British Journal of Hospital Medicine

Br. J. Hosp. Med. (Lond) **2026**; 87(1): 50380
https://doi.org/10.31083/BJHM50380

Article

Survival Prediction and Treatment Decisions in Hepatocellular Carcinoma: A Deep Learning-Based Radiomics Approach

Xiaoqin Wei[1], Jun Xiao[1], Ying Liu[2], Chaofeng Yang[3], Ziren Luo[4], Mingyue Tang[4,*], Xiaowen Chen[1,*]

[1]School of Medical Imaging, North Sichuan Medical College, 637000 Nanchong, Sichuan, China
[2]Department of Radiology, The First Affiliated Hospital of Chengdu Medical College, 610500 Chengdu, Sichuan, China
[3]Department of Radiology, Affiliated Hospital of North Sichuan Medical College, 637000 Nanchong, Sichuan ,China
[4]School of Basic Medicine and Forensic Medicine, North Sichuan Medical College, 637000 Nanchong, Sichuan, China
*Correspondence: 463540287@qq.com (Mingyue Tang); xw_chen_nsmc@126.com (Xiaowen Chen)
Academic Editor: John Alcolado
Submitted: 18 December 2024 Revised: 14 June 2025 Accepted: 10 July 2025 Published: 26 January 2026

Abstract

Aims/Background: Deep learning radiomics (DLRadiomics) can capture a wide range of tumor and lesion characteristics, providing valuable insights into biological behavior, pathophysiological status, and patient prognosis. This study integrated clinical data with deep learning-derived features into a machine learning survival model to assess the effectiveness of hepatectomy and transarterial chemoembolization (TACE) treatments in hepatocellular carcinoma (HCC) patients. **Methods**: This study included pathologically confirmed HCC patients who received either hepatectomy or TACE between January 2013 and December 2022. We utilized three deep learning-based algorithms (ResNet50, ResNet18, and DenseNet121) with contrast-enhanced computed tomography (CT) images to predict the overall survival time. Deep learning features were extracted from these predictive models. Furthermore, a combined survival model was developed by incorporating clinical factors with the deep learning features for two treatment regimens separately. The areas under the curves (AUC) of the receiver operating characteristic (ROC) curves were used to assess the discrimination of the model at different time points. Additionally, nomograms were constructed to predict patient prognosis undergoing different treatment regimens, and their survival risk was evaluated using the Kaplan-Meier analysis. **Results**: This study recruited 409 HCC patients who received either hepatectomy (n = 278, 57 [49–66]; 239 men) or TACE (n = 131, 62 [51–69.5]; 111 men). ResNet50 achieved the highest AUC of 0.866 (95% confidence interval (CI): 0.815–0.917) in the training set and 0.793 (95% CI: 0.675–0.912) in the testing cohort. Overall, six models were constructed to assess overall survival for hepatectomy and TACE treatments, with the combined models exhibiting superior discriminative performance. The C-index for the combined hepatectomy model was 0.836 (95% CI: 0.776–0.897) in the training cohort and 0.861 (95% CI: 0.755–0.967) in the testing cohort. The C-index for the combined TACE model was 0.840 (95% CI: 0.792–0.888) in the training cohort and 0.834 (95% CI: 0.759–0.910) in the testing cohort. Two nomograms were created to help clinicians in selecting a treatment method by examining the difference scores between treatments. **Conclusion**: The machine learning models can potentially predict differences in outcomes between hepatectomy and TACE. Furthermore, prognostic models using deep learning-based features can effectively predict survival risk in HCC patients.

Keywords: deep learning; hepatocellular carcinoma; prognosis; hepatectomy; TACE

1. Introduction

Primary liver cancer is the third most common cause of cancer-related mortality worldwide, with hepatocellular carcinoma (HCC) representing about 75%–85% of cases [1]. Various treatment guidelines have been developed to guide the management of HCC, including the China Liver Cancer (CNLC) staging system [2] and the Barcelona Clinic Liver Cancer (BCLC) staging system [3]. Hepatectomy and transarterial chemoembolization (TACE) are two primary curative interventions for different stages of HCC [4,5]; however, their ability to inform clinical decision-making remains limited. Advances in surgical approaches and the expanding indications for TACE have led to an overlap in the application of these treatments [6–9]. Furthermore, patient heterogeneity within the same disease stage leads to different survival outcomes among those receiving the same treatment. Hence, the optimal treatment strategies for HCC patients remain poorly defined [10]. Moreover, there is a paucity of studies specifically addressing treatment selection and prognosis prediction for patients diagnosed with HCC [11,12]. To evaluate the causal effect of hepatectomy versus TACE, outcome comparisons within the same patient population are essential [13].

Radiomics is an emerging field of quantitative imaging analysis that aims to correlate extensive extracted imaging features with clinical and biological outcomes [14]. Advances in quantitative imaging approaches combined with machine learning have provided an opportunity for translating data science research into more personalized cancer treatments [15].

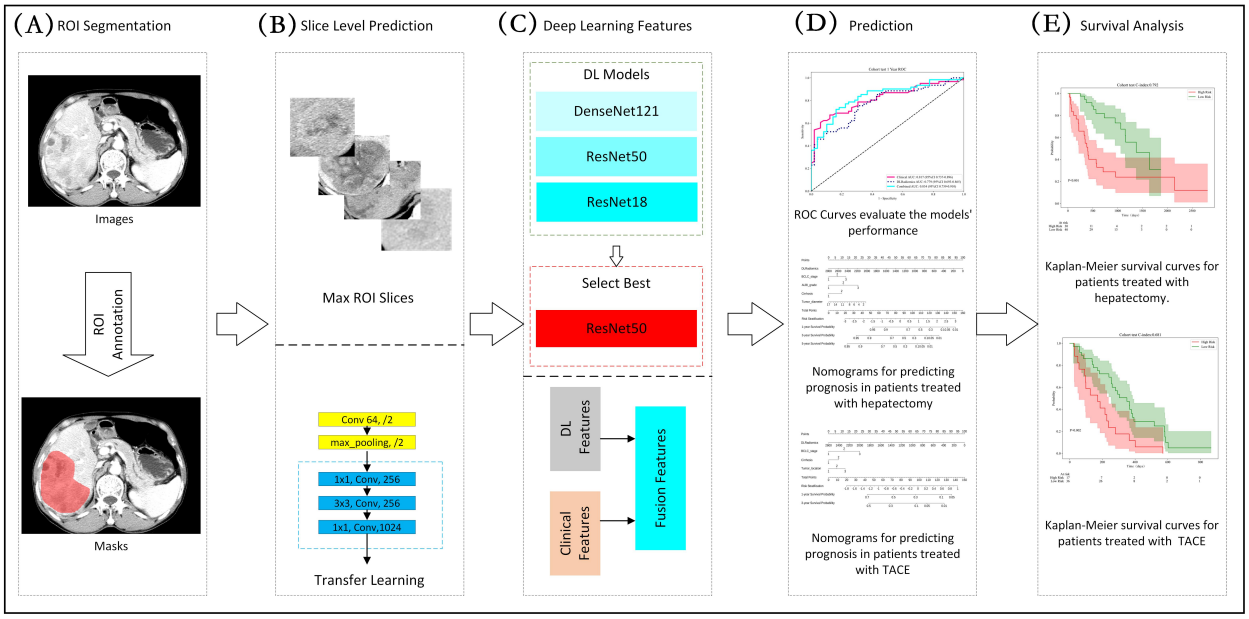

Fig. 1. A flowchart of the integrated study approach. (A) Collecting clinical features and contrast-enhanced computed tomography (CT) images, and delineation of the region of interest (ROI) by using 3D Slicer. (B) Constructing deep learning-based networks. (C) Extracting deep learning radiomics (DLRadiomics) features and developing a combined survival model using clinical and DLRadiomics features individually for two treatments. (D) Comparing the performance of the models and Constructed nomograms to predict the prognosis of patients. (E) Analyzing the predictive ability of the model in different subgroups.

Currently, deep learning-based radiomics (DLRadiomics) can capture various aspects of tumors and other lesions compared to traditional radiomics, offering valuable insights into biological behavior, pathophysiological conditions, and potential prognoses [16]. DLRadiomics models can further improve the predictive accuracy of traditional radiomic approaches. Emerging evidence suggests that DLRadiomics holds significant potential for predicting prognosis and treatment response in various cancer types, such as nasopharyngeal carcinoma, HCC, and non-small cell lung cancer [17]. However, there is a scarcity of studies addressing how DLRadiomics can inform therapeutic decisions between TACE and hepatectomy to enhance patient outcomes.

Computed tomography (CT) imaging-based survival predictions can support response evaluation in clinical trials, precision medicine, and customized clinical interventions. Therefore, we obtained CT images during both the arterial and portal venous phases. To enhance treatment decision-making for patients with HCC, we have devised a DLRadiomics strategy using CT imaging data. The complete workflow of our study design, including feature extraction, model development, and validation, is illustrated in Fig. 1.

2. Methods

2.1 Recruitment of Study Participants

This retrospective cohort study included pathologically confirmed HCC patients treated at Affiliated Hospital of North Sichuan Medical College, China, between January 2013 and December 2022. Initially, patients received either hepatectomy or TACE, and their dataset was acquired from the medical record system of the hospital, with patient identity anonymized.

The inclusion criteria were as follows: (1) diagnosis of HCC confirmed through histopathology or according to the 2018 EASL/AASLD treatment guidelines [18,19]; (2) those who underwent hepatectomy or TACE followed by subsequent monitoring; (3) patients aged 18 years or above; (4) those with no other malignancies. The exclusion criteria included (1) patients with prior systemic or local antitumor treatments (such as transcatheter arterial chemoembolization, immunotherapy, or ablation) before the initial treatment; (2) patients with no CT scans performed before surgery or scans conducted more than 1 month before treatment; (3) patients with poor quality CT images; (4) patients with missing laboratory data, demographic information, or follow-up records.

The primary endpoint was overall survival (OS), defined as the time from the initiation of the first treatment (hepatectomy or TACE) to death from any cause or the last follow-up, which was performed in December 2023. Clinical characteristics, including demographic information, laboratory results, and radiological characteristics, were analyzed in this study. The clinical characteristics of the study participants are detailed in **Supplementary Text 1**.

2.2 Acquisition of Images

All patients received a routine contrast-enhanced abdominal CT within one month before treatment. CT scans were conducted using the Philips Brilliance (Brilliance CT 64 Channel, Philips, Netherlands), GE Lightspeed VCT (GE Healthcare, Chicago, IL, USA), and Siemens Definition AS (Siemens Healthineers, Erlangen, Germany). Each patient received a 60 mL of contrast agent administered intravenously at a rate of 3.0 to 3.5 mL/s using a high-pressure injector. Subsequently, arterial-phase images were captured 25–30 seconds after contrast injection, and portal-phase images were obtained 50–60 seconds post-injection.

2.3 Segmentation of Images

The regions of interests (ROIs) were delineated using 3D Slicer (Version 5.8.1, Kitware, Inc., Albany, NY, USA). Two experienced radiologists independently annotated the ROIs in a blinded and sequential manner to ensure unbiased assessments. Any discrepancies between their annotations were resolved by a senior radiologist with 20 years of clinical experience, who established the final consensus on the region of interest (ROI) boundaries. This meticulous approach ensured accurate and reliable ROI identification, which is crucial for subsequent analyses. All images underwent Z-score normalization to standardize pixel intensities.

2.4 DLRadiomics Procedure

2.4.1 Model Training

Patients were randomly divided into training and testing groups. They were further classified into two groups based on three-year survival outcomes: those who survived longer than three years were categorized as low risk, while those with shorter survival times (less than three years) were designated as high risk. A convolutional neural network (CNN) model was then trained to differentiate between the high- and low-risk groups. Transfer learning was implemented using three architectures—ResNet50, ResNet18, and DenseNet121—each initialized with ImageNet pre-trained weights to leverage previously acquired features. The learning rate was adaptively adjusted using a cosine decay strategy to ensure optimal convergence during training. Detailed training parameters and procedures are provided in **Supplementary Text 2**. The model demonstrating the highest accuracy and generalization performance on the test set was selected. The output probabilities generated by the CNN were defined as DLRadiomics.

Gradient-weighted Class Activation Mapping (Grad-CAM) is a visualization technique that identifies the regions within an image that contribute most significantly to the model's predictions. It works by computing the gradients of the target output (e.g., high-risk vs. low-risk classification) with respect to the final convolutional layer, generating a coarse localization map that indicates the region most relevant to the prediction. This technique enables visualization of how the model focuses on specific areas in the CT images, providing valuable insights into its decision-making process and improving the interpretability and transparency of the deep learning model.

2.4.2 Feature Extraction and Selection for Survival Analysis

After developing deep learning models, we extracted DLRadiomics features from the penultimate layer for survival analysis. Feature selection was conducted in two primary steps to refine the feature set and minimize overfitting. Initially, a correlation coefficient filter was applied to reduce multicollinearity by retaining only one feature from each pair with a correlation coefficient above 0.8. Subsequently, univariate Cox proportional hazards regression was performed to identify features significantly associated with survival outcomes (p-value < 0.05), ensuring that only meaningful predictors were included. The final radiomic signature was then established using LASSO Cox regression. These selected features were then used to construct the final Cox survival model for both treatment groups.

2.5 Signature Building

Patients were classified into either a hepatectomy group or a TACE group. Separate individualized models were established for each treatment using the following steps. First, we conducted univariate and multivariate Cox proportional hazards regression analyses. Features with a multivariate p-value below 0.05 were selected to develop the clinical survival prediction model. Next, for each group, we constructed a combined survival model that integrated the DLRadiomics signature with the selected clinical features. The areas under the curves (AUC) of the receiver operating characteristic curves (ROC) were used to assess the discriminative performance of each model at different time points. Based on these models, nomograms were developed by incorporating both DLRadiomics and clinical variables to predict patient prognosis for each treatment regimen and guide optimal treatment recommendations. Finally, Kaplan-Meier analysis was performed to evaluate the survival risk of the patients within the groups.

2.6 Statistical Analysis

The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Variables following normal distribution were analyzed using t-tests and expressed as mean $\pm$ SD (standard deviation). Non-normally distributed variables were analyzed using the Mann-Whitney U-test and expressed as median (interquartile range). Moreover, categorical variables were analyzed using the chi-square test and reported as counts (%). A $p < 0.05$ was considered statistically significant.

Cox proportional hazards regression analyses were conducted using R software (Version 4.5.1, R Core Team, https://www.r-project.org). Radiomics features were extracted with PyRadiomics version 3.0.1. Machine learning

algorithms, such as Logistic Regression (LR), were implemented with Scikit-learn version 1.0.2 (open source). Furthermore, deep learning models were developed in PyTorch version 1.11.0 (open-source deep learning framework) and optimized using CUDA version 11.3.1 and cuDNN version 8.2.1 (NVIDIA Corporation, Santa Clara, CA, USA).

3. Results

3.1 Patients

This study included 409 patients, comprising 278 individuals (age: 57 [49–66] years; 239 men) who received hepatectomy and 131 (age: 62 [51–69.5]; 111 men) who underwent TACE. In the hepatectomy group, the median survival time for the entire cohort was 517 days (interquartile range [IQR]: 260–1098), with 144 patients alive at the last follow-up. In the TACE group, the median survival time was 218 days (IQR: 129–361), with 23 patients alive at the last follow-up. Patients in both groups were randomly divided into training and testing cohorts: the training cohort included 208 hepatectomy and 78 TACE patients, while the testing cohort included 70 hepatectomy and 53 TACE patients. The preoperative clinical characteristics of the hepatectomy and TACE groups are detailed in Table 1 and **Supplementary Table 1**.

3.2 Association of Clinical Characteristics With Overall Survival of the Study Participants

For both the hepatectomy and TACE groups, clinical characteristics were individually analyzed within the training cohort (n = 208 for hepatectomy, n = 78 for TACE). In the hepatectomy group, univariate Cox proportional hazards regression analysis revealed 13 factors significantly associated with overall survival (OS) ($p < 0.1$), including Child Pugh class, arterial peritumoral enhancement, glutamyl transpeptidase (GGT), HCC capsule, serum albumin (ALB), albumin-bilirubin (ALBI) grade, BCLC stage, cirrhosis, tumor number, international normalized ratio (INR), fusion lesions, intratumoral necrosis, and tumor diameter (**Supplementary Table 2**). Multivariate analysis further confirmed that BCLC stage, cirrhosis, ALBI grade, and tumor diameter were independent predictors of OS ($p < 0.05$). In the TACE group, univariate analysis demonstrated 6 factors (GGT, ALB, ALBI grade, BCLC stage, cirrhosis, and tumor location) were significantly associated with OS ($p < 0.1$) (**Supplementary Table 3**). Moreover, multivariate analysis revealed that BCLC stage, cirrhosis, and tumor location were independent risk factors for survival ($p < 0.05$).

3.3 Development of Deep Learning-Based Radiomics Models and Extracting Features

We stratified patients into low-risk and high-risk subgroups based on a three-year survival prognosis (low-risk: survival >3 years; high-risk: survival ≤ 3 years). Using these subgroups, we developed three deep learning-based radiomics models to predict risk classification. The models were trained on the training cohort and evaluated on the testing cohorts to assess their performance. Among the models, ResNet50 achieved the highest AUC in both the training and testing cohorts, with values of 0.866 (95% confidence interval (CI): 0.815–0.917) and 0.793 (95% CI: 0.675–0.912), respectively.

Furthermore, ResNet50 demonstrated superior performance across multiple evaluation metrics. In the training cohort, it attained an accuracy of 0.766 (sensitivity: 0.724, specificity: 0.877), while in the testing cohort, accuracy was 0.767 (sensitivity: 0.787, specificity: 0.667). These findings surpassed those of DenseNet121 (training accuracy: 0.670, testing accuracy: 0.656) and ResNet18 (training accuracy: 0.761, testing accuracy: 0.600), indicating ResNet50's discriminative power and generalizability. Detailed performance metrics for all models are summarized in **Supplementary Table 4**. DenseNet121 displayed moderate performance, with an AUC of 0.798 (95% CI: 0.732–0.864) in the training cohort, which declined to 0.748 (95% CI: 0.602–0.894) in the testing cohort, suggesting a decrease in predictive capability. ResNet18 demonstrated stable but slightly lower performance, with an AUC of 0.782 (95% CI: 0.717–0.848) in the training set and 0.771 (95% CI: 0.651–0.891) in the test set. These findings indicate that the ResNet50 model outperformed the other models in both the training and testing cohorts (**Supplementary Table 4**, **Supplementary Fig. 1**). Therefore, ResNet50 was selected as the primary model for extracting features in the subsequent analyses.

A total of 1024 prognostically relevant deep radiomic features were extracted from the penultimate convolutional layer of ResNet50. Features were filtered using a correlation coefficient threshold ($|r| \geq 0.80$) and univariate Cox proportional hazards regression ($p < 0.05$), retaining 588 spatially distinct features. Using 10-fold cross-validation, the optimal regularization parameter ($\lambda = 0.173$) was identified in the LASSO Cox regression model, achieving a C-index of 0.679 (95% CI: 0.638–0.713) with 26 non-zero coefficients retained, comprising the final prognostic signature (**Supplementary Figs. 2–4**).

3.4 Visualization of Deep Learning-Based Radiomics Models

To investigate how the deep learning models recognize various samples, we used the Gradient-weighted Class Activation Mapping (Grad-CAM) technique for visualization. As illustrated in Fig. 2, Grad-CAM highlights the activation areas in the final convolutional layer that predict relevant risk categories. This visualization identifies the image regions that significantly impact the model's decision-making, offering valuable insights into its interpretability.

Table 1. Preoperative clinical characteristics of the hepatectomy and TACE groups.

Features	Hepatectomy (n = 278)	TACE (n = 131)	*p*-value
Survival time (days), [Median (Q1, Q3)]	517 (260, 1098)	218 (129, 361)	<0.001
Survival mode, No. (%)			<0.001
Survival	144 (51.8%)	23 (17.6%)	
Non-Survival	134 (48.2%)	108 (82.4%)	
Child Pugh class, No. (%)			<0.001
A	251 (90.3%)	99 (75.6%)	
B	27 (9.7%)	32 (24.4%)	
ECOG-PS, No. (%)			0.291
0	56 (20.1%)	18 (13.7%)	
1	210 (75.5%)	107 (81.7%)	
2	12 (4.4%)	6 (4.6%)	
Gender, No. (%)			0.739
Male	239 (86.0%)	111 (84.7%)	
Female	39 (14.0%)	20 (15.3%)	
HBV infection, No. (%)			0.828
Absent	61 (21.9%)	30 (22.9%)	
Present	217 (78.1%)	101 (77.1%)	
DM, No. (%)			0.946
Absent	251 (90.3%)	118 (90.1%)	
Present	27 (9.7%)	13 (9.9%)	
ALBI grade, No. (%)			<0.001
1	142 (51.1%)	32 (24.4%)	
2	128 (46.0%)	91 (69.5%)	
3	8 (2.9%)	8 (6.1%)	
Cirrhosis, No. (%)			0.254
Absent	202 (72.7%)	88 (67.2%)	
Present	76 (27.3%)	43 (32.8%)	
Tumor location, No. (%)			<0.001
Left	97 (34.9%)	22 (16.8%)	
Right	166 (59.7%)	65 (49.6%)	
Across	15 (5.4%)	44 (33.6%)	
BCLC, No. (%)			<0.001
A stage	142 (51.1%)	33 (25.2%)	
B stage	64 (23.0%)	40 (30.5%)	
C stage	72 (25.9%)	58 (44.3%)	
Fusion lesions, No. (%)			<0.001
Absent	178 (64.0%)	55 (42.0%)	
Present	100 (36.0%)	76 (58.0%)	
HCC capsule, No. (%)			<0.001
Absent	78 (28.1%)	47 (35.9%)	
Integral	133 (47.8%)	25 (19.1%)	
Unintegral	42 (15.1%)	45 (34.4%)	
HCC capsule breakthrough	25 (9.0%)	14 (10.7%)	
Intratumoral necrosis, No. (%)			0.749
Absent	100 (36.0%)	45 (34.4%)	
Present	178 (64.0%)	86 (65.6%)	
Arterial peritumoral enhancement, No. (%)			0.021
Absent	118 (42.4%)	40 (30.5%)	
Present	160 (57.6%)	91 (69.5%)	
Tumor number, No. (Median [Q1, Q3])	1.00 (1.00, 2.00)	2.00 (1.00, 6.00)	<0.001
Tumor diameter (mm) (Median [Q1, Q3])	5.50 (3.73, 8.00)	9.80 (7.20, 12.70)	<0.001

Table 1. Continued.

Features	Hepatectomy (n = 278)	TACE (n = 131)	*p*-value
Age (years) (Median [Q1, Q3])	57.00 (49.00, 66.00)	62.00 (51.00, 69.50)	0.014
PLT (10^9/L), (Median [Q1, Q3])	135.00 (96.00, 191.75)	152.70 (111.50, 209.50)	0.021
AST(U/L), (Median [Q1, Q3])	45.70 (30.33, 68.00)	75.00 (46.45, 130.50)	<0.001
ALT (U/L), (Median [Q1, Q3])	35.50 (24.00, 61.68)	39.00 (26.00, 66.00)	0.227
GGT (U/L), (Median [Q1, Q3])	62.85 (34.00, 139.75)	189.00 (105.50, 363.50)	<0.001
ALB (g/L), (Median [Q1, Q3])	40.05 (36.73, 43.55)	36.80 (33.70, 40.45)	<0.001
TBIL (μmol/L), (Median [Q1, Q3])	16.40 (12.60, 23.20)	21.00 (14.50, 30.15)	<0.001
Scr (μmol/L), (Median [Q1, Q3])	67.85 (57.75, 76.38)	67.40 (57.55, 76.90)	0.780
PT (sec), (Median [Q1, Q3])	13.80 (13.10, 14.30)	13.90 (13.40, 14.70)	0.026
INR (Median [Q1, Q3])	1.07 (1.01, 1.13)	1.09 (1.03, 1.16)	0.014

Abbreviations: TACE, transarterial chemoembolization; HBV, hepatitis B virus; DM, diabetes mellitus; PLT, platelet count; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, glutamyl transpeptidase; TBIL, total bilirubin; Scr, serum creatinine; PT, prothrombin time; INR, international normalized ratio; ECOG-PS, Eastern Cooperative Oncology Group performance status; ALB, albumin; ALBI, albumin-bilirubin; BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma.

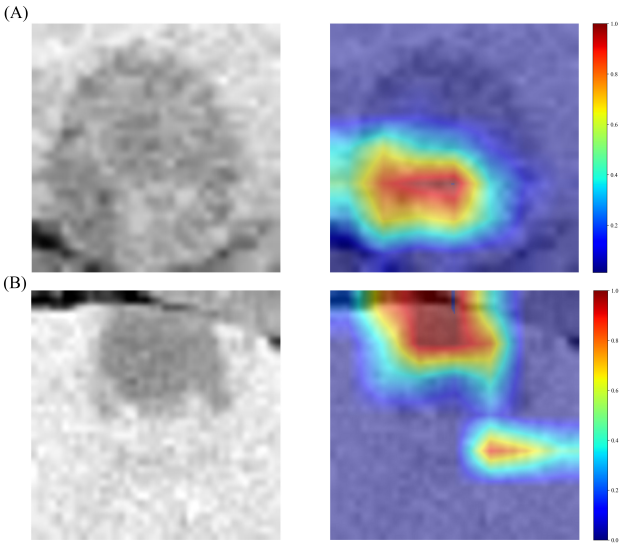

Fig. 2. Grad-CAM visualizations for two representative samples. These visualizations are instrumental in demonstrating how the model focuses on different regions of the images for making predictions. (A) Grad-CAM visualization for a high-risk HCC patient. (B) Grad-CAM visualization for a low-risk HCC patient.

3.5 Development and Validation of Models for Two Treatments Separately

We developed six models to assess the risks of overall survival for each treatment group individually. Performance metrics for these models in the training and testing cohorts are shown in **Supplementary Tables 5,6**. The combined model consistently outperforms the individual Clinical and DLRadiomics models, particularly in the testing cohort, where their C-index values closely match those of the training cohort.

For the hepatectomy group, the combined model yielded a C-index of 0.836 (95% CI: 0.776–0.897) in training and 0.861 (95% CI: 0.755–0.967) in the testing cohorts at 1-year points. Likewise, the combined model for the TACE group achieved a C-index of 0.840 (95% CI: 0.792–0.888) for the training cohort and 0.834 (95% CI: 0.759–0.910) for the testing cohort. Time-dependent receiver operating characteristic curve further demonstrated that the combined model outperforms other models across multiple time points in both the training and testing cohorts. Additional findings are shown in **Supplementary Figs. 5,6**.

In the training cohort, patients were divided into low- and high-risk groups using a cutoff value of 1.03, based on the median risk score. In the validation cohort, a similar classification was conducted using a median cutoff of 1.25. Kaplan-Meier curves (Figs. 3,4) indicated a significant distinction in the overall survival between the low- and high-risk subgroups in both cohorts ($p < 0.05$). Based on these results, two prognostic nomograms were constructed to predict survival outcomes for patients undergoing TACE and hepatectomy, respectively (Fig. 5).

4. Discussion

In this study, the performance of three deep learning algorithms, ResNet18, ResNet50, and DenseNet121, was compared in predicting the prognosis of HCC patients using contrast-enhanced CT images. Among these models, ResNet50 showed the best predictive performance. Deep learning features extracted from ResNet50 were then combined with clinical variables to develop nomograms for predicting the prognosis in patients undergoing hepatectomy or TACE. These nomograms provided recommendations for optimal treatment selection.

Compared to the study by Fu *et al*. [20], which combined traditional radiomics features with clinical data to predict progression-free survival (PFS) differences between hepatectomy and TACE, our approach leverages DLRadiomics to automatically extract high-dimensional imag-

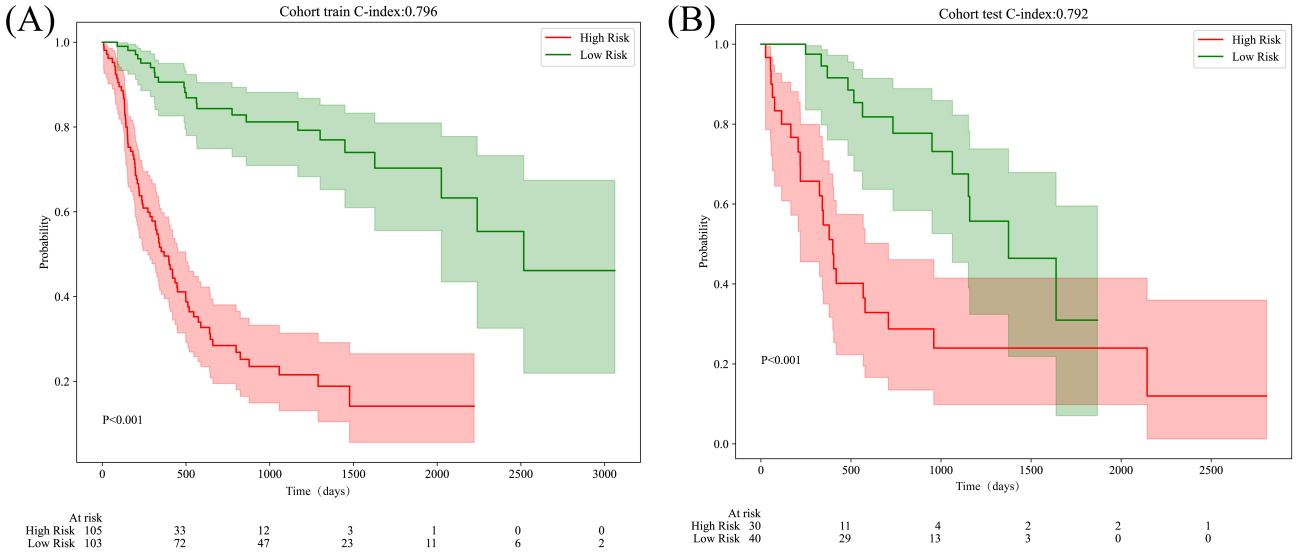

Fig. 3. Kaplan-Meier Curves of survival risk based on combined models treated with hepatectomy. (A) Kaplan-Meier curves for the training cohort treated with hepatectomy. The green line represents the low-risk subgroup, and the red line represents the high-risk subgroup, as predicted by the combined model. (B) Kaplan-Meier curves for the testing cohort treated with hepatectomy. The green line represents the low-risk subgroup, and the red line represents the high-risk subgroup, as predicted by the combined model.

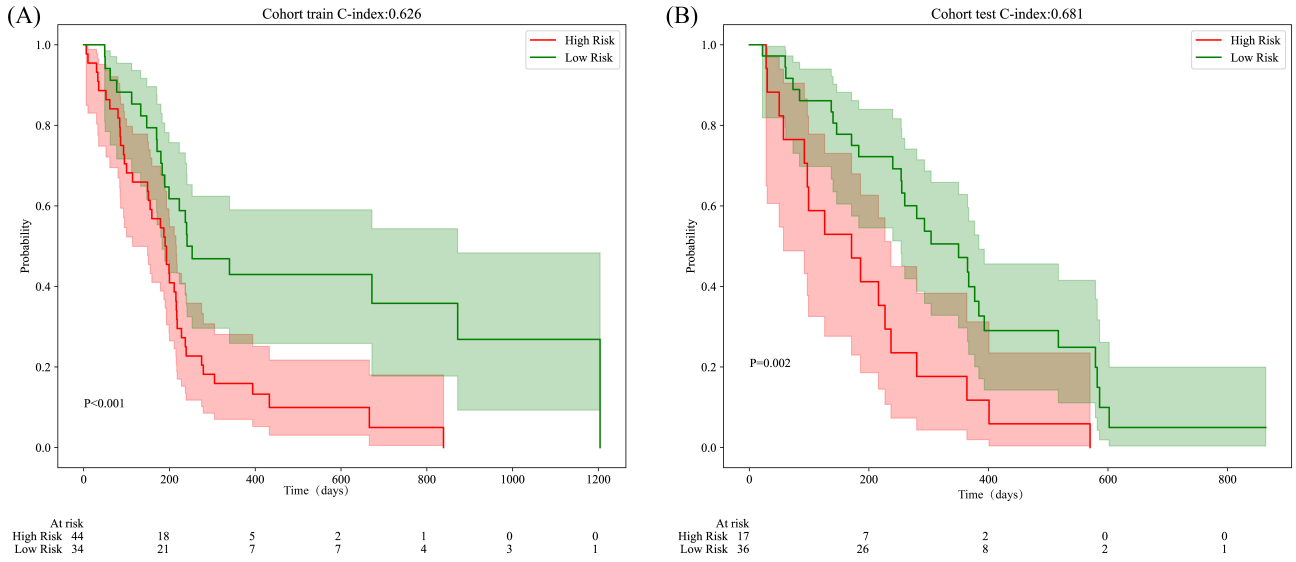

Fig. 4. Kaplan-Meier Curves of survival risk associated with combined models, which were treated with TACE. (A) Kaplan-Meier curves for the training cohort treated with TACE. The green line represents the low-risk subgroup, and the red line represents the high-risk subgroup, as predicted by the combined model. (B) Kaplan-Meier curves for the testing cohort treated with TACE. The green line represents the low-risk subgroup, and the red line represents the high-risk subgroup, as predicted by the combined model.

ing features. This approach overcomes the limitations of manual feature engineering while efficiently capturing tumor heterogeneity and biological behavior. Our combined model achieved a C-index of 0.861 in the testing cohort for predicting the risk of survival in the hepatectomy group, significantly outperforming the C-index of 0.75 observed in Fu *et al.*'s study [20]. Furthermore, the nomograms developed in our study provide an intuitive and individualized scoring system to facilitate treatment decision-making, thereby streamlining clinical workflows and enhancing practical utility.

Treatment guidelines for patients with HCC are often not followed, with approximately 50% of patients deviating from the recommended treatments [21,22]. Existing guidelines, along with clinical and laboratory indicators, may have limited roles in guiding the choice between hepatectomy or TACE therapy [23]. A previous study by Li *et al.* [24] suggested treatment recommendations based on imag-

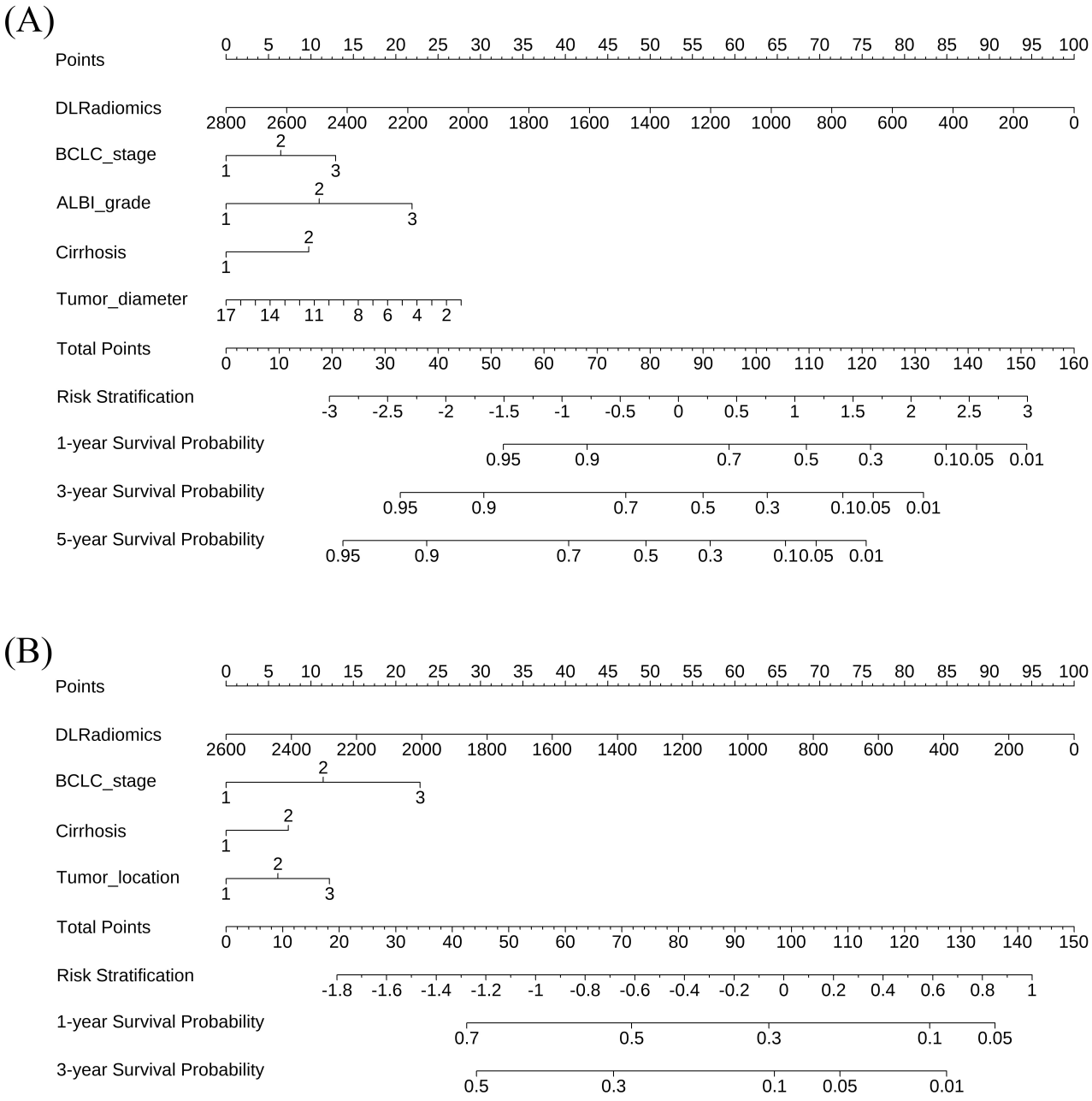

Fig. 5. Constructed nomograms to predict the prognosis of patients with different treatment regimens using DLRadiomics and clinical features. (A) The nomogram, when used in conjunction with hepatectomy, identifies three risk factors associated with the risk of survival. (B) The nomogram, when used in conjunction with TACE, identifies four risk factors associated with the risk of survival.

ing texture characteristics for patients with a single large HCC, though its conclusions were limited by a small sample size. Given this, our study expanded the sample size and compared multiple deep learning models to create a robust prognostic model for patients undergoing hepatectomy or TACE. This strategy aims to help clinicians in formulating tailored treatment decisions based on prognostic differences and individual patient characteristics.

Traditional radiomics methods often rely on manually designed algorithms and predefined rules to extract imaging features [25]. In contrast, DLRadiomics can automatically learn the most representative features directly from raw medical imaging data [26]. For instance, a deep learning model trained on a large-scale lung CT dataset can autonomously detect characteristic patterns associated with lung diseases, such as identifying texture changes along nodule edges, without relying on human-defined rules [27–29]. Utilizing multi-level neural network architectures, deep learning is capable of capturing intricate and subtle details within images [30]. This is crucial for accurate disease

diagnosis, staging, prognosis evaluation, and other clinical assessments, thereby enhancing the precision of radiomics analyses [31]. Hence, we integrated deep learning into the radiomics procedure to automate and optimize feature extraction. The fusion of deep learning with radiomics combines the strengths of both approaches, enabling the efficient and accurate detection of minute, otherwise undetectable features in medical images. This model provides strong support for a range of clinical utilities, including disease diagnosis, treatment response prediction, and outcome prognosis [32,33].

Despite several promising results, we acknowledge some limitations in our study. Firstly, it is a single-center retrospective study, and thus, external validation datasets is needed to evaluate the generalizability of the model. Secondly, the sample size of the single-center study may be relatively limited and given the complexity of causal inference and treatment selection, future studies using larger datasets are necessary to validate our results. Thirdly, the model established in this study is only an auxiliary tool and should not be used as a standalone decision-making algorithm. It must be combined with other clinical evaluation metrics to facilitate comprehensive treatment planning. Fourthly, patients were categorized into high-risk and low-risk groups based on a 3-year survival threshold. However, the number of TACE patients who survived beyond three years was small, and variations in treatment modalities across different disease stages may have impacted the outcomes. Future studies should collect an extensive dataset and conduct stratified analyses based on disease stage to facilitate more robust comparisons of treatment strategies. Lastly, the study did not take into account patients' subsequent treatment, such as additional hepatectomy or TACE procedures. In clinical practice, follow-up treatment decisions are affected by several factors, including treatment response, tumor progression, liver function, and overall patient condition. Our future investigation would collect multicenter data, develop a user-friendly interface to support clinical decision-making, and establish evidence-based treatment guidelines to enhance support for clinical diagnosis. Additionally, we aim to optimize our model by incorporating feedback on clinical applicability, clinician adoption, data integration complexity, and other relevant implementation factors.

5. Conclusion

In conclusion, this study established a prognostic model using deep learning-based features to predict risks of survival in HCC patients. Furthermore, we devised personalized models to evaluate prognostic differences between patients receiving hepatic resection and those undergoing TACE. These models can assist clinicians in making informed, personalized treatment decisions for HCC patients.

Key Points

- DLRadiomics, employing ResNet50 on contrast-enhanced CT images, outperformed traditional radiomics by achieving the highest AUCs in both training and testing cohorts.
- This study developed and validated survival models that combine clinical data with DLRadiomics features to predict risks of overall survival (OS) in hepatocellular carcinoma (HCC) patients treated with hepatectomy or TACE.
- DLRadiomics enhances traditional radiomics by automatically extracting features from medical images, thereby improving prognostic accuracy and offering data-driven, personalized treatment approaches for hepatocellular carcinoma (HCC).
- This strategy offers clinicians practical tools for optimizing treatment decisions, marking a significant advancement in precision oncology.

Availability of Data and Materials

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

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by XQW and XWC. The first draft of the manuscript was written by XQW and XWC. All authors contributed to revising the manuscript critically for important intellectual content. JX, YL, CFY, ZRL, and MYT provided key inputs during the experimental design, data interpretation, and final revision stages. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study design adhered to the principles outlined in the Declaration of Helsinki, and was approved by the Ethics Committee of Affiliated Hospital of North Sichuan Medical College (Approval No. 2022ER059-2). As this study analyzed anonymized data from hospital records, the requirement for obtaining informed consent was waived by the Ethics Committee of Affiliated Hospital of North Sichuan Medical College.

Acknowledgement

We would like to express our sincerest gratitude to all those who provided invaluable support and assistance in the completion of this research project.

Funding

This study has received funding by the Natural Science Foundation of Science and Technology Department of Sichuan Province (2023NSFSC0646), Education Information Technology Research of Sichuan Province (No.DSJ2022266).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/BJHM50380.

References

[1] 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: A Cancer Journal for Clinicians. 2021; 71: 209–249. https://doi.org/10.3322/caac.21660.

[2] Xie D, Shi J, Zhou J, Fan J, Gao Q. Clinical practice guidelines and real-life practice in hepatocellular carcinoma: A Chinese perspective. Clinical and Molecular Hepatology. 2023a; 29: 206–216. https://doi.org/10.3350/cmh.2022.0402.

[3] Trevisani F, Vitale A, Kudo M, Kulik L, Park JW, Pinato DJ, *et al*. Merits and boundaries of the BCLC staging and treatment algorithm: Learning from the past to improve the future with a novel proposal. Journal of Hepatology. 2024; 80: 661–669. https://doi.org/10.1016/j.jhep.2024.01.010.

[4] Vitale A, Romano P, Cillo U, Lauterio A, Sangiovanni A, Cabibbo G, *et al*. Liver Resection vs Nonsurgical Treatments for Patients With Early Multinodular Hepatocellular Carcinoma. JAMA Surgery. 2024; 159: 881–889. https://doi.org/10.1001/jamasurg.2024.1184.

[5] Bai Z, Yu X, Tang Q, Zhang R, Shi X, Liu C. Hepatic arterial infusion chemotherapy combined with lenvatinib and PD-1 inhibitor for treating unresectable hepatocellular carcinoma. British Journal of Hospital Medicine. 2024; 85: 1–12. https://doi.org/10.12968/hmed.2024.0159.

[6] Han JE, Cho HJ, Cheong JY, Lim SG, Yang MJ, Noh CK, *et al*. Impact of guideline adherence on the prognosis of Barcelona clinic liver cancer stage B hepatocellular carcinoma. World Journal of Gastroenterology. 2023; 29: 6122–6137. https://doi.org/10.3748/wjg.v29.i47.6122.

[7] Jamtani I, Lalisang TJM, Mulyawan W. Effect of neoadjuvant transarterial chemoembolization followed by resection versus upfront liver resection on the survival of single large hepatocellular carcinoma patients: A systematic review and meta-analysis. Annals of Hepato-Biliary-Pancreatic Surgery. 2024; 28: 325–336. https://doi.org/10.14701/ahbps.24-009.

[8] Schultheiß M, Bengsch B, Thimme R. Hepatocellular Carcinoma. Deutsche Medizinische Wochenschrift (1946). 2021; 146: 1411–1420. https://doi.org/10.1055/a-1226-3047. (In German)

[9] Yang J, Choi WM, Lee D, Shim JH, Kim KM, Lim YS, *et al*. Outcomes of liver resection and transarterial chemoembolization in patients with multinodular BCLC-A hepatocellular carcinoma. Journal of Liver Cancer. 2024; 24: 178–191. https://doi.org/10.17998/jlc.2024.03.25.

[10] Masch WR, Kampalath R, Parikh N, Shampain KA, Aslam A, Chernyak V. Imaging of treatment response during systemic therapy for hepatocellular carcinoma. Abdominal Radiology. 2021; 46: 3625–3633. https://doi.org/10.1007/s00261-021-03100-0.

[11] Zang Y, Long P, Wang M, Huang S, Chen C. Development and validation of prognostic nomograms in patients with hepatocellular carcinoma: a population-based study. Future Oncology. 2021; 17: 5053–5066. https://doi.org/10.2217/fon-2020-1065.

[12] Li X, Xu Y, Ou Y, Li H, Xu W. Optimizing Treatment Selection for Early Hepatocellular Carcinoma Based on Tumor Biology, Liver Function, and Patient Status. Journal of Hepatocellular Carcinoma. 2025; 12: 777–790. https://doi.org/10.2147/JHC.S514248.

[13] Labgaa I, Taffé P, Martin D, Clerc D, Schwartz M, Kokudo N, *et al*. Comparison of Partial Hepatectomy and Transarterial Chemoembolization in Intermediate-Stage Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. Liver Cancer. 2020; 9: 138–147. https://doi.org/10.1159/000505093.

[14] Wen L, Weng S, Yan C, Ye R, Zhu Y, Zhou L, *et al*. A Radiomics Nomogram for Preoperative Prediction of Early Recurrence of Small Hepatocellular Carcinoma After Surgical Resection or Radiofrequency Ablation. Frontiers in Oncology. 2021; 11: 657039. https://doi.org/10.3389/fonc.2021.657039.

[15] Zhong L, Dong D, Fang X, Zhang F, Zhang N, Zhang L, *et al*. A deep learning-based radiomic nomogram for prognosis and treatment decision in advanced nasopharyngeal carcinoma: A multicentre study. eBioMedicine. 2021; 70: 103522. https://doi.org/10.1016/j.ebiom.2021.103522.

[16] Braghetto A, Marturano F, Paiusco M, Baiesi M, Bettinelli A. Radiomics and deep learning methods for the prediction of 2-year overall survival in LUNG1 dataset. Scientific Reports. 2022; 12: 14132. https://doi.org/10.1038/s41598-022-18085-z.

[17] Minoda Y, Ihara E, Komori K, Ogino H, Otsuka Y, Chinen T, *et al*. Efficacy of endoscopic ultrasound with artificial intelligence for the diagnosis of gastrointestinal stromal tumors. Journal of Gastroenterology. 2020; 55: 1119–1126. https://doi.org/10.1007/s00535-020-01725-4.

[18] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. Journal of Hepatology. 2018; 69: 182–236. https://doi.org/10.1016/j.jhep.2018.03.019.

[19] Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, *et al*. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology. 2018; 67: 358–380. https://doi.org/10.1002/hep.29086.

[20] Fu S, Wei J, Zhang J, Dong D, Song J, Li Y, *et al*. Selection Between Liver Resection Versus Transarterial Chemoembolization in Hepatocellular Carcinoma: A Multicenter Study. Clinical and Translational Gastroenterology. 2019; 10: e00070. https://doi.org/10.14309/ctg.0000000000000070.

[21] Brown ZJ, Tsilimigras DI, Ruff SM, Mohseni A, Kamel IR, Cloyd JM, *et al*. Management of Hepatocellular Carcinoma: A Review. JAMA Surgery. 2023; 158: 410–420. https://doi.org/10.1001/jamasurg.2022.7989.

[22] Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. Nature Reviews. Clinical Oncology. 2023; 20: 864–884. https://doi.org/10.1038/s41571-023-00825-3.

[23] Xie DY, Zhu K, Ren ZG, Zhou J, Fan J, Gao Q. A review of 2022 Chinese clinical guidelines on the management of hepatocellular carcinoma: updates and insights. Hepatobiliary Surgery and Nutrition. 2023b; 12: 216–228. https://doi.org/10.21037/hbsn-22-469.

[24] Li M, Fu S, Zhu Y, Liu Z, Chen S, Lu L, *et al*. Computed tomography texture analysis to facilitate therapeutic decision making in hepatocellular carcinoma. Oncotarget. 2016; 7: 13248–13259. https://doi.org/10.18632/oncotarget.7467.

[25] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More

than Pictures, They Are Data. Radiology. 2016; 278: 563–577. https://doi.org/10.1148/radiol.2015151169.

[26] Xie H, Lei Y, Wang T, Tian Z, Roper J, Bradley JD, *et al*. High through-plane resolution CT imaging with self-supervised deep learning. Physics in Medicine and Biology. 2021; 66: 145013. https://doi.org/10.1088/1361-6560/ac0684.

[27] Lin CY, Guo SM, Lien JJJ, Lin WT, Liu YS, Lai CH, *et al*. Combined model integrating deep learning, radiomics, and clinical data to classify lung nodules at chest CT. La Radiologia Medica. 2024; 129: 56–69. https://doi.org/10.1007/s11547-023-01730-6.

[28] Mikhael PG, Wohlwend J, Yala A, Karstens L, Xiang J, Takigami AK, *et al*. Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk From a Single Low-Dose Chest Computed Tomography. Journal of Clinical Oncology. 2023; 41: 2191–2200. https://doi.org/10.1200/JCO.22.01345.

[29] Zhong Y, Cai C, Chen T, Gui H, Deng J, Yang M, *et al*. PET/CT based cross-modal deep learning signature to predict occult nodal metastasis in lung cancer. Nature Communications. 2023; 14: 7513. https://doi.org/10.1038/s41467-023-42811-4.

[30] Serghiou S, Rough K. Deep Learning for Epidemiologists: An Introduction to Neural Networks. American Journal of Epidemiology. 2023; 192: 1904–1916. https://doi.org/10.1093/aje/kwad107.

[31] Rami-Porta R, Nishimura KK, Giroux DJ, Detterbeck F, Cardillo G, Edwards JG, *et al*. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groups in the Forthcoming (Ninth) Edition of the TNM Classification for Lung Cancer. Journal of Thoracic Oncology. 2024; 19: 1007–1027. https://doi.org/10.1016/j.jtho.2024.02.011.

[32] Jiang Y, Zhou K, Sun Z, Wang H, Xie J, Zhang T, *et al*. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. Cell Reports. Medicine. 2023; 4: 101146. https://doi.org/10.1016/j.xcrm.2023.101146.

[33] Xia T, Zhao B, Li B, Lei Y, Song Y, Wang Y, *et al*. MRI-Based Radiomics and Deep Learning in Biological Characteristics and Prognosis of Hepatocellular Carcinoma: Opportunities and Challenges. Journal of Magnetic Resonance Imaging. 2024; 59: 767–783. https://doi.org/10.1002/jmri.28982.