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 Multicenter Validation of a Contrast-Enhanced CT-Based Radiomics Model for Predicting Extrahepatic Metastasis after TACE in Unresectable Hepatocellular Carcinoma

yiting sun
Fourth Hospital of Hebei Medical University
xiang liu
Fourth Hospital of Hebei Medical University
lei gao
Fourth Hospital of Hebei Medical University
zhilei zhang
Fourth Hospital of Hebei Medical University
yuzhou ye
Fourth Hospital of Hebei Medical University
qi wang
ja1109w@163.com

Fourth Hospital of Hebei Medical University

Research Article

Keywords: Radiomics, Hepatocellular carcinoma, Transarterial chemoembolization, Prediction model

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11425322/v1

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

Additional Declarations: No competing interests reported.

Development and Multicenter Validation of a Contrast-Enhanced CT-Based Radiomics Model for Predicting Extrahepatic Metastasis after TACE in Unresectable Hepatocellular Carcinoma

Abstract

Objective: To develop and validate a prediction model integrating pretreatment contrast-enhanced CT (CECT) radiomics with clinical features, and to evaluate its value in predicting extrahepatic metastasis (EHM) after transarterial chemoembolization (TACE) in patients with unresectable hepatocellular carcinoma (HCC).

Methods: Clinical data of 152 patients with unresectable HCC who underwent TACE at the participating centers between January 2015 and December 2024 were retrospectively collected, and the patients were randomly assigned to a training set and an internal test set at a ratio of 7:3. Seventy additional HCC patients treated at other participating centers served as the external test set.Lesions were delineated on multiphase CECT images for radiomic feature extraction; after LASSO-based feature selection, the selected features were combined with independent clinical risk factors to construct prediction models and a nomogram. Model discrimination and calibration were assessed using receiver operating characteristic (ROC) curves, decision curve analysis (DCA), and calibration curves. The predictive performance of the models was validated using data from the external cohort.

Results: Multivariate analysis showed that age, alpha-fetoprotein (AFP), and maximum tumor diameter were independent predictors of EHM after TACE. Among the multiphase CT radiomics models, the venous-phase random forest (RF) model performed best (AUC = 0.899, 95% CI: 0.799–0.998), outperforming the logistic regression (LR) and support vector machine (SVM) models. The combined clinical-radiomics model achieved an AUC of 0.923 in the internal test set, outperforming the clinical model alone. In the multicenter external validation cohort, the combined model yielded an AUC of 0.878 (95% CI: 0.801–0.957), indicating good generalizability.

Conclusion: The clinical-radiomics combined model, constructed from selected CT radiomic features and clinical features, demonstrated robust performance in predicting EHM in both the single-center development cohort and the multicenter external validation cohort. It holds promise for estimating the probability of EHM in patients with HCC after TACE and provides a noninvasive tool for preoperative risk stratification.

Keywords: Radiomics; Hepatocellular carcinoma; Transarterial chemoembolization; Prediction model

1. Introduction

Hepatocellular carcinoma (HCC) is the sixth most common malignant tumor worldwide and the third leading cause of cancer-related death[1]. The incidence of liver cancer in China ranks fifth in the world, with a mortality rate as high as 41.7%[2]. Most patients with HCC are diagnosed at an intermediate or advanced stage and are therefore ineligible for curative treatments such as surgical resection and liver transplantation[3]. Transarterial chemoembolization (TACE) has become the core treatment method for this population. For patients with unresectable HCC treated with TACE, clinical data show that the incidence of extrahepatic metastasis (EHM) ranges from 20% to 35%[4, 5]. Metastatic lesions often involve vital organs such as the lungs, bones, and lymph nodes, which not only increases the difficulty of subsequent treatment, but also causes severe complications such as organ failure. The EHM is a key risk factor which affects long-term survival[6]. However, currently there is a lack of preoperative multivariable prediction tool that integrates radiological imaging and clinical indicators, which makes it difficult for clinicians to identify patients with high risk of EHM before TACE treatment.

In recent years, radiomics is developing rapidly. This technique enables non-invasive quantitative assessment of tumor spatial heterogeneity and captures intrinsic biological phenotypes of tumors that are difficult to identify by conventional visual inspection[7, 8]. Multiple studies have confirmed its value as a prognostic biomarker for the evaluation of liver tumors[9-11]. However, most radiomics studies are focused on the prediction of recurrence of HCC after surgical resection[12, 13], whereas research on predicting the risk of preoperative EHM in patients undergoing TACE is relatively scarce. In addition, most previous studies only focus on imaging features, the clinical and inflammatory indicators such as alpha-fetoprotein (AFP), neutrophil-to-lymphocyte ratio (NLR), and platelet-to-lymphocyte ratio (PLR) are not integrated. It limits the performance of prediction and clinical translation value of the models[14, 15].

In this multicenter retrospective study, we enrolled patients with unresectable HCC treated with TACE and used post-TACE EHM as the primary endpoint. Clinical, radiomics, and combined clinical-radiomics prediction models were constructed separately, and their generalizability was validated in an independent external cohort. This work may help assess the risk of EHM in patients with HCC before treatment and provide a reference for developing individualized, risk-stratified intervention strategies.

In this multicenter retrospective study, patients with unresectable HCC treated with TACE were enrolled and the post-TACE EHM was used as the primary endpoint. Clinical, radiomics, and clinical-radiomics combined prediction models were constructed respectively. The generalization ability of the models was

validated through an independent external cohort. The aim of this study was to facilitate the assessment of EHM risk in HCC patients before TACE treatment and provide a reference for the development of individualized and risk-stratified intervention strategies in clinical practice.

2. Materials and Methods

2.1 Study Population

This retrospective multicenter study was performed in accordance with the principles of the Declaration of Helsinki. The requirement for written informed consent was waived by the institutional ethics committee, and all clinical and imaging data were fully anonymized and de-identified before analysis. Only de-identified data were included in the final analysis.

From January 2015 to December 2024, 1074 patients with newly-diagnosed HCC receiving TACE were screened from the participating centers, and 152 patients were finally enrolled for the development cohort. Patients were randomly split into a training set (n=106) and an internal test set (n=46) at a 7:3 ratio. For external validation, 529 HCC patients treated with TACE from additional participating centers between January 2021 and December 2024 were screened; 70 patients constituted the external validation cohort.

The inclusion criteria were: (1) diagnosis of HCC was confirmed according to the Guidelines for the Diagnosis and Treatment of Primary Liver Cancer (2026 Edition)[16]; (2) age ≥ 18 years; (3) Child-Pugh score A or B; (4) multiphase contrast-enhanced CT (CECT) was performed within 1 month before TACE; (5) with complete imaging and clinical data. Exclusion criteria were as follows: (1) received other treatments in disease progression; (2) with other severe diseases or malignant cancer; (3) with intrahepatic or distant metastasis before TACE; (4) death due to adverse events or loss to follow-up. The flowchart of patient inclusion and exclusion criteria are shown in Figure 1.

2.2 TACE Procedure

All enrolled patients underwent conventional TACE. Under local anesthesia, the femoral artery was punctured and catheterized using the Seldinger technique. Through a 5F introducer sheath and under guidewire guidance, a 5F catheter was advanced into the target vessels, such as the celiac trunk and the superior mesenteric artery, and digital subtraction angiography (DSA) was performed to determine the location, size, and number of tumors and to identify the tumor-feeding arteries. A microcatheter was then superselectively inserted into the tumor-feeding arteries[17], and arterial embolization was first performed with an emulsion of iodized oil and chemotherapeutic agents (oxaliplatin combined with fluorouracil and calcium levofolinate, or anthracyclines such as mitoxantrone hydrochloride or idarubicin). For refractory patients—defined with

reference to the core criteria for "TACE refractoriness" of the 2014 Japan Society of Hepatology–Liver Cancer Study Group of Japan (JSH-LCSGJ) consensus and the 2021 consensus of the Chinese College of Interventionalists (CCI), both widely used in clinical practice[18, 19], namely, ≥ 50% residual viable target lesion in the liver on CECT or MRI 1–3 months after two or more consecutive standardized TACE sessions, persistently rising tumor markers (e.g., AFP) with imaging evidence of increased tumor activity, or the presence of high-risk features such as vascular invasion—an emulsion of albumin-bound (nab-) paclitaxel and iodized oil was infused instead. This regimen requires no antihistamine premedication, increases the local intratumoral drug concentration, and reduces systemic toxicity; embolization was then completed with gelatin sponge particles or embolic microspheres. The endpoint of TACE was defined as disappearance of tumor staining, filling of small peritumoral portal vein branches with iodized oil, and slight reflux of the embolic agent. The operator then carefully searched for any additional tumor-feeding arteries and repeated the chemoembolization procedure described above. The doses of chemotherapeutic agents and iodized oil were determined and adjusted by the interventional radiologist according to the patient's liver function and tumor characteristics. After the first TACE, patients underwent regular follow-up examinations and received on-demand treatment according to tumor progression and clinical assessment until no viable or recurrent active tumor tissue remained.

2.3 Image Acquisition and Preprocessing

CT examinations were performed on multiple scanners, including the GE Revolution CT, the Siemens Somatom Definition series, and the Philips Brilliance iCT (Table 1). To eliminate the influence of differences in equipment and scanning parameters across centers on radiomic features and to ensure feature stability and comparability, all original CT images underwent uniform standardized preprocessing. First, quality control was performed on the original DICOM images, and scans with severe artifacts or poor image quality were excluded. All lesion regions of interest (ROIs) were manually delineated directly on CT images reconstructed at a slice thickness of 5 mm. After delineation, the CT images and the corresponding binary ROI masks were simultaneously resampled to an isotropic voxel size of $1 \times 1 \times 1$ mm³ to eliminate variations arising from differences in scanning parameters. Gray-level intensities were normalized to a range of 0–255 and quantized by linear discretization with a fixed bin count of 64 to balance feature stability and computational efficiency. The entire preprocessing pipeline was implemented using the SimpleITK and PyRadiomics libraries in Python to ensure the standardization and reproducibility of feature extraction.

For image acquisition, patients were placed in the supine position, head first, with the scan range covering the entire liver. Contrast-enhanced scanning was performed after a routine non-contrast scan. A nonionic iodinated contrast agent (iohexol injection or ioversol injection) was injected through an antecubital vein using

a power injector at a rate of 2.5–3 mL/s. To eliminate the effect of different iodine concentrations on the degree of enhancement, the dose was uniformly set at 1.5 mL/kg. Arterial-phase scanning was performed 35 s after the end of injection, and venous-phase images were acquired at 70 s.

2.4 Clinical Data

Baseline data obtained within 1 week before the first TACE were retrieved from the electronic medical record systems of the participating centers (the participating centers, the participating centers, and the participating centers). As this study focused on EHM after TACE, the study starting point (time zero) was defined as the day of the first TACE. The following data were collected: demographic characteristics (age and sex); underlying liver status (history of hepatitis and cirrhosis); liver reserve function, assessed by Barcelona Clinic Liver Cancer (BCLC) stage and Child-Pugh class; and laboratory indicators, including serum alpha-fetoprotein (AFP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), total bilirubin (TBIL), γ-glutamyl transferase (GGT), albumin (ALB), prothrombin time (PT), thrombin time (TT), red blood cell count (RBC), white blood cell count (WBC), neutrophil count (NEUT), lymphocyte count (LYM), and platelet count (PLT).

2.5 Segmentation and Radiomics Feature Extraction

Two radiologists, blinded to patient outcomes, used the Deepwise Multimodal Research Platform (https://keyan.deepwise.com/) to manually delineate ROIs slice by slice on non-contrast, arterial-phase, and venous-phase CT images with a slice thickness of 5 mm. The maximal tumor contour was regarded as the ROI of the entire tumor; for tumors with ill-defined margins, ROIs were carefully drawn with reference to the other phases. When multiple lesions were present, the largest lesion, which was also the target lesion of TACE, was selected for delineation. On this basis, an automatic dilation algorithm was used to expand outward and contract inward by 2 mm from the tumor boundary, generating a 4-mm-wide annular band defined as the "peritumoral region," because this region frequently harbors the pathological changes of microvascular invasion (MVI), an important predictor of recurrence and survival in patients with HCC[20-22]. To assess intra-observer agreement, one radiologist repeated lesion delineation and feature extraction in 30 randomly selected patients within 1 week; to assess inter-observer agreement, the other radiologist performed ROI delineation and feature extraction in 30 randomly selected patients. The Dice similarity coefficient (DSC) was used to evaluate the spatial overlap between ROIs from different delineations, and the intraclass correlation coefficient (ICC) of the radiomic features extracted from different delineations was used to evaluate the reproducibility of feature values. A DSC > 0.70 indicated good agreement in ROI delineation, and an ICC ≥ 0.75 indicated good feature reproducibility. The delineations of one radiologist were used for all subsequent analyses. This agreement analysis was intended

to ensure the standardization and reliability of the delineation process.

After delineation, feature extraction and image preprocessing were performed using PyRadiomics (https://pyradiomics.readthedocs.io/en/latest/index.html). The extracted feature classes included first-order features, gray level dependence matrix (GLDM), gray level size zone matrix (GLSZM), gray level co-occurrence matrix (GLCM), neighboring gray tone difference matrix (NGTDM), gray level run length matrix (GLRLM), 3D shape, and 2D shape features. Feature selection was then performed to reduce feature redundancy and avoid model overfitting. To distinguish the imaging features, "A" denotes the arterial phase, "V" the venous phase, "T" the whole tumor, and "P" the peritumoral region.

2.6 Model Construction and Validation

Univariate logistic regression analysis was performed on the clinical factors, and, in consideration of clinical relevance, variables with $P < 0.05$ were entered into multivariate analysis to identify independent risk factors for EHM and construct the clinical model. For radiomic features, all feature selection steps, including Spearman correlation analysis and least absolute shrinkage and selection operator (LASSO) regression, were performed strictly within the training set to avoid data leakage. Specifically, Spearman correlation analysis was first used to remove features with an absolute correlation coefficient ≥ 0.9; the LASSO algorithm was then applied to further select the most predictive features, and a radiomics score (Radscore) was calculated for each patient.

Three machine learning models—LR, SVM, and RF—were constructed based on the selected radiomic features. After the radiomics scores were calculated, the optimal radiomics model was selected according to ROC curves and the AUC, and was then combined with clinical risk factors to construct the clinical-radiomics combined model. Finally, all models were comprehensively evaluated and compared in terms of AUC, sensitivity, and specificity. Calibration curves were plotted to assess the agreement between predicted and observed probabilities, and DCA was used to evaluate the clinical utility of the prediction models. After the model with the best predictive performance was identified, it was validated using the test set. To evaluate cross-center generalizability, the optimal model identified in the development cohort was applied directly to the external validation cohort, in which the clinical, radiomics, and combined models were assessed in parallel; the DeLong test was used to compare AUCs between models within the same cohort, thereby comprehensively verifying cross-center generalizability (Figure 2).

The nomogram was constructed using a two-stage modeling approach. In the first stage, an RF model was built based on the radiomic features, and its output predicted probability was used as the Radscore. In the second stage, the Radscore, together with maximum tumor diameter, age, and AFP, was entered into a multivariate

logistic regression model to construct the combined prediction nomogram. The logistic regression equation was as follows:

$$\text{logit}(P) = \beta_0 + \beta_1 \times \text{Tumor diameter} + \beta_2 \times \text{Age} + \beta_3 \times \text{AFP} + \beta_4 \times \text{Radscore}$$

2.7 Cross-Validation

To verify the stability of model performance and rule out the effect of chance, 5-fold cross-validation was performed within the training set. Specifically, the training set (n = 106) was randomly divided into five balanced, non-overlapping subsets. In each round, one subset served as the validation set and the remaining four were combined for model training; this process was repeated until each subset had served as the validation set once. The AUC of each fold was recorded, and the mean and standard deviation were calculated to assess model robustness.

2.8 Patient Follow-ups

EHM was defined as new extrahepatic tumor lesions confirmed by biopsy or newly detected on imaging, including lymph node metastasis[23]. After TACE, follow-up examinations (CT, MRI, or ultrasound) were performed every 1–3 months. All enrolled patients were followed up for at least 12 months, and the follow-up cutoff date was December 31, 2025. The primary endpoint was the first diagnosis of EHM after TACE.

2.9 Statistical Analysis

Statistical analyses were performed using the Deepwise Multimodal Research Platform (version 2.0) and Python (version 3.9). Normally distributed continuous variables are expressed as mean ± standard deviation and were compared using the independent-samples t test; non-normally distributed continuous variables are expressed as median (interquartile range) and were compared using the Mann–Whitney U test. Categorical variables are expressed as number (percentage) and were compared using the chi-square test or Fisher's exact test. $P < 0.05$ was considered statistically significant. Model calibration was quantitatively assessed using the Brier score and the Hosmer–Lemeshow (H-L) test, with $P > 0.05$ in the H-L test indicating good agreement between predicted probabilities and observed events. The DeLong test was used to compare AUCs between models within the same cohort, with Bonferroni correction for multiple comparisons (corrected significance level $\alpha = 0.017$); $P < 0.017$ was considered statistically significant after correction. SHapley Additive exPlanations (SHAP) analysis was used to quantify and visualize the contribution of each imaging feature to the model predictions.

3. Results

3.1 Patient Characteristics and Follow-up

A total of 222 patients were included. The model development cohort comprised 152 patients from the participating centers, who were randomly assigned at a ratio of 7:3 to the training set (n = 106) and the internal test set (n = 46). The external validation cohort comprised 70 patients from the participating centers and the participating centers. The baseline clinical characteristics of the training, internal test, and external validation sets are shown in Table 2. There were no statistically significant differences in any variable between the training and internal test sets ($P > 0.05$). During follow-up, 81 patients (36.5%) developed EHM, including 40 (37.7%) in the training set, 15 (32.6%) in the internal test set, and 26 (37.1%) in the external test set.

3.2 Agreement of ROI Delineation

To ensure that the radiomic features were derived from the tumor region, inter-observer agreement analysis was performed on the target delineations of the two radiologists. Based on 30 randomly selected cases, the mean DSC was 0.78 for arterial-phase images and 0.80 for venous-phase images. The median inter-observer ICC across all radiomic features was 0.79. These results indicate good agreement between the two radiologists in delineating tumor boundaries, ensuring high reproducibility and reliability of the ROIs used for subsequent feature extraction.

3.3 Radiomic Feature Selection

A total of 2153 radiomic features were extracted from each of the non-contrast, arterial-phase, and venous-phase CT images, yielding 6459 features in total. After screening by intra- and inter-observer agreement analysis (ICC > 0.75) and further selection by LASSO regression, 6, 9, and 7 features were ultimately retained for the non-contrast, arterial, and venous phases, respectively. The LASSO coefficient profiles and cross-validation error curves for each phase are shown in Figure 3a–b (non-contrast phase), Figure 3c–d (arterial phase), and Figure 3e–f (venous phase).

3.4 Independent Clinical Predictors

Univariate logistic regression analysis in the training set showed that age, AFP, and maximum tumor diameter were associated with the risk of EHM (all $P < 0.05$). Variables with $P < 0.05$ in the univariate analysis were entered into multivariate logistic regression, which identified age, AFP, and maximum tumor diameter as independent predictors of EHM (Table 3). The clinical model was constructed based on these three predictors.

3.5 Construction and Internal Validation of the Radiomics Models

Based on the radiomic features selected for each phase, models were constructed using three machine learning algorithms (LR, SVM, and RF), yielding nine radiomics models in total. The AUCs of each model in the training and internal test sets are shown in Table 4. Overall, the venous-phase models achieved higher AUCs than the non-contrast and arterial-phase models. In the venous phase, the RF model showed the best predictive

performance, with an AUC of 0.937 (95% CI: 0.893–0.981) in the training set and 0.899 (95% CI: 0.799–0.998) in the internal test set, higher than those of the LR model (test set AUC = 0.695, 95% CI: 0.533–0.856) and the SVM model (test set AUC = 0.714, 95% CI: 0.560–0.868) in the same phase. The ROC curves of the three models for each phase are shown in Figure 4a–b (non-contrast phase), Figure 4c–d (arterial phase), and Figure 4e–f (venous phase), and the DCA curves are shown in Figure 5a–b (non-contrast phase), Figure 5c–d (arterial phase), and Figure 5e–f (venous phase).

3.6 Construction and Internal Validation of the Combined Model

The Radscore of the venous-phase RF model was integrated with the independent clinical predictors (age, AFP, and maximum tumor diameter) to construct the clinical-radiomics combined model. The combined model achieved an AUC of 0.951 (95% CI: 0.912–0.990) in the training set and 0.923 (95% CI: 0.823–0.993) in the internal test set.

In the internal test set, the AUC of the combined model was higher than that of the clinical model alone (AUC = 0.785, 95% CI: 0.632–0.908), and the difference was statistically significant (DeLong test, P = 0.001). The AUC of the combined model was also higher than that of the RF model alone (AUC = 0.899, 95% CI: 0.799–0.998), although the difference was not statistically significant (DeLong test, P = 0.244). The AUC, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of each model are shown in Table 5. The ROC curves of the three models for each phase are shown in Figure 6a–c (non-contrast phase), Figure 6d–f (arterial phase), and Figure 6g–i (venous phase). The calibration curve showed good agreement between the predicted and observed probabilities of the combined model (Hosmer–Lemeshow test, P = 0.493). DCA (Figure 7g–i) showed that the combined model provided the highest net clinical benefit across threshold probabilities of approximately 0.01–0.07.

3.7 Cross-Validation Results

In 5-fold cross-validation within the training set, the venous-phase RF-based combined model achieved AUCs of 0.923, 0.967, 0.954, 0.963, and 0.949 across the five folds, with a mean of 0.951 and a standard deviation of only 0.017. This small fluctuation indicates stable discriminative performance across subsamples of the training set.

3.8 Performance Evaluation in the External Validation Cohort

The best-performing venous-phase RF-based combined model (including all hyperparameters and the preprocessing pipeline)—was applied directly to the independent external validation cohort without any modification. In the external cohort, the combined model achieved a higher AUC than the clinical model alone and the RF model alone. The AUCs of each model in the training set, internal test set, and external validation

set of the development cohort are summarized in Table 6. The calibration and DCA curves (Figure 7a–c [non-contrast phase], Figure 7d–f [arterial phase], and Figure 7g–i [venous phase]) showed that the combined model also maintained good calibration and net clinical benefit in the external cohort.

3.9 Contribution Analysis of Radiomic Features (SHAP Analysis)

SHAP analysis was used to quantify the contribution of each radiomic feature to the predictions of the venous-phase RF model in the entire cohort (training set + internal test set). Figure 8 shows the top seven features ranked by mean absolute SHAP value. Among the features constituting the Radscore, wavelet-HLH_glszm_LowGrayLevelZoneEmphasis had the highest mean SHAP value (0.042 ± 0.020) and contributed most to the model predictions. A beeswarm plot of the SHAP value distribution of each feature in the entire cohort is provided in Supplementary Figure S1.

3.10 Nomogram

A nomogram for predicting the risk of EHM after TACE was constructed based on the independent clinical predictors identified by multivariate logistic regression (age, AFP, and maximum tumor diameter) and the venous-phase Radscore (Figure 9). The contribution of each variable to the prediction is quantified on a points scale, and the predicted probability of EHM corresponding to the total points can be read directly from the probability scale at the bottom.

4. Discussion

In this study, machine learning models for predicting the risk of EHM after TACE in patients with unresectable HCC were developed by integrating pretreatment multiphase CECT radiomic features with independent clinical predictors, and validated their generalizability using multicenter external data. The main findings were as follows: (1) the venous-phase radiomics model showed the best predictive performance; (2) the RF algorithm outperformed LR and SVM; and (3) the combined model performed well in both the internal test set (AUC = 0.923) and the external validation set (AUC = 0.878), and in all evaluation sets the AUCs of the combined model and the RF model were higher than that of the clinical model alone.

Among the imaging phases, the venous-phase CT model showed the best predictive performance, which may be related to the "washout" appearance of HCC in the venous phase. The pathological basis of this sign includes intratumoral fibrous stromal proliferation, pseudocapsule formation, and heterogeneous distribution of necrotic areas[24]. Such microenvironmental heterogeneity can be quantified by radiomic texture features and has been shown to be associated with the invasive and metastatic potential of tumors. Arterial-phase images mainly reflect tumor blood supply and have a relatively limited ability to characterize intrinsic tumor heterogeneity, which may explain their inferior predictive performance compared with the venous phase. At the

algorithm level, the RF model outperformed the LR and SVM models, consistent with the findings of Parmar et al.[25]; as a tree-based ensemble model, RF can effectively capture nonlinear interactions in high-dimensional radiomics data. Our results are consistent with radiomics methodology and further confirm its applicability to the specific task of predicting EHM after TACE in HCC.

The independent clinical risk factors identified in this study (age, AFP, and maximum tumor diameter) are consistent with those reported previously[26, 27]. Younger age was a risk factor for EHM, possibly because younger patients have a greater tumor burden at diagnosis, and differences in the tumor microenvironment and hormone levels may create favorable conditions for tumor growth and dissemination[28]. Second, AFP is a classic serological marker of HCC; elevated AFP levels are closely related to pro-metastatic mechanisms such as tumor activity, angiogenesis, and immune evasion, and AFP has been used in clinical practice to screen high-risk patients and to guide prognostic assessment in patients with primary HCC[29, 30]. Finally, maximum tumor diameter directly reflects tumor burden; large tumors (> 5 cm) are characterized by greater genomic instability and heterogeneity and a higher probability of vascular invasion[31, 32]. Some studies have reported that BCLC stage and tumor burden are also associated with EHM[33], whereas these variables did not reach statistical significance in our multivariate analysis. This may be attributable to differences in etiological distribution, liver reserve function, and TACE protocols among study populations, underscoring the importance of cross-study external validation for assessing model reliability.

Building on the radiomics model, we integrated independent clinical predictors to construct a clinical-radiomics combined model. Both the combined model and the RF model alone achieved higher AUCs than the clinical model alone (internal test set AUCs of 0.923, 0.899, and 0.785, respectively), suggesting that CECT-based radiomic features have good independent predictive performance. Compared with previous studies that used MRI-based radiomics to predict the risk of EHM after TACE[7, 34, 35], the present study used multiphase CECT, which is more widely available in clinical practice, thereby filling a research gap regarding multiphase CT radiomics in this field and offering greater clinical practicality and potential for wider adoption. In addition, the Radscore was associated with the risk of EHM in patients with AFP-negative HCC; its predictive value may stem from the indirect reflection by radiomic features of intratumoral heterogeneity and peritumoral microenvironmental changes. For patients with low AFP expression, radiomics may serve as a complementary source of information to serological markers, providing an important supplement for risk stratification after TACE in AFP-negative patients.

The incidence of EHM after TACE in this study was 36.5%, consistent with the range of 20%–35% reported in previous TACE populations[15, 36]. The reported incidences after surgical resection and microwave

ablation are approximately 11.3% and 10.52%, respectively[23, 37]. This difference mainly results from patient selection bias and differences in disease stage across treatment modalities, as TACE is mainly used in patients with intermediate or advanced disease who are ineligible for curative treatment[38-40]. These patients inherently have a higher propensity for metastasis, indicating a clear clinical need for EHM prediction tools tailored to the TACE population.

From a clinical translation perspective, this model can noninvasively identify patients at high risk of EHM before TACE. For such high-risk patients, individualized management strategies may be adopted: (1) shortening the interval of post-TACE imaging follow-up to 1–2 months; (2) combining systemic anti-angiogenic therapy or immunotherapy at the time of the first TACE; and (3) performing comprehensive baseline screening for distant metastasis. The clinical benefits of these strategies require validation in prospective studies. Furthermore, the difference in AUC between the combined model and the RF model alone was not statistically significant ($P = 0.244$), suggesting that radiomic features alone already possess strong predictive ability and may serve as an effective alternative when complete clinical data are unavailable.

In this study, systemic inflammatory indicators such as NLR and PLR were not independently associated with EHM ($P > 0.05$), consistent with observations in TACE populations[41, 42], suggesting that intensive locoregional therapy may mask the predictive value of baseline systemic inflammatory signals. Dynamic changes in systemic inflammatory indicators (e.g., the change in NLR before and after treatment) may have greater predictive value than a single baseline measurement; dynamic indicators were not included in this study, which constitutes one of its limitations. Future studies may further investigate the predictive value of temporal changes in inflammatory markers for EHM.

Current studies on predicting EHM after TACE rely solely on single-center internal validation; their generalizability lacks support from external populations and is susceptible to bias arising from single-center case selection and equipment differences[12, 27]. To further verify cross-center generalizability, we introduced a multicenter external validation cohort. The combined model still achieved good predictive performance in the external cohort (AUC = 0.878), indicating that the features extracted by CT radiomics may reflect a shared imaging phenotype of tumor aggressiveness. This finding further supports the potential of radiomic features as universal prognostic biomarkers across etiologies and populations, although validation in larger HCC cohorts is still needed.

This study had several limitations. Firstly, It was retrospective in design and is therefore subject to selection bias. Manual ROI delineation is time-consuming and subject to inter-observer variability, although this was controlled through blinding and agreement analysis (DSC > 0.70). Only the largest lesion was analyzed in

patients with multiple lesions, which may not fully reflect tumor burden and spatial heterogeneity. Only patients treated with TACE were included, so the applicability of the model to patients undergoing surgery, ablation, or systemic therapy remains unclear. CT equipment and scanning parameters differed among centers (Table 1); although standardized preprocessing was applied, the potential influence of equipment heterogeneity on radiomic features warrants attention. In addition, the study period spanned 2015 to 2024, during which TACE procedural standards and systemic treatment strategies may have evolved. Despite these limitations, the model developed in this study showed good predictive performance. We will conduct large-sample prospective studies to confirm the clinical utility of the prediction model.

The clinical-radiomics combined model developed in this study can effectively predict the risk of EHM in patients with unresectable HCC before TACE, and its generalizability was confirmed by multicenter external validation. The model provides an objective, noninvasive auxiliary tool for formulating individualized follow-up and intervention strategies and has the potential to be adopted in multicenter clinical practice.

References

1. LI Y, OU Z, YU D, et al. The trends in death of primary liver cancer caused by specific etiologies worldwide: results from the Global Burden of Disease Study 2019 and implications for liver cancer management[J]. BMC Cancer, 2023, 23(1): 598.

2. CHEN T, ZHANG Y, LIU J. Trends in liver cancer mortality in China from 1990 to 2019: a systematic analysis based on the Global Burden of Disease Study 2019[J]. BMJ Open, 2023, 13: e074348.

3. OYENUGA M, MCGREW T, HANNAN L M, et al. Top advances of the year: Reappraisal of systemic versus local therapy for hepatocellular carcinoma[J]. Cancer, 2025, 131(21): e70162.

4. PINTO E, PELIZZARO F, SIMEON V, et al. Tumour Burden Score as a Predictor of Extrahepatic Progression After Transarterial Chemoembolization for Hepatocellular Carcinoma: An Observational Multicenter Study[J]. Aliment Pharmacol Ther, 2026.

5. USMAN YOUNAS M, SAEED A, RAMZAN M, et al. Transarterial chemoembolization in hepatocellular carcinoma: exploring its role in vascular invasion and extrahepatic metastasis: A systematic review[J]. Medicine (Baltimore), 2025, 104(8): e41570.

6. ZHOU C, CHANG B, XIANG Z, et al. Transarterial Chemoembolization (TACE) Combined with Lenvatinib versus TACE Alone in Intermediate-Stage Hepatocellular Carcinoma Patients Beyond Up-To-Seven Criteria: A Retrospective, Propensity Score-Matched Analysis[J]. Acad Radiol, 2024, 31(11): 4456-4465.

7. PENG G, CAO X, HUANG X, et al. Radiomics and machine learning based on preoperative

MRI for predicting extrahepatic metastasis in hepatocellular carcinoma patients treated with transarterial chemoembolization[J]. Eur J Radiol Open, 2024, 12: 100551.

8. Multi-phase contrast-enhanced magnetic resonance image-based radiomics-combined machine learning reveals microscopic ultra-early hepatocellular carcinoma lesions[J]. Eur J Nucl Med Mol Imaging, 2022, 49: 2917-2928.

9. NOUGARET S, TARDIEU M, VARGAS H A, et al. Ovarian cancer: An update on imaging in the era of radiomics[J]. Diagn Interv Imaging, 2019, 100(10): 647-655.

10. ZHANG Y P, ZHANG X Y, CHENG Y T, et al. Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling[J]. Mil Med Res, 2023, 10(1): 22.

11. BERA K, BRAMAN N, GUPTA A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology[J]. Nat Rev Clin Oncol, 2022, 19(2): 132-146.

12.YANG H R, ZHANG J T, MA M M, et al. Value of contrast-enhanced MRI radiomics combined with clinical factors in predicting early recurrence after radical resection of hepatocellular carcinoma[J]. Chinese Journal of Magnetic Resonance Imaging, 2022, 13: 49-55. (in Chinese)

13. HE Y, DONG B, HU B, et al. Radiomics-based machine-learning method to predict extrahepatic metastasis in hepatocellular carcinoma after hepatectomy: a multicenter study[J]. Sci Rep, 2025, 15(1): 29784.

14. CAI M, HUANG W, HUANG J, et al. Transarterial Chemoembolization Combined With Lenvatinib Plus PD-1 Inhibitor for Advanced Hepatocellular Carcinoma: A Retrospective Cohort Study[J]. Front Immunol, 2022, 13: 848387.

15. LEE S W, LEE H L, HAN N I, et al. Early treatment response to transcatheter arterial chemoembolization is associated with time to the development of extrahepatic metastasis and overall survival in intermediate-stage hepatocellular carcinoma[J]. Cancer Chemother Pharmacol, 2017, 79(1): 81-88.

16.National Health Commission of the People’s Republic of China. Guidelines for the diagnosis and treatment of primary liver cancer (2026 edition)[J]. Chinese Journal of Magnetic Resonance Imaging: 1-145. (in Chinese)

17. XU J X, QIN S L, WEI H W, et al. Prognostic factors and an innovative nomogram model for patients with hepatocellular carcinoma treated with postoperative adjuvant transarterial chemoembolization[J]. Ann Med, 2023, 55(1): 2199219.

18. ZHONG B Y, WANG W S, ZHU H D. Expert consensus on transarterial chemoembolization refractoriness and subsequent treatment of hepatocellular carcinoma[J]. Journal of Interventional Radiology, 2022, 31(11): 1039-1044. (in Chinese)

19. MASATOSHI K, OSAMU M, NAMIKI I, et al. JSH Consensus-Based Clinical Practice Guidelines for the Management of Hepatocellular Carcinoma: 2014 Update by the Liver Cancer Study Group of Japan[J]. Liver cancer, 2014, 3(3-4): 458-468.

20. FANG W, MING C, BINBIN D, et al. Predicting microvascular invasion in small (≤ 5 cm) hepatocellular carcinomas using radiomics-based peritumoral analysis[J]. Insights into Imaging, 2024, 15(1): 90-90.

21. SHUJI S, RYOKO K, KOJI O, et al. Microvascular invasion in patients with hepatocellular carcinoma and its predictable clinicopathological factors[J]. Annals of surgical oncology, 2008, 15(5): 1375-1382.

22. WEI G, FANG G, GUO P, et al. Preoperative prediction of microvascular invasion risk in hepatocellular carcinoma with MRI: peritumoral versus tumor region[J]. Insights into Imaging, 2024, 15(1): 188-188.

23. MASATSUGU I, OSAMU I, MASAHIRO S, et al. The Clinicopathological Features of Postoperative Extrahepatic Metastasis of Hcc After Hepatectomy[J]. Annals of Oncology, 2014, 25(S5): v81-v81.

24. YOSHIMITSU K, NISHIE A, TAKAYAMA Y, et al. The Washout of Hepatocellular Carcinoma at Portal Venous Phase vs. Equilibrium Phase: Radiological and Clinicopathological Implication[J]. Cancers (Basel), 2025, 17(19).

25. CHINTAN P, PATRICK G, JOHAN B, et al. Machine Learning methods for Quantitative Radiomic Biomarkers[J]. Scientific reports, 2015, 5(1): 13087.

26. SEANTHAN S, KHAIRUDDIN M, J L R, et al. Extrahepatic metastases occur in a minority of hepatocellular carcinoma patients treated with locoregional therapies: analyzing patterns of progression in 285 patients[J]. Hepatology (Baltimore, Md), 2012, 55(5): 1432-1442.

27. LIU J, ZHAI H, SHANG L, et al. Transarterial Chemoembolization plus Systemic Therapy versus Transarterial Chemoembolization alone for Unresectable, Nonmetastatic Hepatocellular Carcinoma: Meta-Analysis and Trial Sequential Analysis from Randomized Controlled Trials[J]. Liver Cancer, 2026.

28. LEE C H, LEE Y C, SEO S Y, et al. Pre- and postoperative predictors of extrahepatic recurrence after curative resection for hepatocellular carcinoma[J]. BMC Cancer, 2025, 25(1): 1368-1368.

29. R G P, FRIEDRICH F, MASATOSHI K, et al. Biology and significance of alpha-fetoprotein in hepatocellular carcinoma[J]. Liver international : official journal of the International Association for the Study of the Liver, 2019, 39(12): 2214-2229.

30. TREVISANI F, GARUTI F, NERI A. Alpha-fetoprotein for Diagnosis, Prognosis, and Transplant Selection[J]. Semin Liver Dis, 2019, 39(2): 163-177.

31. XU Z, YUAN W, ZHOU Y, et al. Development and nomogram prediction of early postoperative recurrence in hepatocellular carcinoma based on preoperative CT imaging radiomic features and serum features related to microvascular infiltration[J]. J Gastrointest Oncol, 2024, 15(6): 2630-2641.

32. SHEEN H, KIM J S, LEE J K, et al. A radiomics nomogram for predicting transcatheter arterial chemoembolization refractoriness of hepatocellular carcinoma without extrahepatic metastasis or macrovascular invasion[J]. Abdom Radiol (NY), 2021, 46(6): 2839-2849.

33. LI S, WANG Q, MEI J. Risk factors of extra-hepatic progression after transarterial chemoembolization for hepatocellular carcinoma patients: a retrospective study in 654 cases[J]. J Cancer, 2019, 10: 5007-5014.

34. YANG C, YANG H C, LUO Y G, et al. Predicting Survival Using Whole-Liver MRI Radiomics in Patients with Hepatocellular Carcinoma After TACE Refractoriness[J]. Cardiovasc Intervent Radiol, 2024, 47(7): 964-977.

35. SONG W, YU X, GUO D, et al. MRI-Based Radiomics: Associations With the Recurrence-Free Survival of Patients With Hepatocellular Carcinoma Treated With Conventional Transcatheter Arterial Chemoembolization[J]. J Magn Reson Imaging, 2020, 52(2): 461-473.

36.PENG G. Prognostic analysis of patients with hepatocellular carcinoma treated with TACE based on preoperative imaging and repeatedly measured tumor markers[D]. Beijing: Peking Union Medical College, 2024. (in Chinese)

37. TIAN X Q, WU J P, PANG C, et al. Risk factors for extrahepatic metastasis after microwave ablation of hepatocellular carcinoma[J]. International journal of hyperthermia : the official journal of European Society for Hyperthermic Oncology, North American Hyperthermia Group, 2025, 42(1): 2568627.

38. LUKE Z, MAO Z, SIYU C. Comparison of surgical resection and transcatheter arterial chemoembolization for large hepatocellular carcinoma: a systematic review and meta-analysis[J]. Annals of hepatology, 2022, 28(2): 100890-100890.

39. EMAN M, ASHRAF A A, MOHAMMED N, et al. Value of microwave ablation in treatment of large lesions of hepatocellular carcinoma[J]. Journal of digestive diseases, 2015, 16(8): 456-463.

40. AKYILDIZ A, YAVUZ B G, ISMAYILOV R, et al. Comprehensive Review of Transarterial Chemoembolization Plus Systemic Therapy in Advanced Hepatocellular Carcinoma[J]. Journal of Gastrointestinal Cancer, 2025, 56(1): 194-194.

41. MINICI R, SICILIANO M A, AMMENDOLA M, et al. Prognostic Role of Neutrophil-to-Lymphocyte Ratio (NLR), Lymphocyte-to-Monocyte Ratio (LMR), Platelet-to-Lymphocyte Ratio (PLR) and Lymphocyte-to-C Reactive Protein Ratio (LCR) in Patients with Hepatocellular Carcinoma (HCC) undergoing Chemoembolizations (TACE) of the Liver: The Unexplored Corner Linking Tumor Microenvironment, Biomarkers and Interventional Radiology[J]. Cancers (Basel), 2022, 15(1).

42. WANG Y, LIU Z, XU H, et al. MRI-based radiomics model and nomogram for predicting the outcome of locoregional treatment in patients with hepatocellular carcinoma[J]. BMC Med Imaging, 2023, 23(1): 67.

Figures

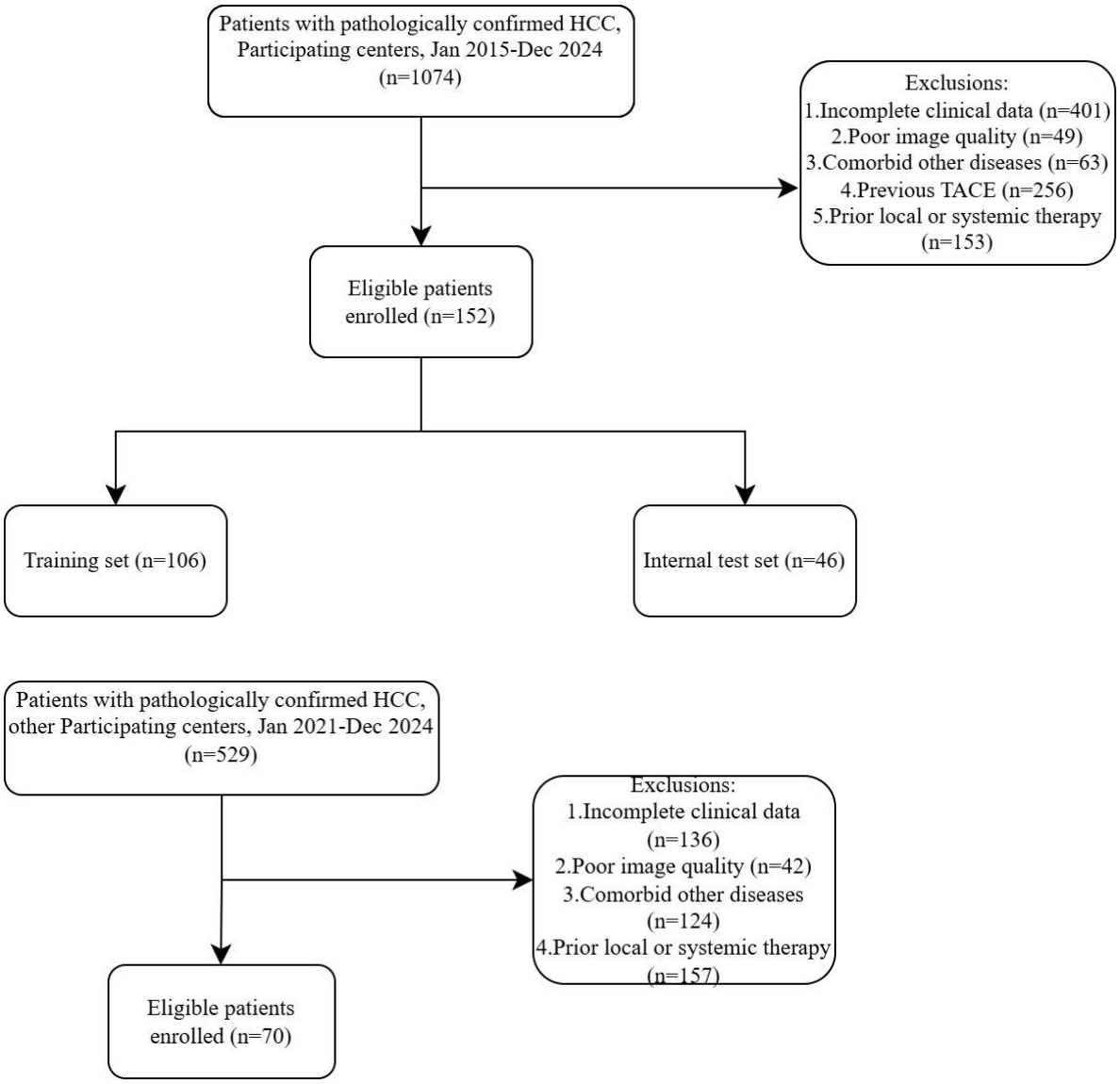


Figure 1. Flowchart of patient enrollment.

Note: "Incomplete medical records" specifically refers to the following: (1) the initial treatment was not performed at the included centers, so key diagnostic and treatment data could not be traced; (2) the core blood tests required for this study (including AFP, liver function, and inflammatory indicators) were not completed at initial admission, so baseline clinical data were unavailable; and (3) only MRI, ultrasound, or other examinations were performed at initial admission, without the multiphase CECT required for this study, so the imaging data needed for radiomics analysis were lacking. No patients were excluded solely because of missing follow-up, and all enrolled patients had complete CECT imaging, clinical, and follow-up data.

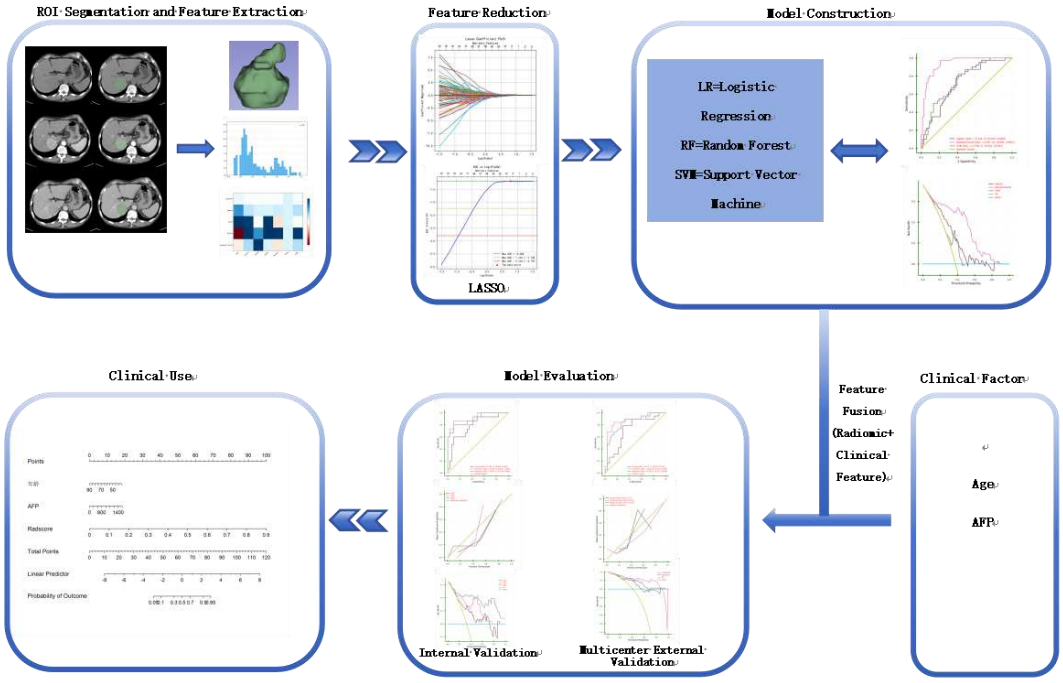


Figure 2. Workflow for the construction of the radiomics prediction models

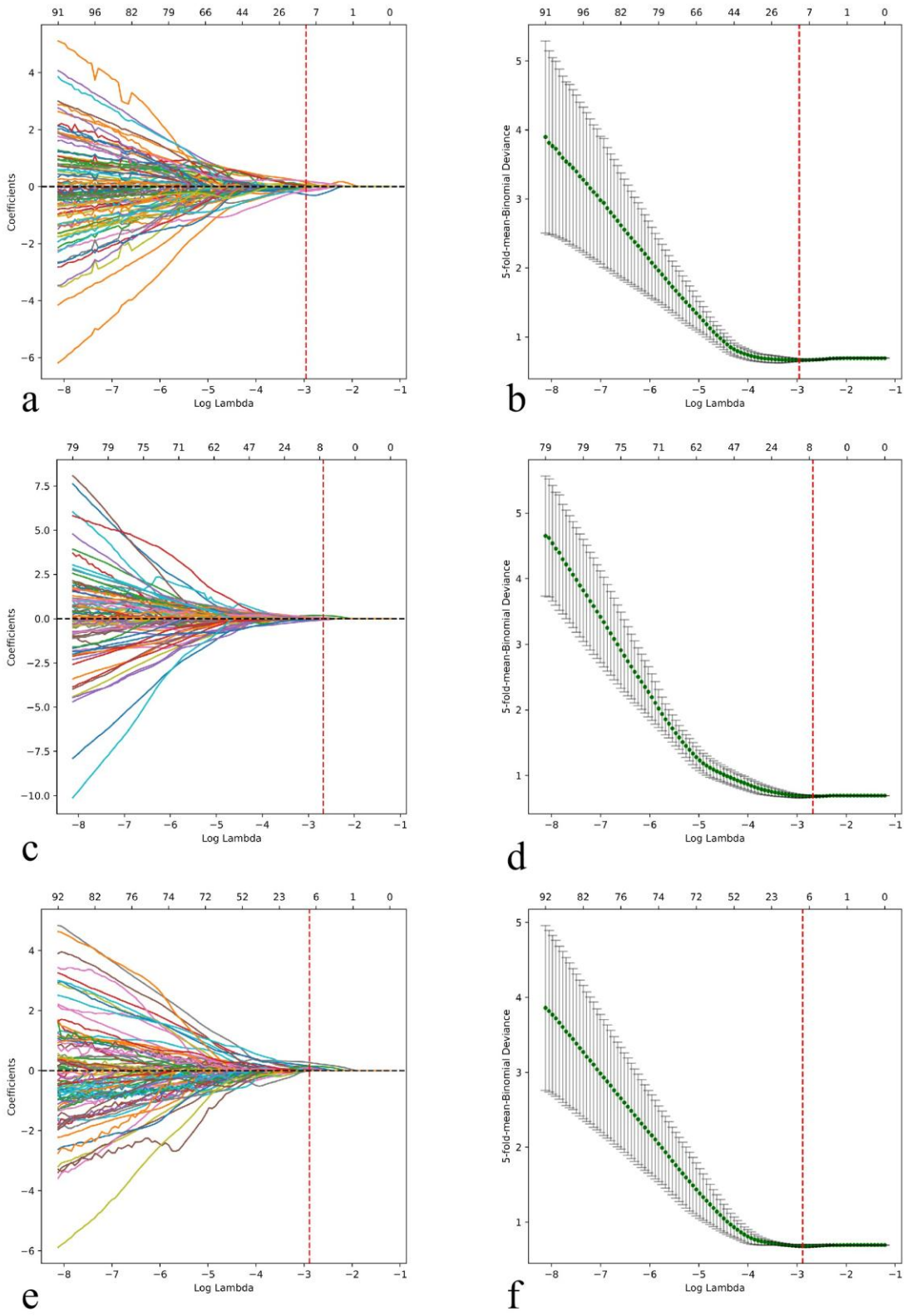


Figure 3. Radiomic feature dimensionality reduction for the non-contrast (a–b), arterial (c–d), and venous (e–f) phases

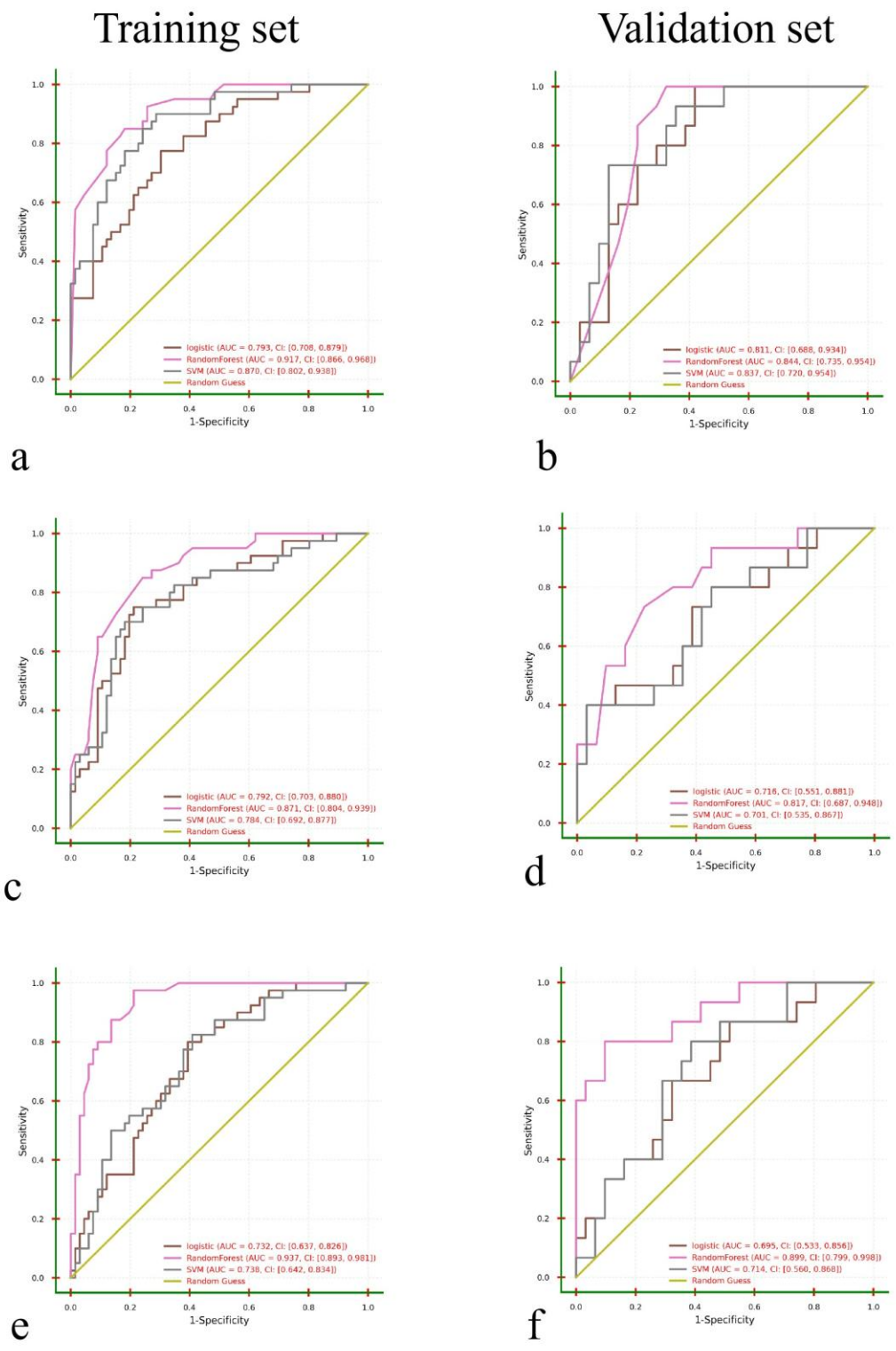


Figure 4. ROC curves of the LR, SVM, and RF radiomics models for the non-contrast (a–b), arterial (c–d), and venous (e–f) phases

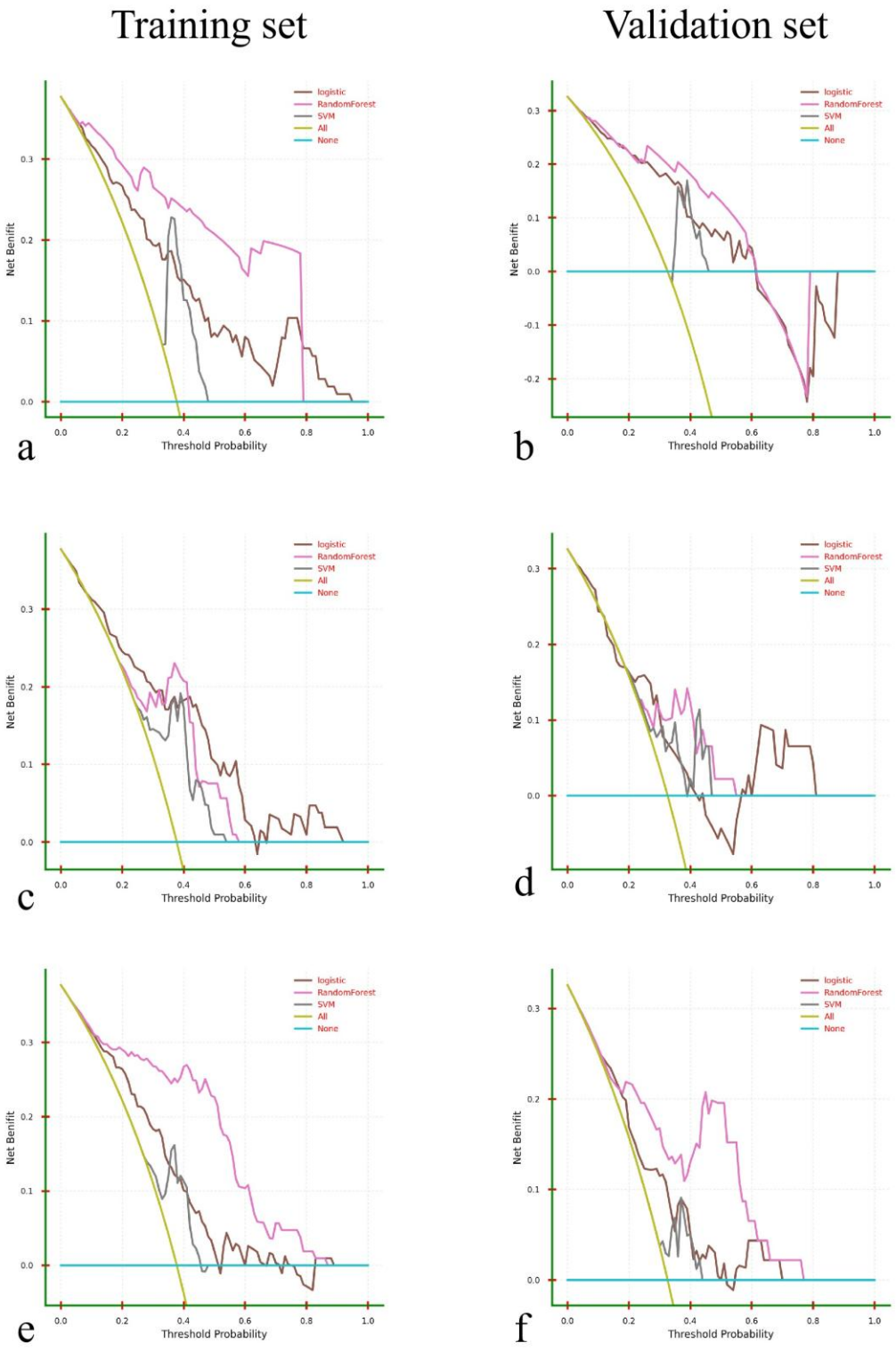


Figure 5. DCA curves of the LR, SVM, and RF radiomics models for the non-contrast (a–b), arterial (c–d), and venous (e–f) phases

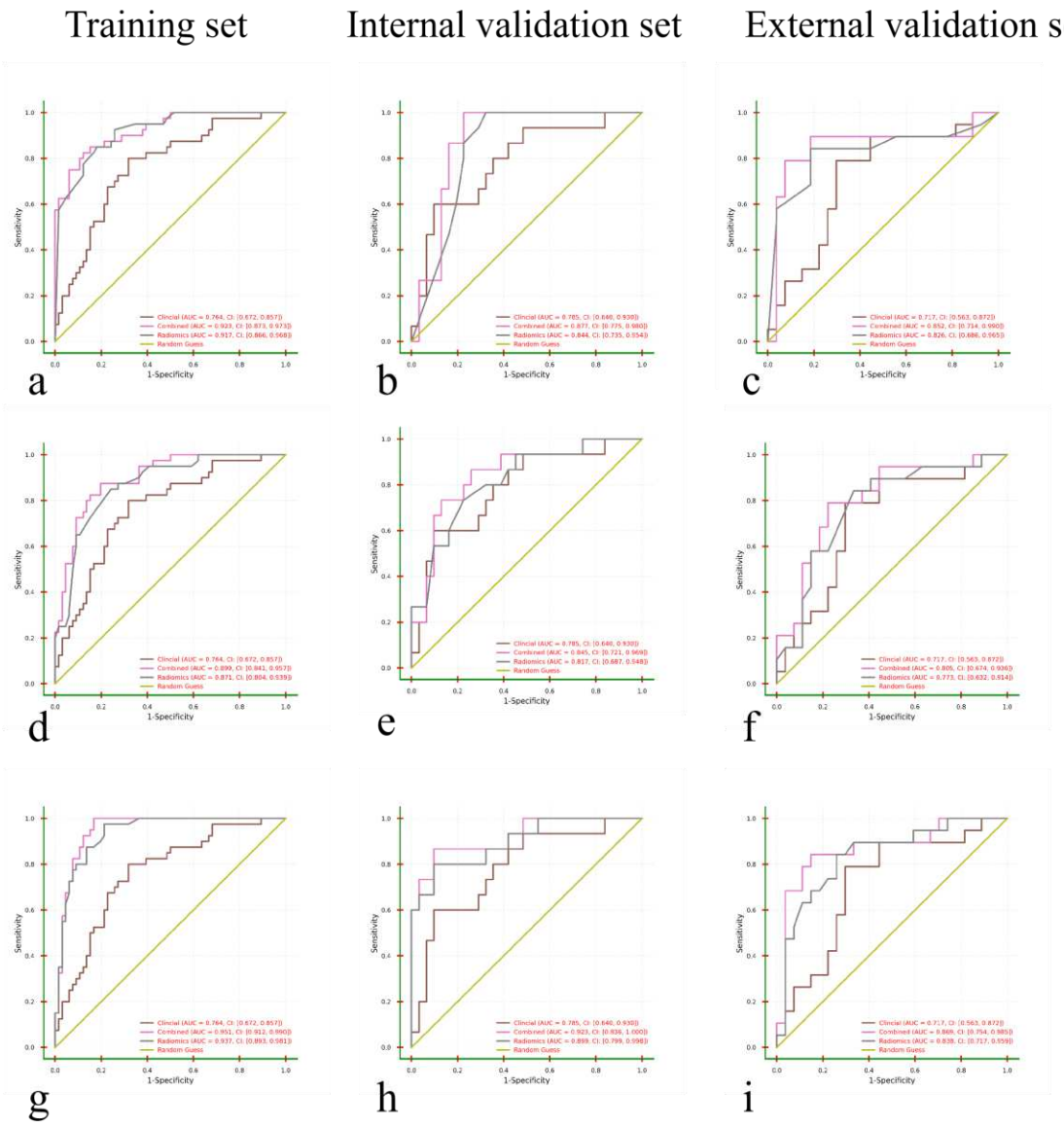


Figure 6. ROC curves of the clinical, radiomics, and combined models for the non-contrast (a–c), arterial (d–f), and venous (g–i) phases

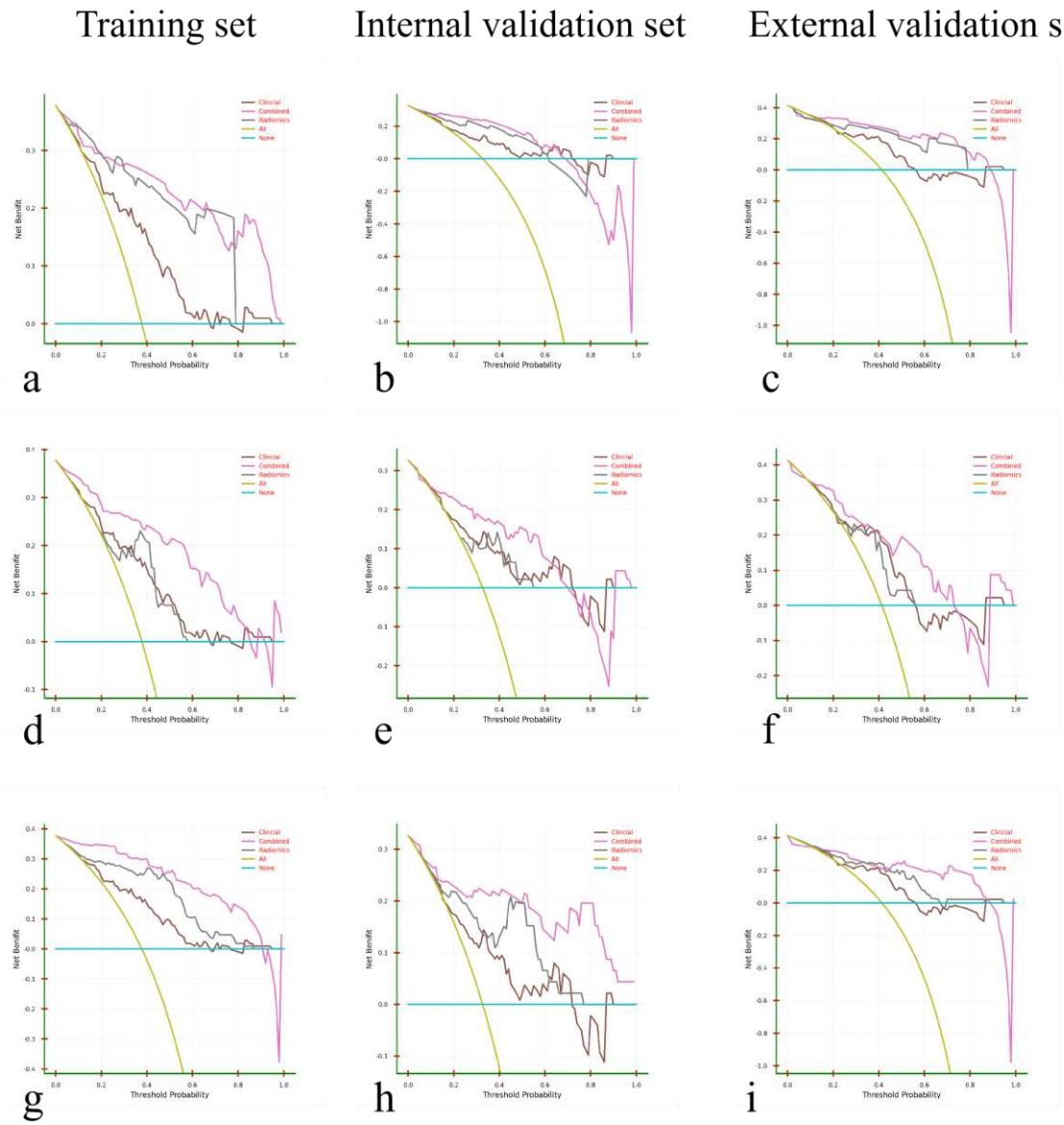


Figure 7. DCA curves of the clinical, radiomics, and combined models for the non-contrast (a–c), arterial (d–f), and venous (g–i) phases

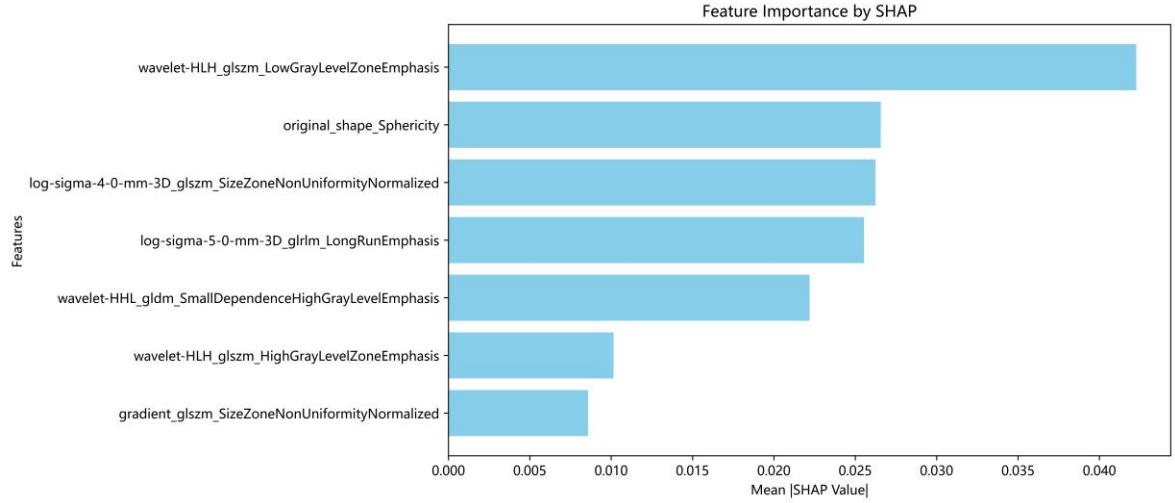


Figure 8. Ranking of radiomic feature importance based on mean absolute SHAP values

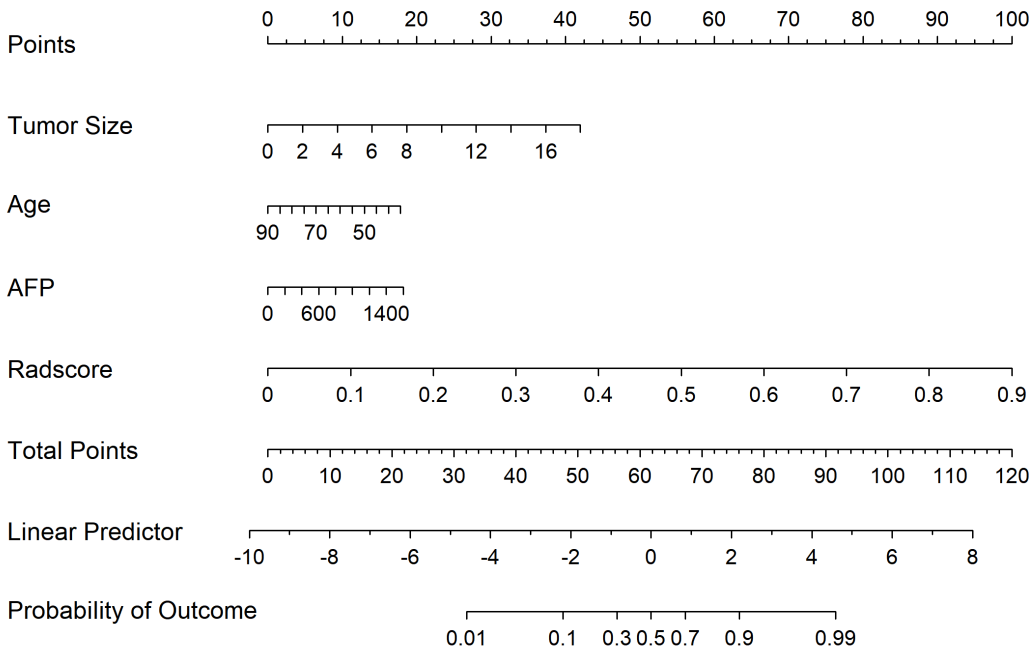


Figure 9. Nomogram for predicting EHM in patients with HCC after TACE.

Tables

Table 1. CT scanning parameters

	Tube voltage (kV)	Tube current (mAs)	Rotation time (s)	Collimation (mm)	Matrix	Slice thickness (mm)	Arterial phase (s)	Venous phase (s)	Contrast flow rate (mL/s)
GE Revolution CT	100	Automatic	0.5	0.625	512×512	5	35	70	2.5
Siemens Somatom Definition Flash	120	150	0.5	0.5	512×512	5	35	70	3.0
Philips Brilliance iCT	120	150	0.5	0.625	512×512	5	35	70	3.0

Table 2. Demographic and clinical characteristics of the 222 patients

Characteristic	Model development cohort			External validation cohort
	Training set	Internal test set	Pvalue	
Sex			0.145	
Male	88(83.0%)	42(91.3%)		61(87.7%)
Female	18(17.0%)	4(8.7%)		9(12.9%)
Age (years)	62.9±9.4	60.6±10.4	0.098	60.6±9.9
Hepatitis			1.000	
Yes	90(84.9%)	39(84.8%)		62(88.6%)
No	16(15.1%)	7(15.2%)		8(11.4%)
Cirrhosis			0.524	
Yes	81(76.4%)	38(82.6%)		60(85.7%)
No	25(23.6%)	8(17.4%)		10(14.3%)
Child-Pugh class			0.937	
A	90(84.9%)	40(87.0%)		52(74.3%)
B	16(15.1%)	6(13.0%)		18(25.7%)
BCLC stage			0.381	
A	36(34.0%)	15(32.6%)		16(22.9%)
B	42(39.6%)	23(50.0%)		36(51.4%)
C	28(26.4%)	8(17.4%)		18(25.7%)
RBC($\times 10^{12}$/L)	4.4±0.6	4.4±0.8	0.923	4.3±0.7
WBC($\times 10^{9}$/L)	6.2±2.7	6.4±2.4	0.446	5.9±2.3
NEUT($\times 10^{9}$/L)	4.2±2.3	4.5±2.4	0.439	4.0±2.1
LYM($\times 10^{9}$/L)	1.3±0.6	1.3±0.4	0.963	1.3±0.6
PLT($\times 10^{9}$/L)	169.4±88.6	185.0±83.0	0.182	156.8±73.4
NLR	3.6±2.5	4.1±3.5	0.341	3.5±2.1
PLR	140.3±78.9	159.2±85.0	0.263	133.2±68.4
INR	1.1±0.2	1.1±0.1	0.193	1.2±0.2
PT(s)	12.6±1.8	12.2±1.5	0.168	12.80±1.8
TT(s)	15.1±1.7	15.0±1.7	0.782	15.6±1.8
ALT(U/L)	69.8±57	72.0±115.3	0.445	62.9±94.4
AST(U/L)	68.2±45.9	68.6±54.9	0.490	70.0±53.8
AFP(ng/mL)	212.6±409.5	245.9±423.6	0.634	328.6±524.8
TBIL(μmol/L)	20.4±9.8	20.7±17.6	0.449	21.0±15.5
ALB(g/L)	38.1±5.1	39.3±5.3	0.208	38.5±6.4
ALP(U/L)	130.3±73.1	129.7±74.8	0.549	142.9±87.5
GGT(U/L)	155.1±129.3	144.3±136.5	0.438	124.4±82.4
Ascites			0.572	
Present	8(8.5%)	2(4.3%)		7(10.0%)
Absent	97(91.5%)	44(95.7%)		63(90.0%)
Maximum tumor diameter (cm)	8.1±4.0	8.9±4.5	0.305	8.2±3.9
Number of tumors			0.673	

Solitary	58(54.7%)	23(50.0%)		30(42.9%)
Multiple	48(45.3%)	23(50.0%)		40(57.1%)
Tumor thrombus			0.380	
Present	27(25.5%)	8(17.4%)		18(25.7%)
Absent	79(74.5%)	38(82.6%)		52(74.3%)

Note: Data are presented as n (%) or mean ± standard deviation.

Table 3. Univariate and multivariate logistic regression analyses in the training set

Variable	Univariate		Multivariate	
	OR(95% CI)	Pvalue	OR(95% CI)	Pvalue
Sex	0.832(0.585-1.183)	0.673		
Age	0.932(0.889-0.977)	0.004	0.932(0.885-0.980)	0.007
BCLC stage	1.731(1.024-2.924)	0.040		
Child-Pugh class	0.988(0.330-2.963)	0.983		
RBC(×10^{12}/L)	1.175(0.639-2.162)	0.603		
WBC(×10^9/L)	1.059(0.914-1.226)	0.445		
NEUT(×10^9/L)	1.090(0.918-1.294)	0.327		
LYM(×10^9/L)	0.746(0.365-1.527)	0.423		
NLR	1.131(0.964-1.328)	0.132		
PLR	1.006(1.001-1.012)	0.023		
INR	1.266(0.104-15.471)	0.853		
PLT(×10^9)	1.004(0.999-1.008)	0.111		
PT(s)	1.010(0.808-1.264)	0.927		
TT(s)	0.915(0.724-1.156)	0.455		
ALT(U/L)	0.997(0.989-1.005)	0.461		
AST(U/L)	1.002(0.994-1.011)	0.544		
AFP(ng/mL)	1.001(1.000-1.002)	0.020	1.002(1.000-1.003)	0.009
TBIL(μmol/L)	0.972(0.932-1.015)	0.208		
ALB(g/L)	1.002(0.928-1.083)	0.953		
ALP(U/L)	1.000(0.995-1.006)	0.858		
GGT(U/L)	1.002(0.999-1.006)	0.111		
Hepatitis	0.744(0.254-2.185)	0.591		
Cirrhosis	1.102(0.434-2.798)	0.838		
Ascites	2.214(0.558-8.789)	0.258		
Number of tumors	1.971(0.915-4.249)	0.083		
Maximum tumor diameter (cm)	1.133(1.021-1.256)	0.018	1.160(1.036-1.300)	0.010
Tumor thrombus	1.457(0.600-3.542)	0.406		

Table 4. AUCs of the LR, SVM, and RF radiomics models

		LR model	SVM model	RF model
Non-contrast	Training set	0.793(0.707-0.879)	0.870(0.802-0.938)	0.917(0.866-0.968)
	Test set	0.811(0.688-0.934)	0.837(0.719-0.955)	0.844(0.735-0.954)
Arterial phase	Training set	0.792(0.703-0.880)	0.784(0.692-0.877)	0.871(0.804-0.939)
	Test set	0.716(0.551-0.881)	0.701(0.535-0.867)	0.817(0.687-0.948)
Venous phase	Training set	0.732(0.638-0.826)	0.738(0.642-0.834)	0.937(0.893-0.981)
	Test set	0.695(0.533-0.856)	0.714(0.560-0.868)	0.899(0.799-0.998)

Table 5. Predictive performance of the clinical, radiomics, and combined models

Phase	Model	Training set						Test set					
		AUC	Accuracy	Sensitivity	Specificity	Positive predictive value (PPV)	Negative predictive value (NPV)	AUC	Accuracy	Sensitivity	Specificity	Positive predictive value (PPV)	Negative predictive value (NPV)
Clinical		0.764 (0.672-0.857)	0.726 (0.631-0.808)	0.800 (0.644-0.910)	0.681 (0.556-0.791)	0.604 (0.460-0.736)	0.849 (0.724-0.933)	0.785 (0.632-0.908)	0.652 (0.498-0.787)	0.933 (0.681-0.998)	0.516 (0.331-0.699)	0.482 (0.295-0.675)	0.941 (0.713-0.999)
Non-contrast	RF model	0.917 (0.866-0.968)	0.830 (0.745-0.896)	0.850 (0.702-0.943)	0.818 (0.704-0.902)	0.739 (0.589-0.857)	0.900 (0.795-0.962)	0.844 (0.735-0.954)	0.783 (0.636-0.891)	1.000 (0.782-1.000)	0.677 (0.486-0.833)	0.600 (0.387-0.789)	1.000 (0.839-1.000)
	Combined model	0.923 (0.874-0.974)	0.858 (0.777-0.918)	0.825 (0.672-0.927)	0.879 (0.775-0.946)	0.804 (0.651-0.912)	0.892 (0.791-0.956)	0.877 (0.775-0.981)	0.804 (0.661-0.906)	1.000 (0.782-1.000)	0.710 (0.520-0.858)	0.625 (0.406-0.812)	1.000 (0.846-1.000)
Arterial phase	RF model	0.871 (0.804-0.939)	0.792 (0.703-0.865)	0.850 (0.702-0.943)	0.758 (0.636-0.855)	0.680 (0.533-0.805)	0.893 (0.781-0.960)	0.817 (0.687-0.948)	0.717 (0.565-0.840)	0.800 (0.519-0.957)	0.677 (0.486-0.833)	0.546 (0.322-0.756)	0.875 (0.796-0.973)
	Combined model	0.899 (0.841-0.957)	0.830 (0.745-0.896)	0.875 (0.732-0.958)	0.803 (0.687-0.891)	0.729 (0.582-0.847)	0.914 (0.810-0.971)	0.845 (0.721-0.969)	0.761 (0.612-0.874)	0.867 (0.595-0.983)	0.710 (0.520-0.858)	0.591 (0.364-0.793)	0.917 (0.730-0.990)
Venous phase	RF model	0.937 (0.893-0.981)	0.858 (0.777-0.919)	0.975 (0.868-0.999)	0.788 (0.670-0.879)	0.746 (0.597-0.847)	0.981 (0.899-0.999)	0.899 (0.799-0.998)	0.761 (0.612-0.874)	0.800 (0.519-0.957)	0.742 (0.554-0.881)	0.600 (0.361-0.809)	0.885 (0.699-0.976)
	Combined model	0.951 (0.912-0.990)	0.896 (0.822-0.947)	1.000 (0.912-1.000)	0.833 (0.721-0.914)	0.784 (0.647-0.887)	1.000 (0.935-1.000)	0.923 (0.823-0.993)	0.739 (0.589-0.857)	0.867 (0.595-0.983)	0.677 (0.486-0.833)	0.565 (0.345-0.768)	0.913 (0.719-0.989)

Table 6. Performance of the prediction models in the training, internal test, and external validation sets

	Training set	Internal test set	External validation set
	AUC	AUC	AUC
Clinical	0.764(0.672-0.857)	0.785(0.632-0.908)	0.688(0.542-0.835)
RF model	0.937(0.893-0.981)	0.899(0.799-0.998)	0.840(0.749-0.932)
Combined model	0.951(0.912-0.990)	0.923(0.823-0.993)	0.878(0.801-0.957)

Supplemental Materials

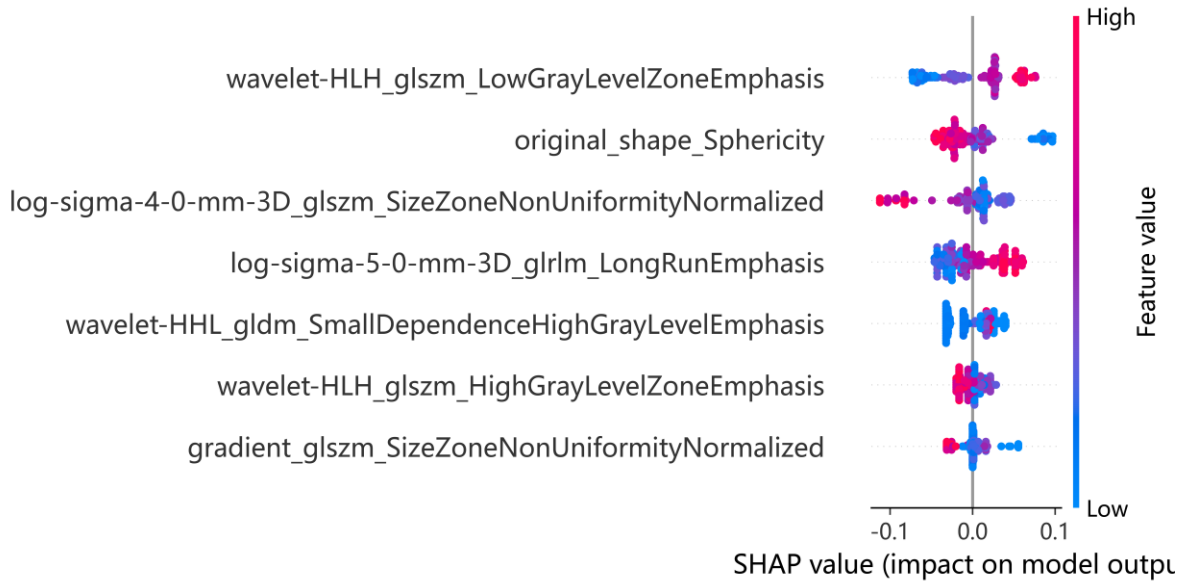
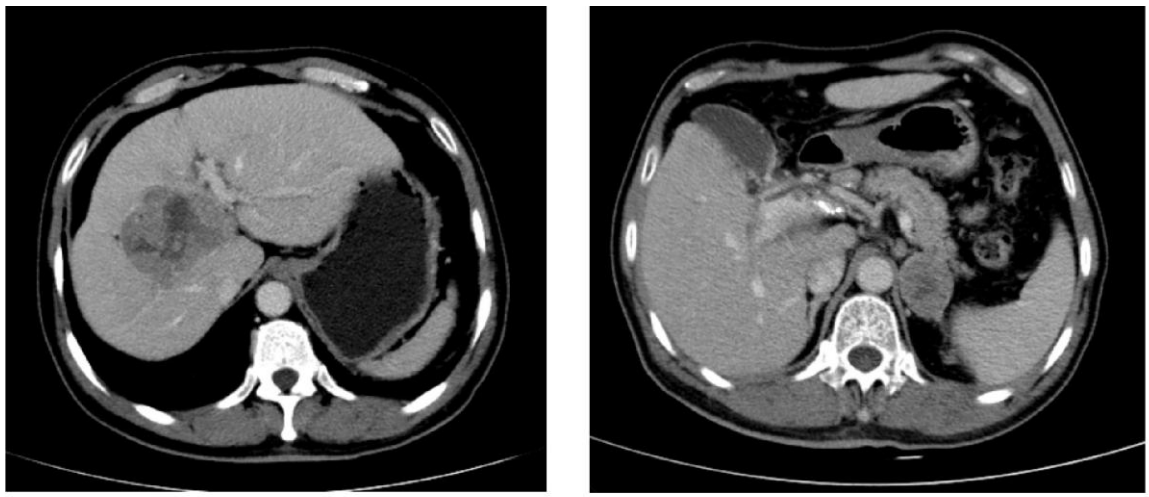
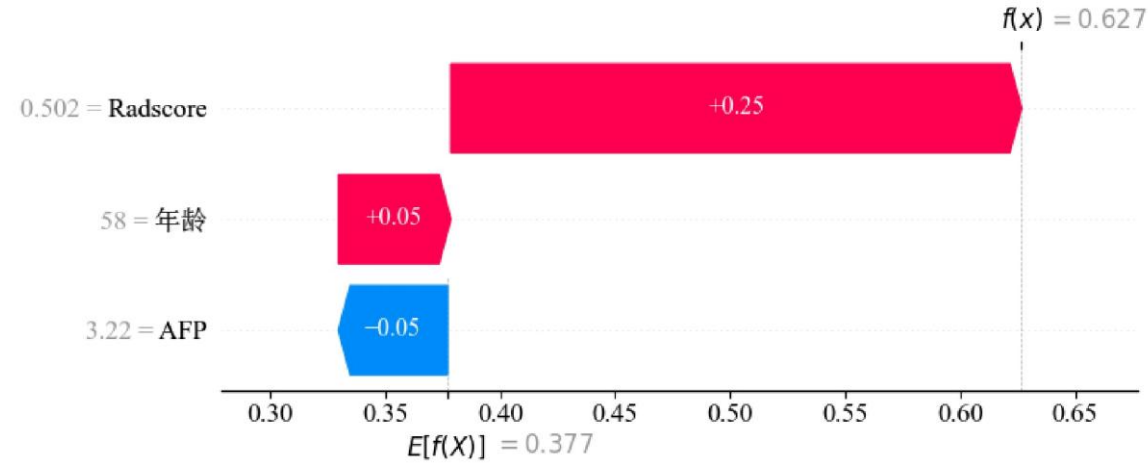


Figure S1. Beeswarm plot of the SHAP value distribution of each radiomic feature



a b



c

Figure S2. Representative case of extrahepatic metastasis in a patient with intermediate-to-advanced HCC Supplementary Figure S2: A 55-year-old man with HCC. The preoperative AFP level was normal (3.2 ng/mL), and liver reserve function was good (Child-Pugh class A). (A) A solitary tumor (maximum diameter, 6.7 cm), consistent with conventional BCLC stage B. However, the SHAP plot generated for this patient (C) showed a predicted probability of 0.627; among all variables, the Radscore made the largest contribution in the SHAP analysis (SHAP value = 0.25), offsetting the negative effect of the low AFP level. Fourteen months after the first TACE, follow-up imaging revealed a metastatic lesion in the left adrenal gland (B). This case demonstrates that, in patients who have not yet exhibited any clinical high-risk features, radiomics can capture metastatic risk not reflected by conventional clinical indicators, enabling early and accurate warning of metastatic potential.