Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

MRI-Based Deep Learning and Radiomics Nomogram for Predicting Hepatocellular Carcinoma Recurrence Within Six Months After Thermal Ablation

Yao Chen[1], Yanan Zhao[1], Weiwei Guan[1], Di Wu[2], Lin Zheng[1], Chengshi Chen[1], Xiang Geng[1], Han Qi[3], Ho-Young Song[1], Hongtao Hu[1]

[1]Department of Interventional Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, Henan, People's Republic of China; [2]Department of Radiology, People's Hospital of Zhengzhou, Zhengzhou, Henan, People's Republic of China; [3]Department of Minimally Invasive Interventional Therapy, Sun Yat-Sen University Cancer Center, Guangzhou, Guangdong, People's Republic of China

Correspondence: Hongtao Hu, Department of Interventional Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, 127 Dongming Road, Zhengzhou, Henan, 450008, People's Republic of China, Email huhongtaogy@163.com

Purpose: Develop a magnetic resonance imaging (MRI)-based deep learning (DL)-radiomics (Rad)-clinical nomogram for predicting early recurrence of hepatocellular carcinoma (HCC) within six months after thermal ablation.

Materials and Methods: Barcelona Clinic Liver Cancer (BCLC) stage 0-A HCC patients who underwent dynamic contrast-enhanced MRI before ablation were retrospectively included. Patients were categorized into non early recurrence and early recurrence groups. A clinical model was constructed through logistic regression analysis of clinical information and radiological features. DL score model and Rad score model were developed using DL features and manual features extracted from dynamic contrast-enhanced MRI, with principal component analysis and least absolute shrinkage and selection operator regression methods. The DL-Rad-Clinical nomogram was constructed through logistic regression analysis. The model performance was primarily evaluated using the area under the receiver operating characteristic curve (AUC).

Results: A total of 224 patients were included in this study (training set: n = 156; test set: n = 68). The DL-Rad-Clinical nomogram was constructed, including Rad score, DL score, natural logarithm alpha-fetoprotein (LnAFP), and multiple low signal lesions as predictive factors. In the training set, the DL-Rad-Clinical nomogram demonstrated better predictive performance (AUC = 0.896, $P < 0.05$). In the test set, the DL-Rad-Clinical nomogram had a higher AUC value compared to other models, although the difference was not statistically significant (AUC = 0.774, $P > 0.05$).

Conclusion: The DL-Rad-Clinical nomogram helped in identifying HCC patients with early recurrence within six months following thermal ablation.

Keywords: HCC, prediction, clinical information, radiological features

Introduction

In 2022, primary liver cancer remained one of the leading causes of new cancer cases and cancer-related deaths in China, ranking fourth for new cases and second for deaths.[1] Hepatocellular carcinoma (HCC) accounts for 75% −85% of primary liver cancer.[2] For very Barcelona Clinic Liver Cancer (BCLC) stage 0-A HCC, resection, ablation, and liver transplantation are the primary methods for achieving radical treatment.[3] Although radical treatments can improve survival, tumor recurrence remains a key factor contributing to increased mortality in HCC patients. Studies show that over 70% of HCC patients experience early or late recurrence within five years after resection or ablation treatment.[4] Recently, some studies suggest that lower circadian rhythm disruption and Ginsenoside Rh1 may help improve the prognosis of HCC patients.[5,6] How to further improve the efficacy of ablation is one of the current research hotspots, and further exploration is urgently needed.

Received: 19 May 2025
Accepted: 23 September 2025
Published: 6 October 2025

Early detection, early diagnosis, and early treatment are crucial for improving the prognosis of HCC patients.[7] In clinical practice, we found that some patients who underwent ablation experienced intrahepatic recurrence within six months. Study showed that HCC patients who experience recurrence within 6 months after resection have a significantly shortened survival time.[8] Tumor characteristics, including size, multifocality, and serum alpha-fetoprotein (AFP) levels, have been identified as strong predictors of recurrence and survival in the HCC population.[9–12] There have been many studies exploring the risk factors for recurrence within 12 or 24 months after ablation. However, research from the East, there is limited knowledge about the predictive factors for recurrence within six months after ablation therapy in HCC patients. As an emerging technology, radiomics, a bridge connecting medical images with precision medicine,[13] can transform potential pathological and physiological information in medical images into high-dimensional quantitative data,[14–16] providing a basis for prognosis prediction. Deep learning (DL) technology is also a recently popular method for extracting internal information from medical images. Leveraging its powerful self-learning capabilities, it can automatically learn features from imaging data based on clinical labels, quantify these features, and enrich the variety of predictive factors.[17] Published studies have shown that deep learning features and radiomics features have a synergistic effect, which can further enhance the model's predictive ability.[18] Accurately assessing the risk of recurrence within six months after ablation for HCC patients by predicting factors holds great promise in reducing the recurrence rate of ablation. However, so far, no relevant research has been published.

Therefore, the purpose of this study was to identify the predictive factors for early recurrence within six months after ablation in BCLC stage 0-A HCC patients and to develop a recurrence risk prediction model incorporating DL, radiomics, and clinical factors.

Materials and Methods

Patients

The study was conducted on consecutive HCC patients who underwent computed tomography (CT) - guided thermal ablation (radiofrequency ablation and microwave ablation) in our hospital from January 2018 to July 2021. The diagnosis of HCC was based on the European Association for the Study of the Liver clinical practice guidelines.[19] The inclusion criteria were as follows: (1) patients with first diagnosis or recurrence after radical treatment and refusing to resection, (2) meeting the Milan criteria: diameter ≤ 3 cm and ≤ 3 tumors; single tumor < 5 cm[20] (3) performance status Eastern Cooperative Oncology Group score 0, (4) Child-Pugh A class, (5) no other malignancies, (6) preoperative magnetic resonance imaging (MRI) within a month, (7) One month after the procedure, confirmed complete ablation of target lesion through multi-phase contrast-enhanced CT, dynamic contrast-enhanced MRI, or contrast-enhanced ultrasound, and (8) post-ablation follow-up until the first recurrence (local and non-local) or at least six months. The exclusion criteria were as follows: (1) lack of complete imaging data, (2) motion artifacts on MRI, (3) concurrent transcatheter arterial chemoembolization. The study recruitment process and workflow were shown in Figures 1 and 2. Patients were randomly divided into the training and test sets at a ratio of 7:3. The study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Medical Ethics Committee of the Affiliated Cancer Hospital of Zhengzhou University (2024–015-002) and individual consent for this retrospective analysis was waived. In this study, we strictly adhered to the principle of patient information confidentiality. When handling information related to the implementation of the study, we assigned each patient a unique research number. In presenting the research results, we ensured that no information that could directly identify the research subjects was included. We did not use any information obtained from the research subjects for any purposes beyond the study.

Clinical Characteristics

Preoperative clinical characteristics, including age, sex, tumor number, maximum tumor diameter, multiple low-signal lesions, the natural logarithm of AFP (LnAFP), alanine aminotransferase, aspartate aminotransferase, albumin, gamma-glutamyl transpeptidase, total bilirubin, prothrombin time, platelet, neutrophil to lymphocyte radio, Child-Pugh class, albumin-bilirubin score,[21] model for end-stage liver disease score,[22] chronic viral hepatitis, portal hypertension, dangerous position,[23] history of treatment and complication, were obtained from electronic medical records. Laboratory examination results were obtained

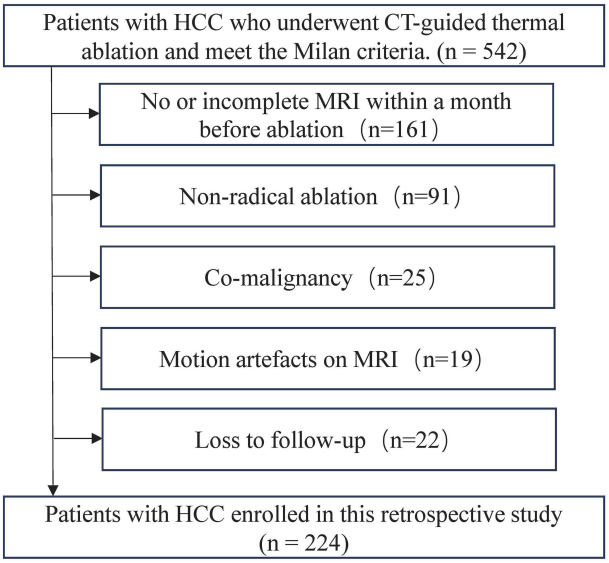

Figure 1 The study recruitment process.
Abbreviations: HCC, hepatocellular carcinoma; CT, computerized tomography; MRI, magnetic resonance imaging.

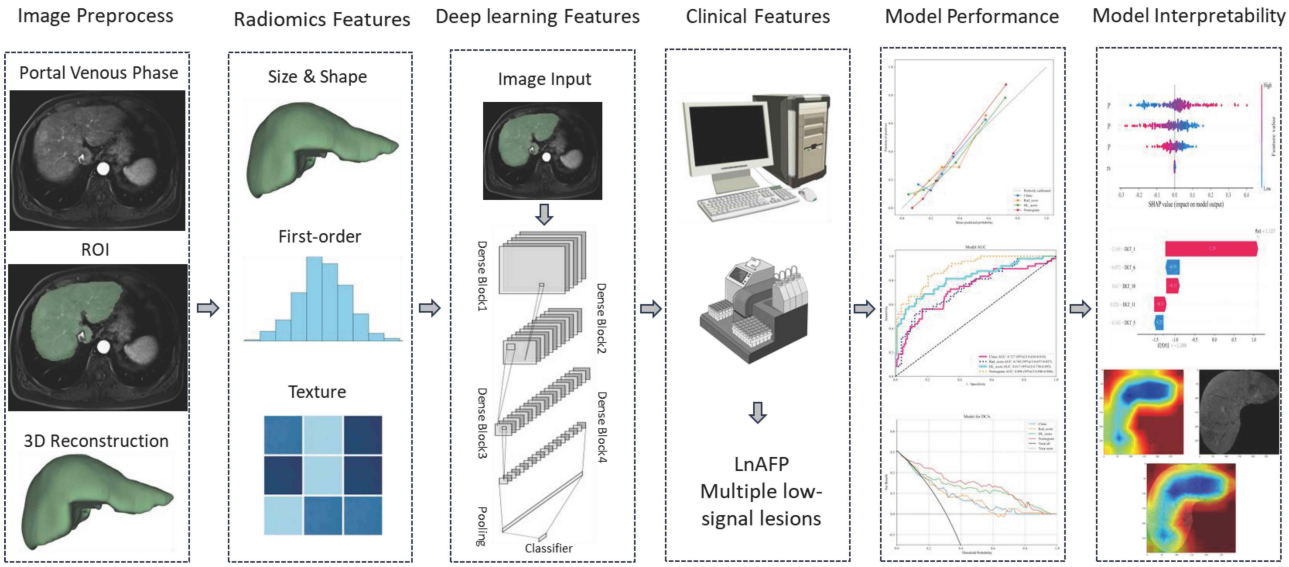

Figure 2 The workflow of study. The process included image preprocess, radiomics feature extraction, deep learning feature extraction, clinical feature collection, model construction and evaluation, and model interpretability.
Abbreviations: ROI, region of interest; 3D, three dimensions; AFP, alpha-fetoprotein.

from blood tests within 2 weeks before ablation. Multiple low-signal lesions were defined as having three or more nodules with low signal intensity in either the arterial or portal phase, despite the absence of typical imaging features of HCC.

Ablation Procedure

Interventional physicians with over five years of experience in thermal ablation developed a tailored treatment plan based on factors such as tumor size, number, and location. The procedure was performed under general anesthesia with CT guidance. Once the patient was anesthetized, a CT scan was used to determine the optimal puncture point, path, angle, and depth for the ablation needle. The treatment employed a single-needle, multi-point overlapping technique. The physician's intent was to completely ablate the tumor and, to the extent possible, cover the surrounding 0.5 cm of peritumoral tissue. After completing the ablation, the needle tract was also ablated to minimize the risk of tumor spread and bleeding.

https://doi.org/10.2147/JHC.S541329

Follow-Up

After ablation, all patients were regularly followed up every one to two months during the first six months, and every two to three months thereafter, based on imaging examinations (multi-phase contrast-enhanced CT, dynamic contrast-enhanced MRI, or contrast-enhanced ultrasound), AFP levels, or biopsy results, until the first recurrence or for at least six months. Early recurrence within six months was defined as the time from the date of ablation treatment to the detection of intrahepatic recurrence, and this time should not exceed six months.

MRI Processing

All patients fasted overnight. MRI scans were conducted with the patient in the supine position, during a breath-hold. Imaging was performed using one of three 3.0 T MRI platforms: two Siemens Healthineers scanners (MAGNETOM Skyra and Prisma, Erlangen, Germany) and one GE Healthcare scanner (SIGNA Architect, Waukesha, WI, USA). The sequences included T2-weighted fast spin echo (fat-suppressed T2-weighted imaging) and fast gradient echo (dynamic contrast-enhanced imaging [DCEI]). For DCEI, an intravenous injection of 0.1 mmol/kg Gd-DTPA was administered at a rate of 2.5 mL/s, followed by 20 mL of saline, using an MR-compatible injector (Medrad, Warrendale, PA, USA).

Considering that MRI scans were performed at different equipment and there was some inhomogeneity between scanners, the image registration, N4ITK MRI bias correction, signal intensity normalization, voxel resampling (with $1\times1\times1$ mm^3) and bin width (with five signal intensity) were performed to minimize the impact.

The DCEI was divided into arterial phase (20–25 seconds), portal venous phase (55–60 seconds) and delay phase (180 seconds). Subsequently, T2WI, arterial phase, portal venous phase and delay phase were exported as Digital Imaging and Communication in Medicine files. The region of interest (ROI) range, defined as the entire liver region excluding the primary branch of the intrahepatic Glisson system in each horizon-slice, was semi-automatically segmented by an interventional radiologist using 3D Slicer (https://www.slicer.org/), and reviewed by a radiologist to minimize potential bias. Disagreements were verified by a senior expert. The radiomics features were extracted using the PyRadiomics package (version 3.0.1).[24]

Based on the annotated ROI mask, each phase of the standardized MRI was cropped layer by layer to retain only the ROI region. DenseNet-169 model was then used for transfer learning, with pre-training on the ImageNet dataset to obtain initial weights.[25] The training process included forward propagation and backpropagation, with real-time data augmentation techniques such as random horizontal flipping and cropping. The model parameters were updated using the Stochastic Gradient Descent optimizer, with an initial learning rate of 0.01, decayed by the cosine annealing algorithm 50 epochs and 244 iteration steps, using a batch size of 32. The schematic diagram of the DenseNet-169 model was shown in Figure S1.

Radiomics Feature and Deep Learning Feature Selection

The Mann–Whitney *U*-test ($p < 0.05$) was used for the initial screening of radiomics features. The correlation coefficients between the remaining features were then calculated, and one feature from each pair with a correlation coefficient above 0.90 was retained. Next, the remaining features were normalized using the z-score and further refined using the least absolute shrinkage and selection operator (lasso) at the minimum lambda value. Additionally, DL features were extracted from the final pooling layer of the DenseNet-169 model and selected using principal component analysis and the lasso regression at the minimum lambda value, which reduced the number of features.

Development and Evaluation of Predictive Models

Key radiomics and DL features from T2WI, arterial phase, portal venous phase, and delay phase were used to create the radiomics (Rad) score and DL score models using logistic regression. Clinical data were analyzed using univariate and multivariate logistic regression to identify predictors of early recurrence within six months and to build a clinical model. The Rad score, DL score, and clinical models were then combined to construct the DL-Rad-Clinical nomogram with logistic regression. The model's predictive performance was assessed using sensitivity, specificity, accuracy, negative predictive value, positive predictive value, and the area under the receiver operating characteristic curve (AUC). Calibration performance was

evaluated by comparing predicted and actual early recurrence within six months using the calibration curve, and clinical usefulness was determined through the net benefit analysis across various threshold probabilities in decision curve analysis.[26]

Model Interpretability

To enhance the transparency of the DL model's decision-making process and to improve its interpretability, we utilized Gradient-weighted Class Activation Mapping (Grad-CAM) for visualization.[27] By using the gradient information from the last convolutional layer of the convolutional neural network, we performed weighted fusion to obtain a class activation map that highlights important regions in the classified target image. At the same time, we further explained the importance of each feature in the model and its impact on the model's prediction probability by calculating Shapley Additive Explanations (SHAP) values.[28] SHAP summary plots offered a global view by quantifying how feature values affect the model's outputs, aiding in the identification of significant features and their trends. Meanwhile, SHAP local bar plots showed the SHAP values for individual test examples, highlighting the influence of each feature on the model's predictions.

Statistical Analysis

The clinical characteristics were analyzed using SPSS 26.0 (IBM Corporation, Armonk, NY, USA), R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria), and Python (version 3.7; Python Software Foundation, Beaverton, USA). Continuous variables were expressed as means ± standard deviations, and categorical variables were expressed as frequencies and percentages. Continuous variables were compared using the Mann–Whitney *U*-test, while categorical variables were compared using Pearson's chi-squared test or Fisher's exact test (2-tailed). Uni- and multivariable logistic regression analyses were performed to identify independent clinical risk factors associated with early recurrence within six months. Variables with a P-value < 0.05 in the univariable analysis were selected as candidates for the multivariable analysis. Odds ratios and 95% confidence intervals were calculated. The DeLong test was used to compare the AUCs. The variance inflation factor was applied to evaluate the collinearity among variables. Differences with P-values < 0.05 were considered statistically significant.

Results

Clinical Characteristics

A total of 224 patients were recruited for this study (training set: n = 156, including 48 with early recurrence and 108 without early recurrence within six months; test set: n = 68, including 21 with early recurrence and 47 without early recurrence within six months). A total of 69 patients (30.8%) were diagnosed with early recurrence within six months. The characteristics of the training and test sets are listed in Table 1. Overall, there were no significant differences in patient characteristics between the two sets.

Feature Selection and Model Construction

A total of 1197 radiomics features were extracted from each sequence, resulting in 4788 features per patient. After screening, seven portal vein phase features were retained. The logistic regression algorithm was then applied to construct the Rad score model. The types and weights of the seven radiomic features were shown in Figure 3a.

Similarly, 1664 DL features were extracted from each image sequence, resulting in a total of 6656 features per patient across four sequences. After filtering, five DL features from the T2WI sequence were retained. These five features were then used as predictors to construct a DL score model using logistic regression. The types and weights of these five DL features were illustrated in Figure 3b.

Uni- and multivariable logistic regression analyses were used to identify preoperative clinical risk factors associated with early recurrence within six months (Table 2). In the multivariable logistic regression model, multiple low-signal lesions and natural logarithm AFP (LnAFP) were included as predictors to construct the clinical model.

Finally, after confirming the absence of collinearity (Table S1), the DL score model, Rad score model, and clinical model were combined into a DL-Rad-Clinical nomogram using the logistic regression algorithm.

Table 1 Baseline Characteristics of Enrolled Patients

Characteristics	All Patients (n=224)	Training Set (n=156)	Test Set (n=68)	P-value
Age (years)	57.9 ± 7.8	58.3 ± 7.4	56.8 ± 8.7	0.11
Sex (male)	157 (70.1%)	108 (69.2%)	49 (72.1%)	0.38
Tumor number (<3)	211 (94.2%)	146 (93.6%)	65 (95.6%)	0.76
Maximum tumor diameter (cm)	2.0 ± 0.8	2.0 ± 0.8	1.9 ± 0.7	0.93
Multiple low-signal lesions (yes)	116 (51.8%)	85 (54.5%)	31 (45.6%)	0.25
LnAFP	2.6 ± 2.0	2.6 ± 2.1	2.7 ± 2.0	0.78
ALT (U/L)	40.7 ± 53.4	41.5 ± 55.5	38.7 ± 48.6	0.39
AST (U/L)	37.6 ± 39.0	39.0 ± 40.9	34.2 ± 34.2	0.10
ALB (g/L)	40.7 ± 4.7	40.6 ± 4.8	41.1 ± 4.7	0.39
GGT (U/L)	58.6 ± 57.0	60.1 ± 59.5	55.1 ± 50.8	0.39
TBIL (μmol/L)	17.6 ± 9.5	17.3 ± 8.7	18.2 ± 11.2	0.94
PT (seconds)	12.7 ± 1.2	12.8 ± 1.3	12.6 ± 1.1	0.27
PLT (10^9)	118.1 ± 49.0	116.7 ± 50.7	121.4 ± 45.2	0.23
NEU/LYM	3.2 ± 3.2	3.3 ± 3.5	2.8 ± 2.3	0.07
Child-Pugh class (A)	220 (98.2%)	154 (98.7%)	66 (97.1%)	0.59
ALBI score	−2.7 ± 0.4	−2.7 ± 0.4	−2.7 ± 0.5	0.46
MELD score	17.4 ± 2.4	17.4 ± 2.3	17.2 ± 2.5	0.50
Chronic viral hepatitis (yes)	216 (96.4%)	150 (96.2%)	66 (97.1%)	1.000
Portal hypertension (yes)	92 (41.1%)	71 (45.5%)	21 (30.9%)	0.054
Dangerous position (yes)	76 (33.9%)	50 (32.1%)	26 (38.2%)	0.44
History of treatment (yes)	217 (96.9%)	151 (96.8%)	66 (97.1%)	0.54
Complication (yes)	8 (3.6%)	3 (1.9%)	5 (7.4%)	0.057

Abbreviations: AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; GGT, gamma-glutamyl transpeptidase; TBIL, total bilirubin; PT, prothrombin time; PLT, platelets; NEU, neutrophil; LYM, lymphocyte; ALBI, albumin-bilirubin; MELD, model for end-stage liver disease.

Performances of Models and Nomogram

The performance of all models in the two sets was shown in Table 3. The AUC values of the models were compared (Figure 4a–d). The DL-Rad-Clinical nomogram demonstrated superior discrimination, with a mean AUC of 0.896 (95% CI, 0.846–0.946) in the training set, which was significantly different from the other models, and an AUC of 0.774 (95% CI, 0.634–0.914) in the test set, which was not significantly different from the other models.

The DL-Rad-Clinical nomogram was shown in Figure 5a. The calibration curve demonstrated good agreement between the predicted and actual early recurrence within six months in the training (Figure 5b) and test (Figure 5c) sets. The decision curve for the DL-Rad-Clinical nomogram showed good performance in terms of clinical application, providing more benefit than either a treat-all or treat-none scheme (Figure 6a and b).

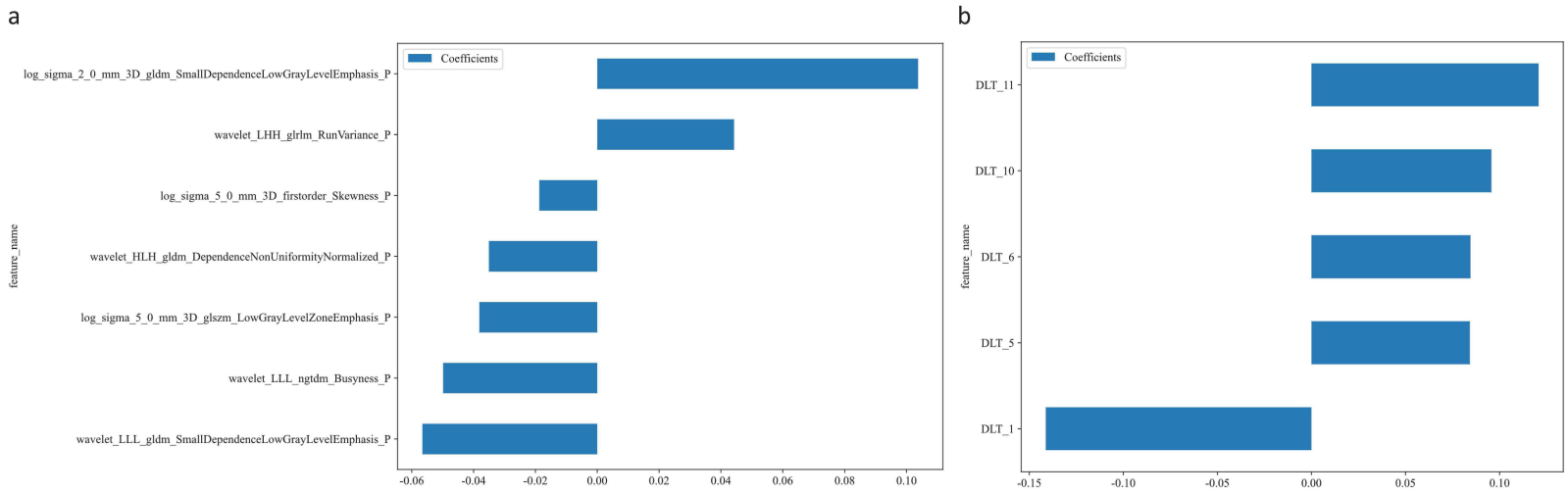

Figure 3 The types and weights of the features. (**a**) showed the variables included in the rad score and the weight coefficient of each variable. (**b**) showed the variables included in the DL score and the weight coefficient of each variable.
Abbreviation: DLT, deep learning transformation.

Table 2 Influencing Clinical Factors of Early Recurrence Within Six Months

Variables	Univariate Analysis		Multivariate Analysis	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Age (years)	0.985 (0.940–1.031)	0.51		
Sex (male)	2.333 (1.139–4.782)	0.02#		
Tumor number (<3)	0.962 (0.238–3.891)	0.96		
Maximum tumor diameter (cm)	1.013 (0.670–1.533)	0.95		
Multiple low-signal lesions (yes)	2.714 (1.311–5.622)	0.007#	2.962 (1.365–6.429)	0.006#
LnAFP	1.390 (1.166–1.657)	<0.001#	1.410 (1.176–1.691)	<0.001#
ALT (U/L)	0.989 (0.976–1.002)	0.10		
AST (U/L)	0.990 (0.976–1.004)	0.17		
ALB (g/L)	0.924 (0.859–0.994)	0.04#		
GGT (U/L)	1.002 (0.997–1.008)	0.45		
TBIL (μmol/L)	0.958 (0.914–1.004)	0.07		
PT (seconds)	0.993 (0.757–1.303)	0.96		
PLT (10^9)	1.002 (0.995–1.009)	0.55		
NEU/LYM	1.005 (0.914–1.106)	0.91		
Child-Pugh class (A)	0.439 (0.027–7.173)	0.56		
ALBI score	1.782 (0.813–3.908)	0.15		
MELD score	0.895 (0.767–1.044)	0.16		
Chronic viral hepatitis (yes)	2.333 (0.454–12.004)	0.31		
Portal hypertension (yes)	1.655 (0.834–3.281)	0.15		
Dangerous position (yes)	1.055 (0.508–2.192)	0.89		
History of treatment (yes)	1.522 (0.246–9.415)	0.65		

Note: #P-value < 0.05.

Abbreviations: AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; GGT, gamma-glutamyl transpeptidase; TBIL, total bilirubin; PT, prothrombin time; PLT, platelets; NEU, neutrophil; LYM, lymphocyte; ALBI, albumin-bilirubin; MELD, model for end-stage liver disease; CI, confidence interval; OR, odds ratio.

Table 3 Performance of the Models in Two Sets

Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy	Set
Clinical	0.727 (0.636–0.818)	0.542	0.833	0.591	0.804	0.744	Train
Rad score	0.742 (0.657–0.827)	0.500	0.880	0.649	0.798	0.763	Train
DL score	0.817 (0.738–0.895)	0.625	0.870	0.682	0.839	0.795	Train
Nomogram	0.896 (0.846–0.946)	0.812	0.796	0.639	0.905	0.801	Train
Clinical	0.703 (0.550–0.857)	0.667	0.681	0.483	0.821	0.676	Test
Rad score	0.711 (0.574–0.848)	0.524	0.830	0.579	0.796	0.735	Test

(*Continued*)

Table 3 (Continued).

Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy	Set
DL score	0.653 (0.490–0.817)	0.524	0.809	0.550	0.792	0.721	Test
Nomogram	0.774 (0.634–0.914)	0.619	0.872	0.684	0.837	0.794	Test

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; DL, deep learning; Rad, radiomics.

Shapley Additive Explanations and Grad-CAM

In the Rad score model, the SHAP summary plot identified wavelet-LHH-glrlm-RunVariance-P, wavelet-LLL-ngtdm-Busyness-P, and log-sigma-5-0-mm-3D-firstorder-Skewness-P as the most important features for predicting early recurrence within six months. A higher value of wavelet-LHH-glrlm-RunVariance-P was associated with a higher SHAP value (Figure 7a), which corresponded to a higher probability of prediction. Conversely, wavelet-LLL-ngtdm-Busyness-P and

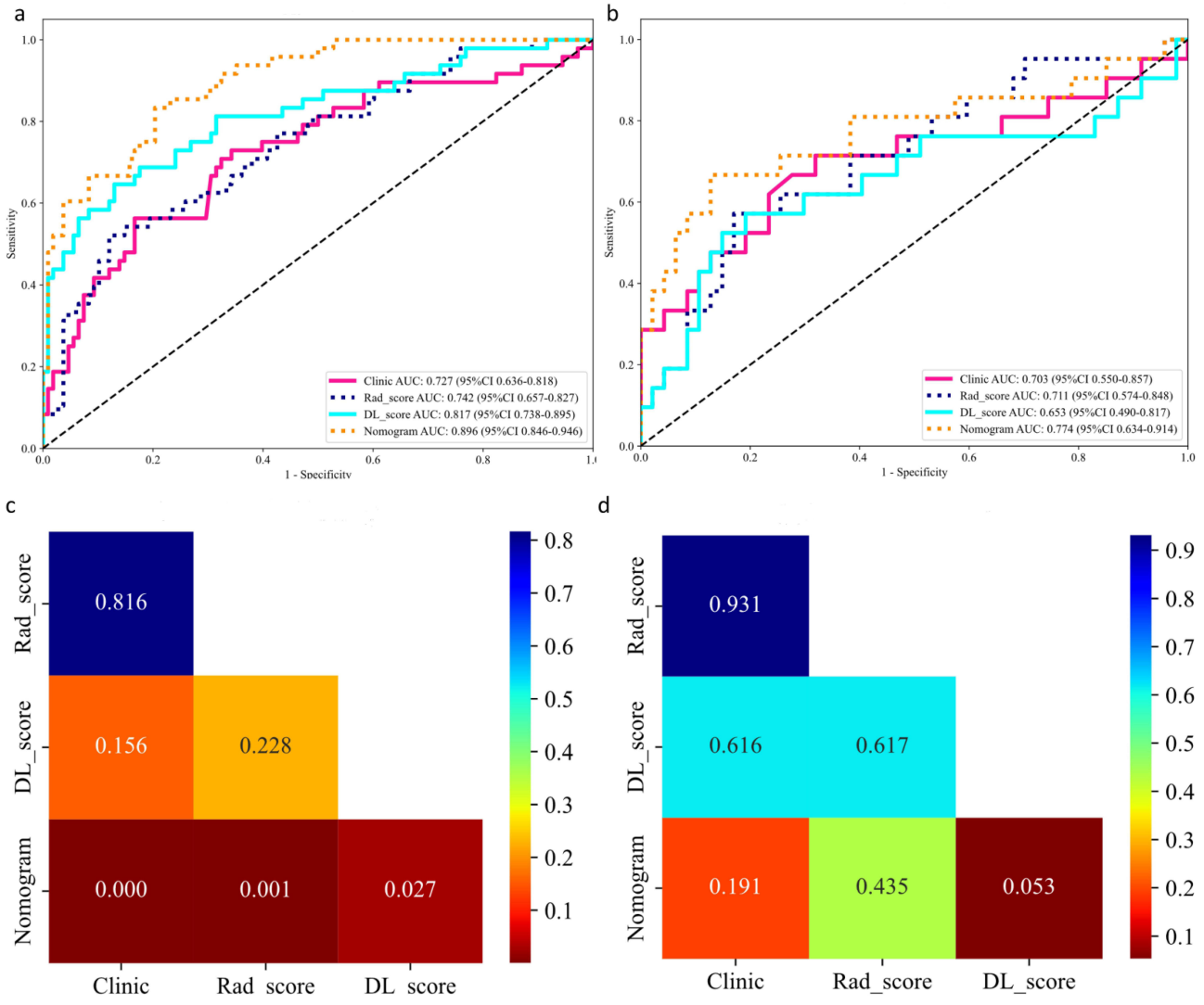

Figure 4 The AUC of the models. (**a**) showed the AUC values and confidence intervals of different models in the training set; (**b**) showed the AUC values and confidence intervals of different models in the test set; (**c**) showed a comparison of AUC values of different models in the training set; (**d**) showed a comparison of AUC values of different models in the test set.

Abbreviations: AUC, area under the receiver operating characteristic curve, Rad, radiomics; DL, deep learning.

https://doi.org/10.2147/JHC.S541329

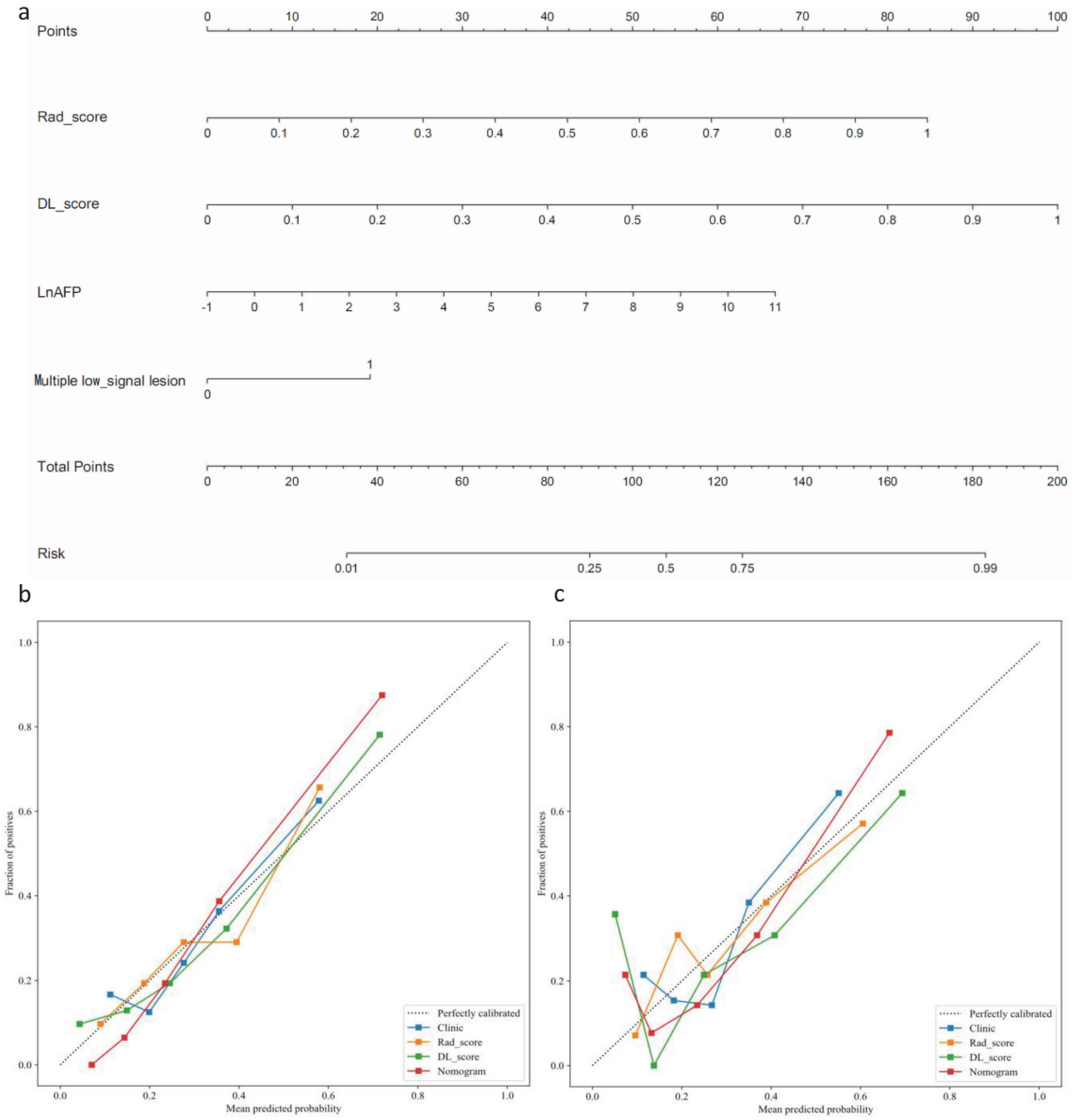

Figure 5 The DL-Rad-Clinical nomogram and calibration curve. (**a**) showed a nomogram with four predictive factors, where the scores of each factor were summed to obtain a total score, and the corresponding risk could be determined from the total score; (**b**) showed the calibration of each model in the training set. The closer the calibration curve of the model was to the dashed line, the better the calibration; (**c**) showed the calibration of each model in the test set.
Abbreviations: Rad, radiomics; DL, deep learning; AFP, alpha-fetoprotein.

log-sigma-5-0-mm-3D-firstorder-Skewness-P showed an inverse relationship. Figure 7b illustrated the SHAP local bar plot, where the case was incorrectly classified, with only log-sigma-5-0-mm-3D-firstorder-Skewness-P contributing correctly.

In the DL score model, the SHAP summary plot revealed that deep learning transformation (DLT) −1, DLT-5, DLT-11, DLT-6, and DLT-10 were the most significant features for predicting early recurrence within six months. All features, except for DLT-1, showed that higher values corresponded to higher SHAP positive values (Figure 7c), which were associated with

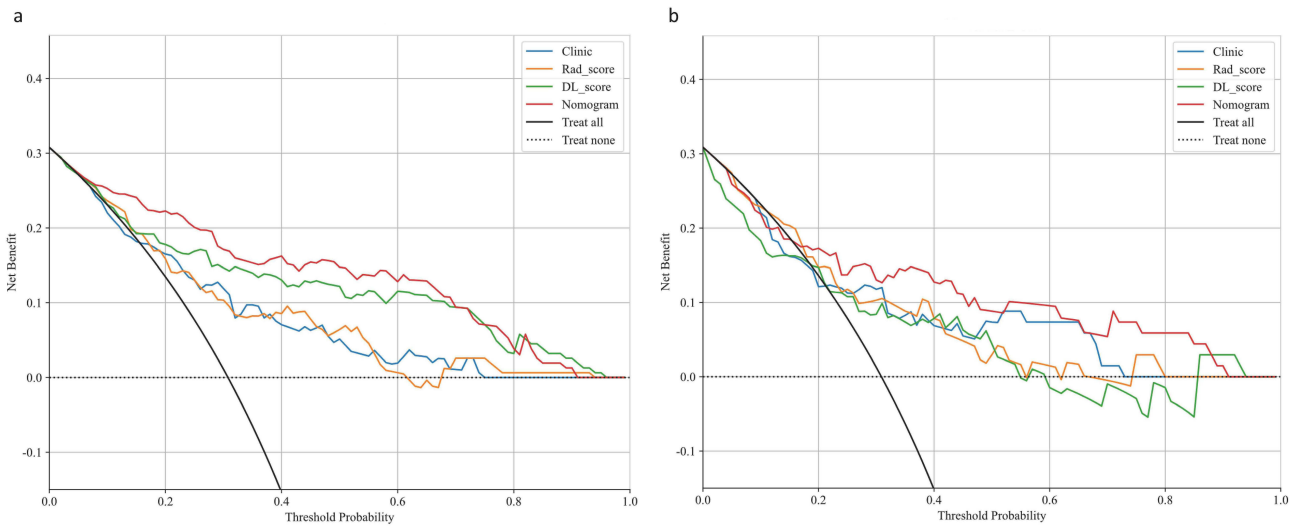

Figure 6 The decision curve analysis. (**a**) showed the net benefit of each model in the training set. The larger the area enclosed by the model curve, treat all, and treat none curves, the greater the net benefit of the model. (**b**) showed the net benefit of each model in the test set.
Abbreviations: DCA, decision curve analysis; Rad, radiomics; DL, deep learning; AFP, alpha-fetoprotein.

increased prediction probabilities. Figure 7d illustrated the SHAP local bar plot, where the case was classified correctly, although DLT-6 and DLT-5 contributed incorrectly.

In terms of Grad-CAM, the attention areas of the DenseNet-169 model were clear, focusing primarily on the boundaries and internal regions of the liver, while the surrounding areas remained inactive (Figure 8).

Discussion

In this study, we identified Rad score, DL score, LnAFP, and multiple low-signal lesions as independent predictors of early recurrence within six months. These four predictors were included in the DL-Rad-Clinical nomogram. Compared to other models, the DL-Rad-Clinical nomogram demonstrated superior predictive performance in both the training and test sets. However, in the test set, no statistically significant differences were found between the AUC of the DL-Rad-Clinical nomogram and those of the other models. A key advantage of the nomogram was that all included parameters were clinically accessible and aligned with clinical guidelines.[19] Moreover, the nomogram showed satisfactory discrimination and calibration abilities and had the potential to stratify HCC patients into low or high-risk groups before thermal ablation, thereby aiding in treatment decisions.

Research on early recurrence within six months after HCC ablation was limited. One study revealed that among Western HCC patients who underwent thermal ablation, the primary risk factors for recurrence within six months post-treatment were the number of tumors and AFP levels.[29] Some studies from Asia also suggest that the size and number of tumors as well as the level of AFP are strong predictors of recurrence in hepatocellular carcinoma.[30,31] In our study, LnAFP was also identified as an independent risk factor for early recurrence within six months, aligning with these findings. However, our study did not find the number of tumors to be an independent risk factor for early recurrence within six months, which contrasted with the published study. This discrepancy could have been due to differences in study populations; the published research had only included BCLC 0-A patients with tumors smaller than 3 cm, while our study had broader inclusion criteria based on Milan criteria, with a maximum tumor diameter of <5 cm. The increase in tumor diameter likely contributed to a higher recurrence rate, which may explain the lack of significant difference in early recurrence within six months rates between single and multiple tumor patients. Additionally, another study indicated that patients with serum AFP levels >20 ng/mL post-treatment had a cumulative recurrence rate of 24.4% within six months, which was similar with our study's result (30.8%). This study had also recommended more comprehensive and intensive monitoring for patients with AFP levels >20 ng/mL to detect HCC recurrence earlier,

https://doi.org/10.2147/JHC.S541329

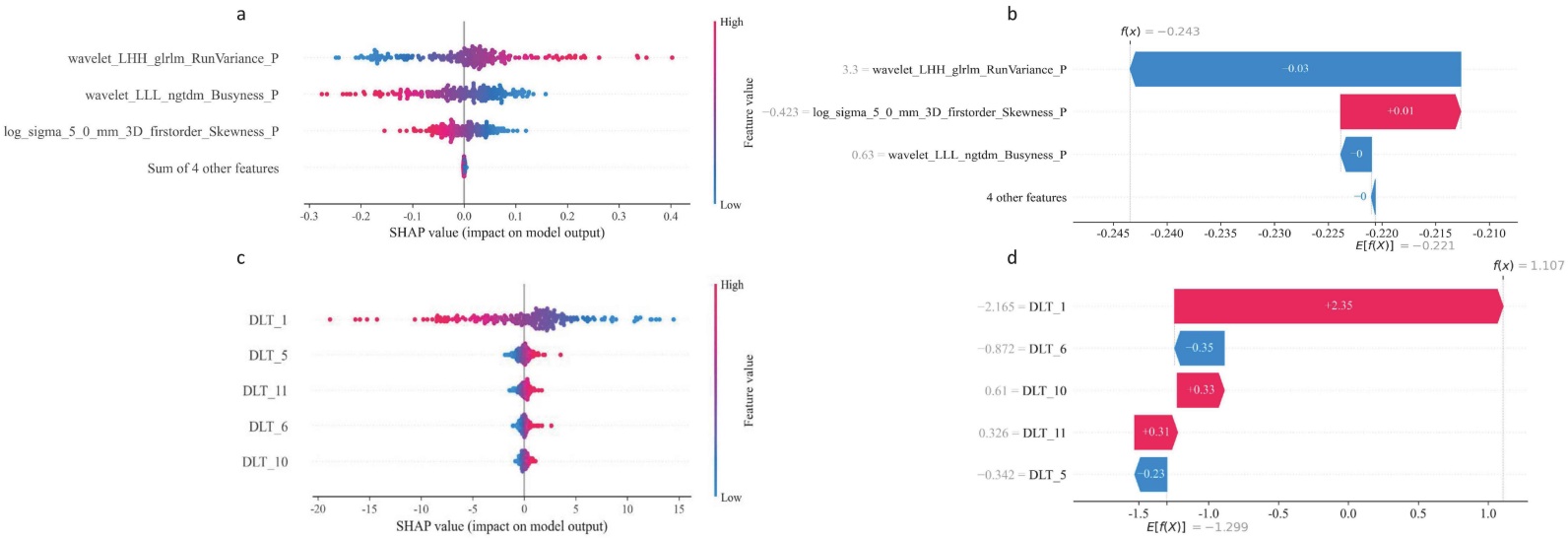

Figure 7 The SHAP summary plot and local bar plot. (**a**) showed the three most important features of the rad score. The value of the first feature positively impacted early recurrence within six months, while the other two features negatively impacted early recurrence within six months. (**b**) showed the magnitude of positive and negative influences of the three most important features in the rad score when predicting the risk of early recurrence within six months for an individual patient. The mean prediction of the model was −0.221, and the final SHAP value for the patient was −0.243. (**c**) showed the five most important features of the DL score. The value of the first feature negatively impacted early recurrence within six months, while the other four features positively impacted early recurrence within six months. (**d**) showed the magnitude of positive and negative influences of the five most important features in the DL score when predicting the risk of early recurrence within six months for an individual patient. The mean prediction of the model was −1.299, and the final SHAP value for the patient was 1.107.
Abbreviations: SHAP, Shapley Additive Explanation; DLT, deep learning transformation.

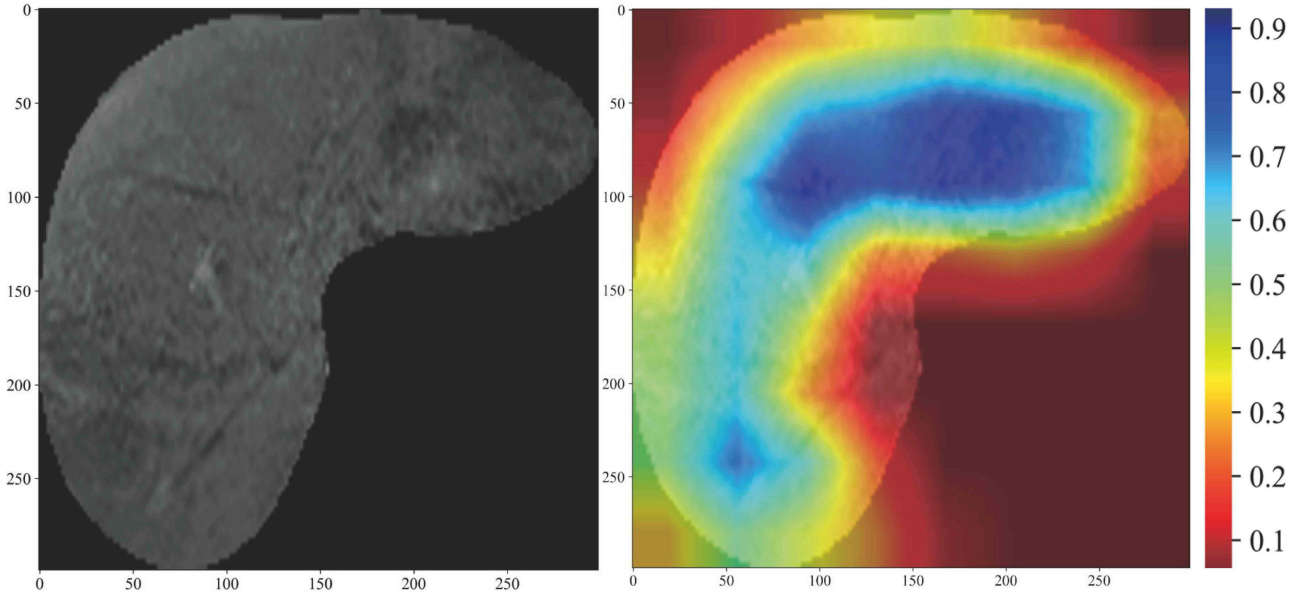

Figure 8 The Grad-CAM of DenseNet-169. The figure showed the contribution distribution of the predicted output by the DenseNet-169 model on this image. Areas with higher scores indicated that the corresponding regions of the image had a stronger response and contribute more to the model.
Abbreviations: Grad-CAM, gradient-weighted class activation mapping.

consistent with our study's aim.[32] However, our study developed a nomogram that provided a more precise assessment of recurrence risk, helping to minimize the waste of medical resources and reduce the economic burden on patients.

Multiple low signal lesions are usually one of the manifestations of cirrhosis or malignancy in portal vein phase of MRI. In fact, chronic hepatitis and cirrhosis are closely related to the development of HCC. During chronic liver disease, dramatic changes in liver microarchitecture and altered function of liver cells create a fertile environment for cancer to develop.[33] At the same time, persistent inflammation, fibrosis, and abnormal liver cell regeneration in patients with chronic hepatitis promote a series of genetic and epigenetic events that eventually lead to the formation of precancerous lesions.[34] In addition, several molecular alterations provide proliferative, invasive, and survival advantages to dysplastic cells, facilitating the transition to HCC.[35] It is well known that most very early HCC (< 2 cm) imaging findings are atypical.[36] However, HCC not showing typical imaging findings cannot be regarded as less aggressive than typical HCC.[37] In summary, these issues pose difficulties in accurately diagnosing HCC, which may ultimately lead to incorrect staging and prognostic evaluation. In the present study, multiple low-signal lesions were used as a parameter to construct the Combined model, which may help screen for patients with potential liver cancer lesions.

There were several limitations in this study. Firstly, this was a retrospective study with a small sample size from a single center, with a high proportion of chronic viral hepatitis patients, no stratification of tumor diameter, non-strict unified of image acquisition, no pathological tissue obtained, a relatively low Radiomics Quality Score, and the clinical usefulness of the model needing further validation in other centers or prospective studies. Secondly, MRI data were obtained from different scanners, which increased variability. Therefore, the image registration, N4ITK MRI bias correction, signal intensity normalization, voxel resampling and bin width (with fixed signal intensity) were performed to minimize the impact. Additionally, during the model training process, beyond the train/test split was not applied for further validation, which reduced the robustness of the model. Finally, several other variables that influence early recurrence within six months were not included in the study, such as postoperative characteristics. As we aimed to develop a pre-procedure model that could be used before ablation.

Conclusion

In conclusion, our research indicated that the DL-Rad-Clinical nomogram was a valuable tool for predicting early recurrence within six months in HCC patients following thermal ablation. However, due to the limitations of this single-center retrospective study, caution was warranted when interpreting the predictive ability of the model.

Abbreviations

MRI, magnetic resonance imaging; DL, deep learning; Rad, radiomics; HCC, hepatocellular carcinoma; BCLC, Barcelona Clinic Liver Cancer; AUC, area under the receiver operating characteristic curve; LnAFP, natural logarithm alpha-fetoprotein; CT, computed tomography; DCEI, dynamic contrast-enhanced imaging; ROI, region of interest; lasso, least absolute shrinkage and selection operator; Grad-CAM, Gradient-weighted Class Activation Mapping; SHAP, Shapley Additive Explanations; DLT, deep learning transformation.

Statement of Ethics

The study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Medical Ethics Committee of the Affiliated Cancer Hospital of Zhengzhou University (No.: 2024-015-002) and individual consent for this retrospective analysis was waived.

Funding

This work was supported by Henan Provincial Medical Science and Technology Key Project Joint Construction Program (Project No.:LHGJ20240119), Medical Education Research Project of Henan Province (Project No.: Wjlx2021334), General Project of Natural Science Foundation of Henan Province (Project No.: 212300410403), Henan Province Outstanding Foreign Expert Workstation Project (Project No.: GZS2022020), Foreign Expert Project of the Ministry of Science and Technology of China (Project No.: G2023026016L), Henan Province Medical Education Research Project (Project No.:Wjlx2021341), Henan Province Health and Health Technology Innovation Talent Program (Project No.:LJRC2024004).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. *J Natl Cancer Cent*. 2024;4(1):47–53. doi:10.1016/j.jncc.2024.01.006
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
3. Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol*. 2022;76(3):681–693. doi:10.1016/j.jhep.2021.11.018
4. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *J Hepatol*. 2006;44(1):217–231. doi:10.1016/j.jhep.2005.10.013
5. Wang XH, Fu YL, Xu YN, et al. Ginsenoside Rh1 regulates the immune microenvironment of hepatocellular carcinoma via the glucocorticoid receptor. *J Integr Med*. 2024;22(6):709–718. doi:10.1016/j.joim.2024.09.004
6. Li XJ, Chang L, Mi Y, et al. Integrated-omics analysis defines subtypes of hepatocellular carcinoma based on circadian rhythm. *J Integr Med*. 2025;23(4):445–456. doi:10.1016/j.joim.2025.06.003
7. Rao NGR, Sethi P, Deokar SS, et al. Potential indicators for the development of hepatocellular carcinoma: a diagnostic strategy. *Curr Top Med Chem*. 2025. doi:10.2174/0115680266349627250626142221
8. Yasuda S, Matsuo Y, Doi S, et al. Preoperative predictors of very early recurrence in patients with hepatocellular carcinoma beyond the Milan criteria. *Langenbecks Arch Surg*. 2024;409(1):283. doi:10.1007/s00423-024-03474-x
9. Doyle A, Gorgen A, Muaddi H, et al. Outcomes of radiofrequency ablation as first-line therapy for hepatocellular carcinoma less than 3 cm in potentially transplantable patients. *J Hepatol*. 2019;70(5):866–873. doi:10.1016/j.jhep.2018.12.027
10. Kim YS, Lim HK, Rhim H, et al. Ten-year outcomes of percutaneous radiofrequency ablation as first-line therapy of early hepatocellular carcinoma: analysis of prognostic factors. *J Hepatol*. 2013;58(1):89–97. doi:10.1016/j.jhep.2012.09.020
11. N'Kontchou G, Mahamoudi A, Aout M, et al. Radiofrequency ablation of hepatocellular carcinoma: long-term results and prognostic factors in 235 Western patients with cirrhosis. *Hepatology*. 2009;50(5):1475–1483. doi:10.1002/hep.23181
12. Rossi S, Ravetta V, Rosa L, et al. Repeated radiofrequency ablation for management of patients with cirrhosis with small hepatocellular carcinomas: a long-term cohort study. *Hepatology*. 2011;53(1):136–147. doi:10.1002/hep.23965
13. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749–762. doi:10.1038/nrclinonc.2017.141
14. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. doi:10.1038/ncomms5006
15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–577. doi:10.1148/radiol.2015151169
16. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441–446. doi:10.1016/j.ejca.2011.11.036

17. Calderaro J, Seraphin TP, Luedde T, Simon TG. Artificial intelligence for the prevention and clinical management of hepatocellular carcinoma. *J Hepatol*. 2022;76(6):1348–1361. doi:10.1016/j.jhep.2022.01.014
18. Dai Y, Zhao S, Wu Q, et al. A CT-based deep learning radiomics scoring system for predicting the prognosis to repeat TACE in patients with hepatocellular carcinoma: a multicenter cohort study. *J Hepatocell Carcinoma*. 2025;12:1647–1659. doi:10.2147/jhc.S525920
19. Galle PR, Forner A, Llovet JM, et al. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol*. 2018;69(1):182–236. doi:10.1016/j.jhep.2018.03.019
20. Mazzaferro V, Bhoori S, Sposito C, et al. Milan criteria in liver transplantation for hepatocellular carcinoma: an evidence-based analysis of 15 years of experience. *Liver Transpl*. 2011;17(Suppl 2):S44–57. doi:10.1002/lt.22365
21. Johnson PJ, Berhane S, Kagebayashi C, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. *J Clin Oncol*. 2015;33(6):550–558. doi:10.1200/jco.2014.57.9151
22. Kamath PS, Kim WR. The model for end-stage liver disease (MELD). *Hepatology*. 2007;45(3):797–805. doi:10.1002/hep.21563
23. Teratani T, Yoshida H, Shiina S, et al. Radiofrequency ablation for hepatocellular carcinoma in so-called high-risk locations. *Hepatology*. 2006;43(5):1101–1108. doi:10.1002/hep.21164
24. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77(21):e104–e107. doi:10.1158/0008-5472.Can-17-0339
25. Huang G, Liu Z, Pleiss G, Maaten LV, Weinberger KQ. Convolutional networks with dense connectivity. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(12):8704–8716. doi:10.1109/tpami.2019.2918284
26. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–574. doi:10.1177/0272989x06295361
27. Selvaraju RRCM, Das A, Vedantam R, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis*. 2020;128:336–359. doi:10.1007/s11263-019-01228-7
28. SM LS-I L. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017;30:4765–4774.
29. Preel A, Hermida M, Allimant C, et al. Uni-, Bi- or Trifocal hepatocellular carcinoma in western patients: recurrence and survival after percutaneous thermal ablation. *Cancers*. 2021;13(11). doi:10.3390/cancers13112700
30. Ngo HTT, Nguyen DD, Dang MX, Doan TTP, Thai TT. Early recurrence of hepatocellular carcinoma in patients without microscopic vascular invasion: clinicopathological characteristics and risk factors. *J Hepatocell Carcinoma*. 2025;12:1167–1175. doi:10.2147/jhc.S524683
31. Tsuji Y, Namisaki T, Takaya H, et al. Risk factors for intrahepatic distant recurrence after radiofrequency ablation for hepatocellular carcinoma. *Dig Dis Sci*. 2025;70(6):2156–2166. doi:10.1007/s10620-025-08884-5
32. Lee J, Joo I, Lee DH, Jeon SK, Lee JM. Clinical outcomes of patients with a high alpha-fetoprotein level but without evident recurrence on CT or MRI in surveillance after curative-intent treatment for hepatocellular carcinoma. *Abdom Radiol*. 2021;46(2):597–606. doi:10.1007/s00261-020-02707-z
33. Affo S, Yu LX, Schwabe RF. The role of cancer-associated fibroblasts and fibrosis in liver cancer. *Annu Rev Pathol*. 2017;12:153–186. doi:10.1146/annurev-pathol-052016-100322
34. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med*. 2019;380(15):1450–1462. doi:10.1056/NEJMra1713263
35. Torrecilla S, Sia D, Harrington AN, et al. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. *J Hepatol*. 2017;67(6):1222–1231. doi:10.1016/j.jhep.2017.08.013
36. Bolondi L, Gaiani S, Celli N, et al. Characterization of small nodules in cirrhosis by assessment of vascularity: the problem of hypovascular hepatocellular carcinoma. *Hepatology*. 2005;42(1):27–34. doi:10.1002/hep.20728
37. Forner A, Vilana R, Bianchi L, et al. Lack of arterial hypervascularity at contrast-enhanced ultrasound should not define the priority for diagnostic work-up of nodules <2 cm. *J Hepatol*. 2015;62(1):150–155. doi:10.1016/j.jhep.2014.08.028