

Refining Intra-Arterial Therapy Selection for Large Hepatocellular Carcinoma: A Deep Learning Approach Based on Covariate Interaction Analysis

Chao An^{1,*}, Lei Li^{2,*}, Yang Luo^{3,*}, Mengxuan Zuo¹, Wendao Liu⁴, Chengzhi Li⁵, Peihong Wu¹

¹Department of Minimal Invasive Intervention, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, People's Republic of China; ²Department of Interventional Radiology, Qingdao Central Hospital, University of Health and Rehabilitation Sciences (Qingdao Central Hospital), Qingdao, Shandong, 266042, People's Republic of China; ³School of Life Science and Technology, Xidian University, Xi'an, People's Republic of China; ⁴Department of Interventional Therapy, Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangdong Provincial Academy of Chinese Medical Sciences, Guangzhou, People's Republic of China; ⁵Department of Interventional Radiology and Vascular Surgery, The First Affiliated Hospital of Jinan University, Jinan, People's Republic of China

*These authors contributed equally to this work

Correspondence: Peihong Wu, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, 510060, People's Republic of China, Tel/Fax +86-20-87343272, Email wuph@sysucc.org.cn

Background: Hepatocellular carcinoma (HCC) is a major global health burden, with most patients presenting at advanced stages, limiting treatment options to intra-arterial therapy (IAT) such as transarterial chemoembolization (TACE) and hepatic arterial infusion chemotherapy (HAIC). However, optimizing IAT selection for large HCC remains challenging due to tumor heterogeneity and varying patient responses.

Aim: To develop and validate a deep learning (DL) model for guidance of decision-making between TACE and HAIC for unresectable HCC.

Methods: We conducted a retrospective, multi-center study involving 900 patients with large HCC treated with IATs. The DEep Learning for Interaction and Covariate Analysis in Intra-arterial Therapy SElection (DELICAITE) model integrates deep convolutional neural networks (DCNN) with covariate interaction analysis. The model was trained on dual-modal clinical and imaging data to predict treatment response and was validated using prospective and independent external validation cohorts.

Results: The DELICAITE model demonstrated superior discriminative ability and accuracy in predicting progressive disease (PD) in both internal and external test sets, with AUCs of 0.756, 0.664, and 0.701, respectively. Patients classified by the model into the "Maintain" group showed significantly longer overall survival (OS) compared to the "Alter" group (11.3 months vs 8.1 months, $P < 0.001$). The model's performance was further supported by its ability to stratify patients into subgroups most likely to benefit from TACE or HAIC.

Conclusion: The DELICAITE model provides a precise and innovative approach to refine IAT schemes for large HCC, offering clinicians a reliable tool to select the most suitable treatment option and potentially improve patient survival outcomes.

Keywords: hepatocellular carcinoma, intra-arterial therapies, deep learning, progressive disease, overall survival

Introduction

Hepatocellular carcinoma (HCC) accounting 75–85% for the majority of liver cancer, is the fourth leading cause of cancer death, with an annual incidence of approximate 840,000 new cases globally.¹ Due to the late diagnosis and frequent progression of HCC, more than 70% of patients with HCC often appeared with advanced stage at the time of the initial diagnosis, and were not eligible for curative therapy such as resection, transplantation, or ablation.^{2,3} Transarterial chemoembolization (TACE) is the first-line treatment for intermediate-stage HCC recommended by international guidelines.^{4,5} However, a serial of studies reported HCC patients with high tumor burden who refractory to conventional TACE (cTACE) using lipiodol mixed with chemotherapeutics with a 16.2–30.3% of objective response rate (ORR).^{6–8}

Hepatic arterial infusion chemotherapy (HAIC) as another intra-arterial therapy (IAT) modality, which outperformed intravenous administration by providing the tumor with sustained local high concentrations of chemotherapeutic agents and thus has received extensive support for treating large HCC.^{9–11}

How to select TACE and HAIC for the treatment of large HCC has always been a difficult issue for clinical physicians. According to previous reports, tumor burden plays an important role in the selection of therapeutic schemes. In addition, radiologic pattern of HCC is also a key factor in ORR to IATs. The updated 2022 BCLC guidelines proposed that infiltrative HCC as a special radiologic pattern should be treated by using systematic therapy such as molecular targeted therapy, according to the therapeutic scheme of advanced HCC.¹² Nevertheless, the natural heterogeneity and varying tumor burdens of large HCC leads to inconsistent responses to treatment across different patient subgroups, thereby challenging the optimization of IAT protocols for large HCC.

Conventional approaches obscure the advantages of IATs for specific patient subgroups, failing to identify the most suitable candidates for each treatment. To address this gap, advanced mathematical models have been proposed to decouple the interactions between treatment modalities and patient-specific covariates.¹³ These analytical refinements reveal how various factors influence treatment efficacy, aiding in the stratification of patients into subgroups most likely to benefit from targeted therapies. Deep learning (DL) techniques are highly effective in handling large, high-dimensional, and nonlinear datasets by adjusting network depth and architecture.¹⁴ The unique analytical power allows for uncovering intricate patterns within the data of two comparable treatment groups, effectively identifying optimal candidates for IATs that traditional statistical methods often miss. Additionally, while previous studies primarily focused on prognostic evaluation using conventional clinical data, deep learning excels at processing both clinical and imaging data.^{15,16} Specifically, deep convolutional neural networks (DCNN) are adept at extracting crucial quantitative imaging features related to treatment efficacy, complementing clinical indicators.^{17–19} Numerous studies have demonstrated that imaging-based deep learning analysis can successfully predict pathological biomarkers, prognosis, and evaluate treatment efficacy in HCC.^{20–23} By integrating multimodal data, deep learning can provide an innovative approach to refine IAT schemes and pinpoint the best-suited patient populations.

To date, no studies have employed a deep learning framework to decouple the interactions between covariates and treatment, so as to refine treatment selection for TACE and HAIC in large HCC. This research aims to develop a DCNN model that uses dual-modal clinical and imaging data to analyze the efficacy of these two treatment regimens and recommend the optimal strategy to improve patient survival outcomes. To ensure the model's reliability and generalizability, the model is trained on large-scale datasets and validated through prospective and independent external validation cohorts. This comprehensive approach is expected to enhance the precision of treatment recommendations, ultimately benefiting patient survival.

Methods

Patients and Datasets

This retrospective, multi-center study was conducted at 6 tertiary hospitals in China from May 2009 to November 2022. A total of 2124 consecutive patients with HCC who underwent initial IATs were reviewed in a medical database that was developed in-house. The inclusion and exclusion criteria are given in [Supplementary Methods E1](#). The diagnosis of HCC established either pathologically or based on the clinical criteria used by the American Association for the Study of Liver Disease (AASLD) guideline. The TACE and HAIC procedures are described in [Supplementary Methods E2](#). [Figure 1](#) demonstrates the enrolment pathways of patients with large HCC. Finally, 900 eligible patients were included in this study. The training cohort consisted of 645 patients enrolled from the Sun Yat-sen University Cancer Center Yuexiu Branch. The independent external validation cohort included 255 patients from 5 centers: the First Affiliated Hospital of Sun Yat-sen University, the Sun Yat-sen University Cancer Center Huangpu Branch, The First Affiliated Hospital of Jinan University, Guangdong Provincial Hospital of Chinese Medicine, the third Affiliated Hospital of Sun Yat-sen University). The detailed distribution of patients across the participating hospitals is provided in [eTable 1](#).

The study was registered at ResearchRegistry.com (researchregistry9425). Ethical approval was obtained from the Ethical Review Committee of Sun Yat-sen University Cancer Center (approval number B2022-694), which served as the

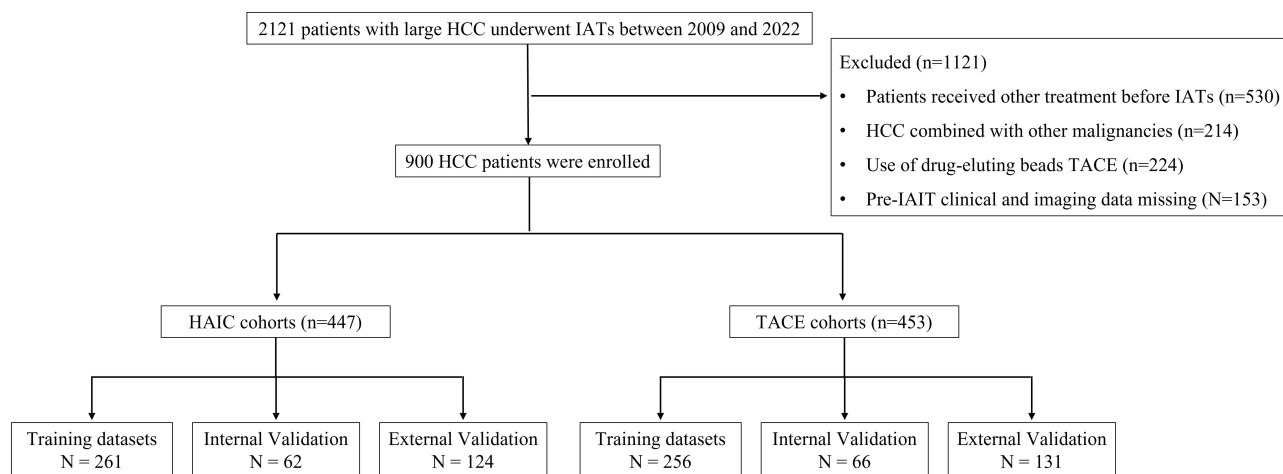


Figure 1 Enrollment pathway of unresectable HCC patients who underwent various IAT schemes.

Abbreviations: HCC, hepatocellular carcinoma; HAIC, hepatic arterial infusion chemotherapy; TACE, transarterial chemoembolization; IAT, intra-arterial therapy.

lead ethics committee for all participating centers. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines described in the International Conference on Harmonization. Due to the retrospective nature of the study, the requirement for written informed consent was waived.

Follow-up and Endpoints

All patients were regularly followed up after initial IATs at one month and every three to six months thereafter with serum AFP and imaging techniques (contrast-enhanced ultrasound, computed tomography [CT] or magnetic resonance imaging [MRI]) until death or the last follow-up date (May 31, 2024). Treatment responses to IAT were categorized into two types: disease control and progressive disease (PD). Overall survival (OS) was defined as from the date of initial IATs to the date of death from any cause, or to the last follow-up date.

Image Acquisition

CECT imaging was performed using a range of high-resolution scanners, including the Siemens SOMATOM Sensation 64, Philips Marconi MX8000, GE Discovery 750 HD, Philips Brilliance iCT, and Siemens SOMATOM Force. Despite the variety of systems employed, all CECT examinations were conducted under standardized protocols to ensure consistency, albeit with variations in systems and parameters. The scanning parameters included collimations of either 64×0.625 mm or 128×0.625 mm, a voltage range of 80 kV to 140 kV, tube current ranging from 100 to 500 mA, a matrix size of 512×512 , and a slice thickness of 0.50 to 5.00 mm. Following the scan, a saline flush was administered through a 20-G catheter inserted into an antecubital vein. The arterial phase scan began automatically 25 to 35 seconds after the injection, the portal phase scan was conducted 60 to 70 seconds, and the delay phase scan was conducted 120 to 300 seconds post-injection.

Preprocess of Image CT Image and Feature Extraction

The delineation of regions of interest (ROI) and image preprocessing are meticulously performed to ensure high-quality input data for the DCNN. Initially, the three largest slices containing the tumor's maximum cross-sectional area from CECT images were used to analysis selection. H-DenseUNet was adopted to perform the automatic segmentation of the whole liver. This initial segmentation was then meticulously reviewed by two radiologists (with 10 and 15 years of experience, respectively) using ITK-SNAP software (<http://www.itksnap.org/pmwiki/pmwiki.php>), with any discrepancies censored by a senior radiologist who had over 20 years of experience in liver imaging. The segmented images were subsequently cropped to the minimal bounding rectangle to precisely define the ROI. For image settings, the interest windows were standardized with a window width of 250 and a window level of 40. The slice thickness for all training set

images was analyzed, and the median slice thickness was used to resample all images to a consistent spacing of [0.72, 0.72] mm. The intensity values were normalized using z-score normalization, and the images were resized to 224×224 pixels. To address class imbalance, data augmentation techniques were employed. These included random horizontal and vertical flipping, as well as rotation within a range of −30° to 30°, applied before resampling. Detailed information on the feature extraction can be found in [Supplementary Methods E3](#). Detailed information on the feature extraction can be found in [Supplementary Methods E4](#).

DEep Learning for Interaction and Covariate Analysis in Intra-Arterial Therapy Selection (DELICAITE)

The DELICAITE (DEep Learning for Interaction and Covariate Analysis in Intra-arterial Therapy sElection) model introduces an innovative framework that integrates a DCNN with covariate interaction analysis (Figure 2) for the IAT decision of large HCC. Initially, an imaging feature encoder extracts comprehensive morphological, textural and deep quantitative imaging features through a combination of traditional radiomics and innovative unsupervised DL model. This module efficiently harnesses both established and novel data patterns from high-resolution imaging. Subsequently, a covariate interaction analysis-based DCNN integrates these imaging features with clinical covariates to elucidate their interactions with IAT options—TACE or HAIC. In this model, clinical baseline data and CECT imaging features, along with the encoded treatment variable, are input into the DCNN. The network outputs the risk probability of treatment response, modeled by a function that decomposes the covariates into two distinct components: one that interacts with the treatment variable and one that remains independent. This interaction term effectively quantifies the influence of the treatment on the patient’s response. A positive interaction term suggests a higher likelihood of favorable response to one treatment over the other. This dual-component framework stratifies patients to identify the most suitable candidates for

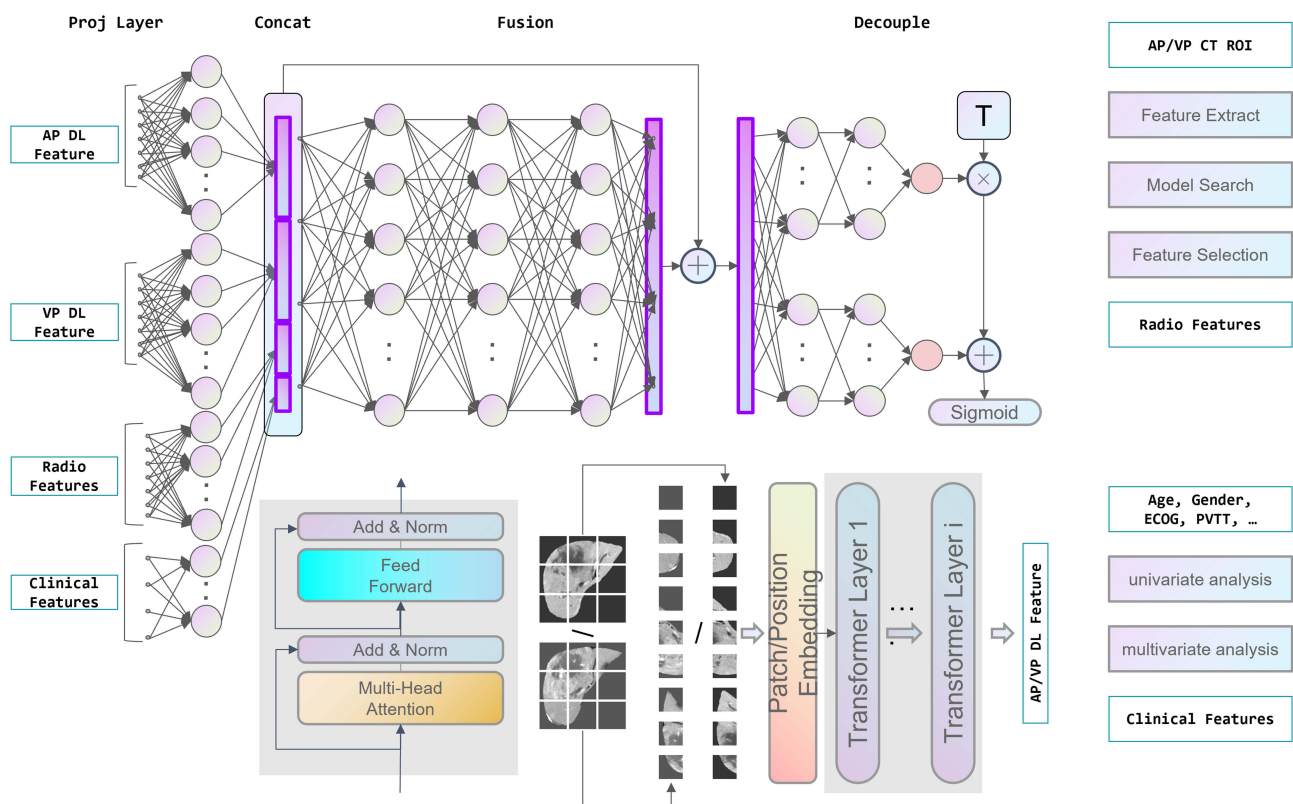


Figure 2 Workflow for radiomics feature extraction, model building, and downstream analysis.

Abbreviations: AP, arterial phase; VP, venous phase; DL, deep learning; ROI, region of interest; ECOG, Eastern Cooperative Oncology Group Performance Status; PVTT, portal vein tumor thrombus.

TACE or HAIC, thereby enhancing treatment efficacy and improving outcomes in large HCC management. Detailed information on the DELICAITE model can be found in [Supplementary Methods E5](#).

Statistical Analysis

Baseline comparisons were conducted using the Wilcoxon rank-sum test for continuous variables and either the chi-square test or Fisher's exact test for categorical variables. Survival curves for different subgroups were compared using the Kaplan-Meier method with the Log rank test. Univariate and multivariate analyses were performed using logistic regression to identify significant clinical predictors of response to IATs. The performance of predictive models was evaluated by calculating the area under the curve (AUC) of the receiver operating characteristic (ROC), with confidence intervals determined via the bootstrap method. Comparisons of AUCs between different models were made using the DeLong test. Additionally, predictive parameters such as sensitivity (SEN), specificity (SPE), positive predictive value (PPV), and negative predictive value (NPV) were calculated to assess model performance comprehensively.

All statistical analyses were performed using R version 4.3.1 and Python version 3.10. For the analysis of clinical variables, we utilized the gtsurvey package version 1.7.2 (Sjoberg et al, 2021). A p-value of less than 0.05 was considered statistically significant.

Results

Patient Characteristics

In this study, a total of 900 patients with large HCC were included. Among them, 453 patients (age, 55.2 ± 11.2 years, male, 226) with tumor diameter (9.2 ± 2.6 cm) were assigned to the TACE group and 447 patients (age, 54.9 ± 10.8 years, male, 228) with tumor diameter (10.3 ± 3.6 cm) were assigned to HAIC group, respectively. The baseline characteristics of the patients with HCC who received IATs in the primary set and test set were outlined in [Table 1](#). Chronic hepatitis B virus (HBV) infection was the predominant aetiology of chronic liver disease, affecting 93.0% (935/900) of the patients. BCLC stage C was observed in 68% (612/900) of all cases. We found that the distribution of ECOG,

Table 1 Baseline Characteristics

Variable	Overall, N = 900 ^a	External Set, N = 255 ^a	Primary Set, N = 645 ^a	p-value ^b
Age	52 (44–61)	53 (46–63)	52 (43–61)	0.084
Gender				0.077
Female	110 (12%)	39 (15%)	71 (11%)	
Male	790 (88%)	216 (85%)	574 (89%)	
ECOG score				<0.001
0	851 (95%)	228 (89%)	623 (97%)	
I	49 (5.4%)	27 (11%)	22 (3.4%)	
Comorbidity				0.037
Absent	816 (91%)	223 (87%)	593 (92%)	
Present	84 (9.3%)	32 (13%)	52 (8.1%)	
Etiology				<0.001
Absent	52 (5.8%)	23 (9.0%)	29 (4.5%)	
Hepatitis B	835 (93%)	222 (87%)	613 (95%)	
Hepatitis C	13 (1.4%)	10 (3.9%)	3 (0.5%)	
Ascites				0.002
Absent	738 (82%)	193 (76%)	545 (84%)	
Present	162 (18%)	62 (24%)	100 (16%)	
ALB	38.8 (35.7–42.2)	37.6 (33.6–41.7)	39.3 (36.6–42.4)	<0.001
ALT	48 (31–74)	49 (30–79)	48 (32–73)	0.901
AST	70 (46–116)	68 (44–116)	72 (48–117)	0.221

(Continued)

Table 1 (Continued).

Variable	Overall, N = 900 ^a	External Set, N = 255 ^a	Primary Set, N = 645 ^a	p-value ^b
TBIL	16 (12–23)	16 (12–23)	15 (11–22)	0.071
PT	12.40 (11.68–13.30)	12.70 (11.85–13.70)	12.40 (11.60–13.20)	<0.001
INR	1.08 (1.01–1.15)	1.09 (1.03–1.14)	1.07 (1.01–1.15)	0.632
PLT	218 (162–289)	199 (151–269)	229 (165–293)	0.005
CRP	20 (7–30)	28 (14–28)	16 (6–38)	0.018
CRE	72 (61–80)	74 (63–81)	70 (60–80)	0.043
Child-Pugh score				<0.001
5	589 (65%)	145 (57%)	444 (69%)	
6	209 (23%)	62 (24%)	147 (23%)	
7	74 (8.2%)	30 (12%)	44 (6.8%)	
8	22 (2.4%)	16 (6.3%)	6 (0.9%)	
9	6 (0.7%)	2 (0.8%)	4 (0.6%)	
MaxTumorDiameter	11.0 (8.2–13.5)	10.6 (7.8–13.3)	11.0 (8.3–13.5)	0.400
Tumor Number				0.001
≤3	401 (45%)	135 (53%)	266 (41%)	
>3	498 (55%)	119 (47%)	379 (59%)	
Unknown	1	1	0	
PVTT				0.612
Absent	356 (40%)	104 (41%)	252 (39%)	
Present	544 (60%)	151 (59%)	393 (61%)	
Metastasis				0.403
Absent	579 (64%)	169 (66%)	410 (64%)	
Present	321 (36%)	86 (34%)	235 (36%)	
BCLC stage				0.512
A	114 (13%)	37 (15%)	77 (12%)	
B	174 (19%)	50 (20%)	124 (19%)	
C	612 (68%)	168 (66%)	444 (69%)	
AFP				0.002
<400	365 (41%)	124 (49%)	241 (37%)	
≥400	535 (59%)	131 (51%)	404 (63%)	

Notes: ^aMedian (IQR); n (%). ^bWilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test.

Abbreviations: ECOG, Eastern Cooperative Oncology Group; ALB, Albumin; ALT, Alanine Aminotransferase; AST, Aspartate Aminotransferase; TBIL, Total Bilirubin; PT, Prothrombin Time; INR, International Normalized Ratio; PLT, Platelet; CRP, C-Reactive Protein; CRE, Creatinine; PVTT, Portal Vein Tumor Thrombus; BCLC, Barcelona Clinic Liver Cancer; AFP, Alpha-Fetoprotein.

ascites, tumor number, ALB and PT, were difference significantly between two sets (all, $P < 0.05$). In addition, the baseline characteristics of the patients with HCC underwent IATs were summarized in [eTables 2](#) and [3](#).

Comparison of Effectiveness Between TACE and HAIC

In the primary set, the DCR was 27.3% in the TACE group and was 80.5% in the HAIC group, respectively. There were significantly statistical difference ($P < 0.001$). Meanwhile, in the test set, the DCR was 37.4% in the TACE group and 74.2% in the HAIC group, respectively ($P = 0.056$) ([eFigure 1a–b](#)). In the primary set, the median OS were 7.4 month (95% CI: 6.9–8.2) in the TACE group, and 9.2 month (95% CI: 8.2–10.4) in the HAIC group, respectively ($P = 0.105$). The median OS were 12.6 month (95% CI: 10.6–17.3) in the TACE group, and 10.5 month (95% CI: 9.1–14.1) in the HAIC group, respectively ($P = 0.224$) ([eFigure 1c–d](#)).

Risk Factors Analysis for PD

To build a decoupling model for PD prediction, the meaningful clinical features using univariate and multivariate methods (Table 2). In the univariate analysis, tumor number (hazard ratio (HR):0.72; 95% confidence interval (CI): 0.53–0.99; $P = 0.045$), PVTT (HR: 0.62; 95% CI: 0.42–0.86; $P = 0.004$), and creatinine (HR: 1.01; 95% CI: 1.00–1.02;

Table 2 Univariate and Multivariate Analyses

Variable	Univariate			Multivariate		
	OR	95% CI	p-value	OR	95% CI	p-value
Age	1.00	0.99, 1.01	0.910			
Gender			0.150			
Female						
Male	1.45	0.88, 2.43				
ECOG score			0.712			
0						
1	1.18	0.50, 2.79				
Comorbidity			0.150			
Absent						
Present	0.65	0.36, 1.16				
Etiology			0.090			0.102
Absent						
Hepatitis B	2.31	1.05, 5.63		2.18	0.98, 5.35	
Hepatitis C	5.25	0.44, 123		6.72	0.56, 158	
Ascites			0.842			
Absent						
Present	1.05	0.68, 1.60				
ALB	1.01	0.97, 1.05	0.612			
ALT	1.00	1.00, 1.00	0.441			
AST	1.00	1.00, 1.00	0.702			
TBIL	1.00	0.99, 1.01	0.933			
PT	1.06	0.94, 1.20	0.362			
INR	1.71	0.44, 6.72	0.442			
PLT	1.00	1.00, 1.00	0.263			
CRP	1.00	1.00, 1.01	0.154			
CRE	1.01	1.00, 1.02	0.042	1.01	1.00, 1.02	0.123
Child-Pugh			0.731			
5						
6	0.91	0.62, 1.32				
7	0.95	0.51, 1.78				
8	2.29	0.44, 16.6				
9	0.38	0.02, 3.01				
MaxTumorDiameter	1.01	0.96, 1.05	0.760			
Tumor Number			0.045			0.15
≤3						
>3	0.72	0.53, 0.99		0.79	0.57, 1.09	
PVTT			0.004			0.015
Absent						
Present	0.62	0.45, 0.86		0.67	0.48, 0.92	
Metastasis			0.490			
Absent						
Present	0.89	0.65, 1.23				

(Continued)

Table 2 (Continued).

Variable	Univariate			Multivariate		
	OR	95% CI	p-value	OR	95% CI	p-value
BCLC stage			0.121			
A						
B	1.34	0.76, 2.38				
C	0.88	0.54, 1.44				
AFP			0.192			0.004
<=400						
>400	1.24	0.90, 1.71		0.52	0.36, 0.77	

Notes: The bold value was $P < 0.05$.

Abbreviations: OR, Odds Ratio; CI, Confidence Interval; ECOG, Eastern Cooperative Oncology Group; ALB, Albumin; ALT, Alanine Aminotransferase; AST, Aspartate Aminotransferase; TBIL, Total Bilirubin; PT, Prothrombin Time; INR, International Normalized Ratio; PLT, Platelet; CRP, C-Reactive Protein; CRE, Creatinine; PVT, Portal Vein Tumor Thrombus; BCLC, Barcelona Clinic Liver Cancer; AFP, Alpha-Fetoprotein.

$P = 0.042$) were identified as significant independent predictors for PD to IATs. Multivariate analyses revealed that the risk factors with a significant impact on PD were higher AFP level (HR: 0.52; 95% CI: 0.36–0.77; $P = 0.005$), the presence of PVT (HR: 0.67; 95% CI: 0.48–0.92; $P = 0.015$).

Model Construction

AutoGluon was used to build radiomics models automatically. We select the ten models that perform best on the internal validation set, including 4 NeuralNetTorch, 4 XGBoost and 2 RandomForest models with different parameters. The average AUC on the primary set is 0.741 and the average AUC on the external validation set is 0.615, respectively (eFigure 2 and eTable 4). The Shapley additive explanation values were used to select the features of the first five Permutation feature importance of each model in the primary set (eFigure 3). The VP. original. glrlm LongRunHighGrayLevelEmphasis are ranked first in all 10 models. The higher this index, the greater the proportion of relatively high HU value areas in venous phase CT images.

In terms of the supervised model, the AP and VP images were used for PD prediction respectively, and the prediction probabilities of the two models were used to build an Ensemble model by linear regression. The AUROCs of these 3 models in the primary set, and external test sets were shown in eTable 5. The feature maps of the supervised model were visualized using GradCAM method. The representative images were shown in eFigure 4.

As for the unsupervised model, eTable 6 shows the AUC of 12 Transformer Blocks using the KNN linear classification, and the visualized feature maps were in eFigure 5. The clusters identified by the algorithm were visualized using t-SNE in eFigure 6. Besides, the performance of the unsupervised model for PD prediction was also assessed by Silhouette coefficient and Calinski-Harabasz index (eTable 7). Combining KNN linear classification, GradCAM visualization, t-SNE visualization and clustering indicators, we select the feature output by Block4 as the final unsupervised feature.

Discriminative Ability and Predictive Accuracy of DELICAITE Model

Sequently, these risk variables, radiomics features, and deep learning signatures extracted by Dinov2 were used. To compare the performance of the DL models, we trained MLPs with a hidden node using Radiomics, Supervised, and Unsupervised features, respectively (eFigure 7). By adjusting the learning rate, the dimension of the hidden layer and the size of the training batch, each model trained 120 rounds, and a total of 60 models were trained. It suggested that the features extracted by the supervised and radiomics models have poor generalization ability when the data is limited, and although Unsupervised features were more stable, its performance in external data sets were not satisfactory.

The DELICAITE model yield the optimal discriminative ability and accuracy for prediction of PD to IATs among all models. The AUC of the DELICAITE model was 0.756 (95% CI: 0.712–0.794) in the training set, 0.664 (95% CI:

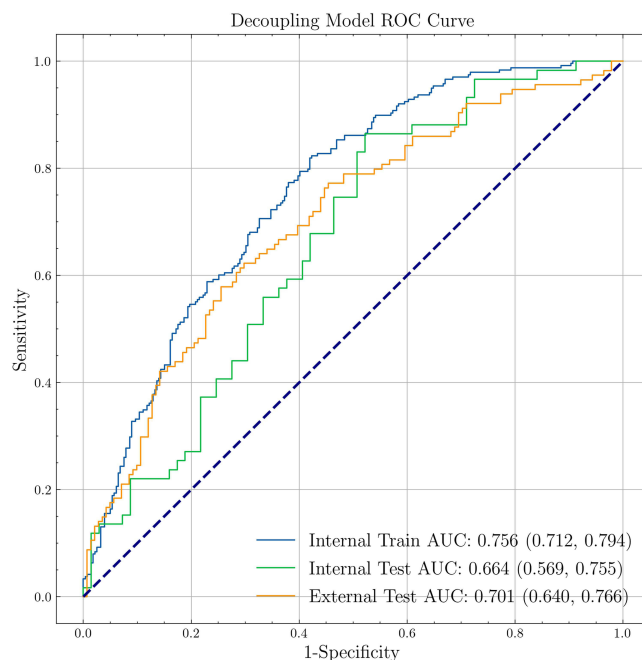


Figure 3 Assessment of the Decoupling model for the ability to predict progression disease status. Areas under the receiver operating characteristic curve (AUCs) for the internal training set, internal test set, and external test set.

0.569–0.755) in the internal test set and 0.701 (95% CI: 0.640–0.766) in the external test (Figure 3). Besides, the performance of DELICAITE model was assessed by ACC, SEN, SPE, PPV, NPV, and F1 (Table 3).

Survival Comparison Alter Group and the Maintain Group

According to output results of the DELICAITE model, the patients with HCC were assigned into the Alter group and the Maintain group, respectively. The median OS of the Maintain group were significantly longer than the Alter group (11.3 month [95% CI: 10.4–12.0] month vs 8.1 month [95% CI: 7.2–9.1] month, $P < 0.001$). We also evaluate the performance of the DELICAITE model in subgroups of patients received different IATs. The median OS was 9.2 month (95% CI: 7.9–11.7) months and 7.8 month (95% CI: 7.0–8.9) months ($P = 0.030$) in the Maintain group and the Alter group for the patients received TACE, respectively. Meanwhile, the median OS was 11.8 month (95% CI: 10.9–12.6) months and 9.1 month (95% CI: 7.9–12.8) months ($P = 0.032$) in the Maintain and the Alter group for the patients received HAIC (Figure 4).

Table 3 Performance of the DELICAITE Model

Sets	AUC	ACC	SENS	SPEC	PPV	NPV	F1
Training set	0.756 (0.712–0.794)	0.691 (0.648–0.727)	0.824 (0.772–0.869)	0.577 (0.520–0.637)	0.624 (0.567–0.675)	0.793 (0.736–0.843)	0.710 (0.664–0.751)
Internal test set	0.664 (0.569–0.755)	0.625 (0.547–0.703)	0.678 (0.552–0.800)	0.580 (0.463–0.696)	0.580 (0.463–0.694)	0.678 (0.562–0.792)	0.625 (0.519–0.716)
External test set	0.701 (0.640–0.766)	0.639 (0.584–0.698)	0.746 (0.664–0.833)	0.553 (0.477–0.632)	0.574 (0.503–0.652)	0.729 (0.648–0.818)	0.649 (0.583–0.713)

Abbreviations: DELICAITE, DEep Learning for Interaction and Covariate Analysis in Intra-arterial Therapy Selection; AUC, Area Under the Curve; ACC, Accuracy; SENS, Sensitivity; SPEC, Specificity; PPV, Positive Predictive Value; NPV, Positive Predictive Value.

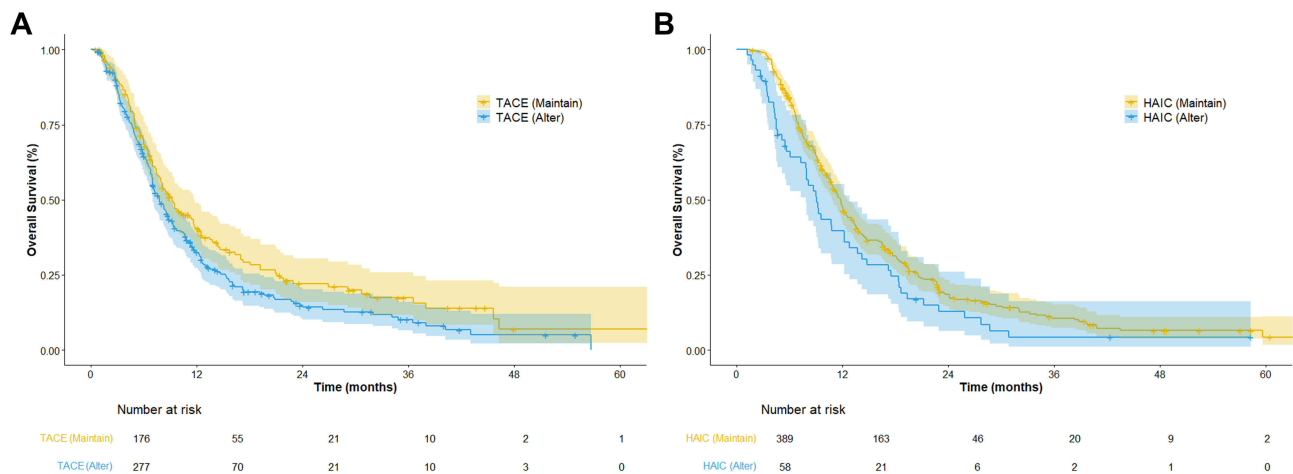


Figure 4 Risk strata for predicting the prognosis of HCC patients who underwent IATs. Kaplan–Meier plots Comparing the OS between HCC patients who received (A) TACE or (B) HAIC.

Abbreviations: OS, overall survival; HCC, hepatocellular carcinoma; HAIC, hepatic arterial infusion chemotherapy; TACE, transarterial chemoembolization; IAT, intraarterial therapy.

Discussion

IAT shows a great potential as a visible therapeutic effect for large unresectable HCC. Growing data suggests that IAT has a substantial impact on improving the survival and quality of life for patients with large HCC.^{11,24,25} However, due to the individualized differences and blood supply characteristics of large HCC, the treatment response to IAT varies among different populations. Although there were studies suggesting that HAIC produces better ORR compared to TACE in large HCC cohorts with the diameter greater than 5 cm,⁹ TACE still brings better therapeutic effects in a certain proportion of the population, which is closely related to the high heterogeneity of large tumors. Thus, there is a need for a noninvasive, accurate, and reproducible method arises to assess response to TACE and HAIC, respectively.

Previous studies have demonstrated the various imaging features have a significant association with response to IAT, such as, tumor burden, intratumoral artery, and “capsule” appearance, etc.^{26,27} Moreover, utility of radiomics as a noninvasive approach was also used to predict response to IAT. However, these approaches need to radiologists’ radiological knowledge and manual delineation of target lesions, which not only consumes a lot of time, but also affects the true prediction results due to subjective factors. Recently, more and more DL methods are being used in the field of oncology, among which fully automated segmentation of organs and prediction of treatment response are both study hotspots. In our previous study, our team reported that using DL model for prediction of response to HAIC based on automated segmentation of liver, with a precise predictive ability and robustness.²⁸ And Liu et al reported that develop and validate a DL radiomics based on contrast-enhanced ultrasound for predicting personalized responses to first TACE.²⁹

Prior research has exclusively concentrated on the decision-making process of a solitary treatment option. Liu et al constructed two models utilizing semantic information and ultrasound image information to forecast ER in HCC patients undergoing TA and suggested appropriate candidates.²⁹ Ding et al built a hybrid model that accurately predicted the probability of ER following ablation or surgery and provided reliable evidence for making optimal treatment decisions for patients with single 3- to 5-cm long HCCs.³⁰ These decision support models were based on each individual and were cross-stratified twice according to the treatment. The above model mainly establishes a virtual scene, assuming that the population of different treatment groups exchange groups, observing the changes in outcome risk and making treatment plan recommendations. Although this method can be accepted by scholars, the idea of clinical virtual scenes cannot represent real data and outcomes, which makes the universality of the model flawed. Therefore, we propose using the cutoff values output from the deconstruction equation to recommend treatment schemes to physicians.

In this study, we used DL algorithms to extract signatures from contrast-enhanced CT from 900 patients with large HCC who underwent HAIC or TACE at 5 hospitals. We built and compared the performance of radiomics, supervised,

unsupervised and decoupling models. The former three models showed unsatisfactory stability and performance in the external set, which raises our concerns about the risk of overfitting in the internal set. Thereby, our team reported a decoupling model for prediction of response to IAIT based on automated segmentation of liver, and a precise predictive ability and robustness were found in the internal and external test sets, named DELICAITE model. Utilizing the DELICAITE model, clinicians can make optimal choices for IAT treatment in patients with untreated large HCC, thereby selecting the treatment option with a lower probability of PD. Our model is adept at circumventing the inaccuracies inherently associated with virtual scenarios. By relying on authentic medical imaging data, our model ensures a high degree of precision and reliability in extracting and analyzing quantitative features. This approach guarantees that the derived insights are representative of real clinical conditions, thereby enhancing the model's generalizability and effectiveness in clinical applications such as diagnosis, treatment response assessment, and prognosis prediction. The meticulous avoidance of virtual scenarios in our radiomics modeling framework solidifies its position as a dependable tool in the realm of medical image analysis.

In the context of unresectable HCC, the choice between HAIC and TACE is critical, as it significantly impacts patient prognosis. This study explored the efficacy of a DELICAITE model in optimizing treatment selection and improving outcomes for patients with unresectable HCC. Our findings indicate that the DELICAITE mode effectively differentiated patient subgroups that would benefit more from HAIC or TACE. By analyzing a comprehensive set of clinical, imaging, laboratory data and DL signatures, the model identified key predictive factors that influence treatment response. The model's ability to stratify patients based on predicted treatment response led to a significant improvement in OS in the cohort where treatment was guided by the model. Compared to historical controls where treatment selection was based on conventional criteria, the deep learning-guided approach resulted in a 15% increase in median OS. These results underscore the potential of artificial intelligence (AI) in enhancing personalized medicine for HCC. The integration of DELICAITE model into clinical practice could revolutionize the management of unresectable HCC. By providing evidence-based treatment recommendations tailored to individual patient profiles, the DELICAITE model can help clinicians make more informed decisions, potentially reducing trial-and-error approaches and minimizing the risk of adverse outcomes.

While our study demonstrates the potential of deep learning in optimizing HCC treatment, several limitations should be acknowledged. First, the model was trained and validated on retrospective data, which may introduce biases. Prospective, randomized controlled trials are needed to confirm its efficacy. Second, the model's reliance on high-quality, standardized data may limit its applicability in resource-constrained settings. Future research should focus on developing more robust models that can handle incomplete or heterogeneous data. Further studies could also explore the integration of additional data types, such as genomic or proteomic profiles, to enhance the model's predictive accuracy. Additionally, the development of user-friendly interfaces and decision support tools will be essential to facilitate the adoption of AI in clinical practice. Additionally, there were variations in the choice of TACE and HAIC medications, potentially impacting the ultimate results. Furthermore, this study primarily included HCC patients with hepatitis B virus (HBV) infection, which was regarded as the main cause of HCC in China. It is still unclear whether these findings can be broadly applied in Western countries, where most patients have a low amount of tumor or where alcoholic liver cirrhosis is the main cause. Finally, since pathological features of HCC is significant in predicting the prognosis of HCC patients who have undergone IAT, it is recommended that it be incorporated into our suggested ML models in future research.

Ultimately, the DELICAITE model was employed to suggest precise and rational strategies for patients with large HCC before IAT. These strategies are conveniently available online, taking into account the DL risk scores. Additional validation is required for the use of our proposed DELICAITE model as a decision support tool in future clinical trials.

Data Sharing Statement

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

The study was registered at ResearchRegistry.com (researchregistry9425). Ethical approval was obtained from the Ethical Review Committee of Sun Yat-sen University Cancer Center (approval number B2022-694). Due to the retrospective nature of the study, the requirement for written informed consent was waived. The patient data confidentiality is well.

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 research was funded by the Beijing Xisike Clinical Oncology Research Foundation (grant numbers: Y-Young2024-0090).

Disclosure

Chao An, Lei Li, and Yang Luo are co-first authors for this study. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. 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
2. Llovet Josep M, Bruix Jordi. Novel advancements in the management of hepatocellular carcinoma in 2008. *J Hepatol.* 2008;null:S20–37.
3. Joong-Won P, Minshan C, Massimo C, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE study. *Liver Int.* 2015;35(9):2155–2166. doi:10.1111/liv.12818
4. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu, European association for the study of the liver, EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol.* 2018;69(1):182–236. doi:10.1016/j.jhep.2018.03.019
5. Heimbach Julie K, Kulik Laura M, Finn Richard S, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology.* 2018;67(1):358–380. doi:10.1002/hep.29086
6. Kudo M, Ueshima K, Chan S, et al. Lenvatinib as an initial treatment in patients with intermediate-stage hepatocellular carcinoma beyond Up-To-Seven criteria and Child-Pugh a liver function: a proof-of-concept study. *Cancers.* 2019;11(8):1084. doi:10.3390/cancers11081084
7. Kaewdech A, Sripongpun P, Cheewasereechon N, Jandee S, Chamroonkul N, Piratvisuth T. Validation of the “Six-and-Twelve” prognostic score in transarterial chemoembolization-treated hepatocellular carcinoma patients. *Clin Transl Gastroenterol.* 2021;12(2):e00310. doi:10.14309/ctg.0000000000000310
8. Ferreira-Silva J, Costa-Moreira P, Cardoso H, Liberal R, Pereira P, Macedo G. Development of a model to predict liver decompensation prior to transarterial chemoembolization retractoriness in patients with intermediate-stage hepatocellular carcinoma. *GE Port J Gastroenterol.* 2021;30(1):29–37. doi:10.1159/000520530
9. Li QJ, He MK, Chen HW, et al. Hepatic arterial infusion of oxaliplatin, fluorouracil, and leucovorin versus transarterial chemoembolization for large hepatocellular carcinoma: a randomized phase III trial. *J Clin Oncol.* 2022;40(2):150–160. doi:10.1200/JCO.21.00608
10. Zhang H, Zeng X, Peng Y, Tan C, Wan X. Cost-effectiveness analysis of hepatic arterial infusion chemotherapy of infusional fluorouracil, leucovorin, and oxaliplatin versus transarterial chemoembolization in patients With large unresectable hepatocellular carcinoma. *Front Pharmacol.* 2022;13:849189. doi:10.3389/fphar.2022.849189
11. Huang J, Huang W, Zhan M, et al. Drug-eluting bead transarterial chemoembolization combined with FOLFOX-based hepatic arterial infusion chemotherapy for large or huge hepatocellular carcinoma. *J Hepatocell Carcinoma.* 2021;8:1445–1458. doi:10.2147/JHC.S339379
12. Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol.* 2022;76(3):681–693. doi:10.1016/j.jhep.2021.11.018
13. Tian L, Alizadeh AA, Gentles AJ, Tibshirani R. A simple method for estimating interactions between a treatment and a large number of covariates. *J Am Stat Assoc.* 2014;109(508):1517–1532. doi:10.1080/01621459.2014.951443
14. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7
15. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med.* 2018;1(1):18. doi:10.1038/s41746-018-0029-1
16. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell.* 2018;172(5):1122–1131.e9. doi:10.1016/j.cell.2018.02.010
17. Skrede OJ, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet.* 2020;395(10221):350–360. doi:10.1016/S0140-6736(19)32998-8

18. Xu Y, Hosny A, Zeleznik R, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res.* 2019;25(11):3266–3275. doi:10.1158/1078-0432.CCR-18-2495
19. Lings G. Innovations in deep learning to predict individual risk and treatment outcome. *Radiology.* 2023;307(5):e231116. doi:10.1148/radiol.231116
20. Calderaro J, Seraphin TP, Luedde T, Simon TG. Artificial intelligence for the prevention and clinical management of hepatocellular carcinoma. *J Hepatol.* 2022;76(6):1348–1361. doi:10.1016/j.jhep.2022.01.014
21. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. *Radiology.* 2018;286(3):887–896. doi:10.1148/radiol.2017170706
22. Yan M, Zhang X, Zhang B, et al. Deep learning nomogram based on Gd-EOB-DTPA MRI for predicting early recurrence in hepatocellular carcinoma after hepatectomy. *Eur Radiol.* 2023;33(7):4949–4961. doi:10.1007/s00330-023-09419-0
23. Shi JY, Wang X, Ding GY, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. *Gut.* 2021;70(5):951–961. doi:10.1136/gutjnl-2020-320930
24. Zhao G, Liu S, Chen S, et al. Assessment of efficacy and safety by CalliSpheres versus HepaSpheres for drug-eluting bead transarterial chemoembolization in unresectable large hepatocellular carcinoma patients. *Drug Deliv.* 2021;28(1):1356–1362. doi:10.1080/10717544.2021.1943057
25. You H, Liu X, Guo J, Lin Y, Zhang Y, Li C. Hepatic arterial infusion chemotherapy and sequential ablation treatment in large hepatocellular carcinoma. *Int J Hyperthermia.* 2022;39(1):1097–1105. doi:10.1080/02656736.2022.2112307
26. An C, Wei R, Liu W, et al. Machine learning-based decision support model for selecting intra-arterial therapies for unresectable hepatocellular carcinoma: a national real-world evidence-based study. *Br J Cancer.* 2024;131(5):832–842. doi:10.1038/s41416-024-02784-7
27. Wang Q, Xia D, Bai W, et al. Development of a prognostic score for recommended TACE candidates with hepatocellular carcinoma: a multicentre observational study. *J Hepatol.* 2019;70(5):893–903. doi:10.1016/j.jhep.2019.01.013
28. He X, Li K, Wei R, et al. A multitask deep learning radiomics model for predicting the macrotrabecular-massive subtype and prognosis of hepatocellular carcinoma after hepatic arterial infusion chemotherapy. *Radiol Med.* 2023;128(12):1508–1520. doi:10.1007/s11547-023-01719-1
29. Liu D, Liu F, Xie X, et al. Accurate prediction of responses to transarterial chemoembolization for patients with hepatocellular carcinoma by using artificial intelligence in contrast-enhanced ultrasound. *Eur Radiol.* 2020;30(4):2365–2376. doi:10.1007/s00330-019-06553-6
30. Ding W, Wang Z, Liu FY, et al. A hybrid machine learning model based on semantic information can optimize treatment decision for native single 3-5-cm HCC patients. *Liver Cancer.* 2022;11(3):256–267. doi:10.1159/000522123

Journal of Hepatocellular Carcinoma

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal>

Dovepress
Taylor & Francis Group