invasion;
2. (b) Poorly controlled ascites or hepatic encephalopathy;
3. (c) Significant dysfunction of liver, kidney, or other vital organs;
4. (d) Active infectious disease around the peri-treatment period;
5. (e) History of other malignant neoplasms;
6. (f) Loss to follow-up.

Data Collection

Baseline information was extracted from electronic medical records and encompassed demographic variables (sex, age, income, education, insurance type), clinical characteristics (etiology, presence of portal hypertension, ascites, subcapsular or paravascular involvement, Child–Pugh class, ALBI grade), tumor features (number of lesions, maximum diameter, BCLC stage), preoperative and postoperative treatments, and serial laboratory parameters (albumin, liver enzymes, bilirubin fractions, creatinine, platelet count, INR, prothrombin time, tumor markers including AFP, CEA, CA19-9, CA125, and D-dimer). Follow-up was conducted until recurrence or last known contact, with recurrence defined as radiological or histopathological confirmation of intrahepatic or extrahepatic HCC relapse.

Sample Size Estimation

Sample size was determined based on published recommendations for developing multivariable prognostic models for binary outcomes[11]. To ensure stable estimation with up to 20 predictor variables and an anticipated optimism-corrected R^2 of ≥ 0.1, a minimum of ~ 900 patients in the training set was targeted. Allowing for potential exclusions and imbalances, the final included sample provided sufficient power for internal validation and external verification.

Missing Data Handlings

For variables with missing values < 20%, multiple imputation by chained equations was applied under the missing-at-random assumption. An imputation model incorporated all candidate predictors and the recurrence outcome. Five imputed datasets were generated, and pooled estimates were used in downstream modeling. Variables with > 20% missingness were excluded from analysis.

Prediction Model Development

Two types of models were constructed: the Cox proportional hazards model and the gradient-boosting XGBoost model. All available covariates were included as inputs to predict the probability of HCC recurrence within the follow-up period. The Cox model was fitted using backward stepwise selection, and the proportional hazards assumption was checked using Schoenfeld residuals. The XGBoost model was tuned via grid search (learning rate, max depth, subsample, colsample_bytree) and trained with 5-fold cross-validation on the training cohort. Performance was assessed separately in the training and validation sets using the area under the receiver operating characteristic curve (AUC), with 95% confidence intervals derived by bootstrapping (1,000 resamples). Given superior discriminative ability and robustness in both cohorts, XGBoost was selected as the final prediction model.

Model Interpretation

To enhance interpretability of the XGBoost model, SHapley Additive exPlanations (SHAP) analysis was performed. Global interpretation was visualized through (A) a bar plot of mean absolute SHAP values ranking feature importance, (B) a beeswarm plot displaying the distribution of SHAP values colored by feature magnitude, and (C) SHAP dependence plots for key predictors to reveal nonlinear associations with predicted recurrence risk. Individualized explanations were provided via waterfall plots for representative high-, intermediate-, and low-risk patients, illustrating each feature's additive contribution relative to the baseline prediction.

Risk Stratification and Decision Curve Analysis

Patients were stratified into low, intermediate, and high recurrence risk groups based on quantiles of predicted probabilities from the final XGBoost model. Decision curve analysis (DCA) was conducted to evaluate the clinical utility of the model across a range of high-risk thresholds (0–0.5) in both training and validation cohorts. Standardized net benefit was calculated and compared with "treat-all" and "treat-none" strategies.

Statistical Analysis

Continuous variables were reported as mean ± standard deviation or median (interquartile range) depending on normality, and compared using Student's t-test or Mann–Whitney U test as appropriate. Categorical variables were summarized as frequencies (percentages) and compared using χ^2 or Fisher's

exact test. All statistical analyses were performed using R version 4.3.2. Two-sided $P < 0.05$ was considered statistically significant.

Results

Basic information

A total of 1,481 HCC patients who underwent RFA between January 2018 and December 2022 were initially screened, of whom 1,115 were ultimately eligible for analysis (Fig. 1). The final cohort comprised 905 patients in the training set and 210 in the validation set. The median age was 60.07 years (SD = 11.35), with 85.3% being male. Hepatitis B virus infection was present in 81.3%, portal hypertension in 38.5%, and ascites in 17.5%. Tumors were predominantly solitary (BCLC A stage, 81.9%) with a mean maximal diameter of 1.75 cm (SD = 0.68), and recurrence occurred in 10.1% during follow-up (Table 1). Baseline characteristics were well-balanced between the training and validation sets (all $P > 0.05$) .

Table 1
Summary characteristics of study cohort

Sex = Male (%)	**Overall (n = 1115)**	**Training set (n = 905)**	**Validation set (n = 210)**
	951 (85.3)	**768 (84.9)**	**183 (87.1)**
Age (mean (SD))	60.07 (11.35)	60.04 (11.23)	60.22 (11.90)
Income (%)			
Less than 5000	48 (4.3)	39 (4.3)	9 (4.3)
5000–8000	194 (17.4)	158 (17.5)	36 (17.1)
8000–10000	503 (45.1)	410 (45.3)	93 (44.3)
More than 10000	370 (33.2)	298 (32.9)	72 (34.3)
Education (%)			
Junior high school and below	77 (6.9)	64 (7.1)	13 (6.2)
Senior high school	212 (19.0)	183 (20.2)	29 (13.8)
University	483 (43.3)	390 (43.1)	93 (44.3)
Master degree or above	343 (30.8)	268 (29.6)	75 (35.7)
Insurance (%)			
Commercial insurance	132 (11.8)	106 (11.7)	26 (12.4)
Medical insurance for urban and rural residents	497 (44.6)	403 (44.5)	94 (44.8)
Employee insurance	427 (38.3)	346 (38.2)	81 (38.6)
Other	59 (5.3)	50 (5.5)	9 (4.3)
Number (%)			
1	944 (84.7)	766 (84.6)	178 (84.8)
2	156 (14.0)	126 (13.9)	30 (14.3)
3	15 (1.3)	13 (1.4)	2 (1.0)
Maximum tumor diameter (mean (SD))	1.75 (0.68)	1.78 (0.69)	1.64 (0.61)
BCLC = B (%)	28 (2.5)	23 (2.5)	5 (2.4)
Portal hypertension = Yes (%)	429 (38.5)	346 (38.2)	83 (39.5)
Ascites = Yes (%)	195 (17.5)	151 (16.7)	44 (21.0)

Sex = Male (%)	Overall (n = 1115)	Training set (n = 905)	Validation set (n = 210)
	951 (85.3)	**768 (84.9)**	**183 (87.1)**
Subcapsular = Yes (%)	114 (10.2)	93 (10.3)	21 (10.0)
Paravascular = Yes (%)	46 (4.1)	40 (4.4)	6 (2.9)
Etoligy(%)			
HBV	907 (81.3)	735 (81.2)	172 (81.9)
HCV	10 (0.9)	9 (1.0)	1 (0.5)
Other	198 (17.8)	161 (17.8)	37 (17.6)
Preoperative treatment = Yes (%)	602 (54.0)	469 (51.8)	133 (63.3)
Postoperative treatment = Yes (%)	101 (9.1)	98 (10.8)	3 (1.4)
Recurrence = Yes (%)	113 (10.1)	103 (11.4)	10 (4.8)
ALB (mean (SD))	42.02 (5.79)	41.95 (5.89)	42.33 (5.34)
ALT (mean (SD))	30.88 (27.05)	30.99 (27.12)	30.43 (26.79)
AST (mean (SD))	34.38 (39.86)	34.40 (37.95)	34.30 (47.34)
Tbil (mean (SD))	17.08 (13.06)	17.21 (13.42)	16.54 (11.36)
Dbil (mean (SD))	7.61 (9.70)	7.70 (10.20)	7.22 (7.17)
Ibil (mean (SD))	9.54 (6.33)	9.61 (6.41)	9.26 (5.96)
Cr (mean (SD))	78.77 (33.76)	79.00 (35.04)	77.78 (27.59)
PLT (mean (SD))	131.07 (63.86)	130.89 (64.58)	131.83 (60.79)
INR (mean (SD))	1.15 (0.98)	1.17 (1.03)	1.10 (0.72)
PT (mean (SD))	18.53 (1.78)	18.54 (1.76)	18.47 (1.88)
AFP (mean (SD))	85.10 (307.67)	87.87 (316.20)	73.14 (268.16)
CEA (mean (SD))	4.32 (19.87)	4.50 (21.91)	3.57 (5.23)
CA 199 (mean (SD))	20.58 (80.23)	19.17 (75.77)	26.64 (97.10)
CA 125 (mean (SD))	37.40 (145.14)	37.15 (137.75)	38.48 (173.82)
D dimer (mean (SD))	698.28 (1147.01)	705.11 (1173.05)	668.84 (1029.31)
Child = B (%)	278 (24.9)	223 (24.6)	55 (26.2)

Sex = Male (%)	Overall (n = 1115)	Training set (n = 905)	Validation set (n = 210)
	951 (85.3)	768 (84.9)	183 (87.1)
ALBI (%)			
1	78 (7.0)	58 (6.4)	20 (9.5)
2	786 (70.5)	637 (70.4)	149 (71.0)
3	251 (22.5)	210 (23.2)	41 (19.5)

SD = standard deviation, HBV = Hepatitis B Virus, HCV = Hepatitis C Virus, ALB = Albumin, ALT = Alanine aminotransferase, AST = Aspartate aminotransferase, Tbil = Total Bilirubin, Dbil = Direct Bilirubin, Ibil = Indirect Bilirubin, Cr = Creatinine, PLT = Platelet Count, INR = International Normalized Ratio, PT = Prothrombin Time, AFP = Alpha–fetoprotein, CEA = Carcinoembryonic Antigen, CA199 = Carbohydrate Antigen 199, CA125 = Carbohydrate Antigen 125, ALBI = Albumin–Bilirubin, BCLC = Barcelona Clinic Liver Cancer.

In the training set, the Cox model achieved an AUC of 0.742 (95%CI: 0.685–0.799), while the XGBoost model reached a higher AUC of 0.827 (95%CI: 0.782–0.872), indicating that XGBoost has a stronger ability to capture complex nonlinear relationships in the data. Similarly, in the validation set, the Cox model had an AUC of 0.629 (95%CI: 0.426–0.831), whereas the XGBoost model improved to 0.738 (95%CI: 0.549–0.927), further demonstrating the advantages of XGBoost in terms of generalization capability and predictive accuracy. Therefore, we ultimately selected the XGBoost model as the predictive model for this study, as it exhibited higher predictive precision and stability in both the training and validation sets.

To elucidate the decision mechanism of the XGBoost model and the directional contributions of individual predictors, we performed SHAP (SHapley Additive exPlanations) analysis. Figure 3A presents the ranked feature importance based on mean SHAP values. The bar chart indicates that Postoperative treatment has the highest mean SHAP value (0.162) and is the most influential predictor, whereas Income exhibits the lowest mean SHAP value (0.001), contributing minimally to model output. The beeswarm plot arranges features in descending order of importance and uses color coding to reflect the magnitude and direction of SHAP values, thereby visualizing how different feature values drive predictions upward or downward. Figure 3B displays SHAP dependence plots for key variables—Postoperative treatment, Maximum tumor diameter, number, and AFP—illustrating the relationship between feature values and their corresponding SHAP values. For instance, specific values of Postoperative treatment are associated with positive SHAP values, suggesting that this treatment modality may markedly increase predicted recurrence risk. Maximum tumor diameter shows an approximately linear positive association with SHAP values, indicating that larger tumor size corresponds to higher predicted probability of recurrence. Similar monotonic trends are observed for number and AFP, reinforcing their independent predictive relevance within the model. Collectively, these

SHAP analyses not only uncover the rationale underlying model predictions but also provide interpretable, quantitative insights for identifying clinically actionable risk factors.

Figure 4A, representing a high - risk sample (Patient ID: 4238941, Actual Outcome: Recurrence, Predicted Probability: 0.594), postoperative treatment makes a substantial positive contribution to the prediction with a SHAP value of + 0.724. The presence of three tumors, also contributes positively with a value of + 0.397. Additionally, AFP level contributes a positive value of + 0.203. These positive - contributing features are represented by yellow bars. Other features have relatively minor contributions or negative impacts, with the baseline contribution being E(x) = 1.07. Figure 4B corresponds to a medium - risk sample (Patient ID: 5652092, Actual Outcome: No Recurrence, Predicted Probability: 0.231). Here, the feature “non-postoperative treatment (postoperative_treatment_1 = 0)” has a negative impact on the prediction, with a SHAP value of − 0.122, as shown by the purple bar. The AFP level makes a small positive contribution with a value of + 0.0357. Other features, such as the absence of ascites (ascites_1 = 0), have negative contributions (e.g., − 0.0227), and some features have minimal effects. The baseline contribution remains at E(x) = 1.07. Figure 4C, which is related to a low - risk sample (Patient ID: 1040008, Actual Outcome: No Recurrence, Predicted Probability: 0.221), both “non-postoperative treatment (postoperative_treatment_1 = 0)” and the presence of ascites (ascites_1 = 1) contribute negatively to the prediction, with SHAP values of − 0.0832 and − 0.0325 respectively. The AFP level also has a negative contribution of − 0.0249. Similar to the other plots, there are features with minor positive or negative impacts, and the baseline contribution is E(x) = 1.07.

In both training and validation sets, the XGBoost model consistently demonstrates higher standardized net benefit across all risk thresholds compared to the "All" and "None" benchmarks. The net benefit of XGBoost decreases gradually as the high-risk threshold increases but remains superior to alternative strategies (Fig. 5A). Notably, in the validation set (Fig. 5B), the XGBoost curve exhibits steeper initial declines at lower thresholds but stabilizes at moderate thresholds, suggesting robust generalization.

Discussions

In this study, we developed and validated an explainable XGBoost model to predict recurrence after RFA in a large, long-term cohort of 1,115 HCC patients. The XGBoost model demonstrated superior discriminative performance compared with the Cox proportional hazards model, achieving AUCs of 0.827 in the training set and 0.738 in the validation set. SHAP-based interpretation revealed that postoperative treatment, number of tumors, maximum tumor diameter, and AFP were the most influential predictors. Individualized waterfall plots illustrated distinct feature contributions across high-, intermediate-, and low-risk patients. Decision curve analysis further confirmed that the model provided greater standardized net benefit than treat-all or treat-none strategies in both training and validation phases.

Previous prognostic models for HCC recurrence after RFA have largely relied on traditional statistical methods or limited-scale datasets, often constrained by linear assumptions and insufficient

transparency [12, 13]. Our findings align with emerging evidence that machine learning approaches can achieve higher predictive accuracy by capturing complex, nonlinear relationships among variables[14]. While some studies have reported comparable AUC values using Cox models, few have integrated interpretability frameworks such as SHAP to elucidate clinical drivers of predictions[15, 16]. Moreover, existing models rarely leverage long-term follow-up data from large cohorts[17]; our study addresses this gap by demonstrating robust performance and generalizability in a sample spanning five years of enrollment and prolonged observation. The incorporation of both global and local explanations offers a level of clinical interpretability that enhances trust and potential uptake in practice.

An unexpected observation was that postoperative treatment appeared as a positive contributor to recurrence risk in certain high-risk cases. This seemingly paradoxical finding may reflect indication bias: patients receiving postoperative treatment likely had more aggressive tumor characteristics or higher perceived recurrence risk, prompting clinicians to initiate additional therapy. Thus, the apparent association does not imply causation but rather underscores the model's ability to capture underlying clinical complexity. Similarly, the influence of multiple tumors and larger tumor diameter accords with established oncologic principles, as these factors inherently increase the likelihood of residual microscopic disease. The model's capacity to assign appropriate weight to such features reinforces its fidelity to real-world pathophysiology.

Our model offers a practical tool for clinicians to stratify HCC patients post-RFA into distinct recurrence risk categories, thereby facilitating individualized surveillance intervals and adjuvant treatment decisions. By highlighting modifiable and non-modifiable risk factors, it can guide preoperative counseling and postoperative management, potentially reducing unnecessary interventions in low-risk patients and ensuring intensive monitoring for high-risk cases. The integration of SHAP analysis supports transparent communication with patients regarding why certain factors elevate their risk profile. In future clinical workflows, embedding such models into electronic health record systems could enable real-time risk estimation and support shared decision-making.

Strengths of this study include the large, long-term single-center cohort, robust model performance across training and validation sets, and comprehensive interpretability via SHAP and decision curve analysis. These elements address key limitations of prior work, namely small sample size, lack of transparency, and short follow-up. However, our study is limited by its single-center design, which may restrict generalizability to populations with different etiologic profiles or treatment protocols. Additionally, residual confounding from unmeasured variables remains possible.

In conclusion, we have developed an accurate, interpretable machine learning model for predicting HCC recurrence after RFA, with demonstrated clinical utility. Its deployment may enhance precision management and improve patient outcomes in routine hepatology and oncology practice.

Declarations

Ethics Approval and Consent to Participate

This study was approved by the The Second Affiliated Hospital of Nanchang University Nanchang Ethics Committee (Approval No. 2024154). Written informed consent was obtained from all individual participants prior to their participation in the study.

Declaration of interests

The authors declare that they have no competing interests.

Data sharing

The raw data supporting the conclusions of this article and the analysis code will be made available by the authors, without undue reservation.

Fundings

This study was supported by the National Natural Science Foundation of China (82160576), the Natural Science Foundation of Jiangxi Province of China (20212BAB216052, 20232ACB206038).

Author Contribution

(I) Conception and design: CHC and JWY.

(II) Administrative support: CHC, JWY, TT, and JW.

(III) Provision of study materials or patients: CHC, JWY, HJW, XLH and BHZ.

(IV) Collection and assembly of data: CHC, JWY, TT, HJW, and BHZ.

(V) Data analysis and interpretation: CHC and JWY.

(VI) Manuscript writing: All authors.

References

1. Llovet JM, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. Nat Rev Dis Primers. 2021;7:6.
2. Ganesan P, Kulik LM. Hepatocellular Carcinoma: New Developments. Clin Liver Dis. 2023;27:85–102.

3. Yeow M, Zhao JJ, Fong KY, et al. Radiofrequency Ablation Versus Repeat Hepatectomy for Recurrent Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. World J Surg. 2022;46:2778–87.
4. Yuan BH, Zhu YK, Zou XM et al. Repeat hepatic resection versus percutaneous ablation for the treatment of recurrent hepatocellular carcinoma: meta-analysis. BJS Open 2022;6.
5. Lu N, Yuan C, Sun B, et al. Machine Learning-Based Model for Predicting Recurrence-Free Survival After Interventional Therapy in Malnourished Hepatocellular Carcinoma Patients. Cancer Med. 2025;14:e71157.
6. Yang J, Cui L, Zhang W, et al. Risk Models for Predicting the Recurrence and Survival in Patients With Hepatocellular Carcinoma Undergoing Radio-Frequency Ablation. Clin Med Insights Oncol. 2024;18:11795549231225409.
7. Huang J, Huang W, Guo Y, et al. Risk Factors, Patterns, and Long-Term Survival of Recurrence After Radiofrequency Ablation With or Without Transarterial Chemoembolization for Hepatocellular Carcinoma. Front Oncol. 2021;11:638428.
8. Xin Y, Zhang X, Li Y, et al. A Systemic Inflammation Response Index (SIRI)-Based Nomogram for Predicting the Recurrence of Early Stage Hepatocellular Carcinoma After Radiofrequency Ablation. Cardiovasc Intervent Radiol. 2022;45:43–53.
9. Li YH, Qian GX, Zhu Y, et al. An Integrated Model Combined Conventional Radiomics and Deep Learning Features to Predict Early Recurrence of Hepatocellular Carcinoma Eligible for Curative Ablation: A Multicenter Cohort Study. J Comput Assist Tomogr. 2025;49:860–71.
10. Ni Z, Wu B, Li M, et al. Prediction Model and Nomogram of Early Recurrence of Hepatocellular Carcinoma after Radiofrequency Ablation Based on Logistic Regression Analysis. Ultrasound Med Biol. 2022;48:1733–44.
11. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2019;38:1276–96.
12. Chen PC, Chiu NC, Su CW, et al. Albumin-bilirubin grade may determine the outcomes of patients with very early stage hepatocellular carcinoma after radiofrequency ablation therapy. J Chin Med Assoc. 2019;82:2–10.
13. Pan J, Chen S, Tian G, et al. Preoperative Albumin-Bilirubin Grade With Prognostic Nutritional Index Predicts the Outcome of Patients With Early-Stage Hepatocellular Carcinoma After Percutaneous Radiofrequency Ablation. Front Med (Lausanne). 2020;7:584871.
14. Hamed AA, Muhammed A, Abdelbary EAM, et al. Can Machine Learning Predict Favorable Outcome After Radiofrequency Ablation of Hepatocellular Carcinoma? JCO Clin Cancer Inf. 2024;8:e2300216.
15. Wu KC, Lee IC, Chi CT, et al. Impact of HCV eradication on recurrence pattern and long-term outcomes in patients with HCV-related hepatocellular carcinoma undergoing radiofrequency ablation. Aliment Pharmacol Ther. 2024;60:940–52.
16. Zhang K, Ru J, Wang W, et al. ViT-based quantification of intratumoral heterogeneity for predicting the early recurrence in HCC following multiple ablation. Liver Int. 2025;45:e16051.

17. Chu MO, Shen CH, Chang TS, et al. Pretreatment Inflammation-Based Markers Predict Survival Outcomes in Patients with Early Stage Hepatocellular Carcinoma After Radiofrequency Ablation. Sci Rep. 2018;8:16611.

Figures

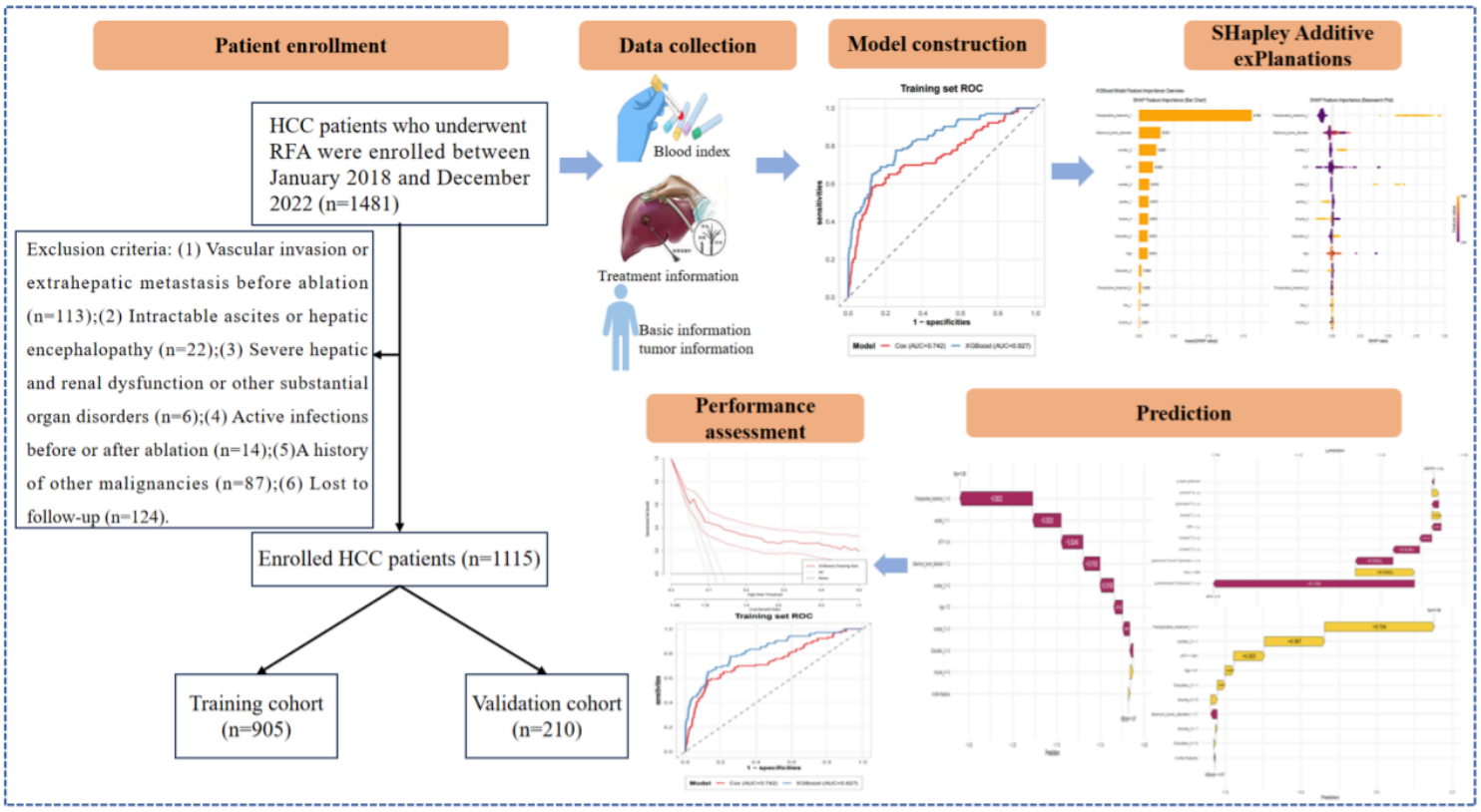
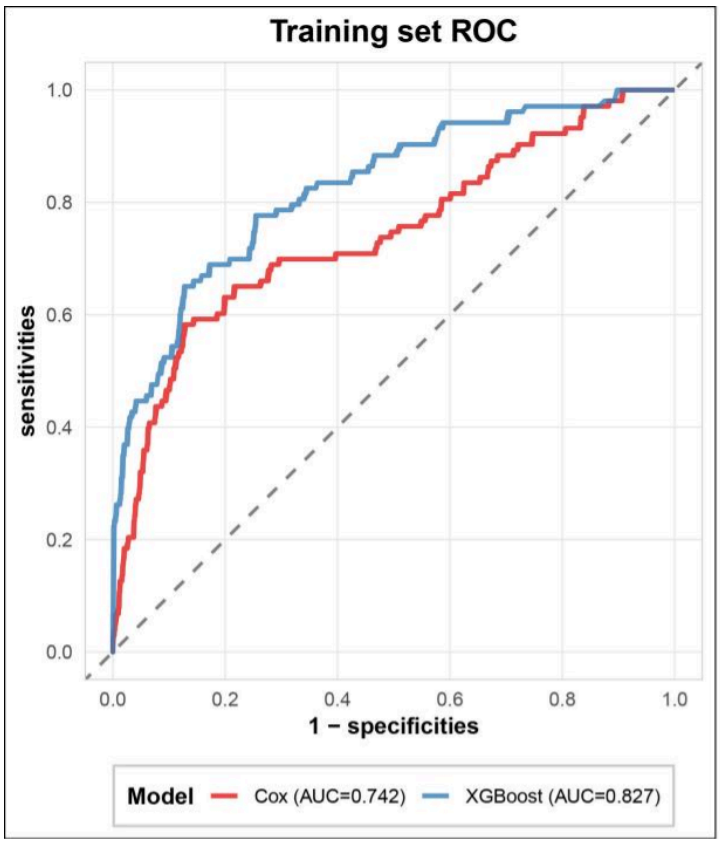


Figure 1

Flowchart of patient enrollment, data collection and analysis process.



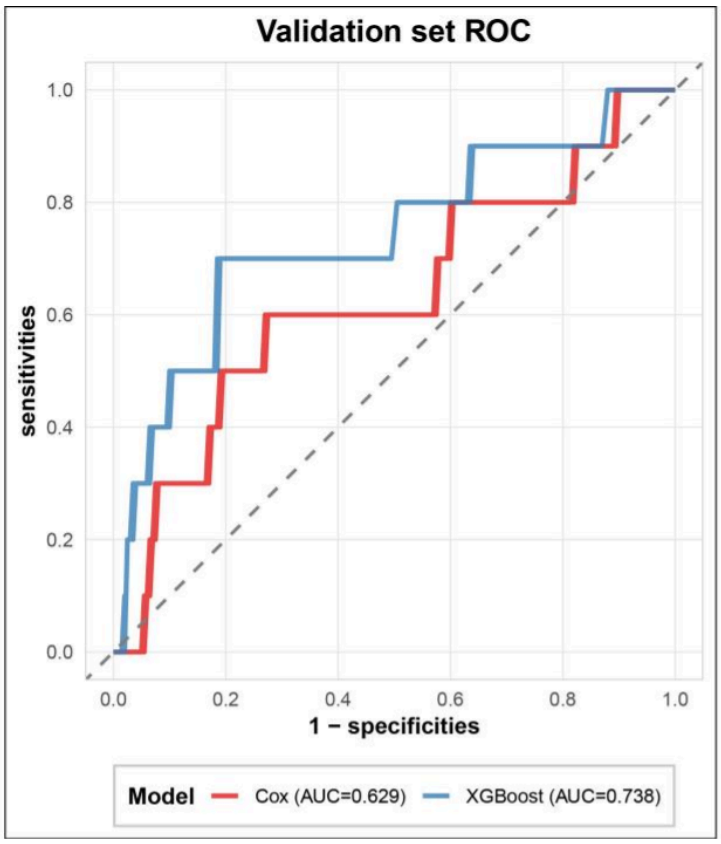


Figure 2

The time-dependent AUCs for the Cox model and XGBoost model



Figure 3

Global and local interpretation of the XGBoost model via the SHapley Additive exPlanations (SHAP) method.(A) SHAP feature importance bar chart: Features are ranked by mean(|SHAP value|), highlighting their overall influence on model predictions. (B) SHAP beeswarm plot: Visualizes the distribution of SHAP values for each feature, where dot position shows SHAP magnitude (x - axis) and feature value (y - axis, colored from low [purple] to high [yellow]); dot density reflects how frequently that feature value

occurs. (C) SHAP dependence plots: Illustrates the relationship between key features and model output. Each point corresponds to an individual sample, with SHAP value on the y - axis and the feature's actual value on the x - axis, revealing how changes in a single feature drive predictions up or down.

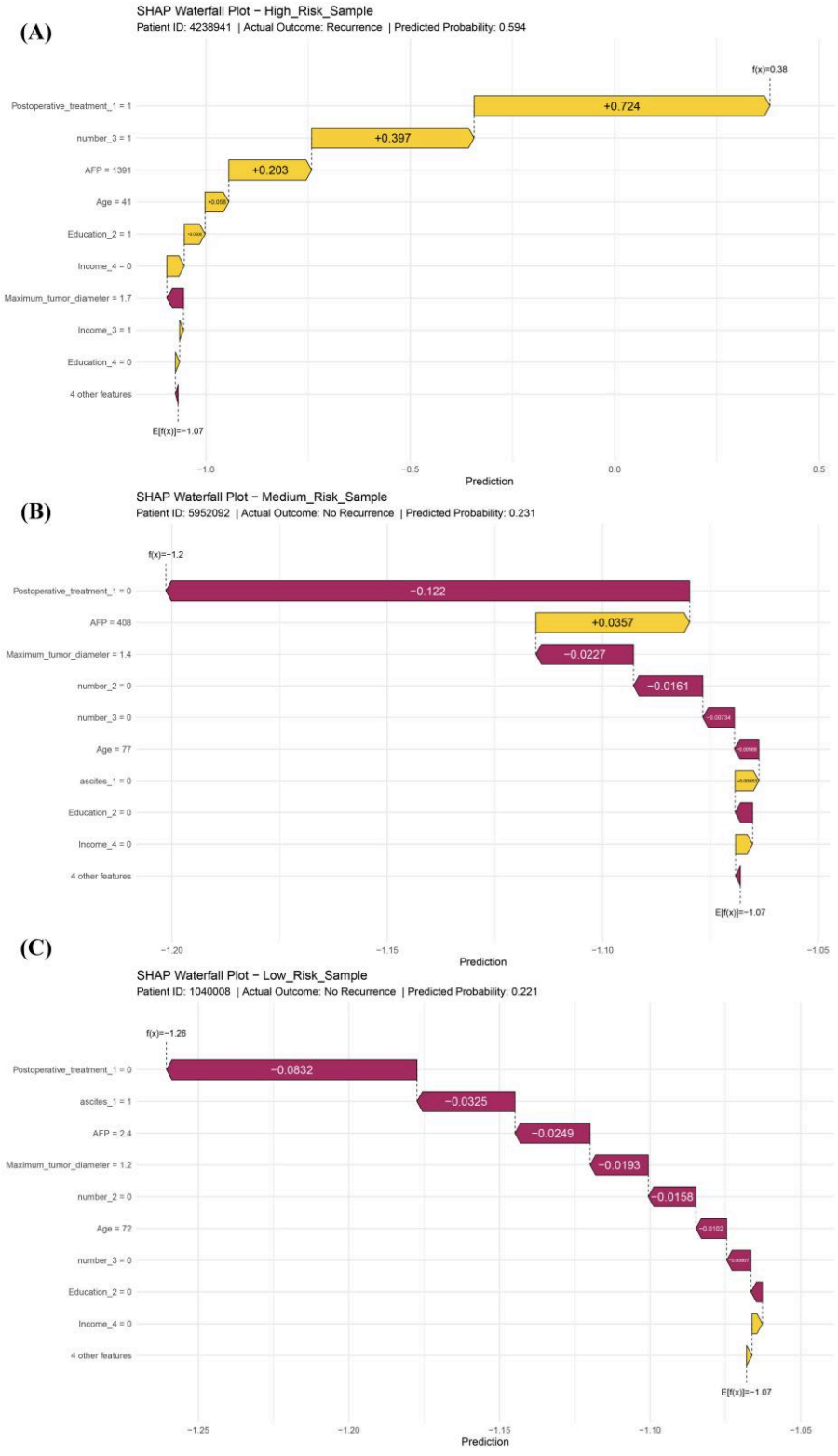


Figure 4

SHAP Waterfall Plot for Risk-Stratified Samples. This set of plots shows the contribution of various features to the prediction results of different risk - level samples using a specific model. In the high - risk sample (A), in the medium - risk sample (B) and low - risk sample (C), different features such as postoperative treatment and ascites have negative (purple bars) or positive contributions. The overall contributions for each sample are also indicated, with the high - risk sample having a predicted probability of recurrence of 0.594 and overall contribution of 0.38, the medium - risk sample having a predicted probability of no recurrence of 0.231 and overall contribution of - 0.122, and the low - risk sample having a predicted probability of no recurrence of 0.221 and overall contribution of - 0.128.

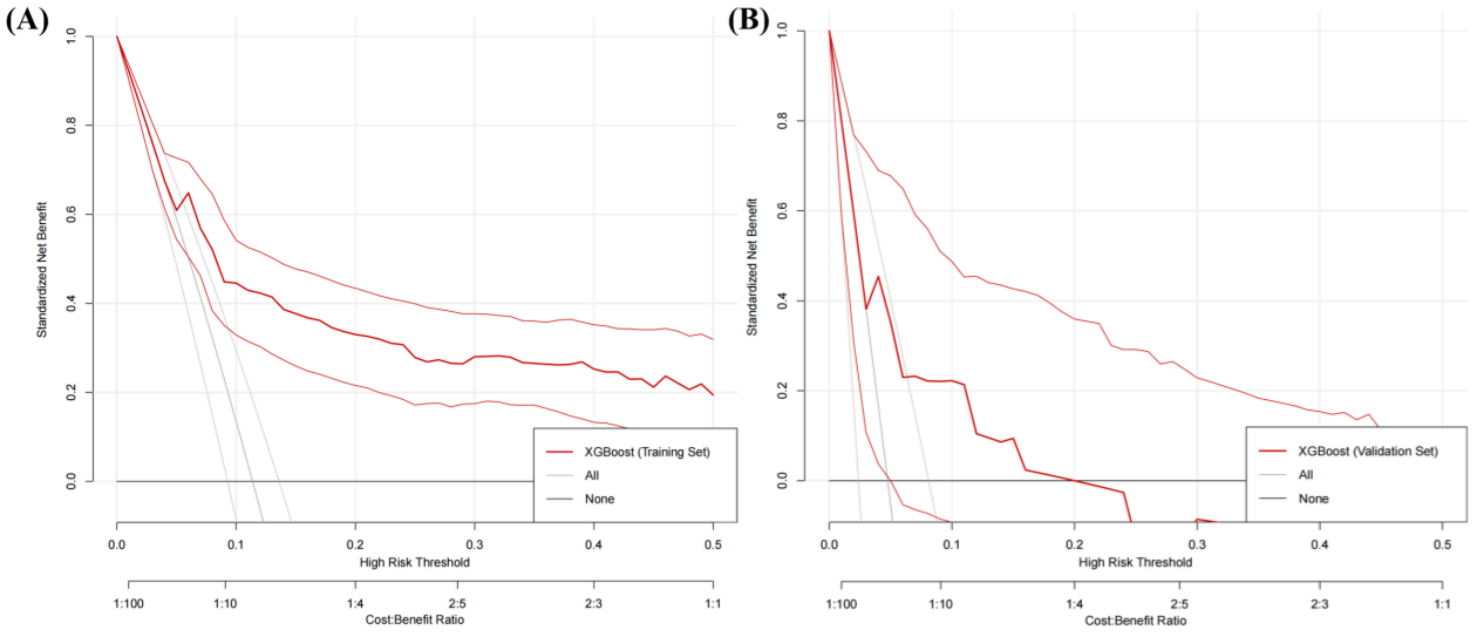


Figure 5

Net benefits plot of the XGBoost model by decision curve analysis for predicting post-RFA recurrence in Training cohort (A) and Validation cohort (B).