Journal of Hepatocellular Carcinoma

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Multicenter Integration of MR Radiomics, Deep Learning, and Clinical Indicators for Predicting Hepatocellular Carcinoma Recurrence After Thermal Ablation

Yandan Wang[1], Yong Zhang[2], Jincheng Xiao[3], Xiang Geng[3], Lujun Han[4], Junpeng Luo[5]

[1]Department of Otorhinolaryngology, Huaihe Hospital of Henan University, Kaifeng, 475000, People's Republic of China; [2]Department of Immunotherapy, Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital, Zhengzhou, 450003, People's Republic of China; [3]Department of Minimally Invasive Intervention, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, 450000, People's Republic of China; [4]Department of Radiology, State Key Laboratory of Oncology in South China, Collaborative Innovation Cancer for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, 510030, People's Republic of China; [5]Translational Medical Center of Huaihe Hospital, Henan University, Kaifeng, 475000, People's Republic of China

Correspondence: Junpeng Luo, Translational Medical Center of Huaihe Hospital, Henan University, Kaifeng, 475000, People's Republic of China, Email hiccp@henu.edu.cn

Background: To develop and validate an innovative predictive model that integrates multisequence magnetic resonance (MR) radiomics, deep learning features, and clinical indicators to accurately predict the recurrence of hepatocellular carcinoma (HCC) after thermal ablation.

Methods: This retrospective multicenter cohort study enrolled patients who were diagnosed with HCC and treated via thermal ablation. We extracted radiomic features from multisequence 3T MR images, analyzed these images using a 3D convolutional neural network (3D CNN), and incorporated clinical data into the model. Model performance was evaluated using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.

Results: The study included 535 patients from three hospitals, comprising 462 males and 43 females. The RDC model, which stands for the Radiomics-Deep Learning-Clinical data model, demonstrated high predictive accuracy, achieving AUCs of 0.794 in the training set, 0.777 in the validation set, and 0.787 in the test set. Statistical analysis confirmed the model's robustness and the significant contribution of the integrated features to its predictive capabilities.

Conclusion: The RDC model effectively predicts HCC recurrence after thermal ablation by synergistically combining advanced imaging analysis and clinical parameters. This study highlights the potential of such integrative approaches to enhance prognostic assessments in HCC patients and offers a promising tool for clinical decision-making.

Keywords: hepatocellular carcinoma, thermal ablation, radiomics, deep learning, MRI, prognostic modeling

Background

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer and poses a significant global health burden. Each year, liver cancer is responsible for approximately 906,000 new cases and 830,000 deaths worldwide, with HCC accounting for 85% to 90% of these cases.[1,2] Recent advancements in thermal ablation techniques, such as radiofrequency[3,4] and microwave ablation,[5] have become widely used due to their convenience, efficacy, and reduced hospital stays. However, a major drawback remains: a notably high recurrence rate.[6]

Recurrence after thermal ablation typically manifests in three forms: local recurrence at the ablation site, intrahepatic distant recurrence at sites within the liver that were not initially targeted, and extrahepatic metastasis forming new lesions outside the liver.[7] Research shows that within three years post-ablation, local recurrence rates can range from 3.2% to 28.5% and may increase to between 10% and 32.1% within five years.[8] Similarly, distant intrahepatic metastasis occurs

Received: 31 July 2024
Accepted: 28 September 2024
Published: 2 October 2024

in 49% to 63.3% of patients within three years and between 24.8% and 81% within five years.[9] People focus on the first two types of recurrence—local recurrence at the ablation site and intrahepatic distant recurrence—because they are more common, occur within the liver where treatment is more feasible, and have a higher potential for curative intervention compared to extrahepatic metastasis, which typically indicates more advanced disease and poorer prognosis.[10]

Given that most recurrences occur within the first two years after thermal ablation, this study specifically focuses on predicting recurrence within this clinically significant period.[11,12] Early recurrence is often associated with tumor aggressiveness and poorer outcomes, making the 2-year time frame critical for timely interventions and prognosis improvement.[13] Focusing on this window allows for a more targeted risk assessment, enabling clinicians to optimize follow-up and treatment strategies during this high-risk period.

Traditionally, studies predicting HCC recurrence post-ablation have primarily utilized clinical indicators.[14] These papers discuss tumor size, number, location, and degree of differentiation as well-known risk factors for recurrence after thermal ablation,[15] but these factors cannot fully capture tumor heterogeneity or dynamics, leading to generalized risk assessments. Radiomics and deep learning offer objective and comprehensive analyses by integrating complex data, leading to more accurate and personalized recurrence predictions. For example, Wang's study leveraging CT and deep learning reported AUCs of 0.75 and 0.825,[16] respectively, while Gao's work using MRI-based radiomics and deep learning features achieved AUCs of 0.91 in training and 0.84 in testing,[17] demonstrating superior performance over models based on radiomics alone.

Despite advances, many models still rely on single imaging modalities and lack multicenter validation. In this study, local and intrahepatic distant recurrence are combined as intrahepatic recurrences due to shared underlying mechanisms and similar liver-directed treatment strategies. This grouping allows for better analysis of liver-specific recurrence and treatment outcomes.

Our study addresses gaps by combining multisequence MRI radiomics, deep learning, and a wide range of clinical indicators in a multicenter framework to improve predictive accuracy and outcomes for HCC patients.

Methods

Patient Selection

This multicenter retrospective study included patients from January 1, 2015, to December 31, 2020, across three institutions: Sun Yat-sen University Cancer Center, Henan Provincial Cancer Hospital, and Huaihe Hospital of Henan University. A total of 3,228 liver cancer patients were screened, and 535 patients who underwent hepatocellular carcinoma ablation were included in the study (Figure 1). The study adhered to the Declaration of Helsinki and received ethical approval from the Ethics Committees of Sun Yat-sen University Cancer Center, Henan Provincial Cancer Hospital, and Huaihe Hospital of Henan University (ethics approval number: SL-B2022-047-02). Since the study was conducted across multiple institutions, with the bulk of the study performed at Sun Yat-sen University Cancer Center, the appropriate authors have been affiliated with this institution. Due to the retrospective nature of the study, informed consent was waived. Detailed inclusion and exclusion criteria are documented in Supplementary Table 1. In compliance with institutional regulations, patient data confidentiality was maintained throughout the study. All patient data used in this research were anonymized, and no identifiable patient information was accessed or disclosed. We have addressed potential biases introduced by the retrospective design and specified that recurrence was studied at a 2-year time point. The generalizability of the model is being tested, and future prospective studies are planned to validate its performance over longer time periods.

Data Acquisition and Transformation

Clinical and demographic data were collected retrospectively from the three hospitals electronic medical records. These included variables such as Age, Sex, Eastern Cooperative Oncology Group (ECOG) performance status, Hepatitis type, and various laboratory measures. MRI scans followed standardized protocols across the three centers, as detailed in Supplementary Table 2. Imaging was conducted in a breath-hold state using meglumine gadobutrol contrast medium, with sequential imaging phases captured and stored in DICOM format. Due to the use of different blood testing

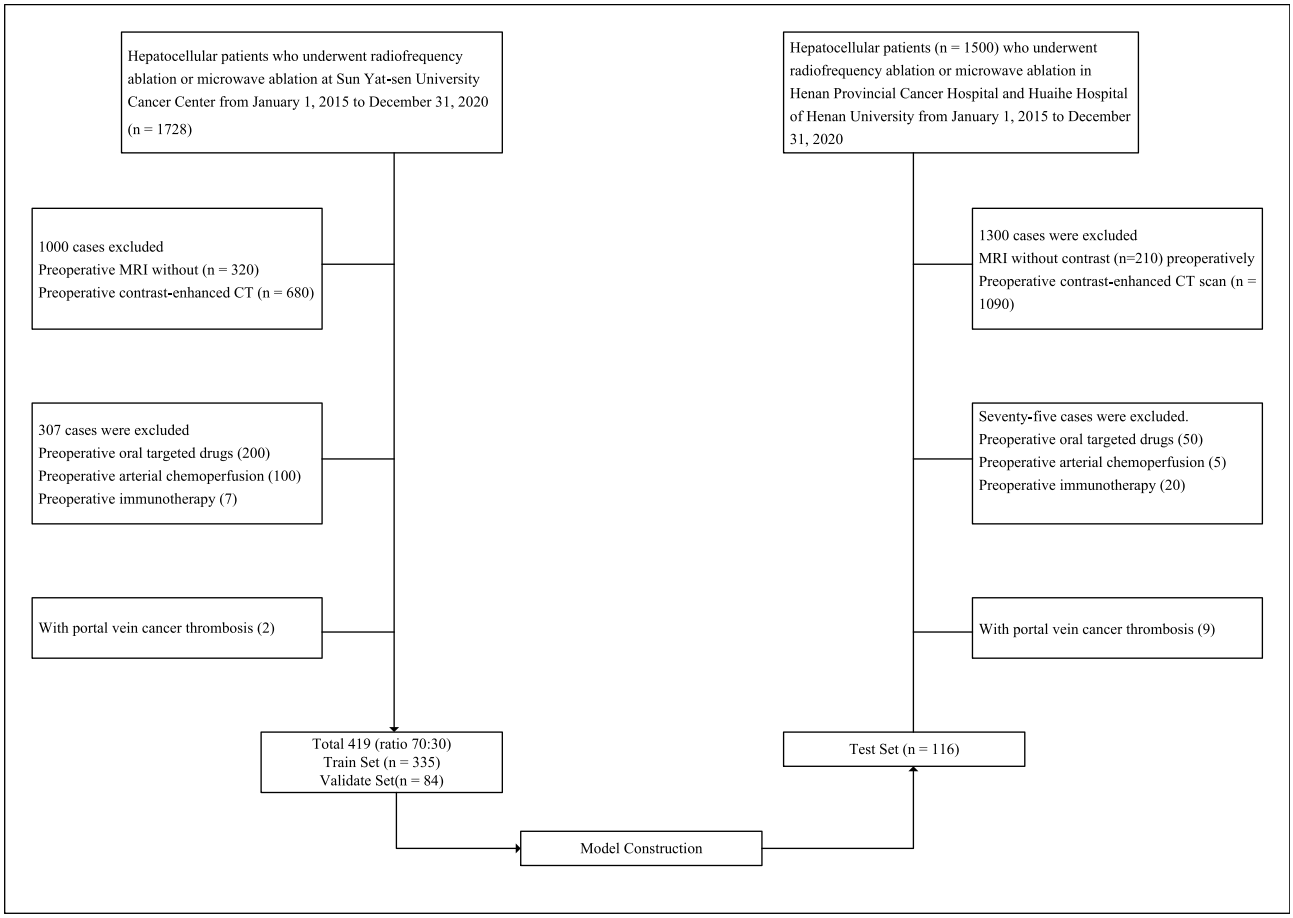

Figure 1 Patient Enrollment Flowchart. From January 1, 2015 to December 31, 2020, 1728 patients with liver cancer were selected from the Cancer Center of Sun Yat-sen University, 1000 cases were only undergoing MR or CT non-contrast scans before surgery, 307 cases were subjected to targeted, immunological or arterial chemotherapy perfusion, and 2 cases of portal cancer thrombosis were eliminated, and 419 cases were finally enrolled. From January 1, 2015 to December 31, 2020, a total of 1500 patients with liver cancer were selected from the two centers of Henan Provincial Cancer Hospital and Huaihe Hospital of Henan University, 1300 cases only underwent MR or CT non-contrast scan before surgery, 75 cases underwent targeted, immunological or arterial chemotherapy perfusion, and 9 cases of portal cancer thrombosis were excluded, and 116 cases were finally enrolled.

equipment across the three hospitals, the reference ranges for platelet count (PLT), gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP), aspartate aminotransferase (AST), and alanine aminotransferase (ALT) varied. To standardize the data and facilitate comparison, we categorized the values for each of these parameters based on whether they were lower, within, or higher than the respective reference range at each institution. Specifically, results below the lower limit of the reference range were classified as "Lower", results within the reference range as "Normal", and results exceeding the upper limit as "Higher".

Ablation Procedure

Thermal ablation was performed under CT guidance by seasoned interventional radiologists, typically 2–4 weeks after completing the MRI scan. The approach was adjusted based on the tumor location and patient anatomy to optimize needle placement and minimize risks. The ablation parameters, described in Supplementary Table 3, ensured a consistent safety margin across treatments. Post-procedure, patients were monitored for immediate complications.

Image Segmentation

Manual segmentation of liver tumors was performed using 3D Slicer software on both arterial and T2-weighted MR images to generate tumor masks. The initial segmentation was carried out independently by two experienced radiologists, each with over five years of expertise in liver imaging. To ensure consistency, the tumor masks from the arterial contrast

(AC) phase were carefully registered to the non-contrast (FLEX), portal venous (VC), and delayed contrast (DC) phase images, allowing for precise mask correspondence across different imaging sequences. In cases where discrepancies arose between the initial segmentations, a third senior radiologist with over 10 years of clinical experience supervised the review process. This senior radiologist facilitated consensus and adjudicated any disagreements to ensure the highest accuracy and reliability in the final tumor masks. Following segmentation, all images and corresponding tumor masks were subject to a final independent review by the two radiologists, with the senior radiologist providing oversight to resolve any remaining differences. To provide visual validation of the radiomics analysis process, representative MRI tumor segmentation figures are shown in Supplementary Figure 1, illustrating the manual segmentation of tumor boundaries on different MRI sequences. This multi-step approach ensured both precision and consistency in the segmentation process.

Radiomics Feature Extraction and Screening

To standardize imaging data across centers and devices, voxel discretization and size resampling were applied to ensure consistency in image quality. Using manually segmented tumor masks from the arterial and T2-weighted MR images, radiomics feature extraction was performed on both tumor regions and surrounding tissue to capture comprehensive spatial information. Following the IBSI guidelines,[18] PyRadiomics[19] was used to extract 1,337 radiomic features per sequence, including first-order statistics, shape descriptors, and higher-order texture features, resulting in 6685-dimensional features across five imaging sequences. Features with an intraclass correlation coefficient (ICC) below 0.7 were retained to ensure reproducibility and robustness, while those exceeding this threshold were removed. A decision tree model was used to rank the remaining features by importance, and the top 20 were selected for further refinement. The model was developed and internally validated using MRI data from Sun Yat-sen University Cancer Center, with a training set consisting of 335 patients for model development and 84 patients for internal validation. To ensure generalizability and reduce overfitting, 5-fold cross-validation was applied, further refining the feature set. Ultimately, 15 key features across five imaging sequences were selected as the most predictive of recurrence risk (Supplementary Table 4). For external validation, 116 patients from Henan Provincial Cancer Hospital and Huaihe Hospital of Henan University were included as the test set. This independent cohort served to evaluate the model's performance outside the development dataset. These features were then validated using univariate Cox proportional hazards analysis. Full details of the selected features and their performance are provided in the Supplementary Data 1.

3D-CNN Network and Preprocessing

For input into the 3D-CNN network, a 30-voxel expanded bounding box was applied around the tumor boundary, ensuring that both the tumor and surrounding tissues were included for feature extraction. MRI images from the same dataset used for radiomics analysis, comprising the training, validation, and test sets, were preprocessed to standardize variations arising from different manufacturers and devices. This included resizing all images to a uniform dimension of 256×256 pixels to ensure consistency across the network inputs. The training patch size was set to [128, 128, 64], with each epoch consisting of 250 patches, evenly balanced between tumor and background regions to maintain the network's ability to differentiate between the two. To optimize performance, the 3D-CNN network was trained over 300 epochs, with an initial learning rate of 0.1. The learning rate was progressively reduced by 20% whenever the improvement in loss function dropped below 10^{-4} for ten consecutive epochs. This adaptive learning rate approach allowed the network to fine-tune its performance incrementally, avoiding overfitting. The network architecture was designed with a final fully connected layer that outputs a 128-dimensional feature vector, which serves as the input for the classification task. Following the initial training phase, the network was fine-tuned using the validation set, which contained 84 patients. Fine-tuning was performed over 60 epochs, with a reduced learning rate of 10^{-5} to enhance the precision of the model. After fine-tuning, the final model was validated externally on the test set of 116 patients from Henan Provincial Cancer Hospital and Huaihe Hospital of Henan University, ensuring that the network's performance generalized well to unseen data. The 3D-CNN network, along with its preprocessing steps, worked synergistically with the radiomics analysis, sharing the same standardized dataset across all phases of model development. The entire workflow of the study is detailed in Figure 2.

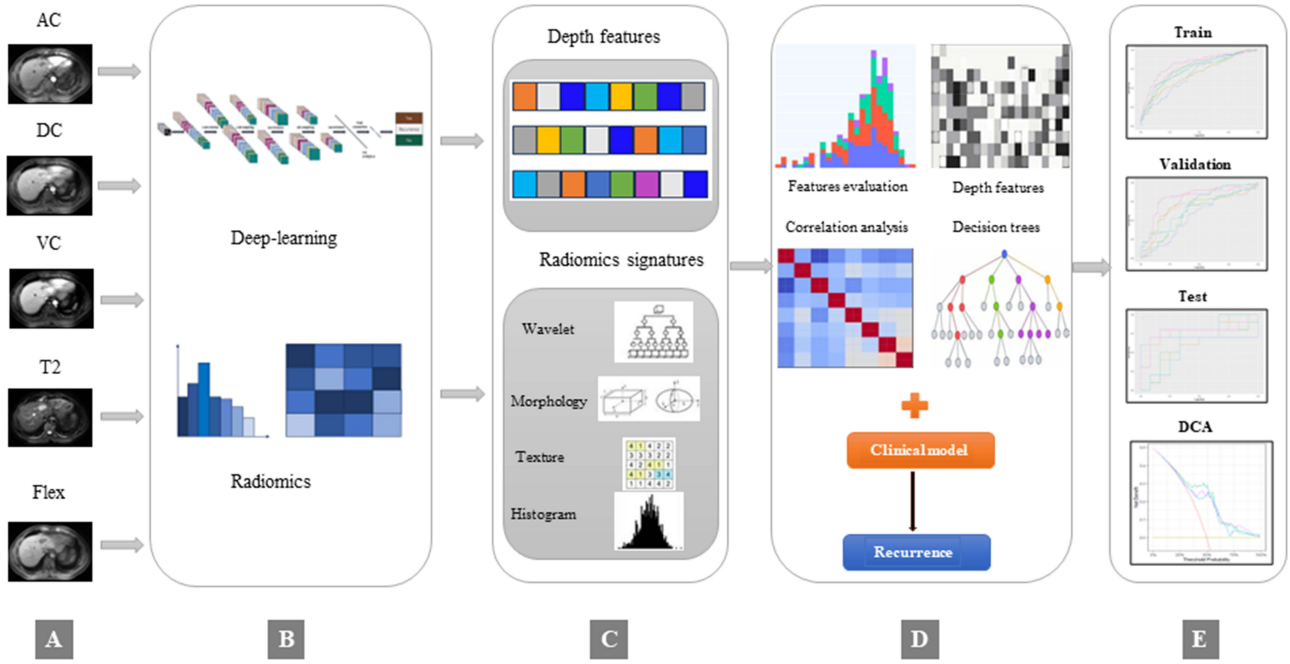

Figure 2 Comprehensive MRI-Based Imaging Analysis for Tumor Characterization and Prediction Model Construction. (A) Annotated image showcasing phases of MRI scans including plain scan, arterial phase, portal venous phase, delay phase, and T2WI sequence. (B) Feature extraction using radiomics and 3D-CNN methods based on the MRI images. (C) Feature selection process for the extracted features. (D) Integration of selected radiomic features, deep learning features, and clinical features for the construction of a predictive model. (E) Evaluation of the predictive model's performance.

Model Construction

The predictive model used in this study integrates radiomic features, deep learning outputs, and clinical indicators to predict the recurrence of hepatocellular carcinoma within two years following thermal ablation. The model was constructed using a multistep approach:

1. Radiomic Feature Extraction: Radiomic features were extracted from multiple MRI sequences (FLEX, AC, VC, DC, and T2) using the PyRadiomics software. Features included first-order statistics, texture features, and shape descriptors.
2. Deep Learning Model: A 3D-CNN was employed to capture spatial information from the MRI images. The model was trained with 250 patches per epoch over 300 epochs, incorporating a final fully connected layer that outputs a 128-dimensional feature vector.
3. Clinical Indicators: Clinical variables included tumor size, number of tumors, albumin levels, AFP, and ECOG score, all of which were selected based on their relevance to recurrence risk.
4. Model Integration: The radiomics, deep learning outputs, and clinical indicators were integrated into a combined model (RDC-model) using logistic regression. The RDC-model was trained and validated on the train and validate datasets, and its performance was tested on an independent test set.

Computational Formula The final RDC-model was built using the following logistic regression equation:

$$P(Recurrence) = \frac{1}{1 + e^{-(\beta_0 + \sum / \beta_1 X_i)}}$$

where P(Recurrence) is the probability of recurrence within two years, β_o is the intercept, β_i are the coefficients for the radiomic features, deep learning outputs, and clinical indicators, and X_i are the respective input values.

The model's coefficients were optimized using cross-validation to minimize overfitting and improve generalizability.

Statistical analysis

The cohort was randomly divided into training and validation sets, maintaining a stratified distribution of key variables. The baseline characteristics of these sets were compared using the chi-square test or Fisher's exact test for categorical variables.[20] Continuous variables were analyzed using independent samples *t* tests or Mann–Whitney *U*-tests.[21] All the statistical analyses were conducted using R (R Foundation for Statistical Computing, version 4.1.2, https://www.r-project.org/). The ICC was computed using the "irr" package.[22] Data cleaning was performed with the "knnImputation" and "DMwR" packages. The "glm" package facilitated univariate and multivariate logistic regression analyses. ROC curves and AUC calculations were performed using the "pROC" package. Additionally, decision curve analysis (DCA) was performed utilizing the "ggDCA" package in R.[23]

Results

Patient Baseline Characteristics

Significant differences were observed in patient characteristics across the cohorts. The mean age was similar (around 59 years, P = 0.704), but the validation and test sets had more female patients (19.0%) compared to the training set (10.4%) (P = 0.021). ECOG performance varied, with most patients in the training (99.7%) and validation (100%) sets scoring 0, while the test set had higher proportions of ECOG 1 (25.0%) and 2 (1.72%) (P < 0.001). The test set had more patients with multiple tumors (35.3%, P = 0.005) and lower platelet counts (65.5%, P < 0.001). Hepatitis C was more prevalent in the test set (6.03%, P = 0.002), and Child-Pugh grade B was more frequent in the test set (18.1%, P < 0.001). Full details are provided in Table 1.

Table 1 Baseline Distribution of Training Sets, Validation Sets, and Test Sets

Parameters	Training Set	Validation Set	Test Set	*P-value*
	n=335	n=84	n=116	
Age (years)	59.0±11.8	58.7±13.1	59.9±9.48	0.704
Gender				0.021
Female	35 (10.4%)	16 (19.0%)	22 (19.0%)	
Male	300 (89.6%)	68 (81.0%)	94 (81.0%)	
ECOG score				<0.001
0	333 (99.7%)	83 (100%)	85 (73.3%)	
1	1 (0.30%)	0 (0.00%)	29 (25.0%)	
2	0 (0.00%)	0 (0.00%)	2 (1.72%)	
Number of tumors				0.005
Single	261 (77.9%)	69 (82.1%)	75 (64.7%)	
Multiple	74 (22.1%)	15 (17.9%)	41 (35.3%)	
Maximum tumor diameter (cm)	2.14±1.30	2.19±1.16	2.42±1.17	0.119
Types of hepatitis				0.002
None	42 (12.5%)	13 (15.5%)	12 (10.3%)	
Hepatitis B	292 (87.2%)	71 (84.5%)	97 (83.6%)	
C	1 (0.30%)	0 (0.00%)	7 (6.03%)	
Alpha-fetoprotein				0.059
<400μg/L	295 (88.1%)	69 (82.1%)	108 (93.1%)	
>400μg/L	40 (11.9%)	15 (17.9%)	8 (6.90%)	
Albumin(g/L)	42.6±4.04	41.9±4.47	38.6±3.89 PM	<0.001
Total bilirubin(μmol/L)	15.3±8.35	13.6±8.45	16.1±8.59	0.102
Platelet				<0.001
Lower than normal	78 (23.3%)	11 (13.1%)	76 (65.5%)	
Normal	254 (75.8%)	73 (86.9%)	39 (33.6%)	
Higher than normal	3 (0.90%)	0 (0.00%)	1 (0.86%)	

(*Continued*)

Table 1 (Continued).

Parameters	Training Set	Validation Set	Test Set	*P-value*
	n=335	n=84	n=116	
Glutamyl transferase				0.384
Lower than normal	2 (0.60%)	0 (0.00%)	2 (0.7%)	
Normal	210 (62.7%)	60 (71.4%)	74 (63.8%)	
Higher than normal	123 (36.7%)	24 (28.6%)	40 (34.5%)	
Alkaline phosphatase				0.217
Lower than normal	13 (3.88%)	1 (1.19%)	1 (0.86%)	
Normal	289 (86.3%)	73 (86.9%)	97 (83.6%)	
Higher than normal	33 (9.85%)	10 (11.9%)	18 (15.5%)	
Aspartate aminotransferase				0.179
Lower than normal	3 (0.90%)	4 (4.76%)	2 (1.72%)	
Normal	250 (74.6%)	62 (73.8%)	83 (71.6%)	
Higher than normal	82 (24.5%)	18 (21.4%)	31 (26.7%)	
Alanine aminotransferase				0.156
Lower than normal	4 (1.19%)	2 (2.38%)	1 (0.86%)	
Normal	271 (80.9%)	69 (82.1%)	84 (72.4%)	
Higher than normal	60 (17.9%)	13 (15.5%)	31 (26.7%)	
BCLC staging				0.001
0	121 (36.1%)	29 (34.5%)	26 (22.4%)	
A	182 (54.3%)	46 (54.8%)	63 (54.3%)	
B	32 (9.55%)	9 (10.7%)	27 (23.3%)	
Child-Pugh				<0.001
A	332 (99.1%)	81 (96.4%)	95 (81.9%)	
B	3 (0.9%)	3 (3.6%)	21 (18.1%)	

Evaluation of The Consistency of Tumor Segmentation

The consistency of manual tumor segmentation was evaluated using the intraclass correlation coefficient (ICC). Intra-observer reliability was high, with ICCs of 0.925 (95% CI: 0.890–0.968) for physician A and 0.934 (95% CI: 0.902–0.954) for physician B. The inter-observer ICC between the two physicians was similarly robust at 0.930 (95% CI: 0.925–0.962), indicating strong agreement in tumor segmentation across observers (Table 2).

Distribution of Thermal Ablation Recurrence

The analysis of recurrence patterns following thermal ablation across the three hospitals reveals distinct differences in demographic and clinical parameters. At Sun Yat-sen University Cancer Center, recurrence was observed in 88.0% of males and 12.0% of females, with an average age of 58.6 years in patients with recurrence. Tumor recurrence was associated with multiple tumors (26.4%) and larger tumor size (2.2 cm), while single tumors were more common in patients without recurrence (83.9%). At Henan Provincial Cancer Hospital, recurrence was more common in males (87.5%) and patients with multiple tumors (50%), with an average tumor size of 2.1 cm in those with recurrence. Similarly, at Huaihe Hospital of Henan University, recurrence was observed in 100% of males and patients with larger tumors (average size 3.7 cm) and

Table 2 Consistency Assessment Results of Liver Tumor Manual Segmentation

Substance	ICC	95% CI
A physician Intra-observe	0.925	0.890–0.968
B physician Intra-observe	0.934	0.902–0.954
A and B physicians Inter-observer	0.930	0.925–0.962

multiple tumors (40%). The distribution of Child-Pugh grade B and BCLC stage B was higher in patients with recurrence, particularly at the external validation hospitals, highlighting more advanced disease in these groups (Table 3).

Univariate and Multivariate Analysis

The univariate and multivariate analyses, based on the train set, revealed significant clinical risk factors influencing recurrence after thermal ablation. Notably, the number of tumors was identified as a strong predictor, with a hazard ratio (HR) of 1.424 (95% CI: 1.085–1.868, P = 0.011) in the multivariate analysis, indicating that patients with multiple tumors had a significantly higher risk of recurrence. Additionally, serum albumin (ALB) level was found to be protective, with an HR of 0.96 (95% CI: 0.925–0.997, P = 0.035), underscoring its importance in reducing recurrence risk. These findings highlight the key prognostic factors identified from the training cohort (Table 4).

Performance of Predictive Models

The Radiomics Model (R-model) achieved the highest AUC of 0.753 in the train set, indicating good predictive power, while the Deep Learning Model (D-model) showed its best AUC of 0.643 with variable performance across different sequences. The Combined RDC-Model outperformed other models with an AUC of 0.787 in the test set, demonstrating robustness in clinical predictions, along with high accuracy (0.784), sensitivity (0.721), and specificity (0.855) as detailed in Table 5. The ROC curves in Figure 3 confirm that the RDC-model has superior discriminative ability across all sets, consistently

Table 3 Distribution of Ablation Recurrence in Three Hospitals (n, %)

Parameter	Sun Yat-sen University Cancer Center		Henan Provincial Cancer Hospital		Huaihe Hospital, Henan University	
	No recurrence	Recurrence	No recurrence	Recurrence	No recurrence	Recurrence
Age (years)	59.3±12.6	58.6±11.5	60.8±10.0	57.8±8.6	62.4±9.7	61.0±13.0
Gender						
Male	185 (87.7)	183 (88.0)	8 (52.9)	49 (87.5)	32 (84.2)	5(100)
Female	26 (12.3)	25 (12.0)	9 (47.1)	7 (12.5)	6 (15.8)	0
ECOG score						
0	211 (100)	205 (99.5)	17 (100)	40 (71.4)	24 (63.2)	4 (80.0)
1	0	1 (0.49)	0	16 (28.6)	12 (31.6)	1 (20.0)
2	0	0	0	0	2 (5.3)	0
Tumor diameter (cm)	2.1±1.2	2.2±1.4	2.7±1.2	2.1±1.0	2.6±1.2	3.7±1.0
Number of tumors						
Single	177 (83.9)	153 (73.6)	14 (82.4)	28 (50.0)	30 (78.9)	3 (60.0)
Multiple	34 (16.1)	55 (26.4)	3 (17.6)	28 (50.0)	8 (21.1)	2 (40.0)
Hepatitis						
None	28 (13.3)	27 (13.0)	0	4 (7.1)	7 (18.4)	1 (20.0)
Hepatitis B	182 (86.3)	181 (87.0)	17 (100)	49 (87.5)	27 (71.1)	4 (80.0)
C	1 (0.47)	0	0	3 (5.4)	4 (10.5)	0
Alpha-fetoprotein						
<400 ng/mL	184 (87.2)	180 (86.5)	17 (100)	52 (92.9)	34 (89.5)	5 (100)
≥400 ng/mL	27 (12.8)	28 (13.5)	0	4 (7.1)	4 (10.5)	0
Albumin (g/L)	42.7±4.1	42.1±4.0	38.8±2.8	39.3±2.3	37.7±5.7	36.2±2.2
Total bilirubin	15.3±8.5	14.6±8.3	15.6±0.8	15.9±1.4	15.9±14.7	21.5±6.9
BCLC staging						
0	84 (39.8)	66 (31.7)	2 (11.8)	14 (25.0)	10 (26.3)	0
A	98 (46.4)	130 (62.5)	4 (23.5)	27 (48.2)	27 (71.1)	5 (100)
B	29 (13.7)	12 (5.77)	11 (64.7)	15 (26.8)	1 (2.6)	0
ChildPugh score						
A	207 (98.1)	206 (99.0)	17 (100)	47 (83.9)	27 (71.1)	4 (80.0)
B	4 (1.9)	2 (1.0)	0	9 (16.1)	11 (28.9)	1 (20.0)

Table 4 Univariate and Multivariate Analysis of Ablation Recurrence of HCC in Training Set

Clinical factors	Univariate			Multivariate		
	HR	95% CI	P value	HR	95% CI	P value
Age	0.997	(0.985–1.009)	0.623			
Gender	1.602	(0.942–2.726)	0.082			
ECOG	5.56	(0.769–40.314)	0.089			
Tumor Size	1.078	(0.958–1.214)	0.213			
Tumor number	1.501	(1.171–1.923)	0.001	1.424	(1.085–1.868)	0.011
Hepatitis	0.894	(0.574–1.392)	0.619			
AFP	1.349	(0.882–2.065)	0.168			
Albumin	0.956	(0.921–0.992)	0.017	0.96	(0.925–0.997)	0.035
Total bilirubin	0.998	(0.981–1.015)	0.789			
Platelet	1.01	(0.712–1.432)	0.956			
Glutamyl transpeptidase	1.315	(0.965–1.793)	0.083			
Alkaline phosphatase	1.067	(0.706–1.614)	0.757			
Aspartate aminotransferase	1.09	(0.779–1.524)	0.616			
Alanine aminotransferase	1.505	(1.039–2.180)	0.030	1.429	(0.986–2.072)	0.059
Child-Pugh	3.56	(0.498–25.456)	0.206			
BCLC	2.115	(1.068–4.186)	0.032	1.858	(0.934–3.697)	0.206

Table 5 Prediction Performance of MRI Sequence-Specific Models for Recurrence Based on Radiomics, Deep Learning, and Clinical Factors Integration

Models	Series	AUC	Accuracy	Sensitivity	Specificity
Radiomics	AC	0.721	0.643	0.690	0.622
	VC	0.661	0.672	0.713	0.641
	DC	0.692	0.654	0.682	0.642
	T2	0.721	0.693	0.711	0.672
	Flex	0.563	0.442	0.422	0.451
	R-model	0.753	0.774	0.667	0.881
Deep Learning	AC	0.623	0.670	0.690	0.650
	VC	0.552	0.613	0.625	0.600
	DC	0.582	0.641	0.657	0.625
	T2	0.613	0.667	0.684	0.650
	Flex	0.562	0.680	0.801	0.780
	D-model	0.643	0.631	0.571	0.690
Radiomics + Deep Learning + Clinical Factors	AC	0.791	0.782	0.781	0.772
	VC	0.605	0.595	0.562	0.621
	DC	0.635	0.596	0.552	0.642
	T2	0.674	0.731	0.673	0.803
	Flex	0.695	0.182	0.213	0.134
	RDC-model	0.787	0.784	0.721	0.855

outperforming the R-model and D-model. Figure 4 Lift Charts reveal that the RDC-model provides the highest lift at higher predicted probabilities in the train set, proving more effective in identifying true positives. Additionally, Figure 5 Decision Curve Analysis highlights the net benefit of the RDC-model across various threshold probabilities in the train set, underscoring its clinical usefulness in accurately predicting liver cancer recurrence more effectively than competing models.

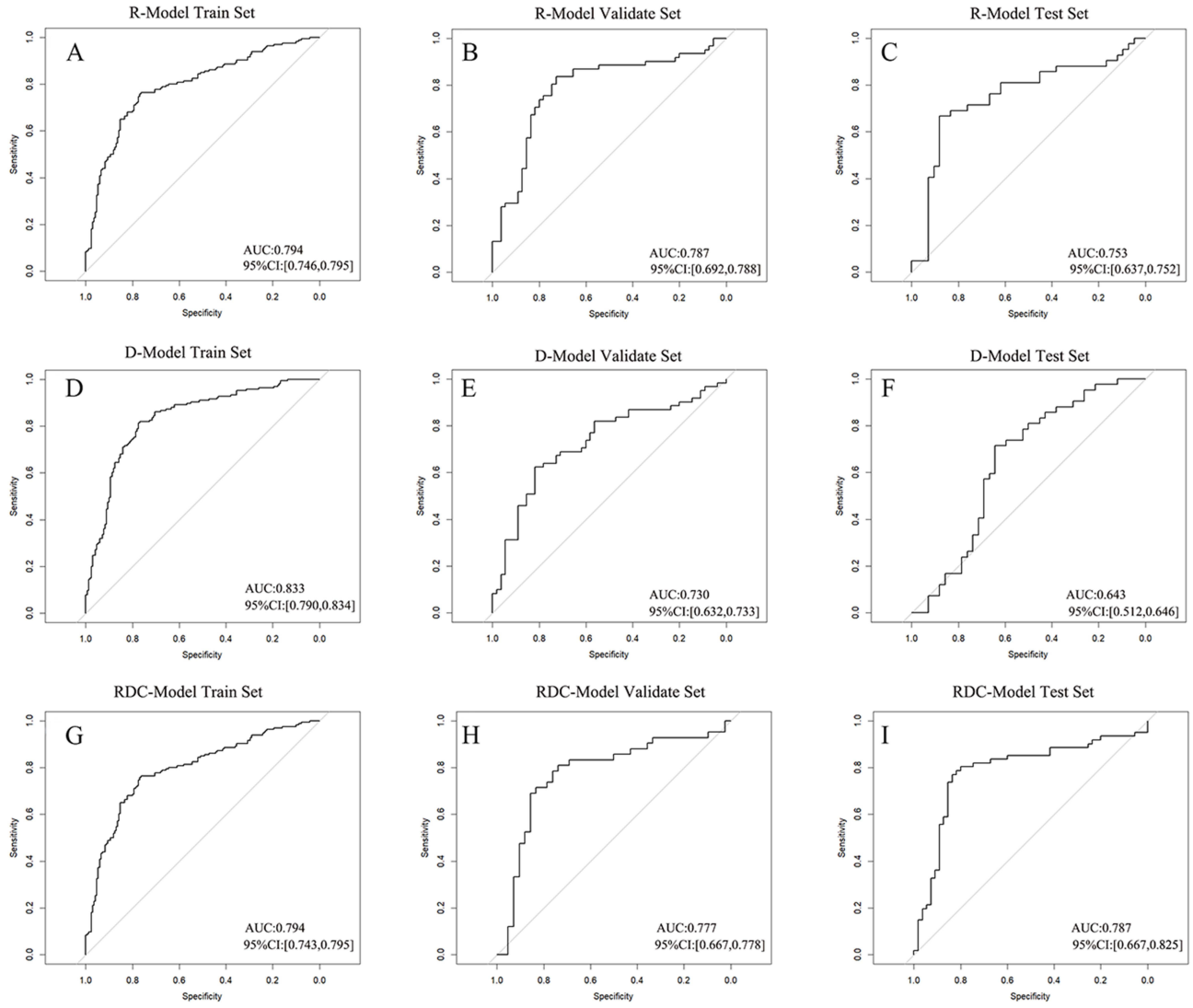

Figure 3 Comparative Performance of R-Model, D-Model, and RDC-Model (**A**-**C**) R-Model: The R-Model exhibits AUCs of 0.794 (train), 0.787 (validate), and 0.753 (test), demonstrating a consistent performance across different datasets. (**D**-**F**) D-Model: The D-Model shows varying AUCs with 0.833 (train), 0.730 (validate), and 0.643 (test), indicating a potential overfitting issue as the performance drops significantly in the test set. (**G**-**I**) RDC-Model: The RDC-Model maintains a high AUC of 0.794 (train), 0.777 (validate), and 0.787 (test), suggesting the most stable and reliable performance among the three models.

Discussion

This multicenter, retrospective cohort study successfully developed and validated a predictive model that integrates MRI-based radiomics, deep learning features, and clinical data to estimate the risk of HCC recurrence following thermal ablation. The RDC-model, which combines radiomic and deep learning features with clinical indicators, demonstrated robust predictive performance, achieving AUC values of 0.794 in the train set, 0.777 in the validate set, and 0.787 in the test set. These findings highlight the strength of combining imaging data with clinical variables, as this integrated approach improved the model's ability to predict recurrence compared to models based solely on radiomics or deep learning features.

Recent advancements in medical imaging prediction have increasingly integrated radiomics and deep learning. For instance, Zhao's research on early postsurgical recurrence prediction using multisequence MRI achieved AUCs of 0.757 and 0.758, in the training and validation phases, respectively. Performance improved significantly to 0.878 and 0.873 with the incorporation of image semantics.[24] Despite these findings, Zhao's study lacked external validation, which brings its generalizability into question. Conversely, Gao's research, which included gadobenate dimeglumine-enhanced

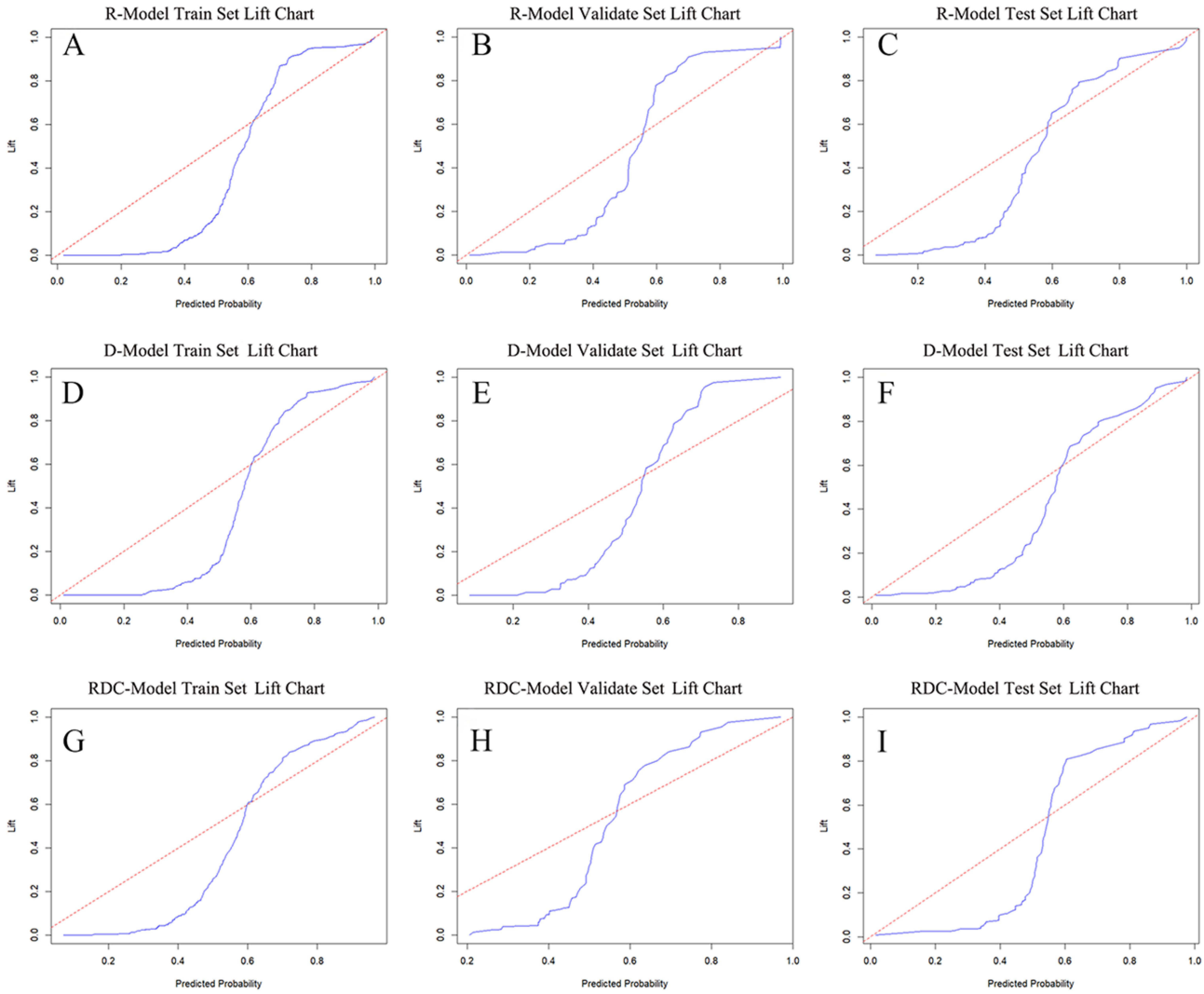

Figure 4 Lift Chart Analysis for R-Model, D-Model, and RDC-Model (**A-C**) R-Model Lift Charts: The R-Model's lift charts for the train, validate, and test sets show the model's ability to rank predicted probabilities of positive outcomes. The train set lift chart illustrates how well the model ranks higher-risk cases compared to the baseline random selection, with higher lift indicating better model performance. The validate set lift chart serves as a key evaluation of the model's generalizability to unseen data, reflecting its consistency in ranking positive outcomes. (**D-F**) D-Model Lift Charts: The D-Model's lift charts for the train, validate, and test sets display the predicted probabilities, with the train set chart indicating the model's effectiveness in prioritizing high-risk cases. The validate set lift chart assesses the model's performance on unseen data, demonstrating its predictive stability and potential improvement over random chance. (**G-I**) RDC-Model Lift Charts: The RDC-Model's lift charts for the train, validate, and test sets illustrate its ranking effectiveness. The train set lift chart shows the model's ability to separate high-risk from low-risk cases, highlighting its superior lift compared to the baseline. The validate set chart, on the other hand, validates the model's predictive accuracy on new data, confirming its practical utility in identifying true positives.

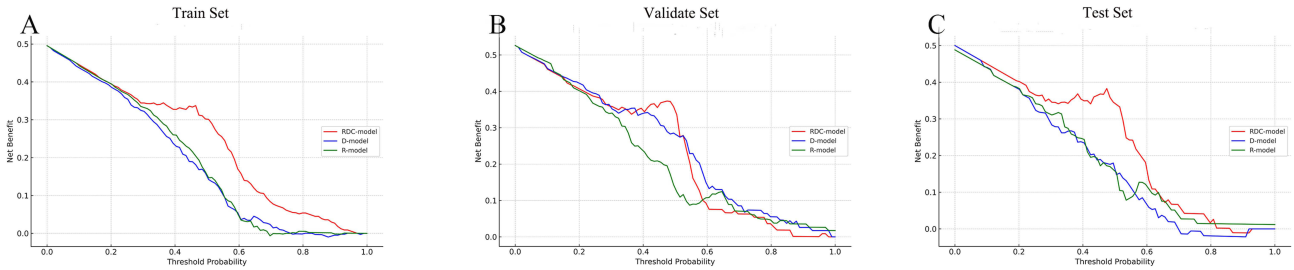

Figure 5 Decision Curve Analysis for R-Model, D-Model, and RDC-Model in Train, Validate, and Test Sets (**A**) Train Set DCA Comparison: The DCA curves for the R-Model, D-Model, and RDC-Model in the training set are compared. (**B**) Validate Set DCA Comparison: The validation set DCA curves offer an independent evaluation of the models' predictive capabilities. (**C**) Test Set DCA Comparison: The test set DCA curves serve as a final evaluation of the models' performance.

MRI radiomics and deep learning, surpassed the performance of our RDC model in the early phases of the study.[17] This discrepancy could be attributed to several factors: (1) The mean tumor size in Gao's study was larger, facilitating easier feature extraction; (2) A liver-specific contrast agent was used, enhancing tumor visibility beyond what was achieved with the standard agents in our study; and (3) The inclusion of additional imaging sequences such as ADC, DWI, and hepatocyte phase sequences provided a more comprehensive analysis.

While these studies underscore the benefits of integrating radiomics and deep learning, they did not demonstrate significant improvements from adding clinical data. In contrast, our study revealed improvements in model performance with the integration of clinically relevant variables such as albumin levels and tumor count, which were rigorously analyzed through univariate and multivariate methods, resulting in a test set AUC of 0.787.

In our analysis, we conducted a comparative evaluation of three predictive models, the R-model (based on radiomics), the D-model (based on deep learning), and the RDC-model, which integrate both radiomics and deep learning with clinical data. While both the R-model and the D-model demonstrated substantial potential, achieving AUCs of 0.75 and 0.72, respectively, they were outperformed by the RDC model, which achieved greater efficacy with an AUC of 0.787.

The R-model focuses on extracting and analyzing image-based features but has limited ability to capture the complex biological variability of HCC. Similarly, the D-model utilizes 3D-CNN to analyze imaging data but fails to provide a comprehensive prognosis due to the lack of clinical context. In contrast, the RDC model's incorporation of clinical parameters such as Alb concentration and tumor count significantly enhanced its predictive power. The integration of these clinical factors improved the model's AUC from 0.75 in the R-model and 0.72 in the D-model to 0.787 in the RDC-model.

This comprehensive approach underscores the importance of combining radiomics and deep learning with clinical insights to form a holistic view of the patient's condition. By embedding clinical data into the model, the RDC model not only achieved higher accuracy but also provided a more reliable tool for clinicians to predict HCC recurrence. This finding supports the broader notion that while radiomics and deep learning alone offer valuable insights, their fusion with clinically relevant data is essential for enhancing prognostic accuracy and tailoring patient-specific therapeutic strategies.

The predictive model developed in this study has potential clinical applications by aiding in the early identification of patients at higher risk of recurrence following thermal ablation. By incorporating personalized risk assessments, the tool can guide decisions such as more frequent follow-up imaging or earlier interventions for high-risk patients, potentially improving outcomes by enabling earlier detection of recurrence.

Integrating this model into clinical workflows could streamline decision-making and resource allocation, providing objective, consistent risk assessments. Future prospective studies will be necessary to validate the model's real-world impact on patient outcomes, but these findings suggest that it can enhance patient care through more personalized and data-driven management strategies.

Despite its insights, our study has limitations. Its retrospective design and the inherent associated biases demand prospective, multicenter studies for broader validation of the model's applicability. Moreover, the performance of deep learning models tends to improve with larger datasets, which were not available in this study. Additionally, the influence of tumor location on recurrence rates remains an underexplored area that requires further investigation. We acknowledge potential biases introduced by the retrospective design and have specifically focused on a 2-year time point for studying recurrence. To further assess the generalizability of the model, we plan future prospective studies that will validate the model's applicability over different time horizons and across diverse patient populations.

Conclusions

In summary, a hybrid model that integrates radiomics, deep learning characteristics, and clinical information shows promise as an effective tool for predicting recurrence following HCC thermal ablation.

Abbreviations

MRI, Magnetic resonance imaging; HCC, Hepatocellular carcinoma; AUC, Area under the curve; CT, computed tomography; ECOG, Eastern Cooperative Oncology Group; PACS, Picture archiving and communication system; DICOM, Digital imaging and communications in medicine; ICC, Intraclass correlation coefficient; IBSI, imaging biomarker standardization initiative;

ROC, Receiver operating characteristic; DCA, Decision curve analysis; HR, Hazard ratio; ADC, Analog-to-digital converter; DWI, Diffusion weighted imaging.

Data Sharing Statement

To protect patient confidentiality, the raw data are not publicly accessible. However, they can be requested from Junpeng Luo at hiccp@henu.edu.cn for legitimate inquiries.

Ethics Approval and Consent to Participate

The study design was approved by the Ethics Committees of Sun Yat-sen University Cancer Center, Henan Provincial Cancer Hospital, and Huaihe Hospital of Henan University (ethics approval number: SL-B2022-047-02). Due to the retrospective nature of the study, informed consent was waived.

Acknowledgments

The authors kindly thank Lujun Han (Sun Yat-sen University Cancer Center) and Xiang Geng (Henan Provincial Cancer Hospital) for providing image data and conducting quality inspection on the image quality.

Funding

This work was supported by the following grants: Henan University Interdisciplinary Advanced Research Institute Construction Project (CX3070A0970005) Henan Provincial Medical Science and Technology Public Relations Program Provincial Ministerial Co-Construction Key Project (No: SBGJ202302093) Henan Provincial Medical Science and Technology Research Project (Grant Number: LHGJ20210206, YZ) Henan Provincial Scientific and Technological Project (Grant Number: 232102310102, YZ).

Disclosure

The authors declare that they have no competing interests.

References

1. Petrick JL, Florio AA, Znaor A, et al. International trends in hepatocellular carcinoma incidence, 1978-2012. *Int J Canc*. 2020;147(2):317–330. doi:10.1002/ijc.32723
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. Lencioni R, Cioni D, Crocetti L, et al. Early-stage hepatocellular carcinoma in patients with cirrhosis: long-term results of percutaneous image-guided radiofrequency ablation. *Radiology*. 2005;234(3):961–967. doi:10.1148/radiol.2343040350
4. Lee MW, Kang D, Lim HK, et al.: Updated 10-year outcomes of percutaneous radiofrequency ablation as first-line therapy for single hepatocellular carcinoma < 3 cm: emphasis on association of local tumor progression and overall survival. *Eur Radiol* 2020, **30**(4):2391–2400.
5. Abdelaziz AO, Nabeel MM, Elbaz TM, et al. Microwave ablation versus transarterial chemoembolization in large hepatocellular carcinoma: prospective analysis. *Scand J Gastroenterol*. 2015;50(4):479–484. doi:10.3109/00365521.2014.1003397
6. Hermida M, Cassinotto C, Piron L, et al. Multimodal Percutaneous Thermal Ablation of Small Hepatocellular Carcinoma: predictive Factors of Recurrence and Survival in Western Patients. *Cancers*. 2020;12(2):313. doi:10.3390/cancers12020313
7. Bo XW, Sun LP, Yu SY, Xu HX. Thermal ablation and immunotherapy for hepatocellular carcinoma: recent advances and future directions. *World J Gastrointest Oncol*. 2021;13(10):1397–1411. doi:10.4251/wjgo.v13.i10.1397
8. 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
9. Shiina S, Tateishi R, Arano T, et al. Radiofrequency ablation for hepatocellular carcinoma: 10-year outcome and prognostic factors. *Am J Gastroenterol*. 2012;107(4):569–577. doi:10.1038/ajg.2011.425
10. Fan F, Dong G, Han C, et al. Peripheral immune factors aiding clinical parameter for better early recurrence prediction of hepatocellular carcinoma after thermal ablation. *Int J Hyperthermia*. 2023;40(1):2172219. doi:10.1080/02656736.2023.2172219
11. Chen QF, Liu S, Lyu N, Jia Z, Chen M, Zhao M. Surveillance Strategy after Complete Ablation of Initial Recurrent Hepatocellular Carcinoma: a Risk-Based Machine Learning Study. *J Vasc Interv Radiol*. 2021;32(11):1548–1557e1542. doi:10.1016/j.jvir.2021.07.025
12. Abdelhamed W, El-Kassas M. Non-surgical Treatment Options in Managing Recurrent Hepatocellular Carcinoma. *Gene Expression*. 2023;22(3):211–221. doi:10.14218/GE.2023.00010
13. Chinnaratha MA, Sathananthan D, Pateria P, et al. High local recurrence of early-stage hepatocellular carcinoma after percutaneous thermal ablation in routine clinical practice. *Eur J Gastroenterol Hepatol*. 2015;27(3):349–354. doi:10.1097/MEG.0000000000000270
14. Bai XM, Cui M, Yang W, et al. The 10-year Survival Analysis of Radiofrequency Ablation for Solitary Hepatocellular Carcinoma 5 cm or Smaller: primary versus Recurrent HCC. *Radiology*. 2021;300(2):458–469. doi:10.1148/radiol.2021200153

15. Dong TT, Wang L, Li M, Yin C, Li YY, Nie F. Clinical Results, Risk Factors, and Future Directions of Ultrasound-Guided Percutaneous Microwave Ablation for Hepatocellular Carcinoma. *J Hepatocell Carcinoma*. 2023;10:733–743. doi:10.2147/JHC.S409011
16. Wang W, Chen Q, Iwamoto Y, et al. Deep Learning-Based Radiomics Models for Early Recurrence Prediction of Hepatocellular Carcinoma with Multi-phase CT Images and Clinical Data. *Annu Int Conf IEEE Eng Med Biol Soc*. 2019;2019:4881–4884. doi:10.1109/EMBC.2019.8856356
17. Gao W, Wang W, Song D, et al. A predictive model integrating deep and radiomics features based on gadobenate dimeglumine-enhanced MRI for postoperative early recurrence of hepatocellular carcinoma. *Radiol Med*. 2022;127(3):259–271. doi:10.1007/s11547-021-01445-6
18. Zwanenburg A, Vallieres M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020;295(2):328–338.
19. 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
20. Nowacki A. Chi-square and Fisher's exact tests. *Clevel Clinic J Med*. 2017;84(9 suppl 2):e20–e25. doi:10.3949/ccjm.84.s2.04
21. McKnight PE, Najab J. Mann-Whitney U Test. *Corsini Ency Psych*. 1.
22. Zhukovska MOO VV. Statistical Software R in Corpus-Driven Research and Machine Learning. *Info Techn Learn Tool*. 2021;86(6):1–18.
23. Jones E, Harden S, Crawley MJ. *The R book*. Third edition. edn ed. Hoboken, NJ: Wiley; 2022.
24. Zhao Y, Wu J, Zhang Q, et al. Radiomics Analysis Based on Multiparametric MRI for Predicting Early Recurrence in Hepatocellular Carcinoma After Partial Hepatectomy. *J Magn Reson Imagi*. 2021;53(4):1066–1079. doi:10.1002/jmri.27424