Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

A CT-Based Deep Learning Radiomics Scoring System for Predicting the Prognosis to Repeat TACE in Patients with Hepatocellular Carcinoma: A Multicenter Cohort Study

Yanmei Dai[1,2], Sheng Zhao[2], Qiong Wu[2], Jin Zhang[3], Xu Zeng[2], Huijie Jiang[2]

[1]Department of Radiology, Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, Affiliated Cancer Hospital of University of Electronic Science and Technology of China, Chengdu, 610041, People's Republic of China; [2]Department of Radiology, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, 150086, People's Republic of China; [3]Department of Radiology, Xijing Hospital, Air Force Medical University, Xi'an, Shaanxi, People's Republic of China

Correspondence: Huijie Jiang, Email jianghuijie@hrbmu.edu.cn

Purpose: This study aimed to construct a novel retreatment scoring system to screen patients with hepatocellular carcinoma (HCC) who could benefit further after transarterial chemoembolization (TACE).

Patients and Methods: 310 patients with HCC were retrospectively recruited from three hospitals. The training and validation cohort were randomly selected from Center 1, and two external testing cohorts comprised from Center 2 and Center 3, respectively. Deep learning score and handcrafted radiomics signatures were constructed from the pretreatment arterial-phase and venous-phase CT images. The optimal features were screened using SelectKBest and LASSO regression. The AUC of the optimal combined model, consisting of HBsAg, five radiomics features, and DLscore, was 0.97, 0.89, 0.76, and 0.84 in the four cohorts, respectively. The optimal model was well calibrated. The prediction performance was assessed with respect to receiver operating characteristics, calibration, and decision curve analysis. Kaplan-Meier survival curves based on the scoring system were used to estimate the overall survival (OS).

Results: The optimal combined model consisted of HBsAg, 5 radiomics signatures, and DLscore, which AUC in four cohorts was 0.97, 0.89, 0.76, and 0.84, respectively, with good calibration. Decision curve analysis confirmed that the combined model was clinically useful. After Cox regression analysis of these characteristics, the scoring system (HBsAg-Radscore-DLscore, HRD) was significantly associated with OS in patients with HCC, and was superior to the traditional ART score and ABCR score between high and low-risk patients.

Conclusion: Deep learning and radiomics had good performance in predicting the OS of patients with HCC treated with repeated TACE. The HRD score is a potentially valuable and intelligent prognostic scoring system better than the traditional score.

Keywords: hepatocellular carcinoma, transarterial chemoembolization, prognostic score, radiomics, deep learning

Introduction

With the improvement of precision medicine, the characteristics of hepatocellular carcinoma (HCC) treatment include multidisciplinary participation in the combination of multiple treatment options and the selection of reasonable treatments for patients with HCC of different stages that can maximize the efficacy of treatment.[1,2] Transarterial chemoembolization (TACE) is one of the common methods of locoregional therapies, especially for patients with HCC in stage B of Barcelona Clinic Liver Cancer (BCLC), and most guidelines recognize it as the preferred treatment option.[3] TACE for patients with initially unresectable HCC can achieve transformation and create opportunities for surgical resection and ablation.[4,5] In addition, TACE can be used as an adjuvant therapy after surgical resection, and some studies have shown that postoperative TACE has the effect of reducing recurrence and prolonging survival time in patients with a high risk of recurrence.[6] In the clinic, it is necessary to determine whether patients need retreatment with TACE according to the follow-up results,

Received: 2 March 2025
Accepted: 18 July 2025
Published: 29 July 2025

especially in patients with giant HCC, who often need three or more TACE treatments to achieve a complete response. However, repeated ineffective TACE not only fails to control tumor progression but also aggravates liver function injury, and affects the prognosis of patients.[7] It is then crucial to establish a criterion that enables the timely termination of repeated TACE (re-TACE). Previously, some studies[8,9] constructed retreatment scores (such as ABCR score, ART score) to assess the survival of patients with HCC after two consecutive TACE sessions, identifying patients with dismal prognoses who could not benefit from re-TACE. Nevertheless, these scores, composed of basic preoperative and postoperative clinical and imaging characteristics, are single and low performance for survival prediction. In recent years, the application of radiomics in the field of liver cancer has gradually deepened and matured.[10] Through high-throughput extraction of quantitative features in medical images, machine learning (ML) algorithms are applied to screen potentially valuable features and construct prediction models,[11] which are used for diagnosis, staging, efficacy, and prognosis of liver cancer.[12] Additionally, deep learning (DL), as a branch of ML, has considerable value in constructing prediction models, especially based on the background of big data, DL models may have higher and more stable performance than ML models.

This study aims to develop a new, comprehensive, stable, and efficient retreatment scoring system using multi-dimensional features to assist clinical judgment of whether patients with HCC treated with initial TACE are suitable for re-TACE and whether they can achieve long-term efficacy, to guide the clinic to adjust the subsequent treatment plan promptly, improve the overall efficacy and prognosis, and ultimately prolong overall survival (OS) in patients with HCC.

Materials and Methods

Patients

The subjects were obtained from three centers, namely, Guangdong Provincial People's Hospital (Center 1), the First Affiliated Hospital of Sun Yat-sen University (Center 2), and the Second Affiliated Hospital of Harbin Medical University (Center 3), and the study was approved by the Institutional Ethics Committee and the research cooperation agreement. Due to the data being retrospective, the requirement for informed consent from patients was waived. This retrospective study included 145 patients with HCC from September 2007 to September 2019 in Center 1, 101 patients from January 2012 to November 2019 in Center 2, and 64 patients from February 2013 to September 2020 in Center 3. All patients in Center 1 were randomly divided into a training cohort and validation cohort in a ratio of 8:2, and both Center 2 and Center 3 were used as independent external testing cohorts, respectively. The inclusion criteria: a) patients with HCC were diagnosed by histological or noninvasive radiologic criteria;[3] b) patients received initial and sequential second TACE; c) before initial TACE, patients were approved for contrast-enhanced computed tomography (CECT); d) clinical information is complete. The exclusion criteria: a) patient has a history of other cancer; b) patients received other invasive treatments before initial TACE, such as radiofrequency ablation, surgery, radiotherapy, etc.; c) poor quality of CT images that might affect the analysis. The detailed selection flow chart is shown in Figure 1.

Treatment Procedure and Follow-Up

Percutaneous puncture of the right femoral artery was carried out using the Seldinger technique. After a successful puncture, a 5F arterial sheath was inserted. According to the angiographic findings, a 2.7F microcatheter was used to enter the tumor supply branch with super-selective intubation guided by the guidewire followed by injection of embolization drugs and chemotherapy drugs. The dose of embolization drugs and chemotherapeutic drugs was based on tumor burden and patients' characteristics. When the blood supply to the tumor is reduced, the operation is completed.

CECT was performed within 4–6 weeks after the operation to check the embolization of the tumor and determine whether there was complete necrosis or a "viable tumor".[13] According to the modified Response Evaluation Criteria in Solid Tumors (mRECIST),[14] the postoperative radiological responses of tumors were divided into four categories: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). Radiological responses were independently assessed by two radiologists with 5–10 years of experience and were blinded to clinical information. After the disease stabilized, patients were followed up every 8–12 weeks. OS was defined as the interval between the time of the second TACE and the time of death from causes relevant to this study. Data on patients lost to follow-up or alive as of June 30, 2022, were defined as censored data.

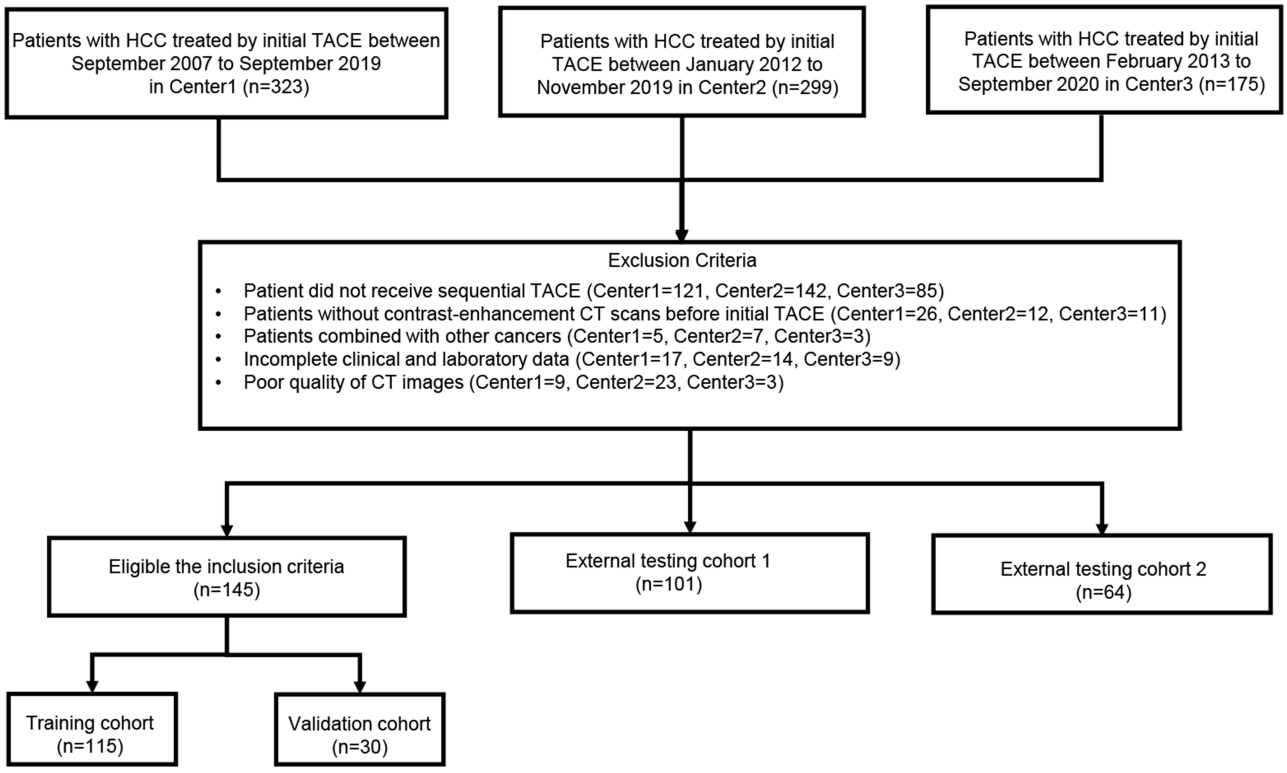

Figure 1 Flowchart of patients' enrollment.

Retreatment Prognosis Prediction Model

Traditional Radiomics Models

Image preprocessing and radiomics feature extraction methods were described in the supplementary material S1.1–S1.4. The whole flow chart is shown in Figure 2. After screening the radiomics signatures for arterial phase, venous phase, and combined arterial-venous phase, the optimal features are used to construct a conventional radiomics model using a support vector machine (SVM). Subsequent clinical models and each combined model were modeled using the SVM classifier.

Deep Learning Models

3D ResNeXt is a deep learning model based on 2D ResNeXt,[15] which can better process 3D data by introducing 3D convolutional layers and residual connectivity and has been widely used in medical image analysis with good performance.[16–18] In this study, the attention mechanism is added to improve the learning ability of the model (Figure 3). Details are in the supplementary material S1.5.

The experiments in this part were conducted in a Linux environment (Ubuntu 7.5.0), with a basic configuration of a CPU Intel E5-1650 3.50 GHz, 128G DDR4 memory, and four RTX 2080Ti graphics cards. The experiments were programmed using the Python language version 3.6.13 (Python Software Foundation), the deep learning framework used was PyTorch (version 1.8.2, https://pytorch.org/), and the main packages used were scikit-learn package version 0.20.4, torch package version 1.8.2, torchvision package version 0.9.2, and SimpleITK package version 1.2.4.

Clinical Models

Basic clinical features and CT imaging characteristics of all patients were collected, including gender, age, alpha-fetoprotein (AFP), HBsAg, cirrhosis (CT), vascularization, tumor thrombus, tumor number, longest diameter of tumors, BCLC stage, Child-Pugh class base, albumin bilirubin (ALBI), change in Child-Pugh score increase, first postoperative radiologic response and AST increase of 25%. Potential variables capable of predicting OS after re-TACE were screened by univariate analysis ($P<0.1$).

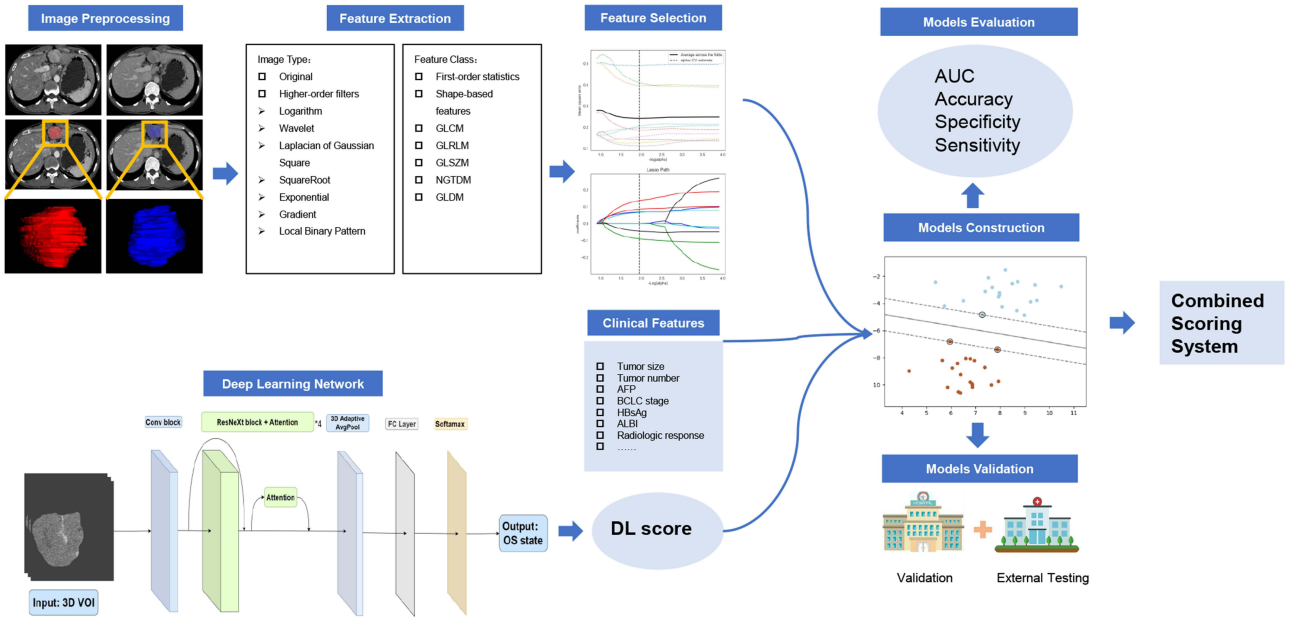

Figure 2 Overview of the flowchart in this study.

Combined Models

A total of six combined prediction models were included in this study, including the arterial-venous phase radiomics model, clinical model + arterial phase radiomics model, clinical model + venous phase radiomics model, clinical model + arterial-venous phase radiomics model, arterial-venous phase deep learning model, and arterial-venous phase radiomics model + arterial-venous phase deep learning + clinical model. Among them, the arterial-venous phase deep learning model takes the CT images in the arterial phase and venous phase together as the input of the deep learning network, and the predicted probability of its output is used as the DLscore.

Retreatment Score

The scores of patients with HCC treated with re-TACE, including the ART score and the ABCR score, were proposed according to previous studies.[8,9] The development of the novel score in this study was also based on the features with potential predictive value, and the hazard ratio (HR) value and weight B value of each feature were calculated by multivariable Cox regression analysis, and the B value of each feature was multiplied by 2 and then summed to obtain the novel score for each patient.

Statistical Analysis

Categorical variables were summarized as numbers (percent), and continuous variables were summarized as mean (standard deviation, Sd) or median (interquartile range, IQR) based on the underlying distribution of the data. The Chi-square test or Fisher's exact test (when the theoretical frequency is less than 5) was used for categorical variables, and the independent samples *t*-test or Mann–Whitney *U*-test (if the data did not satisfy normal distribution or the variance was not uniform) was used for continuous variables. The OS probabilities were evaluated by Kaplan-Meier survival analysis and the Log rank test. The optimal cutoff value was determined using the maximally selected rank statistics method, and patients were classified into high-risk or low-risk groups. Differences were considered statistically significant when $P<0.05$.

Results

Patient Characteristics

In the training set (n=115), the mean age of patients with HCC was 53.09±15.17 years, and there were 40 (34.78%), 46 (40.00%), and 29 (25.22%) patients with BCLC stages A, B, and C, respectively, and predominantly male patients (n=101, 87.83%). A total of 73 (63.48%) patients died up to the termination of follow-up in this dataset, and 42 (36.52%)

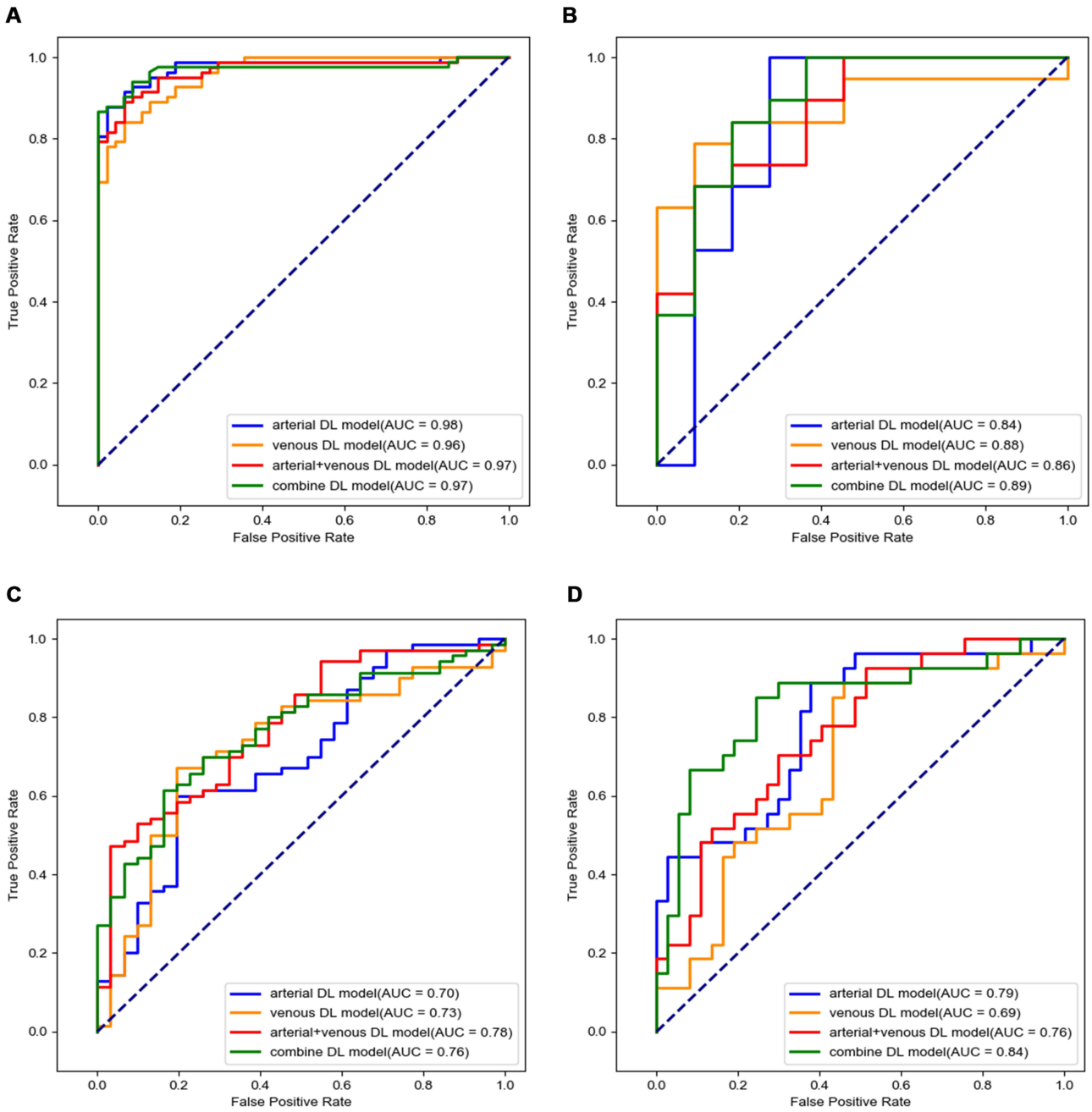

Figure 3 Receiver operating characteristic analysis of the predicting models. (**A**–**D**) are results from the training cohort, validation cohort, external testing cohort 1, and external testing cohort 2, respectively.

patients were censored. In the validation set, the mean age of patients with HCC was 52.80 ± 13.04 years, and there were 8 (26.67%), 12 (40.00%), and 10 (33.33%) patients with BCLC stages A, B, and C, respectively, and predominantly male patients (n=29, 96.67%). A total of 19 (63.33%) patients in this dataset died by the end of follow-up, and 11 (36.67%) patients experienced deletion. In external testing cohort 1, the mean age of patients with HCC was 53.02 ± 11.53 years, and there were 31 (30.69%), 22 (21.78%), and 48 (47.52%) patients with BCLC stages A, B, and C, respectively, and predominantly male patients (n = 95, 94.06%). A total of 70 (69.31%) patients died up to the termination of follow-up in this dataset, and 31 (30.69%) patients were censored. In external testing cohort 2, the mean age of patients with HCC was 56.48 ± 7.31 years, and there were 27 (42.19%), 23 (35.94%), and 14 (21.88%) patients with BCLC stages A, B, and C, respectively, and predominantly male patients (n = 58, 90.63%). A total of 27 (42.19%) patients died by the end of follow-up and 37 (57.81%) patients were censored in this dataset. Details in Table 1.

https://doi.org/10.2147/JHC.S525920

Table 1 Baseline Characteristics of Patients with Hepatocellular Carcinoma in Four Cohorts

Characteristic	Training Cohort (N=115, %)	Validation Cohort (N=30, %)	External Testing Cohort 1 (N=101, %)	External Testing Cohort 2 (N=64, %)
Gender				
Male	101 (87.83)	29 (96.67)	95 (94.06)	58 (90.63)
Female	14 (12.17)	1 (3.33)	6 (5.94)	6 (9.38)
Age, years	53.09±15.17	52.80±13.04	53.02±11.53	56.48±7.31
BCLC Stage				
A	40 (34.78)	8 (26.67)	31 (30.69)	27 (42.19)
B	46 (40.00)	12 (40.00)	22 (21.78)	23 (35.94)
C	29 (25.22)	10 (33.33)	48 (47.52)	14 (21.88)
AFP, ng/mL				
≤200	47 (40.87)	14 (46.67)	44 (43.56)	41 (64.06)
>200	68 (59.13)	16 (53.33)	57 (56.44)	23 (35.94)
HBsAg				
Positive	91 (79.13)	28 (93.33)	86 (85.15)	52 (81.25)
Negative	24 (20.87)	2 (6.67)	15 (14.85)	12 (18.75)
Cirrhosis (CT)				
I/II grade	105 (91.30)	27 (90.00)	81 (80.20)	55 (85.94)
III/IV grade	10 (8.70)	3 (10.00)	20 (19.80)	9 (14.06)
Vascularization				
I/II grade	36 (31.30)	11 (36.67)	20 (19.80)	26 (40.63)
III/IV grade	79 (68.70)	19 (63.33)	81 (80.20)	38 (59.38)
Tumor Thrombus				
Yes	86 (74.78)	20 (66.67)	53 (52.48)	50 (78.13)
No	29 (25.22)	10 (33.33)	48 (47.52)	14 (21.88)
Tumor Number				
1	38 (33.04)	6 (20.00)	30 (29.70)	32 (50.00)
≥2	77 (66.96)	24 (80.00)	71 (70.30)	32 (50.00)
Longest Diameter, cm	79.97±39.95	83.52±46.17	83.81±41.75	57.08±32.88
Response				
CR	0 (0)	0 (0)	1 (0.99)	3 (4.69)
PR	47 (40.87)	12 (40.00)	62 (61.39)	32 (50.00)
SD	53 (46.09)	15 (50.00)	23 (22.77)	22 (34.38)
PD	15 (13.04)	3 (10.00)	15 (14.85)	7 (10.94)
Child-Pugh Class Base				
A	85 (73.91)	23 (76.67)	78 (77.23)	52 (81.25)
B	26 (22.61)	6 (20.00)	23 (22.77)	12 (18.75)
C	4 (3.48)	1 (3.33)	0 (0)	0 (0)
ALBI				
1	17 (14.78)	2 (6.67)	18 (17.82)	19 (29.69)
2	90 (78.26)	23 (76.67)	80 (79.21)	45 (70.31)
3	8 (6.96)	5 (16.67)	3 (2.97)	0 (0)

(*Continued*)

Table 1 (Continued).

Characteristic	Training Cohort (N=115, %)	Validation Cohort (N=30, %)	External Testing Cohort 1 (N=101, %)	External Testing Cohort 2 (N=64, %)
Child-Pugh Score Increase				
0	97 (84.35)	22 (73.33)	57 (56.44)	55 (85.94)
+1 point	18 (15.65)	8 (26.67)	44 (43.56)	9 (14.06)
AST Increase				
≤25%	93 (80.87)	22 (73.33)	78 (77.23)	51 (79.69)
>25%	22 (19.13)	8 (26.67)	23 (22.77)	13 (20.31)

Prognosis Model

After comparing the prediction performance of each model, it was found that the combined model consisting of five radiomics signatures in the arterial-venous phase, DLscore, and HBsAg had the optimal prediction ability. The AUC, accuracy, sensitivity, and specificity of the optimal model was 0.97, 0.92, 0.96, and 0.85 in the training cohort, respectively; was 0.89, 0.80, 0.84, and 0.73 in the validation cohort, respectively; was 0.76, 0.70, 0.71, and 0.68 in the external testing cohort 1, respectively; and was 0.84, 0.77, 0.67, and 0.84 in external testing cohort 2, respectively (Figure 3). The AUC of the optimal model was slightly lower both in the external testing cohort 1 and cohort 2 than in the training and validation cohort. The process of constructing each model is described in the supplementary material S2.1–S2.4. The decision curves in the training cohort and validation cohort showed that within a certain range of thresholds, the net clinical benefit of the optimal model is superior to the clinical model and radiomics models (Figure 4). Subsequently, we performed a risk stratification analysis of patients in Center 1 based on features within the optimal model, as described in the supplementary material S2.5.

K-M Analysis of the Prognosis Scoring System

In this study, the retreatment score was obtained by multiplying the B value of each variable in the optimal model after multivariate Cox regression analysis by two (Table 2). The novel score (HBsAg-Radscore-DLscore, HRD) = Radscore (high-risk subgroup) × 2.73 + DLscore (high-risk subgroup) × 12.15 + HBsAg (positive) × 1.37.

In Center 1, the patients were divided into high-risk and low-risk subgroups according to a cutoff value (4.10). We found that the difference between the two subgroups was statistically significant in predicting OS after re-TACE ($P<0.001$) (Figure 5A). In the low-risk subgroup, the median OS could not be calculated because of the small number of patients who experienced death. The median OS for patients in the high-risk subgroup was 224 days (95% CI: 172.44–275.56). In external testing cohort 1, the median OS was 336 days (95% CI: 0–675.99) in the high-risk subgroup of HRD and 744 days (95% CI: 698.82–1089.18) in the low-risk subgroup, and the difference in OS between the two subgroups was statistically significant ($P=0.003$) (Figure 5D). In external testing cohort 2, the median OS was 1019 days (95% CI: 161.89–1876.12) in the high-HRD risk subgroup and 1403 days (95% CI: 773.90–2032.10) in the low-risk subgroup, but the difference in OS between the two subgroups was not statistically significant (P=0.875) (Figure 5G). Figure 6 demonstrates CT images of two patients with HCC (high HRD score and low HRD score) before initial TACE, after initial TACE, and after re-TACE.

In Center 1, 107 patients had an ART score of 0–1.5, with a median OS of 584 days (95% CI: 316.32–851.68); 38 patients had an ART score of ≥2.5, with a median OS of 348 days (95% CI: 286.50–409.50), and the difference in OS between the two subgroups was not statistically significant ($P=0.220$) (Figure 5B). 66 patients had an ABCR score of 0, with a median OS of 564 days (95% CI: 274.77–853.24); 57 patients had an ABCR score of 1–3, with a median OS of 438 days (95% CI: 126.41–749.60); 22 patients had an ABCR score of ≥4, with a median OS of 336 days (95% CI: 209.03–462.97), and the difference in OS between the three subgroups was again not statistically significant ($P=0.775$) (Figure 5C). In the external testing cohort 1, the median OS was 391 days (95% CI: 175.14–606.87) in the high-risk subgroup of ART score and 594 days (95% CI: 520.39–667.62) in the low-risk subgroup, and the difference in OS between the two subgroups was not statistically significant ($P=0.525$) (Figure 5E). The high-risk, intermediate-risk, and low-risk subgroup of ABCR score had a median OS of 155 days (95% CI: 100.99–209.01), 553 days (95% CI: 225.44–880.56), and

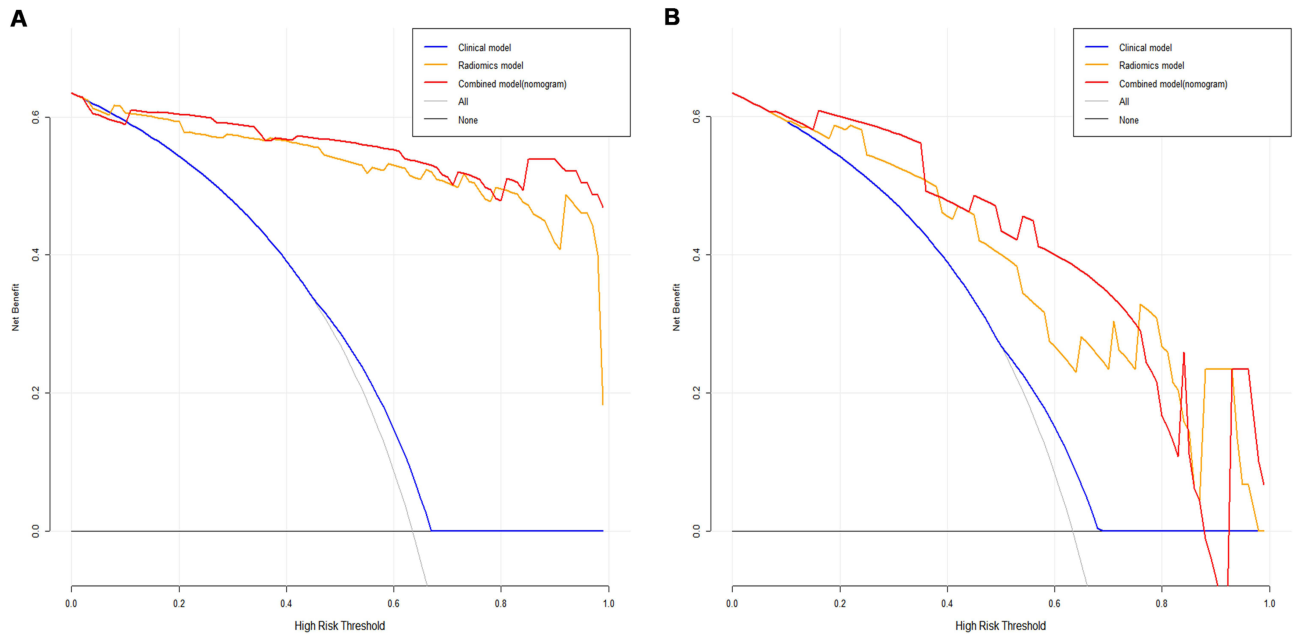

Figure 4 Decision curve analysis of the predicting models. (**A** and **B**) are results from the training cohort, and validation cohort.

691 days (95% CI: 89.59–886.59), respectively, and the difference in OS among the three subgroups was statistical significance (*P*=0.004) (Figure 5F). In external testing cohort 2, the median OS was 863 days (95% CI: 593.421–1132.58) in the high-risk subgroup of ART score and 1430 days (95% CI: 853.46–2006.54) in the low-risk subgroup, and the difference in OS between the two subgroups was not statistically significant (*P*=0.198) (Figure 5H). The median OS of the intermediate-risk subgroup of ABCR score was 632 days (95% CI: 370.30–893.70), and its median OS and 95% CI could not be calculated due to the small number of patients in the high-risk subgroup and low-risk subgroup. However, the difference in OS among the three subgroups was statistically significant (*P*=0.018) (Figure 5I).

Discussion

In this study, we used deep learning and radiomics to construct an optimal model to predict the OS status of patients with HCC after re-TACE and develop a novel HRD score to assess the high-risk and low-risk patients, which is a potential and

Table 2 New Score for Predicting Overall Survival of Patients with Hepatocellular Carcinoma After Repeat Transarterial Chemoembolization

Variable	Overall Survival			HRD Score
	HR	95% CI	B	
HBsAg				
Negative	1			–
Positive	1.24	0.67–2.32	0.685	1.37
Radscore				
Lowscore	1			–
Highscore	1.41	0.86–2.30	1.367	2.73
DLscore				
Lowscore	1			–
Highscore	5.89	3.32–10.43	6.077	12.15

Notes: HRD, HBsAg + Radscore + DLscore. Combined score=B*2.

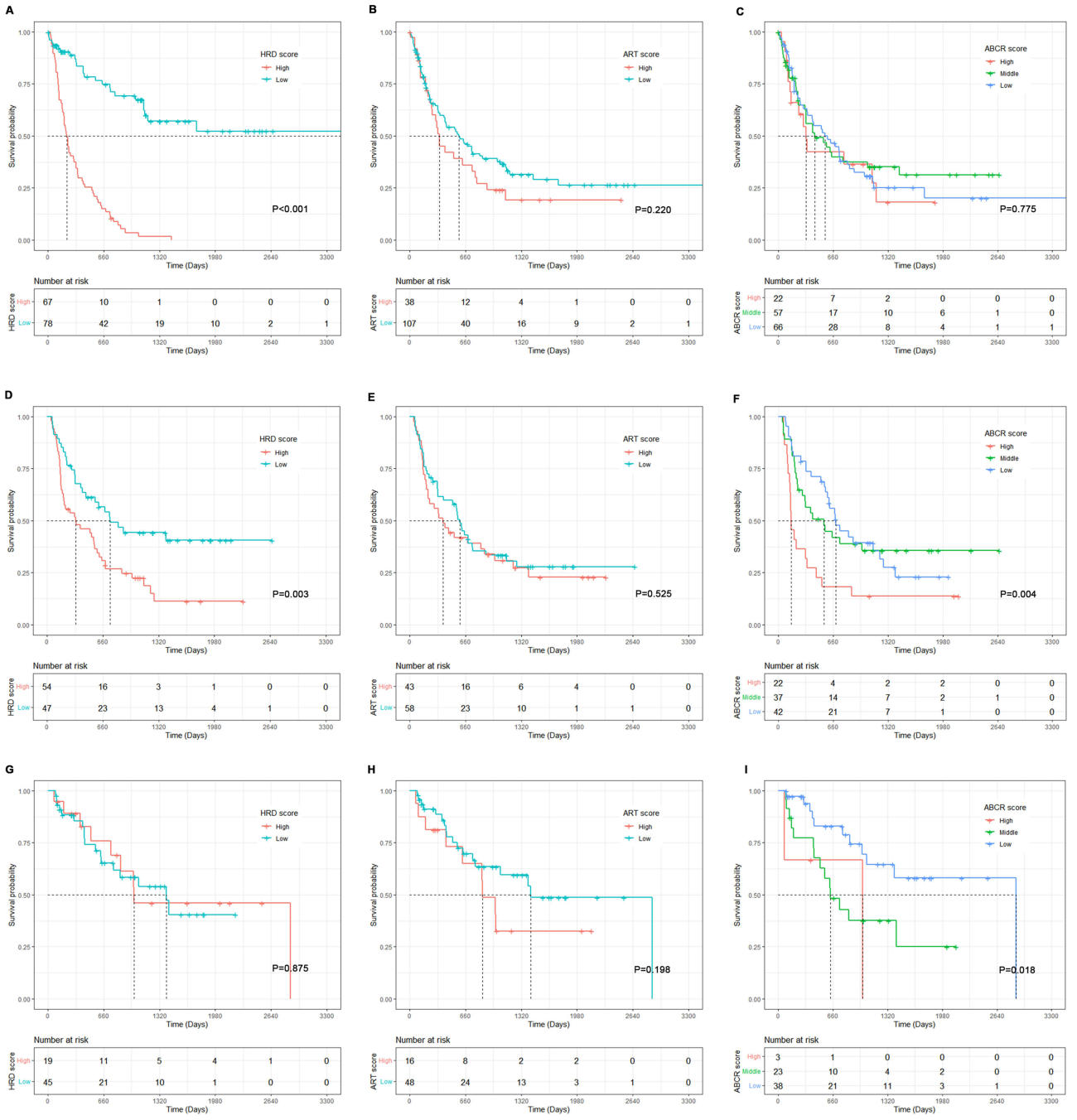

Figure 5 Kaplan-Meier survival analysis of HRD score, ART score, and ABCR score in center 1, external testing cohort 1, and external testing cohort 2. (**A**–**C**) are the results in center 1, the difference in OS between subgroups risk stratified by HRD score was statistically significant (*P*<0.001). (**D**–**F**) are the results in external testing cohort 1; the difference in OS between subgroups stratified by HRD score and ABCR score was statistically significant (*P*=0.003, *P*=0.004). (**G**–**I**) are the results in external testing cohort 2, the difference in OS between subgroups risk stratified by ABCR score was statistically significant (*P*=0.018).

intelligent scoring system, and better than the traditional scores. For patients with HCC who need to receive multiple TACE treatments, it is crucial to accurately assess the treatment efficacy and determine when it is necessary to terminate TACE treatment and switch to other treatment modalities, which will help to improve the treatment efficacy of the patients, protect the liver function on time, and even prolong the overall survival of the patients. Based on this starting point, the present study constructed a predictive model and a novel score for retreatment efficacy assessment based on artificial intelligence to find out the population of patients with HCC who can benefit from TACE and those who are unsuitable for TACE, and who are recommended to switch to other treatments according to the relevant guidelines. For

https://doi.org/10.2147/JHC.S525920

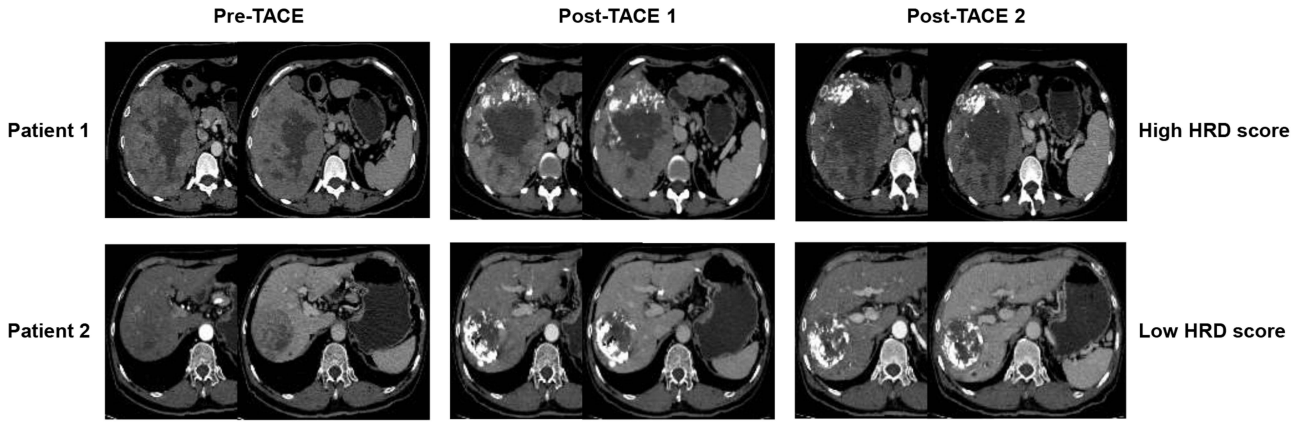

Figure 6 CT images of two patients before transarterial chemoembolization, after initial treatment, and after re-treatment, respectively. From the CT images (both arterial phase and venous phase), it can be seen that patient 1 with a high HRD score had a larger tumor with significant internal necrosis, less iodine oil deposition after both initial and repeated TACE and no significant reduction in the volume of the living tumor; patient 2 with a low HRD score had a smaller tumor, with a gradual increase in iodine oil deposition and a reduction in the volume of the living tumor after two consecutive TACE sessions.

those patients who are not suitable for re-TACE, what treatment modalities (eg, radiation, immunotherapy, targeted therapy, and target-immunotherapy) should be turned to, and which group of patients can obtain the optimal clinical benefit from which subsequent treatment modalities are worth exploring in future studies.

Previous studies have been conducted to analyze the prognostic factors affecting patients with HCC after TACE and various prognostic scores have been developed to find the patient that would derive the greatest benefit from TACE. Sieghart et al[9] developed the ART score which categorized patients into high-risk subgroup and low-risk subgroup, whose median OS was 27.6 and 8.1 months in the validation cohort, respectively. Adhoute et al[8] established the ABCR score which categorized patients into three risk subgroups, and the median OS of patients with HCC in the external validation cohort was 37.8, 17.1, and 7.5 months, respectively, and the median OS of patients in the subgroups differentiated by this score was shorter than that of those scored by ART score. These two scores are prognostic scores developed for patients who have received two consecutive TACE. The novel retreatment score (HRD score) developed in our study divided patients into high-risk and low-risk subgroups, with the median OS for patients in the high-risk subgroup in center 1 being 7.5 months. This result is consistent with the ABCR score. Unlike previous scores, this study did not limit the time to re-TACE but was selected on demand. In addition, the HRD score in this study was developed based on a multidimensional feature that included a Radscore and a deep learning score in addition to a clinical feature.

Few previous studies have utilized radiomics and deep learning networks to assess the prognosis of patients with HCC after TACE. Liu et al[19] developed deep learning scores capable of predicting survival using a multitask deep learning network trained on CT images in 243 patients with HCC treated with TACE and pointed out that the deep learning score was used as an independent risk factor affecting OS after TACE. Based on the above findings, we proposed the idea of applying radiomics signatures and deep learning to the prediction of survival after re-TACE in patients and developed a novel score (HRD score) to assist the clinic in early determination of the population of patients who are suitable for re-TACE, and to be able to provide advice on switching to other treatments for those who are not suitable for re-treatment.

Also, this study compared the value of the ART score and ABCR score in predicting OS. For the ART score, we did not find a statistically significant difference in OS between patients in the high-risk and low-risk subgroups delineated in the three centers (P=0.220; P=0.525; P=0.198). The possible reason for this is that our study did not limit the interval time between two TACE sessions, making the final scores different. However, previous studies have questioned and validated the efficacy of assessing ART scores.[20] Sieghart et al pointed out the reasons for the discrepancy: geographic differences in the study population (Japan vs Europe), and the differences in TACE operations (Japanese clinicians are more meticulous than European clinicians, which results in better outcomes for Japanese patients with HCC who receive TACE, resulting in fewer treatments and potentially longer survival). Terzi et al[21] similarly noted that the ART score may

not be suitable for assessing survival after re-TACE in Italian patients with HCC. In general, the traditional ART score may not be very applicable and the assessment of efficacy. For the ABCR score, although the difference in OS between the three subgroups of patients in center 1 was not statistically significant (P=0.775), the difference between in the two external testing cohorts was statistically significant (P=0.004; P=0.018). This shows that our HRD scores and ABCR scores are in good agreement.

To the best of our knowledge, studies that have used radiomics signatures and deep learning together with clinical features to predict survival after re-TACE are rare up to now, which is what makes this study unique and innovative. Nevertheless, there are some limitations of this study that need to be discussed. First, due to the data volume requirements of radiomics and deep learning, the final amount of data in this study is still small even though this is a multicenter study, and overfitting or underfitting may occur during the process of model training and validation, which makes the model less robust. Second, since the data in this study came from three different centers, the differences in CT scanning equipment and parameters may have an impact on the extracted radiomics signatures and the deep learning network. Third, although the optimal model shows better performance in the training cohort and the internal validation cohort, the AUC decreases in the two external test cohorts, which also indicates that our model generalizes poorly and needs further optimization. Fourth, the HR of DLscore was significantly higher than that of HBsAg and Radscore, which also indicates that the HRD score is not yet stable, and the weights of the three features differ greatly. In the future, further experiments will be needed to validate and improve the HRD scores.

Conclusion

The model based on multidimensional features including clinical feature, radiomics signatures derived from CT images, and deep learning predictive probabilities has good performance in predicting the prognosis of patients with re-TACE. The HRD score developed based on the optimal prediction model was a potentially valuable score in identifying patients with HCC who could benefit further after re-TACE. The optimal model and the novel score will help to screen patients with HCC who are no longer suitable for re-TACE, and can assist the clinic to terminate TACE on time to switch to other therapeutic modalities, so that the patients can obtain the optimal clinical benefits. Future studies should explore and analyze which treatment modality to switch to after failure of re-TACE can improve the outcome and prolong the prognosis of patients with HCC.

Abbreviations

AFP, Alpha-fetoprotein; ALBI, Albumin bilirubin; AUC, Area under the receive; BCLC, Barcelona Clinic Liver Cancer; CECT, Contrast-enhanced computed tomography; CI, Confidence interval; CR, Complete response; DICOM, Digital imaging and communications in medicine; DL, Deep learning; HCC, Hepatocellular carcinoma; HR, Hazard ratio; ICC, Inter/Intra-class correlation coefficient; IQR, Interquartile range; LASSO, Least absolute shrinkage and selection operator algorithm; LoG, Laplacian of Gaussian; LBP, Local Binary Pattern; ML, Machine learning; Mrecist, modified Response Evaluation Criteria in Solid Tumor; OS, Overall survival; PD, Progressive disease; PR, Partial response; ROC, Receiver operating characteristic; ROI, Region of interest; SD, Stable disease; Sd, Standard deviation; SVM, Support vector machine; TACE, Transarterial chemoembolization; VOI, Volume of interest.

Data Sharing Statement

The raw data for this study is available from the corresponding author upon reasonable request.

Ethical Approval

The ethics committee approval of the Second Affiliated Hospital of Harbin Medical University (NO.KY2019-217) was obtained for this study, and informed consent was waived due to the retrospective nature of this study. This study was based on a project led by the Second Affiliated Hospital of Harbin Medical University, with the First Affiliated Hospital of Sun Yat-sen University and the Guangdong Provincial People's Hospital as collaborators and signing of the relevant scientific agreements, and the confidentiality of patient data was protected throughout the study period. We state that we will keep patient data confidentiality and obey the Declaration of Helsinki.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by grants from the National Key Research and Development Program of China (2019YFC0118100), National Natural Science Foundation of China (U23A20478 and 62171167).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Llovet JM, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers*. 2021;7(1):6. doi:10.1038/s41572-020-00240-3
2. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301–1314. doi:10.1016/S0140-6736(18)30010-2
3. Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology*. 2018;67(1):358–380. doi:10.1002/hep.29086
4. Paul A, Sotiropoulos G, Gerken G, Saner FH. How to decide about liver transplantation in patients with hepatocellular carcinoma: size and number of lesions or response to TACE? *J Hepatol*. 2014;60(2):463–464. doi:10.1016/j.jhep.2013.09.032
5. Sapisochin G, Barry A, Doherty M, et al. Stereotactic body radiotherapy vs. TACE or RFA as a bridge to transplant in patients with hepatocellular carcinoma. An intention-to-treat analysis. *J Hepatol*. 2017;67(1):92–99. doi:10.1016/j.jhep.2017.02.022
6. Bureau of Medical Administration. National health commission of the people's republic of china. standardization for diagnosis and treatment of hepatocellular carcinoma (2022 edition)]. *Zhonghua Gan Zang Bing Za Zhi*. 2022;30(4):367–388. doi:10.3760/cma.j.cn501113-20220413-00193.
7. Clinical Guidelines Committee of Chinese College of Interventionalists. Expert consensus on transarterial chemoembolization refractoriness and subsequent therapies in hepatocellular carcinoma. *Zhonghua Nei Ke Za Zhi*. 2022;61(8):860–866. doi:10.3760/cma.j.cn112138-20220615-00452.
8. Adhoute X, Penaranda G, Naude S, et al. Retreatment with TACE: the ABCR SCORE, an aid to the decision-making process. *J Hepatol*. 2015;62(4):855–862. doi:10.1016/j.jhep.2014.11.014
9. Sieghart W, Hucke F, Pinter M, et al. The ART of decision making: retreatment with transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology*. 2013;57(6):2261–2273. doi:10.1002/hep.26256
10. Bo Z, Song J, He Q, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. *Comput Biol Med*. 2024;173:108337. doi:10.1016/j.compbiomed.2024.108337
11. 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
12. Huang Y, Qian H. advancing hepatocellular carcinoma management through peritumoral radiomics: enhancing diagnosis, treatment, and prognosis. *J Hepatocell Carcinoma*. 2024;11:2159–2168. doi:10.2147/JHC.S493227
13. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis*. 2010;30(1):52–60. doi:10.1055/s-0030-1247132
14. Llovet JM, Lencioni R. mRECIST for HCC: performance and novel refinements. *J Hepatol*. 2020;72(2):288–306. doi:10.1016/j.jhep.2019.09.026
15. Zhou T, Zhao Y, Wu J. ResNeXt and Res2Net structures for speaker verification. 2020.
16. G T, Sk K. A deep learning based effective model for brain tumor segmentation and classification using MRI images. *JAIT*. 2023;14(6):1280–1288. doi:10.12720/jait.14.6.1280-1288
17. Soleimani P, Farezi N. Utilizing deep learning via the 3D U-net neural network for the delineation of brain stroke lesions in MRI image. *Sci Rep*. 2023;13(1):19808. doi:10.1038/s41598-023-47107-7
18. Lee JO, Ahn SS, Choi KS, et al. Added prognostic value of 3D deep learning-derived features from preoperative MRI for adult-type diffuse gliomas. *Neuro-Oncology*. 2023:noad202. doi:10.1093/neuonc/noad202.
19. Liu Q-P, Xu X, Zhu F-P, Zhang Y-D, Liu X-S. Prediction of prognostic risk factors in hepatocellular carcinoma with transarterial chemoembolization using multi-modal multi-task deep learning. *EClinicalMedicine*. 2020;23:100379. doi:10.1016/j.eclinm.2020.100379
20. Kudo M, Arizumi T, Ueshima K. Assessment for retreatment (ART) score for repeated transarterial chemoembolization in patients with hepatocellular carcinoma. *Hepatology*. 2014;59(6):2424–2425. doi:10.1002/hep.26760
21. Terzi E, Terenzi L, Venerandi L, et al. The ART score is not effective to select patients for transarterial chemoembolization retreatment in an Italian series. *Dig Dis*. 2014;32(6):711–716. doi:10.1159/000368007