



# Radiomics analysis based on contrast-enhanced MRI for predicting short-term efficacy of drug-eluting beads transarterial chemoembolization in hepatocellular carcinoma

Zihan Xi<sup>1,3</sup> · Yuanxin Ye<sup>1</sup> · Yongbo Yang<sup>1</sup> · Yiwei He<sup>1</sup> · Ziyang Song<sup>1</sup> · Qian Ma<sup>1,3</sup> · Hui Zeng<sup>2</sup> · Guoliang Shao<sup>2</sup> 

Received: 7 October 2023 / Revised: 28 March 2024 / Accepted: 29 March 2024 / Published online: 19 July 2024  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

## Abstract

**Objective** We developed and validated a clinical-radiomics model for preoperative prediction of the short-term efficacy of initial drug-eluting beads transarterial chemoembolization (D-TACE) treatment in patients with hepatocellular carcinoma (HCC).

**Methods** In this retrospective cohort study of 113 patients with intermediate and advanced HCC, 5343 features were extracted based on three sequences of the arterial phase (AP), diffusion-weighted imaging, and T2-weighted images based on contrast-enhanced magnetic resonance imaging, and minimum redundancy maximum correlation and least absolute shrinkage and selection operator (LASSO) regression were applied for feature selection and model construction. Multifactor logistic regression was used to build a clinical-imaging model based on clinical factors and a clinical-radiomics model. The area under the curve (AUC) and calibration curves were used to assess model performance, and the clinical value of the model was analyzed using decision curve analysis. The relationship between the actual and predicted short-term efficacy of the combined model and progression-free survival (PFS) was evaluated using Kaplan–Meier survival curves and log-rank tests.

**Results** A total of 34 radiomics features were selected by LASSO, and the clinical-radiomics model had the best predictive performance (AUC = 0.902 and AUC = 0.845 for the training and testing sets, respectively), and the model based on AP had the best predictive performance among the four radiomics models (AUC = 0.89 for the training set and AUC = 0.85 for the testing set); the multifactorial logistic regression results showed that microsphere type ( $p = 0.042$ ) and AP Rad-score ( $p = 0.01$ ) were associated with short-term efficacy. In addition, a difference in PFS was observed in patients with HCC with different short-term efficacies predicted by the combined model. Moreover, prognosis was better in the objective versus non-objective response group.

**Conclusions** The combined clinical-radiomics model is an effective predictor of the short-term efficacy of initial D-TACE in patients with HCC, contributing to clinical and economic benefits for patients.

**Keywords** Hepatocellular carcinoma · Drug-eluting beads transarterial chemoembolization · MRI · Radiomics · Short-term efficacy

## Introduction

With mortality rates as high as 8.3%, liver cancer is the fifth most common malignant tumor and third leading cause of tumor-related deaths worldwide, posing a serious threat to public health [1]. Surgical resection and liver transplantation are the primary curative treatment options for hepatocellular

carcinoma (HCC). However, most patients are diagnosed in the intermediate to advanced stages, and only approximately 5%–15% of patients with HCC have the opportunity to undergo radical surgery [2]. According to the Barcelona Clinic Liver Cancer (BCLC) staging criteria [3], transarterial chemoembolization (TACE) is the first-line treatment option for patients with intermediate to advanced HCC. In addition, drug-eluting beads transarterial chemoembolization (D-TACE) is an embolization therapy involving drug-eluting microspheres containing chemotherapeutic agents, known as “drug-eluting microspheres” owing to their long duration of

Zihan Xi, Yuanxin Ye, and Yongbo Yang have contributed equally to this work and should be considered co-first authors.

Extended author information available on the last page of the article

action and low systemic toxicity and side effects. Moreover, D-TACE has been widely used in clinical practice.

Related studies have reported that D-TACE has greater advantages than conventional transarterial chemoembolization (c-TACE) in some HCC groups [4–6], such as patients with cirrhosis [6]. Moreover, D-TACE causes less postoperative pain compared with c-TACE [7]. However, as the cost of D-TACE is higher than that of c-TACE, noninvasive methods of efficacy prediction can reduce the financial burden on patients to some extent. In addition, some studies have shown that patients who achieve complete response (CR) after initial TACE treatment have a favorable long-term prognosis [8–10]. Therefore, effective monitoring of the short-term efficacy of TACE (especially after the initial treatment) is partly beneficial for physicians with regard to early adjustments to the therapeutic regimen, leading to longer survival times for these patients. Among imaging examinations, Magnetic Resonance Imaging (MRI) has high research value owing to its advantages in multi-parameter, multi-directional, and multi-functional imaging. In previous studies, some imaging signs were associated with the biological behavior of HCC, such as the maximum tumor diameter or degree of enhancement in the arterial phase [9]. Despite attempts by radiologists to standardize the evaluation of macroscopic tumor biological behaviors, the use of such qualitative imaging features in tumor characterization or efficacy assessment remains subjectively variable, even among senior physicians [11]. Recently, MRI-based radiomics showed improved predictive performance [12–14]. Therefore, our study was conducted to predict the short-term efficacy of initial D-TACE treatment in patients with HCC by constructing a preoperative multiparametric MRI radiomics model and further explore the relationship between the short-term efficacy of initial D-TACE treatment and long-term prognosis (progression-free survival).

## Materials and methods

### Research population

This retrospective analysis was approved by the Ethical Review Board of our hospital (IRB-2021–414), and the requirement for informed consent was waived. Previous studies on D-TACE prognosis-related radiomics included approximately 50–100 patients [15]; Therefore, a total of 113 patients with HCC confirmed by pathology or imaging examinations who received their initial D-TACE treatment between October 2018 and May 2022 at our hospital were included after the selection process.

The exclusion criteria were as follows: (1) preoperative history of targeted immunotherapy, (2) postoperative loss to follow-up, (3) lack of preoperative enhanced MRI or poor

image quality, and (4) presence of diffuse HCC. The patient recruitment process is illustrated in Fig. 1. Finally, 113 patients were randomized into training ( $n = 79$ ) and testing ( $n = 34$ ) groups.

### Evaluation of efficacy

The short-term efficacy and long-term prognosis (PFS) of patients after the initial D-TACE treatment were evaluated according to the modified Response Evaluation Criteria in Solid Tumors (mRECIST). Patients were classified into two groups according to the treatment response to the initial D-TACE treatment. The objective response (OR) group included patients who achieved CR or partial response (PR), whereas the non-objective response (nR) group included patients with progressive disease (PD) or stable disease (SD). Efficacy evaluations were performed by three radiologists: two radiologists with more than 8 years of experience who completed the evaluations independently and another with more than 15 years of experience who was responsible for the efficacy evaluation in the occurrence of conflicting opinions. Figure 2 shows two examples of efficacy evaluations (OR vs. NR group).

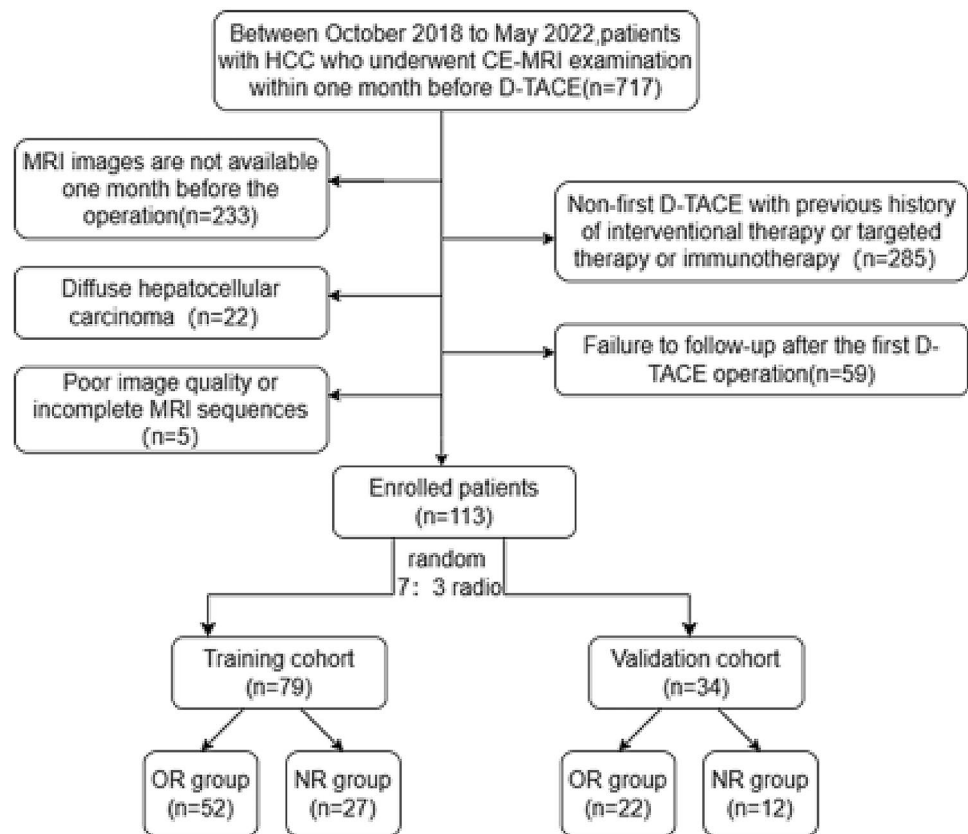
### MRI protocol

All patients underwent CE-MRI of the upper abdomen within 1 month before operation and MRI of the abdomen within 4–6 weeks after operation. Before the scan, they underwent uniform breathing and breath-holding training. CE-MRI was performed in our department using two devices. DWI was performed using a cross-sectional single-shot plane echo imaging sequence with b-values of 50 and 800 s/mm, respectively. Detailed scanning parameters are described in Supplementary data, Table S1.

### D-TACE procedure

The embolizing agents included CalliSpheres drug-loaded microspheres (Jiangsu Suzhou CalliSpheres Biomedical Technology Co., Ltd., China) or DC Beads drug-loaded microspheres (Boston Scientific International Medical Trading Co., Ltd., USA), and the diameter of one microsphere was specified as 70–150, 100–300, or 300–500  $\mu\text{m}$ , which was selected according to the tumor size and blood supply. After sedimentation and discharge of the supernatant, an appropriate amount of the chemotherapeutic drug was loaded and diluted with a contrast agent (volume ratio, 1:1). Under fluoroscopy, the embolic agent was slowly injected into the blood-supplying artery of the tumor through the microcatheter. The endpoint of embolization was complete stagnation of blood flow in the tumor trophoblastic vessels or contrast agent retention.

**Fig. 1** Flowchart of study enrollment. *OR* Objective response, *NR* non-Objective response



### Collection of clinical-imaging features and follow-up

Clinical-imaging characteristics were collected by two radiologists with more than 8 years of experience. The collected clinical information comprised three aspects: patients' baseline data, including demographic information and laboratory indicators; surgical information; and imaging features. The enrolled patients were followed up to assess PFS, mainly by clinical and imaging evaluations at the time of patient review every 1–3 months. PFS was calculated from the time of the initial D-TACE procedure to the time of progression, or if no disease progression was observed at the time of the last follow-up.

### Radiomics feature extraction and analysis

CE-MRI (AP, DWI, and T2WI) results were output in Digital Imaging and Communications in Medicine (DICOM) format. Then We use publicly sourced software (3D Slicer version 5.1.0, <https://www.slicer.org/>) for semi-automatic outlining. Two radiologists from the abdominal subspecialty group outlined the regions of interest (ROIs). The three dimensional ROI included a pseudo-envelope at the outer edge, while avoiding extra-tumoral vessels as much as possible. Intra-observer agreement was assessed by calculating the intra-group correlation coefficient (ICC). To calculate

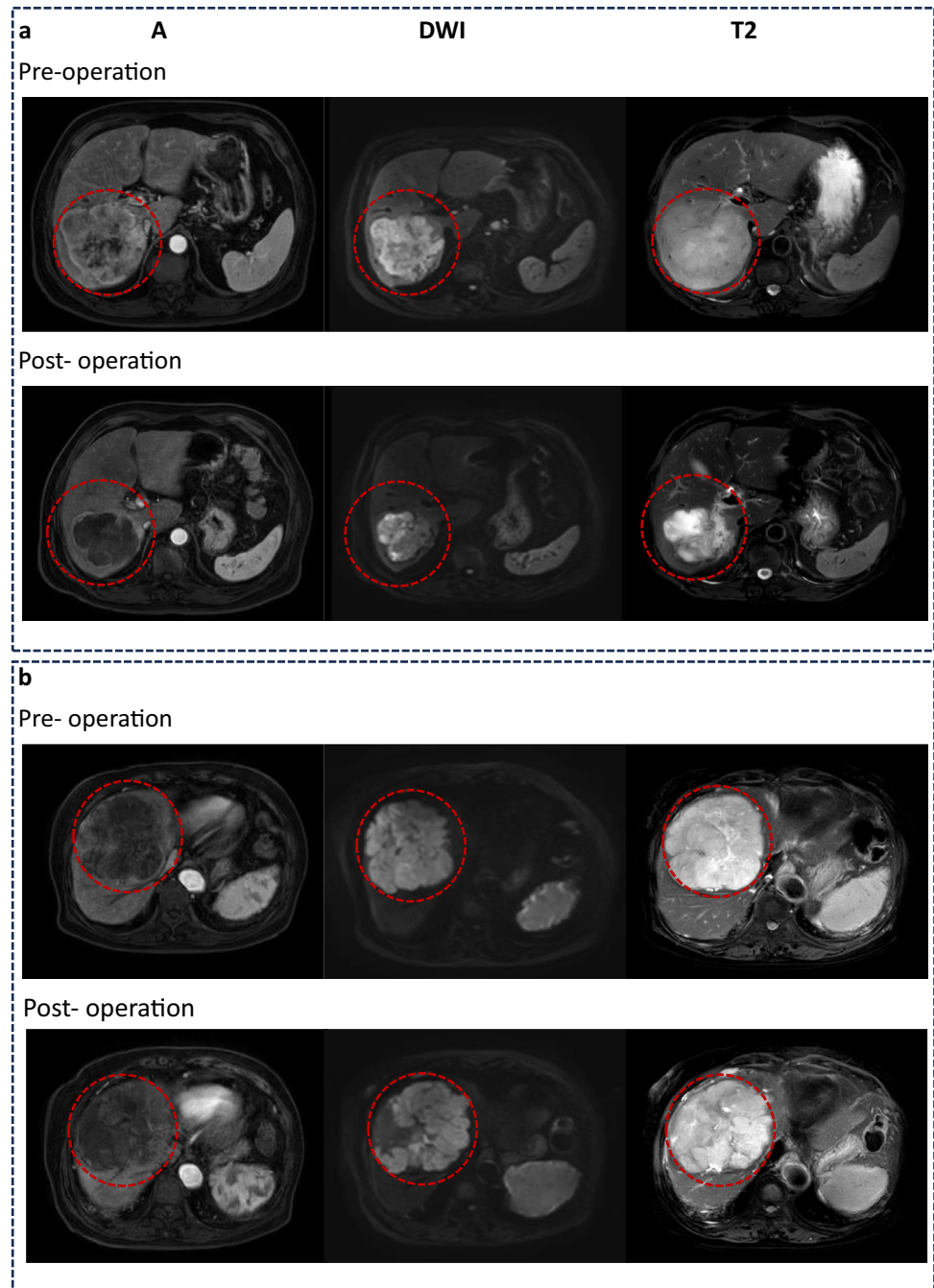
the intra-observer ICC, MRI results from three sequences were randomly selected for 20 patients and outlined twice at 2-month intervals by an outliner. The intra-observer ICC was calculated after extracting the features using ROI segmentation. The reader was considered to have better consistency compared with the other readers when the ICC was > 0.75; therefore, the remaining target areas were completed by this reader.

Image preprocessing and feature extraction were performed using pyradiomics (version 3.1.0). We use two preprocessing methods in our study. The first is a linear interpolation algorithm to resample the images to a voxel size of  $1 \times 1 \times 1$  mm, and the second is a z-score method to normalize the radiomics features. [16, 17]. Then, 1781 radiomics features were extracted from each sequence (detailed data are available in the Supplementary Data, Table S2). A total of 5343 features were extracted from the three sequences. Radiomics feature values were normalized using the z-score method. The workflow of the radiomics analysis is shown in Fig. 3.

### Feature selection and radiomics model construction

We used two methods for feature selection and dimensionality reduction for radiomics. First, minimum redundancy maximum correlation (mRMR) was used to eliminate

**Fig. 2** Representative pre- and post-treatment CE-MRI images of for the OR and NR groups of patients with HCC conforming to the mRECIST. **a** A 70-year-old male HCC patient with a maximum tumor diameter of 10.2 cm underwent CE-MRI 2 days before D-TACE and CE-MRI review 35 days after treatment. AP, DWI, and T2WI images were obtained, and the results indicated CR according to the mRECIST, with no clear enhanced surviving lesions. **b** An 80-year-old female HCC patient with a maximum tumor diameter of 11.7 cm underwent CE-MRI 1 day before D-TACE and CE-MRI review 37 days after treatment. The results indicated SD; the maximum diameter of the surviving tumor lesion was approximately 10.6 cm, and the necrosis rate was approximately 9%. *CE-MRI contrast-enhanced magnetic resonance imaging; OR objective response; NR non-objective response; AP arterial phase; DWI diffusion-weighted imaging; T2WI T2-weighted imaging*

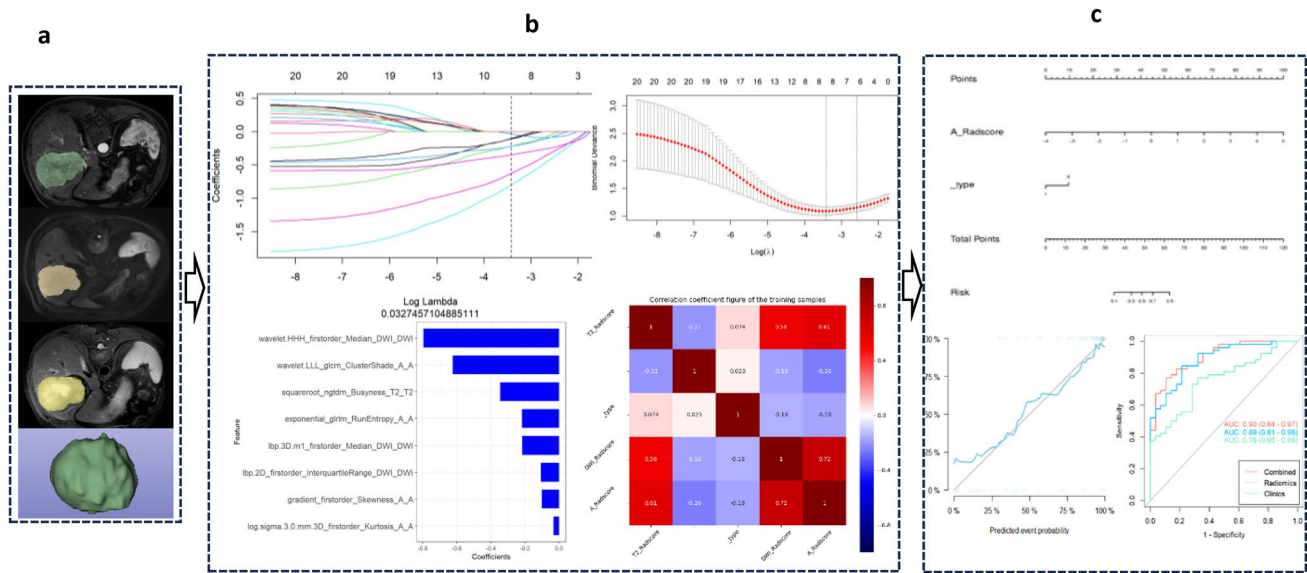


redundant and irrelevant features. Then LASSO was conducted to choose the optimized subset of features to construct the final model. The LASSO includes choosing the regular parameter  $\lambda$ , determining the number of the feature. The basic principle of LASSO is to introduce the L1 regularization term based on ordinary least squares to achieve feature selection and coefficient sparsification of the model by minimizing the objective function. The optimized features were selected to construct the final model, and the best radiomic features were chosen. Subsequently, a radiomics model was constructed based on the features selected for

each sequence using multifactor logistic regression analysis. Rad-scores were calculated using a linear combination of selected radiomic features weighted according to their respective coefficients.

### Clinical-radiomics modeling and evaluation

We further introduced clinical factors that were highly correlated with D-TACE efficacy into the prediction model to improve the prediction performance based on Rad-scores. First, clinical variables ( $p$  values  $< 0.1$  in the univariate



**Fig. 3** Analytical workflow of radiomics in our study. **a** Tumors are indicated by semi-automatic drawings of lines around the entire tumor contour in AP, DWI, and T2WI sequences, and three dimensional segmentations were created. **b** We used two methods of feature selection, mRMR and LASSO to select the features. Correlation coefficients were also calculated. **c** Radiomics models were built

using logistic regression algorithms, and a nomogram was established, evaluated by calibration curve and decision curve analysis. *AP arterial phase; DWI diffusion-weighted imaging; T2WI T2-weighted imaging; mRMR minimum redundancy maximum correlation; LASSO least absolute shrinkage and selection operator*

analysis and the selected radiomics features were subjected to multifactorial logistic regression analysis to identify the independent risk factors affecting short-term efficacy ( $p$  value  $< 0.05$ ). These independent risk factors were then analyzed using a multifactorial logistic regression algorithm to establish a clinical-radiomics model. Finally, a nomogram of clinical-radiomics was created. The area under the curve (AUC) was plotted to assess the predictive performance of the model. DeLong's test was applied to compare the differences in AUCs between the models. In addition, we used calibration and decision curves to assess the calibration and clinical application of the model. Finally, the AUCs were calculated, calibration curves were plotted using the formulas constructed in the training set, and the model was validated in the testing set.

**Statistical analysis**

The two primary endpoints of interest in this study were the actual short-term efficacy of the initial D-TACE treatment and PFS. The former clinical-radiomics model was developed and assessed using R software (version 4.1.3, R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics and assessment of PFS for the different short-term efficacies (actual and predicted) were performed using SPSS software (version 26.0, IBM Corp., Armonk, NY, USA). Normally distributed continuous variables are denoted by the mean  $\pm$  standard deviation (Mean  $\pm$  SD), and

those that are not normally distributed are denoted by the median ( $P_{25}$ ,  $P_{75}$ ). Qualitative data were represented as numbers ( $n$ ) and percentages (%). The t-test or Mann–Whitney  $U$  test was used to compare quantitative data. Categorical variables were analyzed using chi-square or Fisher's exact tests. A rank-sum test was used to analyze the rank data. Risk factors with  $p$  values  $< 0.1$  in the univariate logistic regression analysis were included in the multivariate logistic regression model. Survival curves were constructed using the Kaplan–Meier method, and group comparisons were made using log-rank tests. The performance of our models was assessed by the AUCs and its 95% confidence interval (CI). We compared the AUCs of different models using Delong's test.  $p$  values  $< 0.05$  were considered statistically significant.

**Results**

**Clinical-imaging characteristics and follow-up results**

Table 1 summarizes the clinical-imaging characteristics of the 113 patients. A total of 74 patients were categorized into the OR group and 34 patients into the NR group. The results of multivariate logistic regression showed that microsphere type was a valid predictor ( $p = 0.042$ , OR = 0.224, 95% CI 0.053–0.948).

**Table 1** Baseline clinical characteristics of patients with HCC in the training and testing datasets

Variables	Training set ( <i>n</i> = 74)	Testing set ( <i>n</i> = 39)	<i>p</i> value
Gender ( <i>n</i> ,%)			
Male	64(81.0)	29(85.3)	0.584
Age Mean ± SD	58 ± 13	59 ± 12	0.776
History of disease ( <i>n</i> ,%)			
Hepatitis B	61(77.2)	27(79.4)	0.796
Diabetes	10(12.7)	3(8.8)	0.791
Preoperative antiviral treatment ( <i>n</i> ,%)	17(21.5)	10(29.4)	0.367
Liver function index M(P <sub>25</sub> , P <sub>75</sub> )			
ALP(U/L)	135(103,197)	138(108,184)	0.963
GGT(U/L)	149(84,255)	145(82,234)	0.853
AST(U/L)	53(36,83)	51(38,83)	0.950
ALT(U/L)	36(24,72)	43(29,56)	0.391
Prealbumin(U/L) Mean ± SD	169 ± 75	183 ± 76	0.355
AFP (ng/ml) (P <sub>25</sub> , P <sub>75</sub> )	279(12,6335)	1582(99,27,469)	0.045
Child–pugh ( <i>n</i> ,%)			0.780
A	74(93.7)	33(97.1)	
B	5(6.3)	1(2.9)	
BCLC ( <i>n</i> ,%)			0.128
A	16(20.3)	2(5.9)	
B	19(24.1)	9(26.5)	
C	44(55.7)	23(67.6)	
Drug-loaded drugs ( <i>n</i> ,%)			0.605
Epirubicin	51(64.6)	21(61.8)	
Idarubicin	15(19.0)	9(26.5)	
Irinotecan	13(16.4)	4(11.7)	
Drug loading capacity(mg) (P <sub>25</sub> , P <sub>75</sub> )	50(30,60)	50(10,60)	0.436
Microsphere Type ( <i>n</i> ,%)			0.608
CalliSpheres	39(49.4)	15(44.1)	
DCB	40(50.6)	19(55.9)	
Microsphere size(μm) ( <i>n</i> ,%)			0.416
70–150	2(2.5)	0(0.0)	
100–300	24(30.4)	7(20.6)	
300–500	53(67.1)	27(79.4)	
Maximum tumor diameter (cm) Mean ± SD	8.7 ± 3.7	9.0 ± 3.8	0.717
Tumor/liver volume ratio Mean ± SD	0.207 ± 0.166	0.265 ± 0.193	0.106
Clear tumor margins ( <i>n</i> ,%)	57(72.2)	22(64.7)	0.429
Degree of early tumor enhancement ( <i>n</i> ,%)			0.646
Mild inhomogeneity	13(16.5)	5(14.7)	
Obviously inhomogeneity	57(72.2)	24(70.6)	
Obviously homogeneity	9(11.4)	5(14.7)	
Significant edge enhancement ( <i>n</i> ,%)	34(43)	21(61.8)	0.068
Intra-tumoral lipid ( <i>n</i> ,%)	9(11.4)	0(0.0)	0.094
Tumor with capsule ( <i>n</i> ,%)	63(79.9)	30(88.2)	0.278
Intra-arterial tumor ( <i>n</i> ,%)	63(79.7)	27(79.4)	0.968
With portal thrombosis ( <i>n</i> ,%)	27(34.2)	20(58.8)	0.015
Satellite focal ( <i>n</i> ,%)	43(54.4)	18(52.9)	0.884

ALP alkaline phosphatase, GGT  $\gamma$ -glutamyl transpeptidase, AST aspartate aminotransferase, ALT alanine transaminase, AFP  $\alpha$ -fetoprotein, BCLC Barcelona clinic liver cancer

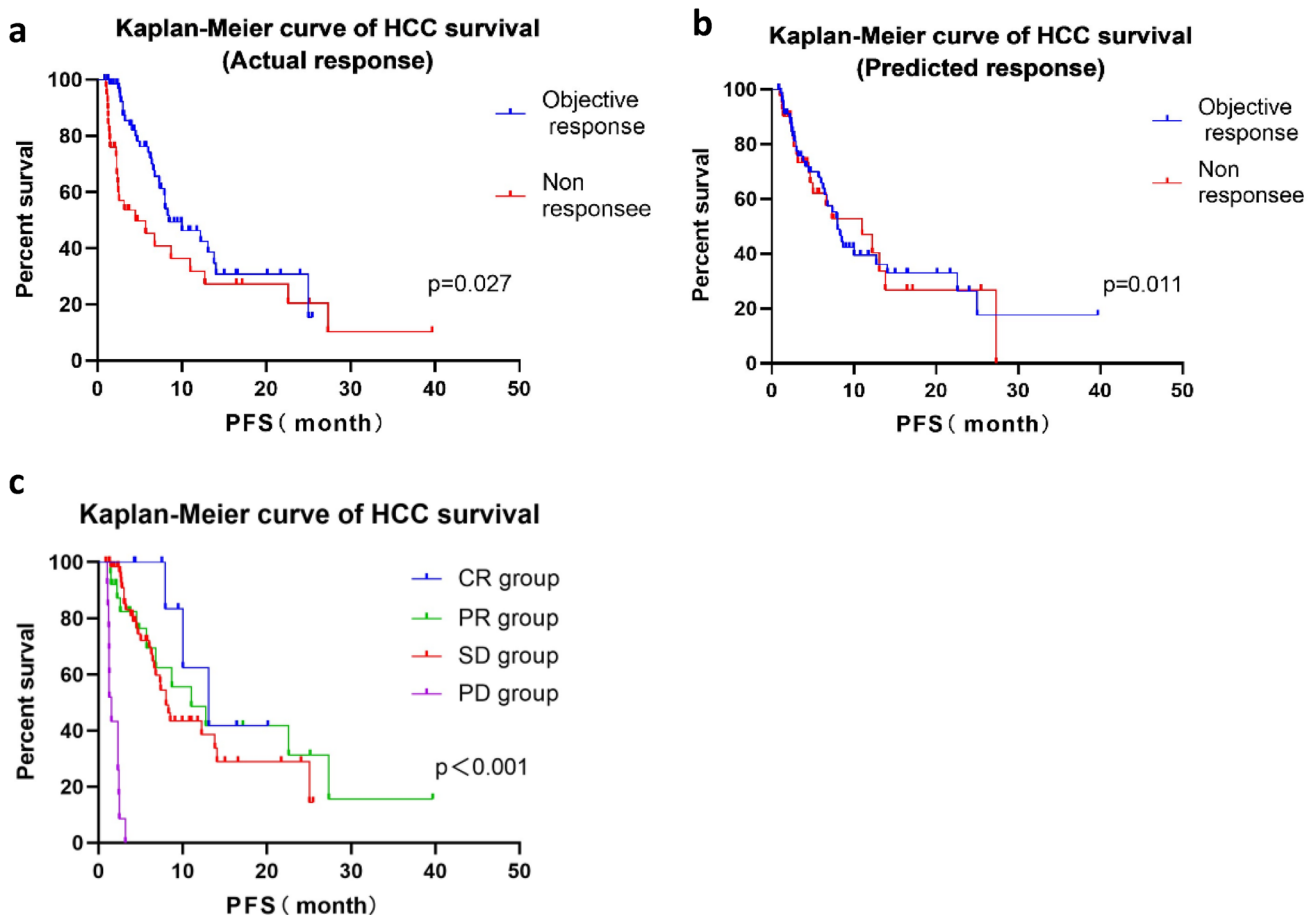
\*Indicates the *p* value < 0.05

More than half of the patients in the training (71%) and testing (53%) sets had OR in the efficacy evaluation of the initial D-TACE treatment, with a median time from initial the D-TACE procedure to final follow-up of approximately 9.5 (95% CI 6.6–12.4) months. The Kaplan–Meier curves of the actual and predicted short-term efficacies and PFS times for the two groups are shown in Fig. 4a, b. The median survival times in the actual OR and NR groups were 8.5 (95% CI 4.8–12.2) and 4.5 (95% CI 0.0–9.0) months, respectively, and the difference was statistically significant ( $p = 0.027$ ). A statistically significant difference was also observed in median survival times between the groups of patients that the best model predicted in the OR vs. NR analysis ( $p = 0.011$ ). Patients were divided into four groups (as shown in Fig. 4c) according to the Kaplan–Meier curve for efficacy based on the mRECIST. The median PFS times were 13.1 (95% CI 6.6–19.6), 8.0 (95% CI 6.7–9.3), 11 (95% CI 3.8–18.2), and 1.5 (95% CI 1.1–1.9) months in the CR, PR, SD, and PD groups, respectively, with a statistically significant difference ( $p < 0.001$ ). This finding indicates that OR versus NR

status after the initial D-TACE treatment was associated with a better long-term prognosis, with the longest PFS observed in the CR group.

### Feature selection and radiomics modeling processes

After ICC analysis, no statistically significant differences were observed in the two outlining results for observer A ( $p = 0.85$ ); therefore, the final outlining assignment was completed by observer A. In total, 5343 features (1781 features per sequence) were extracted from the AP, DWI, and T2WI sequences. Redundant features were eliminated sequentially using the mRMR method, with 20 features retained per sequence. Subsequently, 10 AP, 7 DWI, and 9 T2WI, combined with 8 triple-sequence, radiomics features were screened by LASSO and five-fold cross-validation, for a total of 34 radiomics features (Detailed information is available in Supplementary Data, Fig. S1). The Rad-score was calculated for each patient, and the formula is described in the Supplementary Data, Table S3.



**Fig. 4** Kaplan–Meier curve for HCC survival. **a** Actual response status and **b** predicted response status by the combined model of two groups. **c** Actual response status of four groups

## Clinical-radiomics model construction and evaluation

Univariate analysis showed that three clinical (microsphere size, microsphere type, and drug load type) and two imaging (maximum tumor diameter and tumor-to-liver volume ratio) characteristics were statistically significantly different (Supplementary Data, Table S4). Multifactorial analysis revealed that microsphere type (odds ratio = 0.224, CI = 0.053–0.948,  $p = 0.04$ ) was an independent risk factor for predicting efficacy. Therefore, we used microsphere types to construct the clinical-radiomics models.

We built a total of six predictive models: four radiomics models, a clinical-imaging model, and a clinical-radiomics model. Among these radiomics models, the best performance was observed in the AP phase, with AUCs of 0.89 (95% CI 0.81–0.96) and 0.85 (95% CI 0.69–1.00) in the training and testing sets, respectively (Table 2). The ROC curves for the radiomics models are plotted in Fig. S2 (Supplementary Data). The model combining clinical factors (microsphere type) and radiomics models (AP) performed the best among all models, with AUCs of 0.902 (95% CI 0.832–0.959) and 0.845 (95% CI 0.708–0.944) in the training and testing sets, respectively. Moreover, the addition of the Rad-score to the clinical-imaging model significantly improved the predictive efficiency of the training set (AUC, 0.717–0.902;  $p = 0.004$ ), suggesting that the combined model maximizes the predictive value of efficacy. In the DeLong test (Supplementary Data, Table S5) for both the clinical-imaging and combined models comprised each of the four radiomics models, the combined model had the best performance (AP,  $p = 0.002$ ; DWI,  $p < 0.001$ ; T2WI,  $p = 0.005$ ; triple-sequence,  $p = 0.002$ ). However, the DeLong test for the radiomics and combined models did not show a statistically significant difference between the two groups in terms of predictive performance (AP,  $p = 0.32$ ; DWI,  $p = 0.84$ ; T2WI,  $p = 0.25$ ; triple-sequence,  $p = 0.58$ ). Furthermore, ROC curves showed that radiomics curves based on a single sequence or three sequences frequently overlapped with those of the combined model, with much

greater clinical efficacy than that of a single clinical-imaging model (Fig. 5).

## Establishment and evaluation of nomogram

The nomogram was constructed based on the aforementioned predictive model, which effectively combined Rad-Scores and the selected clinical predictors (Fig. 6a). The calibration curves showed good calibration in the training set, which was slightly worse in the testing set (Fig. 6b, c). The Hosmer-Lemesow test showed no statistically significant difference in either set ( $p = 0.235 < 0.05$ ), indicating good agreement between the nomogram predictions and actual observations.

## Clinical-radiomics model and nomogram for clinical application

Decision curves illustrate how well the radiomics model, the clinical-imaging model, and the combined model of clinical utility perform. Moreover, the addition of the radiomics model resulted in significantly higher clinical efficacy in the combined model, with the combined model achieving the best clinical efficacy among all models.

## Discussion

TACE plays an irreplaceable role in transformation therapy and is a nonsurgical alternative for patients with moderate to advanced HCC. However, owing to the high heterogeneity of HCC = c, the clinical response to TACE varies widely among patients with HCC. In addition, D-TACE is more expensive than c-TACE. Therefore, if the short-term efficacy of D-TACE can be predicted preoperatively, the treatment plan can be adjusted according to the patient's specific condition. This can simultaneously improve the prognosis of the patients, reduce the economic burden, and avoid the wastage of medical resources.

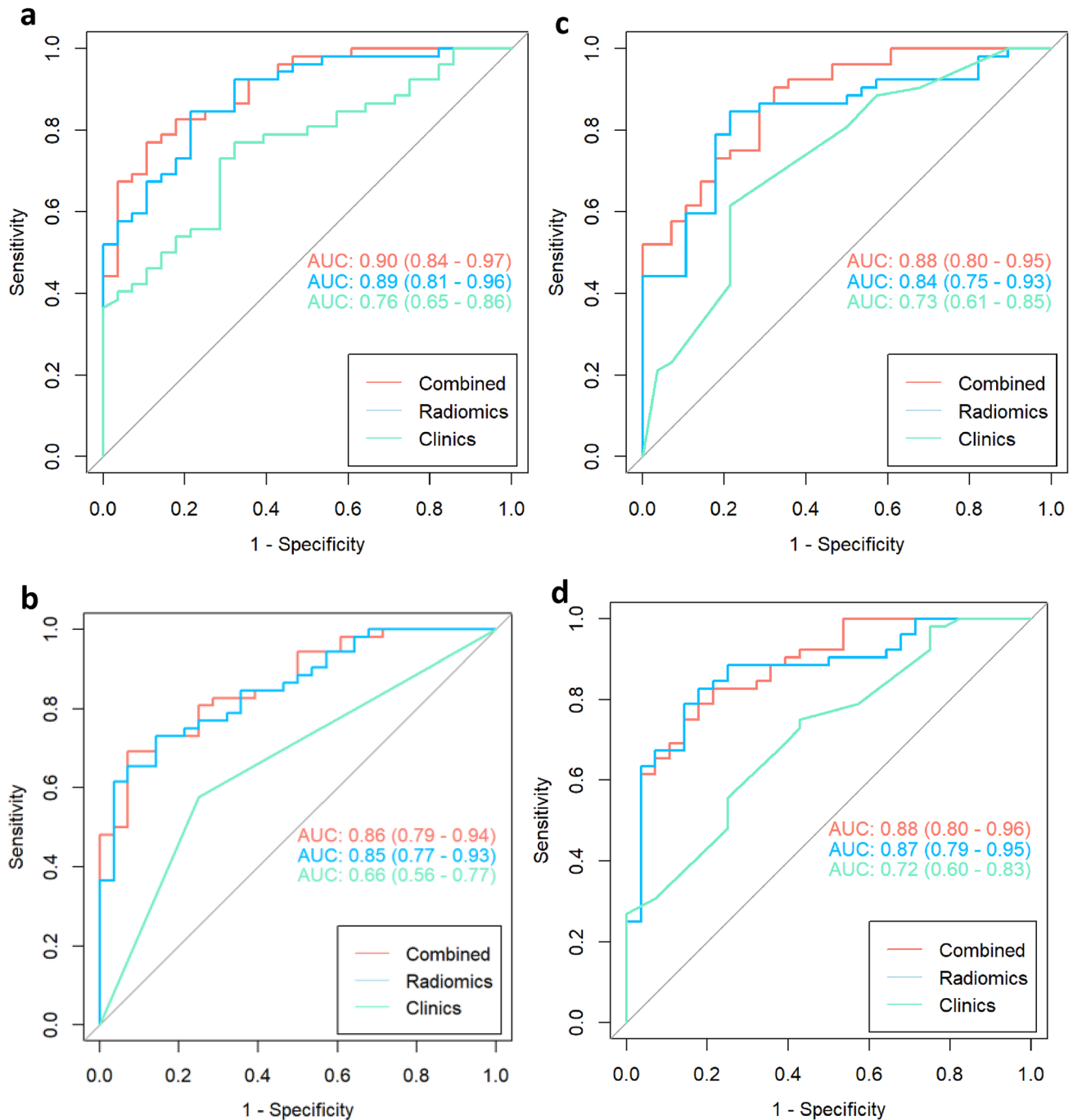
**Table 2** The performance of different models in the training and test sets for treatment response prediction

Models	TR				non-TR			
	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE
AP	0.89	0.83	0.85	0.78	0.85	0.82	0.86	0.72
DWI	0.85	0.78	0.73	0.85	0.83	0.76	0.77	0.73
T2WI	0.84	0.83	0.85	0.79	0.81	0.82	0.91	0.64
Mp-MRI	0.87	0.83	0.83	0.82	0.84	0.70	0.59	0.91
Clinic	0.72	0.63	0.56	0.78	0.64	0.62	0.68	0.50
Combined	0.90	0.81	0.75	0.93	0.85	0.79	0.82	0.75

AUC area under the curve, ACC accuracy, SEN sensitivity, SPE specificity

\*Indicates the  $p$  value  $< 0.05$

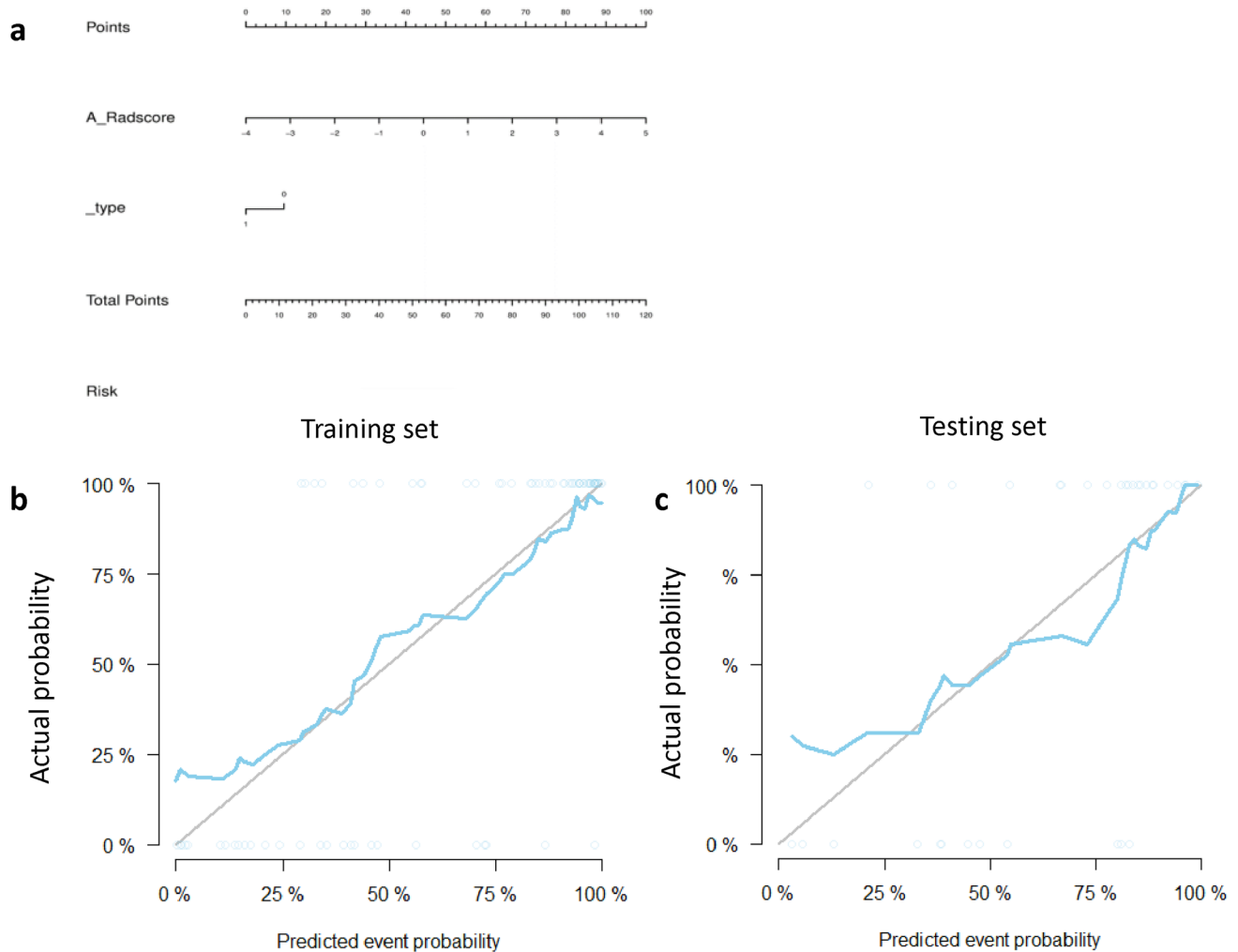




**Fig. 5** ROC curves of the combined model based on different sequences in the training set. **a** AP. **b** DWI. **c** T2WI. **d** Triple-sequence. *AP* arterial phase; *DWI* diffusion-weighted imaging; *T2WI* T2-weighted imaging

D-TACE is an interventional approach that differs from c-TACE. To the best of our knowledge, most studies have applied radiomics to predict the efficacy of c-TACE [13, 19, 20] or a combination of the two TACE approaches [12, 21] rather than D-TACE, resulting in fewer studies applying radiomics to predict D-TACE efficacy alone and even fewer studies based on CE-MRI sequences. Tipaldi et al.

[15] predicted the short-term efficacy of D-TACE based on preoperative CT texture features in 50 patients with HCC, with an optimal texture feature AUC of 0.733. Ultimately, the combined clinical-radiomics model performed the best in our study, which is in agreement with the results of a previous study [19]. However, its predictive performance was lower than that in our study, possibly owing to the following



**Fig. 6** Nomogram and calibration curves of proposed model. **a** Nomogram of combined model integrating Rad-score and significant clinical variables. **b** Calibration of the combined model in the training set. **c** Calibration of the combined model in the testing set

reasons. First, the features extracted in our study included first-, second-, and higher-order features, and were therefore more varied and comprehensive. Second, the soft-tissue contrast in the MRI sequences was better than that in the CT images, which showed the details of the lesions in a subtle manner. A previous study demonstrated that MRI outperformed CT for predicting the presence of microvascular invasion in patients with HCC, with a maximal diameter of 2–5 cm [22]; however, radiomics studies comparing the advantages of CT and MRI for predicting the efficacy of D-TACE are lacking.

We used two different MRI scanners with the same field strength (3.0 T) in our study, and the radiomics results were unaffected. First, the MRI data were subjected to image normalization, gray level standardization, and image alignment before modeling. Second, we used two different MRI scanners of the same field strength (3.0 T) controlled the differences between parameters to some extent. This phenomenon

also better approximates a clinical practice setting, as it is impossible to guarantee that all patients will be imaged using the same scanning equipment.

In our study, a model based on three preoperative CE-MRI sequences and clinical-imaging features was developed and validated to predict the short-term efficacy of D-TACE. The model successfully combined the radiomics Rad-scores with clinical-imaging features to predict short-term efficacy, with patients categorized into OR and NR groups. Our study cohort consisted of 113 patients with HCC, and 79 (70%) patients actually achieved OR, suggesting that the majority of patients with HCC had favorable D-TACE efficacy. Most of the study participants were men with hepatitis B, BCLC stage B or C, similar to previous studies [23] and in line with the epidemiological characteristics of HCC [24], clinical practice, and therapeutic strategies for BCLC stage [3].

Our study included two endpoints. First, to evaluate the short-term efficacy of the initial D-TACE treatment in

patients with HCC, four radiomics models were established using three CE-MRI sequences, which were also combined with the relevant clinical-imaging factors to predict the short-term efficacy of the initial D-TACE treatment in patients with HCC (mainly in the intermediate to advanced stages). Finally, a total of six models were constructed. Among these models, clinical-imaging features combined with a model based on AP radiomics demonstrated the best predictive efficacy. This may be because the AP-based radiomics model performed the best among the radiomics models, showing better accuracy, sensitivity, and specificity. The addition of other sequences in the combined model did not significantly improve the predictive performance ( $p = 0.435$ ). We hypothesized that the AP is more indicative of lesion heterogeneity. Obviously, tumor heterogeneity can be generated by variations in cell density, neovascularization, and the extracellular matrix [25]. Thus, the AP was the best performer among the four radiomics models in our study. Notably, the best AUC value was not obtained for the radiomics model consisting of the three sequences together, with AUC values of 0.87 and 0.84 for the training and testing sets, respectively. This is similar to the findings of Liu et al. [14], with the highest predictive efficacy reported for the augmented delayed-phase radiomics model (AUC = 0.907), which was superior to that of the T2WI, AP, portal phase, and delayed-phase multi-parameter radiomics models consisting of four sequences (AUC = 0.875). In the present study, this may be related to multiparametric feature redundancy and different image resolutions (layer thickness in the AP was different from that in T2WI and DWI in the present study). The radiomics features in our study for the predictive model included seven texture-based features and three intensity-based features in the first-order features, and shape-based features were not incorporated (The features and their definitions in the model are shown in the Supplementary Data S2 and Fig S1). Kurtosis, 10 Percentile, and Maximum are intensity-based features that are capable to study the histogram properties of tumor intensity. Grey Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Gray Level Dependence Matrix (GLDM) are all texture-based features that have the advantage of complementing the spatial relationships of neighboring pixels [26]. And these are not available with shape or intensity-based features. Moreover, texture-based features can reflect the heterogeneity of tumors [26, 27]. This is a valuable reminder for determining the differences in the efficacy of D-TACE between individuals with different HCCs.

These results suggest to radiologists that the homogeneity of tumor intensity and the distribution of enhancement in the tumor in patients with HCC have the potential to help clinicians predict, to a certain extent, the efficacy of D-TACE. Simultaneously, radiomics also demonstrates the advantage

of more subtle biological features of tumors that cannot be distinguished by the radiologist's naked eye.

In addition, shape-based features did not appear to reflect excellent predictive efficacy in our study, which is consistent with the absence of morphological information such as tumor volume in the clinical-imaging portion of the model. This might be explained by the fact that our study population was dominated by patients with massive HCC, with little variability in morphology between individual tumors. It thereby leads to no significant correlation between morphological type of information and D-TACE efficacy.

Among the clinical-imaging characteristics, only microsphere type was used to construct the combined model. The effects of microsphere size and drug-carrying concentration on D-TACE efficacy have been reported in clinical studies [28–30], whereas microsphere and drug-carrying types have been reported less frequently. Regarding drug-carrying types, drugs have toxic strengths. Idarubicin, for example, was the most cytotoxic of the 11 anticancer drugs tested against the three HCC cell lines [31]. In addition, the release rates vary among different drugs, which affects the degree of drug penetration into the tumor tissue and its distribution in the liver tissue [32]. Few studies have explored the effects of microsphere types on D-TACE efficacy. From the perspective of microsphere construction, different chemical structures of microspheres have different drug-carrying capacities and drug release rates [33]. In addition, the mechanical properties of the microspheres (such as compressibility and deformability) may affect their distribution density in the arteries, embolization effect, and suitability for catheterization [34]. Finally, microsphere type proved to be a clinical factor associated with the short-term efficacy of D-TACE in our study, whereas the drug-carrying type and microsphere size were excluded from the multifactorial logistic regression. We hypothesized that this may be due to the wide variety of microsphere sizes and drug-carrying types used in our relatively small cohort, which resulted in unbalanced data.

However, our clinical-imaging model showed low predictive performance in the training and testing sets, possibly because the clinical factors and macroscopic imaging features cannot reflect tumor heterogeneity better than the combined clinical-radiomics model. However, when the Rad-score was added, it outperformed the single clinical-imaging model. We hypothesized that this phenomenon is related to the sample size [35]. In our research, AP-based radiomics combined with microsphere types form a clinical-radiomics model with the best predictive performance. From one side, as mentioned above, radiomics can provide high-dimensional biological information about tumor heterogeneity, which is intrinsic information that determines the variable individual efficacy of D-TACE. As for the other side, microsphere types are selectable during D-TACE treatment, which is extrinsic information that determines the efficacy

of patients with HCC. Thus, the combination of the both can provide the best predictive effect.

The secondary endpoint was PFS after initial D-TACE treatment in patients with HCC. Kaplan–Meier and log-rank tests showed that patients whose actual short-term efficacy after initial D-TACE was classified as OR had a better long-term prognosis compared with patients classified as NR ( $p = 0.03$ ). The group of patients attributed to the best prediction model with an outcome of OR also had a better long-term prognosis than that in patients with predicted NR ( $p = 0.01$ ). This validates the effect of the model established in this study and provides supporting evidence for the clinical significance of predicting the short-term efficacy of initial D-TACE treatment. Moreover, the CR group had the best long-term prognosis ( $p < 0.001$ ). The time to progression differed significantly between the CR and PD groups, with the longest PFS after the initial D-TACE observed in the CR group. The Kaplan–Meier curves indicated that the survival curves for the PR and SD groups overlapped, and the time to progression was longer in the SD versus PR group with the accumulation of follow-up time, which may have been caused by the unbalanced distribution of the number of cases in each group and the highly censored values used in this study. Therefore, this should be explored in a larger prospective study. Related studies have been conducted on this topic. For instance, Kim et al. investigated the relationship between patients who achieved CR after initial TACE treatment and patients who achieved CR after multiple TACE procedures and their respective overall survival (OS) rates among 314 patients with HCC, the results showed that an initial CR was the strongest predictor of a favorable prognosis [36]. However, both studies focused on analyzing OS after c-TACE; therefore, our study is innovative in that we evaluated the relationship between long-term prognosis and the short-term efficacy of D-TACE.

Our study has some limitations. First, this was a single-center, retrospective study with a small sample size. In addition, selection bias was unavoidable during the case-screening process. A larger sample size is required to validate our results. In addition, cases from multiple centers are required for external validation, which should be considered in future studies. Second, not all included patients had pathological results; therefore, a risk of bias existed. We also intend to include HCC cases with different pathological results for a more advanced stratification in future studies. Third, our second study endpoint was PFS, although the gold standard for long-term patient prognosis is OS, which we hope to include in subsequent studies.

Furthermore, with the results of our study, we identify directions that we can continue to explore in the future. We have only utilized one single imaging modality, but other imaging methods such as ultrasound, CT and even positron emission tomography (PET) can provide a more

comprehensive information about the biological behavior of the tumor. In addition, if genomics at the molecular level is combined with radiomics at the imaging level, this multimodal data may provide more comprehensive prognostic information. These deserve to be explored further in the future.

In conclusion, we showed that MRI-based radiomics features can predict the short-term efficacy of initial D-TACE treatment in patients with HCC and played a leading role in the combined clinical-radiomics model.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00261-024-04319-3>.

**Funding** This work was funded by Innovative Research Group Project of the National Natural Science Foundation of China (Grant no. 82072032); Medical Science and Technology Project of Zhejiang Province (Grant nos. 2023KY598, 2021KY567).

## References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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):209-49.
- Anwanwan D, Singh SK, Singh S, Saikam V, Singh R. Challenges in liver cancer and possible treatment approaches. *Biochim Biophys Acta Rev Cancer.* 2020;1873(1):188314.
- Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol.* 2022;76(3):681-93.
- Zhang Z-S, Li H-Z, Ma C, Xiao Y-D. Conventional versus drug-eluting beads chemoembolization for infiltrative hepatocellular carcinoma: a comparison of efficacy and safety. *BMC Cancer.* 2019;19(1):1162.
- Li H, Wu F, Duan M, Zhang G. Drug-eluting bead transarterial chemoembolization (TACE) vs conventional TACE in treating hepatocellular carcinoma patients with multiple conventional TACE treatments history: A comparison of efficacy and safety. *Medicine (Baltimore).* 2019;98(21):e15314.
- Wang T, Du Y-N, Sun J, Song H, Jiang Y, Liu F, et al. Drug-eluting bead transarterial chemoembolization could improve the hepatic hemodynamics of patients with unresectable hepatocellular carcinoma: a retrospective cohort study. *J Gastrointest Oncol.* 2023;14(1):302-11.
- Golfieri R, Giampalma E, Renzulli M, Cioni R, Bargellini I, Bartolozzi C, et al. Randomised controlled trial of doxorubicin-eluting beads vs conventional chemoembolisation for hepatocellular carcinoma. *Br J Cancer.* 2014;111(2):255-64.
- Peng C-W, Teng W, Lui K-W, Hung C-F, Jeng W-J, Huang C-H, et al. Complete response at first transarterial chemoembolization predicts favorable outcome in hepatocellular carcinoma. *Am J Cancer Res.* 2021;11(10):4956-65.
- Chen M, Cao J, Hu J, Topatana W, Li S, Juengpanich S, et al. Clinical-Radiomic Analysis for Pretreatment Prediction of Objective Response to First Transarterial Chemoembolization in Hepatocellular Carcinoma. *Liver Cancer.* 2021;10(1):38-51.
- Llovet JM, Real MI, Montaña X, Planas R, Coll S, Aponte J, et al. Arterial embolisation or chemoembolisation versus symptomatic

- treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet*. 2002;359(9319):1734-9.
11. Min JH, Lee MW, Park HS, Lee DH, Park HJ, Lim S, et al. Interobserver Variability and Diagnostic Performance of Gadoteric Acid-enhanced MRI for Predicting Microvascular Invasion in Hepatocellular Carcinoma. *Radiology*. 2020;297(3):573-81.
  12. Kong C, Zhao Z, Chen W, Lv X, Shu G, Ye M, et al. Prediction of tumor response via a pretreatment MRI radiomics-based nomogram in HCC treated with TACE. *Eur Radiol*. 2021;31(10):7500-11.
  13. Song W, Yu X, Guo D, Liu H, Tang Z, Liu X, et al. MRI-Based Radiomics: Associations With the Recurrence-Free Survival of Patients With Hepatocellular Carcinoma Treated With Conventional Transcatheter Arterial Chemoembolization. *J Magn Reson Imaging*. 2020;52(2):461-73.
  14. Liu Q-P, Yang K-L, Xu X, Liu X-S, Qu J-R, Zhang Y-D. Radiomics analysis of pretreatment MRI in predicting tumor response and outcome in hepatocellular carcinoma with transarterial chemoembolization: a two-center collaborative study. *Abdom Radiol (nY)*. 2022;47(2):651-63.
  15. Tipaldi MA, Ronconi E, Lucertini E, Krokidis M, Zerunian M, Polidori T, et al. Hepatocellular Carcinoma Drug-Eluting Bead Transarterial Chemoembolization (DEB-TACE): Outcome Analysis Using a Model Based On Pre-Treatment CT Texture Features. *Diagnostics (Basel)*. 2021;11(6).
  16. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol*. 2018;19(9):1180-91.
  17. Orlhac F, Lecler A, Savatovski J, Goya-Outi J, Nioche C, Charbonneau F, et al. How can we combat multicenter variability in MR radiomics? Validation of a correction procedure. *Eur Radiol*. 2021;31(4):2272-80.
  18. Calderaro J, Ziol M, Paradis V, Zucman-Rossi J. Molecular and histological correlations in liver cancer. *J Hepatol*. 2019;71(3):616-30.
  19. Sun Y, Bai H, Xia W, Wang D, Zhou B, Zhao X, et al. Predicting the Outcome of Transcatheter Arterial Embolization Therapy for Unresectable Hepatocellular Carcinoma Based on Radiomics of Preoperative Multiparameter MRI. *J Magn Reson Imaging*. 2020;52(4):1083-90.
  20. Ivanics T, Salinas-Miranda E, Abreu P, Khalvati F, Namdar K, Dong X, et al. A Pre-TACE Radiomics Model to Predict HCC Progression and Recurrence in Liver Transplantation: A Pilot Study on a Novel Biomarker. *Transplantation*. 2021;105(11):2435-44.
  21. Dai Y, Jiang H, Feng S-T, Xia Y, Li J, Zhao S, et al. Noninvasive Imaging Evaluation Based on Computed Tomography of the Efficacy of Initial Transarterial Chemoembolization to Predict Outcome in Patients with Hepatocellular Carcinoma. *J Hepatocell Carcinoma*. 2022;9:273-88.
  22. Meng X-P, Wang Y-C, Zhou J-Y, Yu Q, Lu C-Q, Xia C, et al. Comparison of MRI and CT for the Prediction of Microvascular Invasion in Solitary Hepatocellular Carcinoma Based on a Non-Radiomics and Radiomics Method: Which Imaging Modality Is Better? *J Magn Reson Imaging*. 2021;54(2):526-36.
  23. Luo J, Huang Z, Wang M, Li T, Huang J. Prognostic role of multiparameter MRI and radiomics in progression of advanced unresectable hepatocellular carcinoma following combined transcatheter arterial chemoembolization and lenvatinib therapy. *BMC Gastroenterol*. 2022;22(1):108.
  24. Gravitz L. Liver cancer. *Nature*. 2014;516(7529):S1.
  25. Lee JY, Lee BC, Kim HO, Heo SH, Shin SS, Jeong YY. Liver MRI and clinical findings to predict response after drug eluting bead transarterial chemoembolization in hepatocellular carcinoma. *Sci Rep*. 2021;11(1):24076.
  26. Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, et al. Radiomics and radiogenomics in lung cancer: A review for the clinician. *Lung Cancer*. 2018;115:34-41.
  27. Vliegenthart R, Fouras A, Jacobs C, Papanikolaou N. Innovations in thoracic imaging: CT, radiomics, AI and x-ray velocimetry. *Respirology*. 2022;27(10):818-33.
  28. Lee S-Y, Ou H-Y, Yu C-Y, Huang T-L, Tsang LL-C, Cheng Y-F. Drug-eluting bead transarterial chemoembolization for hepatocellular carcinoma: does size really matter? *Diagn Interv Radiol*. 2020;26(3):230-5.
  29. Prajapati HJ, Xing M, Spivey JR, Hanish SI, El-Rayes BF, Kauh JS, et al. Survival, efficacy, and safety of small versus large doxorubicin drug-eluting beads TACE chemoembolization in patients with unresectable HCC. *AJR Am J Roentgenol*. 2014;203(6):W706-W14.
  30. Lin C-Y, Liu Y-S, Pan K-T, Chen C-B, Hung C-F, Chou C-T. The short-term safety and efficacy of TANDEM microspheres of various sizes and doxorubicin loading concentrations for hepatocellular carcinoma treatment. *Sci Rep*. 2021;11(1):12277.
  31. Boulin M, Guiu S, Chauffert B, Aho S, Cercueil J-P, Ghiringhelli F, et al. Screening of anticancer drugs for chemoembolization of hepatocellular carcinoma. *Anticancer Drugs*. 2011;22(8):741-8.
  32. Mikhail AS, Pritchard WF, Negussie AH, Krishnasamy VP, Amchin DB, Thompson JG, et al. Mapping Drug Dose Distribution on CT Images Following Transarterial Chemoembolization with Radiopaque Drug-Eluting Beads in a Rabbit Tumor Model. *Radiology*. 2018;289(2):396-404.
  33. He Y, Xu T, Ding T, Gong Y, Zeng H, Xi Z, et al. In vitro comparative study of multimodal imaging nano-assembled microspheres with two clinical drug-eluting beads loaded with doxorubicin. *Drug Deliv*. 2023;30(1):2197177.
  34. Mikhail AS, Negussie AH, Mauda-Havakuk M, Owen JW, Pritchard WF, Lewis AL, et al. Drug-eluting embolic microspheres: State-of-the-art and emerging clinical applications. *Expert Opin Drug Deliv*. 2021;18(3):383-98.
  35. Zhao Y, Wu J, Zhang Q, Hua Z, Qi W, Wang N, et al. Radiomics Analysis Based on Multiparametric MRI for Predicting Early Recurrence in Hepatocellular Carcinoma After Partial Hepatectomy. *J Magn Reson Imaging*. 2021;53(4):1066-79.
  36. Kim BK, Kim SU, Kim KA, Chung YE, Kim M-J, Park M-S, et al. Complete response at first chemoembolization is still the most robust predictor for favorable outcome in hepatocellular carcinoma. *J Hepatol*. 2015;62(6):1304-10.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Zihan Xi<sup>1,3</sup> · Yuanxin Ye<sup>1</sup> · Yongbo Yang<sup>1</sup> · Yiwei He<sup>1</sup> · Ziyang Song<sup>1</sup> · Qian Ma<sup>1,3</sup> · Hui Zeng<sup>2</sup> · Guoliang Shao<sup>2</sup> 

✉ Zihan Xi  
13931382378@163.com

✉ Guoliang Shao  
shaogl@zjcc.org.cn

Yuanxin Ye  
yeyuanxin\_io@163.com

Yongbo Yang  
dryangyb@163.com

Yiwei He  
heyw@zjcc.org.cn

Ziyang Song  
szy1997522@163.com

Qian Ma  
mqjy1212@163.com

Hui Zeng  
zenghui@zjcc.org.cn

<sup>1</sup> Department of Radiology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou 310022, Zhejiang, China

<sup>2</sup> Department of Intervention, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou 310022, Zhejiang, China

<sup>3</sup> Postgraduate Training Base Alliance of Wenzhou Medical University, Wenzhou 325035, China