Name of Journal: *World Journal of Radiology*
Manuscript NO: 121161
Manuscript Type: ORIGINAL ARTICLE

Retrospective Study

Computed tomography radiomics-based machine learning nomogram for preoperative prediction of glypican-3 expression in hepatocellular carcinoma

Zi-Han Zheng, Chun-Hua Wu, Jin-Bo Hu, Jun-Feng Xu, Xin-Yue Zi, Jin-Hui Chen, Qin He, Wen-Ying Dong

Zi-Han Zheng, Department of Radiology, The Second People's Hospital of Yuxi City, Yuxi 653100, Yunnan Province, China

Chun-Hua Wu, Jin-Bo Hu, Jun-Feng Xu, Xin-Yue Zi, Jin-Hui Chen, Qin He, Department of Radiology, The First Affiliated Hospital of Dali University, Dali 671000, Yunnan Province, China

Wen-Ying Dong, Internal Medicine Teaching and Research, College of Nursing of Dali University, Dali 671003, Yunnan Province, China

Co-first authors: Zi-Han Zheng and Chun-Hua Wu.

Author contributions: Zheng ZH and Wu CH were responsible for investigation, writing – original draft, formal analysis, methodology, software, writing – original draft as co-first authors; Hu JB was responsible for formal analysis, methodology, software, writing – review and editing; Xu JF was responsible for funding acquisition, investigation, supervision, validation, writing – review and editing; Zi XY, Chen JH, and He Q were responsible for data curation, investigation, writing –

review and editing; Dong WY was responsible for conceptualization, investigation, project administration, resources, writing – review and editing.

Supported by the Key Research and Development Program in the Field of Biomedicine of the Dali Science and Technology Bureau, No. 20252903B030003.

Corresponding author: Wen-Ying Dong, Director, Internal Medicine Teaching and Research, College of Nursing of Dali University, No. 22 Wanhua Road, Dali 671003, Yunnan Province, China. 157273210@qq.com

Received: March 23, 2026
Revised: April 16, 2026

ACCEPTED

Abstract

BACKGROUND

Hepatocellular carcinoma (HCC) is among the most common and fatal primary liver malignancies. Glypican-3 (GPC3) is a useful biomarker for HCC diagnosis and targeted therapy, but reliable noninvasive approaches for predicting GPC3 expression before surgery remain limited.

AIM

To develop and validate a computed tomography (CT)-based radiomics model using machine learning for the preoperative prediction of GPC3 expression in HCC.

METHODS

This retrospective study included 103 patients with pathologically confirmed HCC who underwent contrast-enhanced CT at two centers between January 2013 and October 2023. Patients were assigned to training (n = 72) and testing (n = 31) sets using a 7:3 stratified random sampling. Regions of interest were manually delineated on non-contrast, arterial, portal venous, and delayed-phase images, followed by radiomic feature extraction. After feature selection, radiomics models were constructed using eight machine learning algorithms. Independent clinical predictors were identified by univariate and multivariate logistic regression and used to construct a clinical model. A combined model was then developed by integrating the optimal radiomics model with the clinical predictors, and a corresponding nomogram was generated. Model performance was evaluated using the area under the curve (AUC), DeLong test, net reclassification improvement, and integrated discrimination improvement. Calibration curves and decision curve analysis were used to assess the clinical utility of the nomogram.

RESULTS

Alpha-fetoprotein and total bilirubin were independent clinical predictors of GPC3 expression. After feature selection, 10 radiomic features were retained. Among the radiomics models, the random forest classifier showed the strongest predictive

ACCEPTED

performance, with an AUC of 0.959 in the training set and 0.862 in the testing set. Integration of the radiomics signature with alpha-fetoprotein and total bilirubin further improved model performance, yielding AUCs of 0.979 and 0.948 in the training and testing sets, respectively. In both cohorts, the combined model showed positive reclassification gains, with net reclassification improvement > 0 and integrated discrimination improvement > 0. Calibration curves suggested a better goodness-of-fit for the clinical model. Decision curve analysis indicated that the nomogram derived from the combined model provided greater net clinical benefit across a broad range of threshold probabilities.

CONCLUSION

A nomogram combining a random forest-based CT radiomics model with clinical predictors showed strong performance and potential clinical value for preoperative prediction of GPC3 expression in HCC.

Key Words: Hepatocellular carcinoma; Glypican-3; Radiomics; Machine learning; Nomogram

Zheng ZH, Wu CH, Hu JB, Xu JF, Zi XY, Chen JH, He Q, Dong WY. Computed tomography radiomics-based machine learning nomogram for preoperative prediction of glypican-3 expression in hepatocellular carcinoma. *World J Radiol* 2026; In press

DOI: https://doi.org/10.4329/wjr.121161

Core Tip: Glypican-3 is a clinically relevant biomarker in hepatocellular carcinoma, but its expression is usually confirmed only after biopsy or surgical resection. This study developed a computed tomography radiomics-based machine learning nomogram for the preoperative, noninvasive prediction of glypican-3 expression. The combined model integrated a random forest-derived radiomics signature with alpha-fetoprotein and total bilirubin and showed strong predictive performance,

ACCEPTED

with evidence of clinical utility. This imaging-based approach may support preoperative risk stratification and individualized management in patients with hepatocellular carcinoma.

INTRODUCTION

Primary liver cancer is among the most common malignancies in China, and hepatocellular carcinoma (HCC) is its predominant histological subtype[1,2], accounting for 75%-85% of all primary liver cancers[3]. HCC often develops insidiously and remains asymptomatic in the early stages. As a result, many patients are diagnosed at an intermediate or advanced stage, when curative surgical resection is no longer feasible[4]. The identification of specific biomarkers and molecular targets is therefore important for improving early diagnosis and supporting the development of targeted therapies for HCC[5].

Glypican-3 (GPC3), a member of the glypican family consisting of heparan sulfate chains attached to a core protein, participates in several signaling pathways and contributes to the regulation of tumor proliferation and metastasis[6]. GPC3 is highly expressed in HCC tissues but is absent from normal liver tissue. Its overexpression has also been associated with poor prognosis in patients with HCC[5], supporting its value as a biomarker for HCC detection. But preoperative assessment of GPC3 expression remains limited, and currently available approaches do not provide sufficient accuracy[7].

Advances in computational methods have increased the use of radiomics for disease diagnosis and treatment evaluation. In HCC, a previous study used computed tomography (CT)-derived radiomic features to evaluate GPC3 expression and reported moderate predictive performance, with area under the curves (AUCs) of 0.842 in the training cohort and 0.726 in the testing cohort[7]. But studies that integrate multiple machine learning algorithms for preoperative prediction of GPC3 expression remain limited. The present study therefore applied several machine learning algorithms to develop radiomics-based predictive models.

Machine learning methods, including support vector machines (SVMs), random forests (RFs), k-nearest neighbors (KNNs), and logistic regression (LR), can identify

ACCEPTED

patterns within high-dimensional data and generate predictive classifications. These approaches have been used in clinical settings to improve diagnostic accuracy and prognostic assessment across a range of diseases[8]. In oncology, their capacity to process large radiomic feature sets and build stable predictive models has been reported in multiple applications[9]. Although machine learning has increasingly been used in liver cancer research, most existing studies on GPC3 prediction have relied mainly on CT-based or magnetic resonance imaging (MRI)-based radiomics, without systematically comparing or incorporating a broader machine learning framework.

Because contrast-enhanced abdominal CT is reproducible, noninvasive, and widely available in routine clinical practice, it remains central imaging modality for HCC evaluation.

Therefore, the present study aimed to develop machine learning–based radiomics models using contrast-enhanced CT to preoperatively assess GPC3 expression status, thereby providing clinically relevant guidance for individualized treatment planning.

ACCEPTED

MATERIALS AND METHODS

Study population

This study was approved by the Institutional Review Board of our hospital (IRB approval No. DFY20250120001) and conducted in accordance with the Declaration of Helsinki. The requirement for written informed consent was waived due to the study's retrospective design.

We retrospectively reviewed clinical and imaging data of patients with pathologically confirmed HCC from the First Affiliated Hospital of Dali University and Dali First People's Hospital between January 2013 and October 2023.

Inclusion criteria: (1) Postoperative pathological diagnosis of HCC with available immunohistochemical results for GPC3 (GPC3 positivity defined as ≥ 5% GPC3-immunoreactive tumor cells within the same microscopic field[7,10]; (2) Upper-abdominal or whole-abdominal non-contrast and contrast-enhanced CT performed

within one month before surgery; and (3) A single HCC lesion identifiable on CT.

Exclusion criteria: (1) Prior treatments, including biopsy, transcatheter arterial chemoembolization, radiotherapy, or partial hepatectomy; (2) Lesions too small to allow clear delineation, potentially affecting the region of interest (ROI) segmentation, feature extraction, and image evaluation; and (3) Incomplete clinical data.

Ultimately, 103 patients met the inclusion criteria, including 73 GPC3-positive and 30 GPC3-negative cases.

CT examination protocol

All patients underwent abdominal CT, both non-contrast and contrast-enhanced. Scanners included a Philips 16-slice CT, Siemens 64-slice CT, and third-generation Siemens dual-source CT at the First Affiliated Hospital of Dali University, and a Siemens 64-slice CT at Dali First People's Hospital. The acquisition parameters were as follows: (1) Tube voltage, 120 kV; (2) Slice thickness, 5 mm; and (3) Interslice spacing, 5 mm.

Contrast-enhanced CT was performed using a dual-barrel high-pressure injector. Iohexol (50 mL containing 15 g of iodine; Yangtze River Pharmaceutical Group) was injected *via* the antecubital vein at a rate of 5-7 mL/second, followed by an equal volume of saline flush. Arterial phase, portal venous phase, and delayed phase images were acquired at 40 seconds, 70 seconds, and 90 seconds after injection, respectively.

Clinical data and CT imaging features

GPC3 expression status was assessed by a senior pathologist using the ≥ 5% positivity threshold. Collected clinical variables included sex, age, viral hepatitis status, total bilirubin (TBIL), direct bilirubin, indirect bilirubin, alanine aminotransferase, aspartate aminotransferase, total protein, albumin, globulin, alpha-fetoprotein (AFP), and carcinoembryonic antigen.

Two radiologists independently evaluated the CT features using the Picture

ACCEPTED

Archiving and Communication System. Interobserver agreement for imaging features was assessed using the kappa statistic.

The following CT features were recorded: (1) Liver cirrhosis: Changes in overall liver size; disproportional enlargement or atrophy of liver lobes; nodular hepatic surface; loss of normal hepatic architecture; diffuse or heterogeneous low attenuation; and widened fissures; (2) Ascites: Presence of intraperitoneal fluid; and (3) Portal vein tumor thrombus: Hypo-attenuating or iso-attenuating nodular, strip-like, or irregular filling defects within the portal vein on portal venous phase images.

Radiomics analysis

Non-contrast, arterial phase, portal venous phase, and delayed phase CT images were exported from the Picture Archiving and Communication System in Digital Imaging and Communications in Medicine (DICOM) format and imported into 3D Slicer (https://www.slicer.org/). Two radiologists manually delineated two-dimensional regions of interest (ROIs) slice-by-slice along the tumor boundary in all four image phases. These ROIs were then automatically merged into a three-dimensional volume of interest (VOI) (Figure 1). In cases of disagreement, a senior radiologist adjudicated the final VOI. The final VOIs were exported in NIfTI format for further analysis.

ACCEPTED

Radiomics feature standardization and extraction

The entire dataset was randomly split into a training set and a test set at a 7:3 ratio. The training set was used to develop the predictive models, whereas the testing set was used for external performance evaluation. To mitigate the effects of different scanners and acquisition parameters on radiomic features, all CT images were resampled to an isotropic voxel size of $1 \times 1 \times 1$ mm^3 for standardization.

Radiomic feature extraction was subsequently performed. Interobserver consistency of features extracted by the two radiologists was evaluated using the intraclass correlation coefficient. Features with intraclass correlation coefficient < 0.75 were excluded. For subsequent analysis, features extracted by either radiologist were randomly selected. A total of 833 features were extracted from each CT phase

(non-contrast, arterial, portal venous, and delayed), resulting in 3332 radiomic features. The extracted features included shape descriptors, first-order statistics, and texture features.

Feature selection and model construction

All radiomic features were initially normalized by *Z*-score standardization, giving each feature a mean of 0 and a variance of 1. Preliminary screening was then performed with the Mann-Whitney *U* test, and features with $P < 0.05$ were retained. To limit redundancy among highly correlated variables, Spearman correlation coefficients were calculated. For feature pairs with a correlation coefficient > 0.9, one feature was removed through an iterative procedure that preferentially eliminated the more redundant variable.

Final feature selection was carried out using the least absolute shrinkage and selection operator with 10-fold cross-validation. The optimal penalty parameter (λ) was selected according to the minimum cross-validation error. Features with non-zero coefficients were retained, resulting in a final panel of 10 radiomic features.

Radiomics models were then developed using eight machine learning classifiers: LR, SVM, KNN, RF, extra trees (ET), extreme gradient boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and multilayer perceptron (MLP) neural network.

A radiomics score (Radscore) was calculated for each model using the formula: $\text{Radscore} = I + \beta_1 \times W_1 + \beta_2 \times W_2 + \beta_3 \times W_3 + \ldots$, where I represents the intercept, β the feature coefficients, and W the corresponding selected radiomic features.

Statistical analysis

Statistical analyses were performed using SPSS version 26.0 and R version 3.6.1. Normally distributed continuous variables were expressed as the mean ± SD and compared using the independent-samples *t*-test. Non-normally distributed continuous variables were expressed as median and interquartile range and compared using the Mann-Whitney *U* test or the Wilcoxon signed-rank test. Categorical variables were presented as *n* (%) and compared using the Pearson χ^2

ACCEPTED

test or Fisher's exact test.

Receiver operating characteristic (ROC) curves were generated, and the AUC, sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were calculated to evaluate model performance. Delong's test was used to compare AUCs between radiomics models. Net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were used to assess the incremental predictive value of each model.

The optimal radiomics model was integrated with the clinical model to construct a radiomics-clinical combined model. A nomogram was generated based on this combined model. The Hosmer-Lemeshow test was applied to evaluate the goodness-of-fit of the clinical model, the optimal radiomics model, and the nomogram in both the training and testing sets. Calibration curves were plotted to assess the calibration performance of these models, and decision curve analysis (DCA) was performed to evaluate the clinical utility of the nomogram.

ACCEPTED

RESULTS

Clinical and CT imaging characteristics of patients

In this study, 103 patients met the inclusion criteria, including 73 GPC3-positive and 30 GPC3-negative cases. Among them, 58 patients were from the First Affiliated Hospital of Dali University [GPC3 (+) 41; GPC3 (-) 17], and 45 were from Dali First People's Hospital [GPC3 (+) 32; GPC3 (-) 13]. Patients were randomly assigned to a training set (n = 72) and a testing set (n = 31) using stratified sampling at a 7:3 ratio.

All CT features evaluated by the two radiologists demonstrated substantial interobserver agreement, with Kappa coefficients > 0.70, indicating high consistency. The clinical and imaging characteristics of the training and testing cohorts are summarized in Table 1. Between the two cohorts, alanine aminotransferase and aspartate aminotransferase levels differed significantly (P < 0.05). Between the GPC3-positive and GPC3-negative groups, TBIL and AFP showed statistically significant differences in the training cohort (P < 0.05), whereas no variables differed significantly in the testing cohort (P > 0.05).

Radiomics feature analysis and radiomics model construction radiomics feature selection

After multi-step feature selection, 3, 1, 2, and 4 optimal radiomic features were retained from the non-contrast, arterial-phase, portal-venous-phase, and delayed-phase CT images, respectively. The selected features and their relative importance are presented in Table 2 and Figure 2.

Construction and evaluation of radiomics models

The selected radiomics features were used to construct eight machine learning-based radiomics models using LR, SVM, KNN, RF, ET, XGBoost, LightGBM, and MLP classifiers, and their predictive performance was subsequently validated in the testing cohort. The ROC curves for all models are presented in Figure 3, and the corresponding diagnostic metrics, including AUC, sensitivity, specificity, accuracy, positive predictive value, and negative predictive value, for both the training and testing cohorts are summarized in Table 3. DeLong test results (Figure 4) indicated that in the training cohort, significant AUC differences ($P < 0.05$) were observed among several model pairs, including LR *vs* SVM, ET, and XGBoost; SVM *vs* KNN, XGBoost, and MLP; KNN *vs* RF, ET, and XGBoost; RF *vs* XGBoost; and XGBoost *vs* LightGBM and MLP, while all remaining pairwise comparisons showed no significant differences ($P > 0.05$). In contrast, no significant AUC differences were detected among the models in the testing cohort (all $P > 0.05$). Overall, the RF-based radiomics model achieved the highest AUCs in both cohorts and demonstrated superior predictive performance compared with the other radiomics models.

Construction and evaluation of the clinical model

Univariate and multivariate LR analyses were performed in the training cohort (Table 4). AFP and TBIL were identified as independent predictors of GPC3 expression ($P < 0.05$). Based on these variables, a clinical model was constructed using the RF classifier (Table 5).

Construction and evaluation of the combined model

ACCEPTED

A combined model integrating the optimal radiomics model (RF radiomics model) and the clinical model was developed. ROC curves for the clinical, RF radiomics, and combined models are presented in Figure 5, and their corresponding diagnostic metrics are summarized in Table 5.

As shown in Figure 6, DeLong test results indicated that, in the training cohort, the combined model achieved significantly higher AUC values than the clinical model ($P < 0.05$). The RF radiomics model also significantly outperformed the clinical model ($P < 0.05$), whereas no significant difference was observed between the RF radiomics model and the combined model ($P > 0.05$). In the testing cohort, no pairwise AUC comparisons among the three models were statistically significant ($P > 0.05$). Furthermore, NRI and IDI analyses (Table 6) demonstrated that, in the training cohort, the combined model provided greater improvement than both the clinical model and the RF radiomics model (NRI > 0, IDI > 0), and the RF radiomics model also improved prediction relative to the clinical model. In the testing cohort, the combined model showed improvement over the RF radiomics model (NRI > 0, IDI > 0).

Nomogram construction and evaluation

As shown in Figure 7, a nomogram was developed from the combined model. In both the training and testing cohorts, the clinical model, the RF radiomics model, and the nomogram satisfied the Hosmer-Lemeshow goodness-of-fit criterion, with *P* values greater than 0.05, indicating acceptable calibration.

The calibration curves in Figure 8 showed that the clinical model had the closest agreement between predicted and observed probabilities. By contrast, DCA (Figure 9) indicated that the nomogram achieved the highest net clinical benefit across a wider range of threshold probabilities than either the clinical model or the RF radiomics model. These findings suggest that, despite the stronger calibration of the clinical model, the nomogram may provide greater practical value for clinical decision-making.

DISCUSSION

ACCEPTED

HCC is the third leading cause of cancer-related death worldwide and remains a major cause of cancer mortality in China, creating a substantial clinical and socioeconomic burden[11]. Although several treatment options are available, including hepatic resection, liver transplantation, local ablation, transarterial chemoembolization, radiotherapy, systemic chemotherapy, targeted therapy, and immunotherapy, overall outcomes remain unsatisfactory[12]. GPC3 has become a focus of HCC research because of its close association with oncogenesis, tumor biology, and patient prognosis[13,14]. Several studies have also reported that elevated GPC3 expression is associated with poorer survival, supporting its potential role as a prognostic biomarker[15]. In current clinical practice, GPC3 expression is mainly assessed by pathological biopsy. Biopsy provides histological confirmation, but it is invasive and vulnerable to sampling error, especially when tumors show heterogeneous expression patterns[6]. Noninvasive imaging-based methods that can accurately predict GPC3 expression would therefore have clear clinical utility.

In this study, AFP and TBIL emerged as independent clinical predictors of GPC3 expression. After feature selection, ten radiomic features were retained. Among the radiomics models, the RF classifier achieved the best predictive performance, with AUC values of 0.959 and 0.862 in the training and testing cohorts, respectively. When the radiomics signature was combined with AFP and TBIL, predictive accuracy improved further, yielding AUC values of 0.979 and 0.948. The combined model also showed consistent positive reclassification in both cohorts, as reflected by NRI > 0 and IDI > 0. Calibration curves indicated good agreement between predicted and observed outcomes, while DCA showed that the nomogram based on the combined model offered greater net clinical benefit across a broad range of threshold probabilities. Together, these findings support the added value of integrating imaging-derived quantitative features with conventional clinical biomarkers. Such model may improve risk stratification and assist clinical decision-making, particularly in identifying patients who could benefit from targeted or immune-based treatment strategies related to GPC3 expression.

Recent studies have examined imaging-based prediction of GPC3 expression. Xu *et al*[7] analyzed clinical data and contrast-enhanced CT images from 152 patients

ACCEPTED

with solitary HCC and showed that a CT-based radiomics nomogram could differentiate GPC3-positive from GPC3-negative tumors before surgery. Zhang *et al*[16] developed radiomic and clinical models using Gd-EOB-DTPA-enhanced MRI and found that hepatobiliary-phase radiomic features were closely associated with GPC3 positivity. Their combined radiomics-clinical model provided a noninvasive, individualized approach for identifying GPC3-positive HCC. Together, these studies indicate that radiomics may help characterize the molecular phenotype of HCC.

Machine learning has been increasingly applied to medical imaging because it can extract high-dimensional quantitative features and build predictive models from complex image data[8]. But many previous radiomics studies have relied on a single machine learning classifier. To improve generalizability and allow direct comparison across modeling strategies, the present study evaluated eight commonly used algorithms: LR, SVM, KNN, RF, ET, XGBoost, LightGBM, and MLP[9]. All models showed acceptable predictive performance, but the RF-based radiomics model achieved the highest AUC (0.959). This performance may reflect several methodological strengths of RF, including its suitability for high-dimensional feature spaces, relative resistance to overfitting, capacity to handle imbalanced data, and ability to estimate variable importance[17]. Compared with earlier studies, our model achieved stronger predictive performance. For example, the CT-based radiomics nomogram reported by Xu *et al*[7] showed only moderate performance, whereas the radiomics-clinical model developed by Sun *et al*[18], which combined a radiomics score with AFP levels, achieved an AUC of 0.867-0.895. These results suggest that careful selection of machine learning algorithms can improve the performance of radiomics models for predicting molecular biomarkers in HCC.

Our analysis of feature contribution is consistent with Wang *et al*[19], who found that delayed-phase imaging features had stronger predictive value than arterial-phase features. In the present study, delayed-phase radiomic features constituted the largest share of the optimal feature set. One plausible explanation is the pseudocapsule commonly observed in HCC lesions. Because the pseudocapsule is better visualized during the delayed phase, lesion boundaries can be defined more accurately, allowing radiomic features to capture relevant textural and structural

ACCEPTED

heterogeneity[20]. This imaging context may account for prominent contribution of delayed-phase radiomics to GPC3 prediction.

Calibration analysis showed close agreement between the nomogram-predicted probabilities and the observed outcomes, supporting the model calibration. Overall, the combined RF radiomics-clinical model offers noninvasive approach for preoperative prediction of GPC3 expression in HCC and may help support individualized clinical decision-making.

This study has several limitations. First, the data were obtained from two medical centers, which may have introduced heterogeneity related to imaging equipment, scanning protocols, and operator-dependent factors. Although normalization and preprocessing were used to reduce these effects, inter-institutional variability cannot be fully excluded. Second, the retrospective design and relatively small sample size may increase the risk of overfitting and limit the generalizability of the findings. Third, external validation and prospective multicenter verification were not performed. Fourth, manual tumor segmentation may have introduced observer-related variability, and feature reproducibility was not fully assessed. But all segmentations were performed by experienced radiologists who had received standardized professional training. Fifth, the analysis was limited to conventional contrast-enhanced CT images and did not incorporate functional imaging biomarkers or multimodal approaches, such as CT perfusion, CT-derived extracellular volume fraction, or MRI. Finally, the biological interpretation of the selected radiomic features remains incomplete. Future large-scale, prospective multicenter studies should incorporate multimodal imaging data and rigorous validation strategies.

CONCLUSION

Overall, the eight machine learning models and the clinical model showed acceptable accuracy for predicting GPC3 expression in patients with HCC. The combined model, which incorporated the RF-based radiomics model with clinical predictors, yielded the best diagnostic performance. The nomogram developed from this model offers a practical approach for preoperative, noninvasive estimation of

ACCEPTED

GPC3 expression and may help guide more individualized clinical management.

REFERENCES

1 **Wang Z**, Qin H, Liu S, Sheng J, Zhang X. Precision diagnosis of hepatocellular carcinoma. *Chin Med J (Engl)* 2023; **136**: 1155-1165 [PMID: 36939276 DOI: 10.1097/CM9.0000000000002641]

2 **Feng S**, Wang J, Wang L, Qiu Q, Chen D, Su H, Li X, Xiao Y, Lin C. Current Status and Analysis of Machine Learning in Hepatocellular Carcinoma. *J Clin Transl Hepatol* 2023; **11**: 1184-1191 [PMID: 37577233 DOI: 10.14218/JCTH.2022.00077S]

3 **Liu XY**, Jin J, Zhang S, Zhang H, Zhao Y, Cai L, Zhang JZ. Dermoscopy of cutaneous metastases from primary hepatocellular carcinoma. *Chin Med J (Engl)* 2019; **132**: 2131-2132 [PMID: 31449078 DOI: 10.1097/CM9.0000000000000413]

4 **Vell MS**, Loomba R, Krishnan A, Wangensteen KJ, Trebicka J, Creasy KT, Trautwein C, Scorletti E, Seeling KS, Hehl L, Rendel MD, Zandvakili I, Li T, Chen J, Vujkovic M, Alqahtani S, Rader DJ, Schneider KM, Schneider CV. Association of Statin Use With Risk of Liver Disease, Hepatocellular Carcinoma, and Liver-Related Mortality. *JAMA Netw Open* 2023; **6**: e2320222 [PMID: 37358849 DOI: 10.1001/jamanetworkopen.2023.20222]

5 **Zhou F**, Shang W, Yu X, Tian J. Glypican-3: A promising biomarker for hepatocellular carcinoma diagnosis and treatment. *Med Res Rev* 2018; **38**: 741-767 [PMID: 28621802 DOI: 10.1002/med.21455]

6 **Lin F**, Clift R, Ehara T, Yanagida H, Horton S, Noncovich A, Guest M, Kim D, Salvador K, Richardson S, Miller T, Han G, Bhat A, Song K, Li G. Peptide Binder to Glypican-3 as a Theranostic Agent for Hepatocellular Carcinoma. *J Nucl Med* 2024; **65**: 586-592 [PMID: 38423788 DOI: 10.2967/jnumed.123.266766]

7 **Xu YH**, Xie SS, Wang J, Li SC, Liu JX, Zhang YM, Ye ZX, Shen W. [Value of enhanced CT radiomics nomogram in predicting GPC3 expression in single hepatocellular carcinoma]. *Guoji Yixue Fangshexue Zazhi* 2022; **45**: 259-266 [DOI: 10.19300/j.2022.L19410]

8 **Gao J**, Zhang M, Chen Q, Ye K, Wu J, Wang T, Zhang P, Feng G. Integrating machine learning and molecular docking to decipher the molecular network of

ACCEPTED

aflatoxin B1-induced hepatocellular carcinoma. *Int J Surg* 2025; **111**: 4539-4549 [PMID: 40392001 DOI: 10.1097/JS9.0000000000002455]

9 **Feng Z**, Li H, Liu Q, Duan J, Zhou W, Yu X, Chen Q, Liu Z, Wang W, Rong P. CT Radiomics to Predict Macrotrabecular-Massive Subtype and Immune Status in Hepatocellular Carcinoma. *Radiology* 2023; **307**: e221291 [PMID: 36511807 DOI: 10.1148/radiol.221291]

10 **Zhao J**, Gao S, Sun W, Grimm R, Fu C, Han J, Sheng R, Zeng M. Magnetic resonance imaging and diffusion-weighted imaging-based histogram analyses in predicting glypican 3-positive hepatocellular carcinoma. *Eur J Radiol* 2021; **139**: 109732 [PMID: 33905978 DOI: 10.1016/j.ejrad.2021.109732]

11 **Bo Z**, Song J, He Q, Chen B, Chen Z, Xie X, Shu D, Chen K, Wang Y, Chen G. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. *Comput Biol Med* 2024; **173**: 108337 [PMID: 38547656 DOI: 10.1016/j.compbiomed.2024.108337]

12 **Xia T**, Zhao B, Li B, Lei Y, Song Y, Wang Y, Tang T, Ju S. MRI-Based Radiomics and Deep Learning in Biological Characteristics and Prognosis of Hepatocellular Carcinoma: Opportunities and Challenges. *J Magn Reson Imaging* 2024; **59**: 767-783 [PMID: 37647155 DOI: 10.1002/jmri.28982]

13 **Shirakawa H**, Suzuki H, Shimomura M, Kojima M, Gotohda N, Takahashi S, Nakagohri T, Konishi M, Kobayashi N, Kinoshita T, Nakatsura T. Glypican-3 expression is correlated with poor prognosis in hepatocellular carcinoma. *Cancer Sci* 2009; **100**: 1403-1407 [PMID: 19496787 DOI: 10.1111/j.1349-7006.2009.01206.x]

14 **Midorikawa Y**, Ishikawa S, Iwanari H, Imamura T, Sakamoto H, Miyazono K, Kodama T, Makuuchi M, Aburatani H. Glypican-3, overexpressed in hepatocellular carcinoma, modulates FGF2 and BMP-7 signaling. *Int J Cancer* 2003; **103**: 455-465 [PMID: 12478660 DOI: 10.1002/ijc.10856]

15 **Shi D**, Shi Y, Kaseb AO, Qi X, Zhang Y, Chi J, Lu Q, Gao H, Jiang H, Wang H, Yuan D, Ma H, Wang H, Li Z, Zhai B. Chimeric Antigen Receptor-Glypican-3 T-Cell Therapy for Advanced Hepatocellular Carcinoma: Results of Phase I Trials. *Clin Cancer Res* 2020; **26**: 3979-3989 [PMID: 32371538 DOI: 10.1158/1078-0432.CCR-19-3259]

ACCEPTED

16 **Zhang N**, Wu M, Zhou Y, Yu C, Shi D, Wang C, Gao M, Lv Y, Zhu S. Radiomics nomogram for prediction of glypican-3 positive hepatocellular carcinoma based on hepatobiliary phase imaging. *Front Oncol* 2023; **13**: 1209814 [PMID: 37841420 DOI: 10.3389/fonc.2023.1209814]

17 **Wang YY**, Yang WX, Du QJ, Liu ZH, Lu MH, You CG. Construction and evaluation of a liver cancer risk prediction model based on machine learning. *World J Gastrointest Oncol* 2024; **16**: 3839-3850 [PMID: 39350987 DOI: 10.4251/wjgo.v16.i9.3839]

18 **Sun S**, Xiao S, Jiang Z, Xiao J, He Q, Wang M, Fan Y. Radiomic Analysis of Contrast-Enhanced CT Predicts Glypican 3-Positive Hepatocellular Carcinoma. *Curr Med Imaging* 2024; **20**: e15734056277475 [PMID: 38462828 DOI: 10.2174/0115734056277475240215115629]

19 **Wang WJ**, Huang JC, Yang X, Hang JL, Sun J, Fu JX, Ye J, Luo XF. [Value of enhanced MRI radiomics in preoperative prediction for Ki-67 expression in hepatocellular carcinoma]. *Zhongguo Zhongxiyi Jiehe Yingxiang Zazhi* 2024; **22**: 245-249,261 [DOI: 10.3969/j.issn.1672-0512.2024.03.001]

20 **Gu D**, Xie Y, Wei J, Li W, Ye Z, Zhu Z, Tian J, Li X. MRI-Based Radiomics Signature: A Potential Biomarker for Identifying Glypican 3-Positive Hepatocellular Carcinoma. *J Magn Reson Imaging* 2020; **52**: 1679-1687 [PMID: 32491239 DOI: 10.1002/jmri.27199]

ACCEPTED

Footnotes

Institutional review board statement: This study was approved by the Institutional Review Board of the First Affiliated Hospital of Dali University and Dali First People's Hospital (IRB approval No. DFY20250120001) and conducted in accordance with the Declaration of Helsinki.

Informed consent statement: The requirement for written informed consent was waived due to the study's retrospective design.

Conflict-of-interest statement: All authors declare no conflict of interest in publishing the manuscript.

Data sharing statement: The datasets generated and analyzed in the present study are available from the corresponding author upon reasonable request.

AI contribution statement: No AI tools, including ChatGPT, Grammarly, DeepL, or any other AI software, were used at any stage. The entirety of the main text and the response to reviewers was written by the authors without the assistance of AI. No AI tools were used for language polishing, translation, data analysis, or writing assistance of the manuscript or response letter. No AI tools participated in the design of the study or interpretation of its results. No images in the manuscript were generated by AI. In summary, no generative AI was used in the preparation of the manuscript and the response letter.

Open-Access: This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial (CC BY-NC 4.0) license. No commercial re-use. See Permissions. **Published by** Baishideng Publishing Group Inc.

Peer review: Externally peer reviewed.

ACCEPTED

Peer-review model: Single blind

Specialty type: Radiology, nuclear medicine and medical imaging

Country of origin: China

Peer-review report's classification

Scientific quality: Grade B, Grade C, Grade C, Grade C

Novelty: Grade B, Grade C, Grade C, Grade C

Creativity or innovation: Grade B, Grade C, Grade C, Grade C

Scientific significance: Grade B, Grade C, Grade C, Grade C

P-Reviewer: Suresh A, Assistant Professor, India; Xia L, MD, China; Zhang MY, Doctorate Student, China **S-Editor:** Luo ML **L-Editor: P-Editor:**

ACCEPTED

Figures and Tables

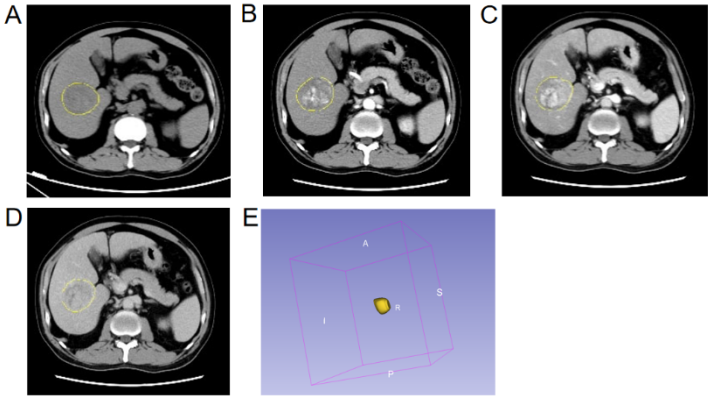


Figure 1 A 40-year-old male patient with hepatocellular carcinoma showing positive glypican-3 expression and an alpha-fetoprotein level of 961.3 ng/mL. A-D: They represent the abdominal computed tomography plain scan, arterial phase, portal venous phase, delayed phase, segmentation of the region of interest; E: The volume of interest in the arterial phase.

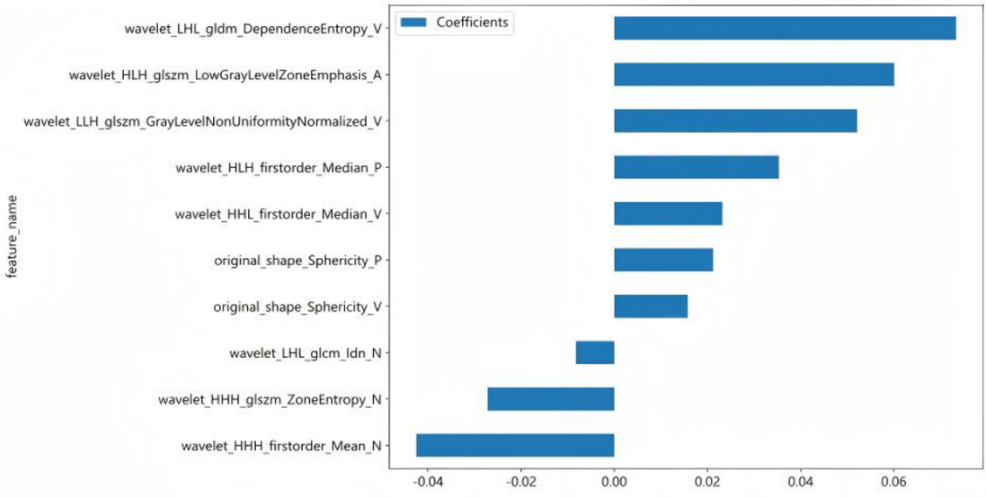


Figure 2 Computed tomography radiomics features and their corresponding correlation coefficients.

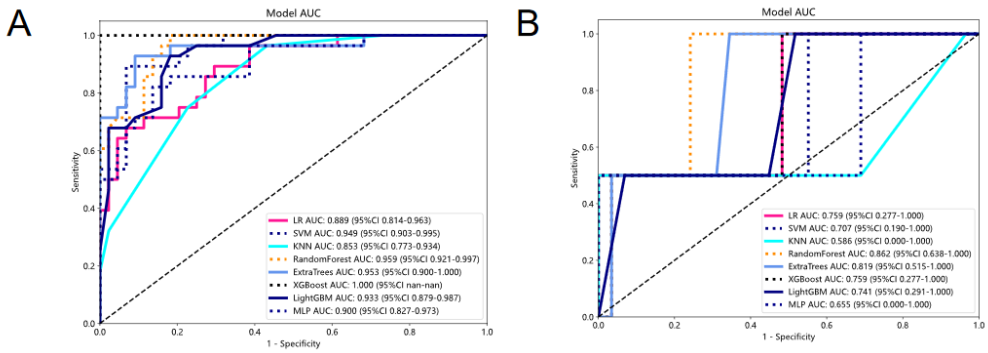


Figure 3 Receiver operating characteristic curves of radiomics models in training

and testing cohorts. A: Receiver operating characteristic curve of the radiomics model in the training cohort, showing the model's discriminatory performance; B: Receiver operating characteristic curve of the radiomics model in the testing cohort, demonstrating the model's predictive performance on unseen data. AUC: Area under the curve; KNN: K-nearest neighbors; LightGBM: Light Gradient Boosting Machine; LR: Logistic regression; MLP: Multilayer perceptron; SVM: Support vector machine; XGBoost: Extreme gradient boosting.

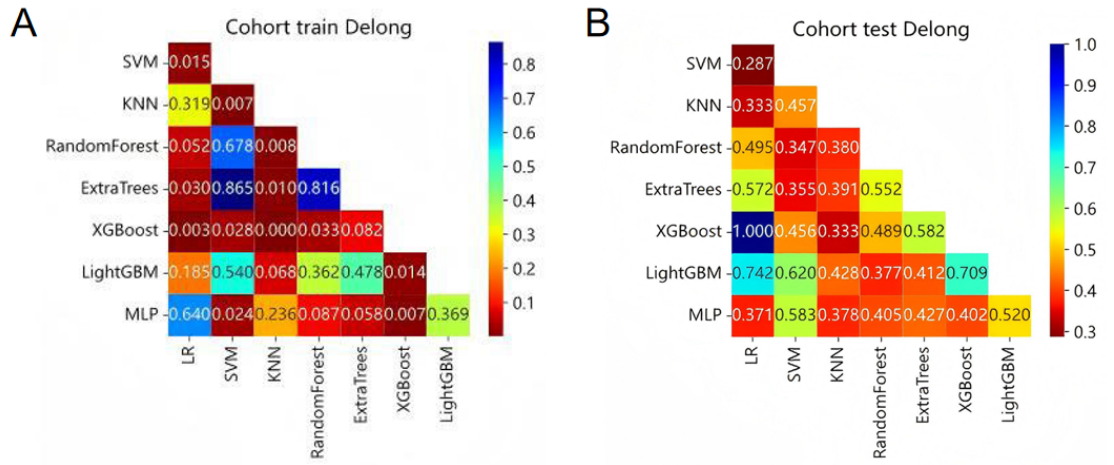


Figure 4 Comparison of area under the curve values for radiomics models using DeLong's test in training and testing cohorts. A: DeLong test results comparing area under the curve values of the radiomics model in the training cohort, indicating statistical differences in model performance; B: DeLong test results comparing area under the curve values of the radiomics model in the testing cohort, assessing the significance of predictive performance differences on independent data. KNN: K-nearest neighbors; LightGBM: Light Gradient Boosting Machine; LR: Logistic regression; MLP: Multilayer perceptron; SVM: Support vector machine; XGBoost: Extreme gradient boosting.

ACCEPTED

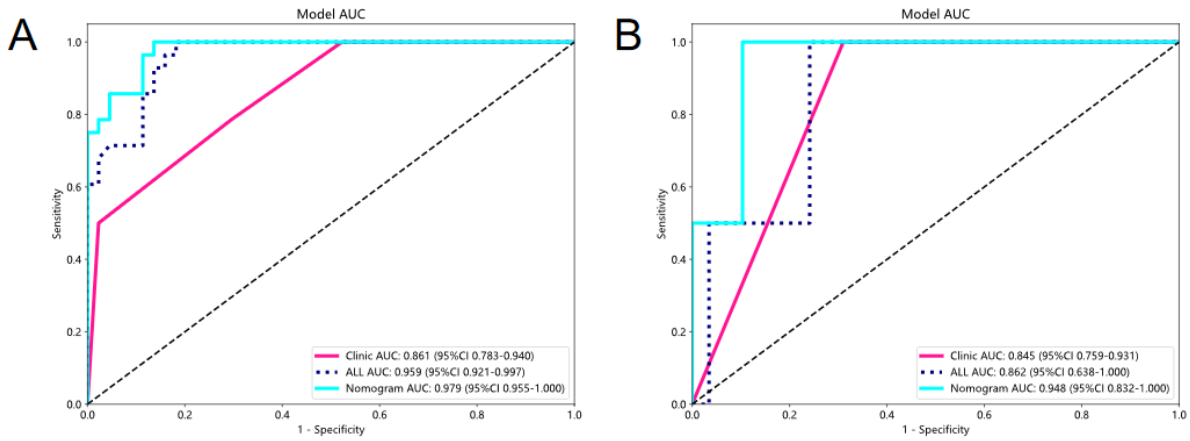


Figure 5 Receiver operating characteristic curves of clinical, radiomics, and combined models in training and testing cohorts. A: Receiver operating characteristic curves for the clinical model, random forest radiomics model, and combined nomogram model in the training cohort, illustrating comparative predictive performance; B: Receiver operating characteristic curves for the clinical model, random forest radiomics model, and combined nomogram model in the testing cohort, demonstrating model performance on independent data. Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model; AUC: Area under the curve.

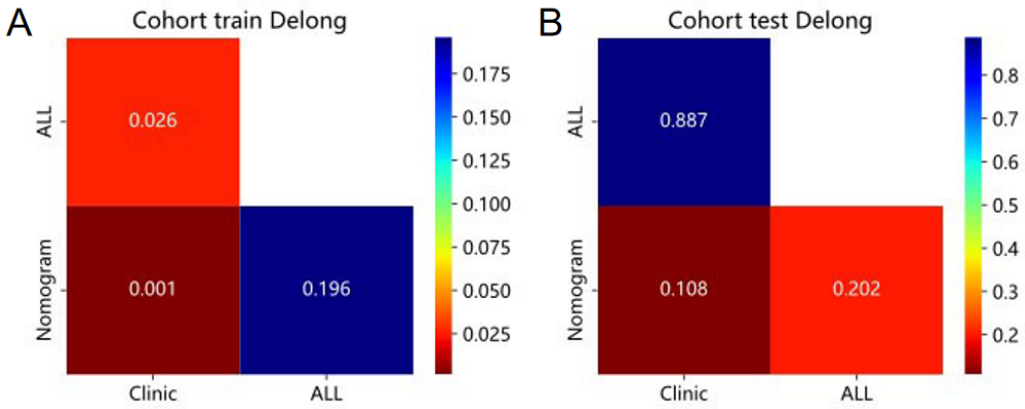


Figure 6 DeLong test comparison of area under the curve values for clinical, radiomics, and combined models in training and testing cohorts. A: DeLong test results comparing area under the curve values of the clinical model, random forest radiomics model, and combined nomogram model in the training cohort, indicating statistical differences in model performance; B: DeLong test results comparing area under the curve values of the clinical model, random forest radiomics model, and combined nomogram model in the testing cohort, assessing the significance of

ACCEPTED

predictive performance differences on independent data. Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model.

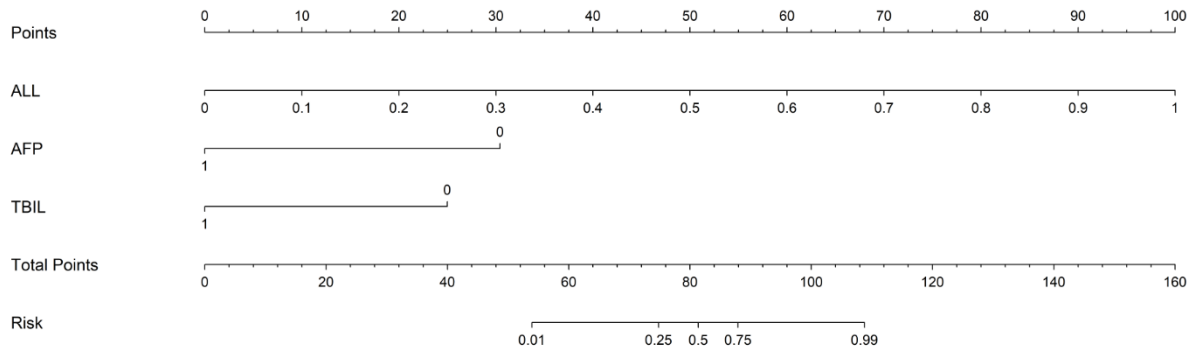


Figure 7 Nomogram constructed based on the combined model. AFP: Alpha-fetoprotein; TBIL: Total bilirubin; ALL: Random forest radiomics model; Points: Item score; Total Points: Overall score; Risk: Predicted probability.

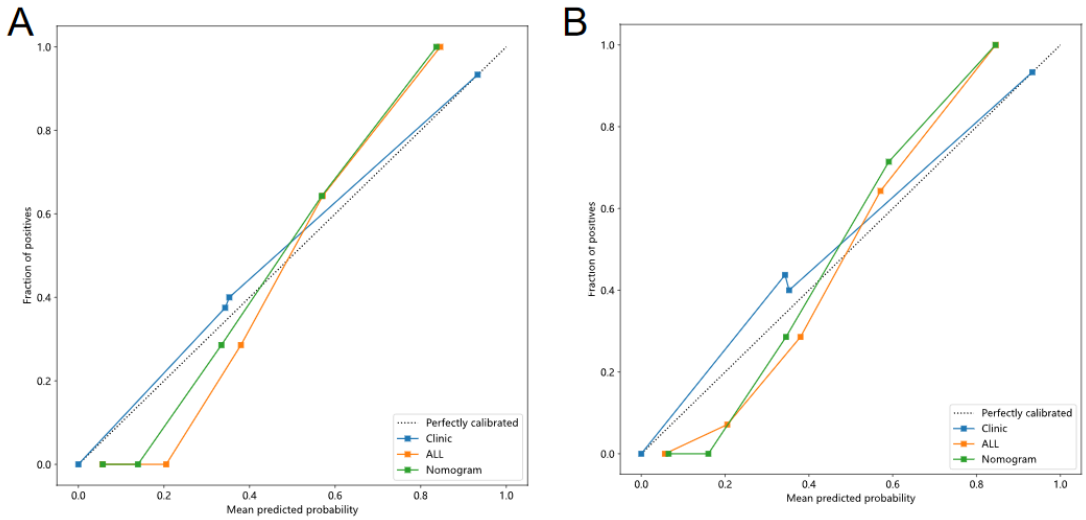


Figure 8 Calibration curves for the training and testing cohorts. A: Calibration curve of the model in the training cohort, illustrating the agreement between predicted and observed probabilities; B: Calibration curve of the model in the testing cohort, demonstrating the model's calibration performance on independent data. Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model.

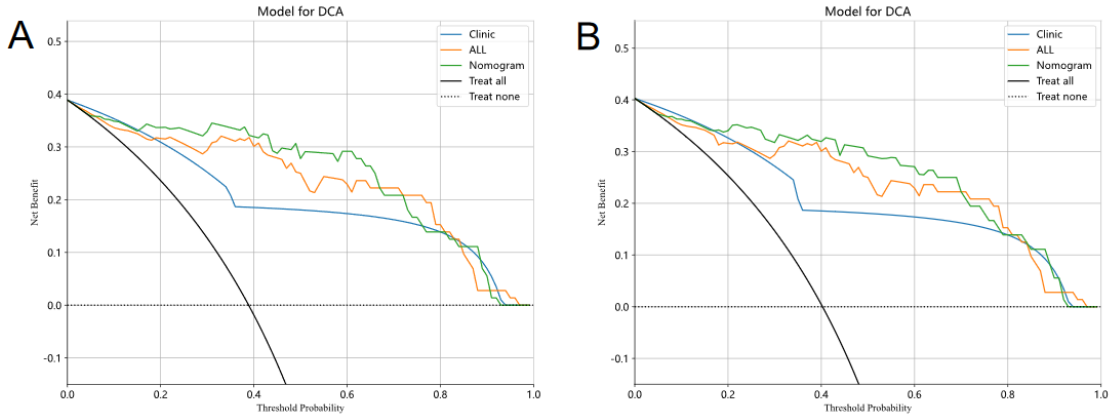


Figure 9 Decision curve analysis for the training and testing cohorts. A: Decision curve of the model in the training cohort, showing the net clinical benefit across a range of threshold probabilities; B: Decision curve of the model in the testing cohort, presenting the model's clinical utility when applied to unseen data. Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model; DCA: Decision curve analysis.

Table 1 Comparison of clinical and imaging characteristics between the training and testing cohorts, *n* (%)/mean ± SD

Demographics	**Training set (*n* = 72)**		***P***	**Testing set (*n* = 31)**		***P***	***P***
	GPC3 (+)	**GPC3 (-)**	**value**	**GPC3 (+)**	**GPC3 (-)**	**value**	**value**
Age	55.05 ± 10.69	57.21 ± 9.08	0.377	55.14 ± 9.25	52.50 ± 3.54	0.695	0.662
CEA (ng/mL)	10.31 ± 42.90	2.76 ± 1.65	0.357	4.75 ± 6.70	1.94 ± 0.83	0.563	0.647
Sex			1.000			0.100	0.421
Male	37 (84.09)	23 (82.14)		23 (79.31)	-		
Female	7 (15.91)	5 (17.86)		6 (20.69)	2 (100.00)		
Viral hepatitis			0.229			1.000	0.639
0	31 (70.45)	15 (53.57)		21 (72.41)	1 (50.00)		
1	13 (29.55)	13 (46.43)		8 (27.59)	1 (50.00)		
TBIL (μmol/L)			< 0.001			0.897	0.057

ACCEPTED

0	11 (25.00)	20 (71.43)		20 (68.97)	2 (100.00)		
1	33 (75.00)	8 (28.57)		9 (31.03)	-		
DBIL (μmol/L)			0.137			0.749	0.530
0	16 (36.36)	16 (57.14)		11 (37.93)	-		
1	28 (63.64)	12 (42.86)		18 (62.07)	2(100.00)		
IBIL (μmol/L)			0.727			0.100	0.070
0	22 (50.00)	16 (57.14)		23 (79.31)	-		
1	22 (50.00)	12 (42.86)		6 (20.69)	2 (100.00)		
ALT (U/L)			0.758			0.749	0.015
0	27 (61.36)	19 (67.86)		11 (37.93)	-		
1	17 (38.64)	9 (32.14)		18 (62.07)	2 (100.00)		
AST (U/L)			0.448			0.897	0.017
0	23 (52.27)	18 (64.29)		9 (31.03)	-		
1	21 (47.73)	10 (35.71)		20 (68.97)	2 (100.00)		
TP (g/L)			0.163			1.000	0.397
0	30 (68.18)	24 (85.71)		19 (65.52)	1 (50.00)		
1	14 (31.82)	4 (14.29)		10 (34.48)	1 (50.00)		
ALB (g/L)			1.000			1.000	0.300
0	31 (70.45)	20 (71.43)		17 (58.62)	1 (50.00)		
1	13 (29.55)	8 (28.57)		12 (41.38)	1 (50.00)		
GLB (g/L)			0.638			0.724	0.704
0	38 (86.36)	26 (92.86)		25 (86.21)	1 (50.00)		
1	6 (13.64)	2 (7.14)		4 (13.79)	1 (50.00)		
AFP (ng/mL)			< 0.001			0.327	0.684
0	13 (29.55)	22 (78.57)		11 (37.93)	2 (100.00)		
1	31 (70.45)	6 (21.43)		18 (62.07)	-		
Liver cirrhosis			0.229			1.000	1.000
0	13 (29.55)	13 (46.43)		10 (34.48)	1 (50.00)		
1	31 (70.45)	15 (53.57)		19 (65.52)	1 (50.00)		
Ascites			0.597			0.724	0.280

ACCEPTED

0	2 (4.55)	3 (10.71)		4 (13.79)	1 (50.00)	
1	42 (95.45)	25 (89.29)		25 (86.21)	1 (50.00)	
Portal vein tumor thrombus			1.000			0.449 0.940
0	5 (11.36)	4 (14.29)		2 (6.90)	1 (50.00)	
1	39 (88.64)	24 (85.71)		27 (93.10)	1 (50.00)	

GPC3: Glypican-3; TBIL: Total bilirubin; DBIL: Direct bilirubin; IBIL: Indirect bilirubin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; TP: Total protein; ALB: Albumin; GLB: Globulin; AFP: Alpha-fetoprotein; CEA: Carcinoembryonic antigen.

Table 2 Final selected computed tomography radiomics features

Feature category and count	Radiomics feature name
Shape features (2)	original _ shape _ Sphericity _ P
	original _ shape _ Sphericity _ V
First-order features (3)	wavelet _ HLH _ firstorder _ Median _ P
	wavelet _ HHL _ firstorder _ Median _ V
	wavelet _ HHH _ firstorder _ Mean _ N
Texture features (5)	Wavelet _ LHL _ gldm _ DependenceEntropy _ V
	wavelet _ HLH _ glszm _ LowGrayLevelZoneEmphasis _ A
	wavelet _ LLH _ glszm _ GrayLevelNonUniformityNormalized _ V
	Wavelet _ LHL _ glcm _ ldn _ N
	wavelet _ HHH _ glszm _ ZoneEntropy _ N

N: Non-contrast phase; A: Arterial phase; P: Portal venous phase; V: Delayed phase.

Table 3 Performance of radiomics models for predicting glypican-3 expression status in hepatocellular carcinoma

Set	Model	AUC	95%CI	Sensitivity	Specificity	PPV	NPV	Accuracy
Training set	LR	0.889	0.814-0.963	0.643	0.932	0.857	0.804	0.819
	SVM	0.949	0.903-0.995	0.857	0.932	0.889	0.911	0.903

ACCEPTED

	KNN	0.853	0.773-0.934	0.750	0.773	0.677	0.829	0.764
	RF	0.959	0.921-0.997	0.964	0.818	0.771	0.973	0.875
	ET	0.953	0.900-1.000	0.893	0.909	0.862	0.930	0.903
	XGBoost	1.000	1.000-1.000	0.964	1.000	1.000	0.978	0.986
	LightGBM	0.933	0.879-0.987	0.857	0.841	0.774	0.902	0.847
	MLP	0.900	0.827-0.973	0.786	0.864	0.786	0.864	0.833
Testing set	LR	0.759	0.277-1.000	0.500	0.517	0.067	0.937	0.516
	SVM	0.707	0.190-1.000	0.000	0.966	0.000	0.933	0.903
	KNN	0.586	0.000-1.000	0.000	1.000	0.000	0.935	0.935
	RF	0.862	0.638-1.000	0.705	0.759	0.725	0.957	0.742
	ET	0.819	0.515-1.000	0.500	0.690	0.100	0.952	0.677
	XGBoost	0.759	0.277-1.000	0.500	0.517	0.067	0.937	0.516
	LightGBM	0.741	0.291-1.000	0.500	0.552	0.071	0.941	0.548
	MLP	0.655	0.000-1.000	0.000	1.000	0.000	0.935	0.935

LR: Logistic regression; SVM: Support vector machine; KNN: K-nearest neighbors; RF: Random forest; ET: Extra trees; MLP: Multilayer perceptron; AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value; XGBoost: Extreme gradient boosting; LightGBM: Light Gradient Boosting Machine.

Table 4 Univariate and multivariate analysis of clinical and computed tomography imaging features

Characteristics	**Univariate analysis**			**Multivariate analysis**		
	OR	**95%CI**	***P* value**	**OR**	**95%CI**	***P* value**
Age	1.005	0.995-1.015	0.377			
CEA (ng/mL)	0.998	0.995-1.001	0.357			
Sex	1.034	0.797-1.342	0.832			
Viral hepatitis	1.190	0.975-1.452	0.150			
TBIL (μmol/L)	0.638	0.535-0.759	< 0.001	0.636	0.549-0.738	< 0.001
DBIL (μmol/L)	0.819	0.676-0.991	0.086			
IBIL (μmol/L)	0.934	0.770-1.134	0.560			
ALT (U/L)	0.935	0.764-1.145	0.582			

ACCEPTED

AST (U/L)	0.890	0.733-1.081	0.322			
TP (g/L)	0.801	0.643-0.998	0.097			
ALB (g/L)	0.989	0.799-1.224	0.931			
GLB (g/L)	0.855	0.629-1.163	0.400			
AFP (ng/mL)	0.627	0.529-0.744	< 0.001	0.626	0.541-0.725	< 0.001
Liver cirrhosis	0.840	0.689-1.025	1.150			
Ascites	0.797	0.546-1.165	0.322			
Portal vein tumor thrombus	0.938	0.700-1.259	0.719			

OR: Odds ratio; CEA: Carcinoembryonic antigen; TBIL: Total bilirubin; DBIL: Direct bilirubin; IBIL: Indirect bilirubin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; TP: Total protein; ALB: Albumin; GLB: Globulin; AFP: Alpha-fetoprotein.

Table 5 Performance of the combined model for predicting glypican-3 expression status in hepatocellular carcinoma

Set	Model	AUC	95%CI	Sensitivity	Specificity	PPV	NPV	Accuracy
Training set	Clinic	0.861	0.783-0.940	0.500	0.977	0.933	0.754	0.792
	All	0.959	0.921-0.997	0.964	0.818	0.771	0.973	0.875
	Nomogram	0.979	0.955-1.000	0.964	0.864	0.818	0.974	0.903
Testing set	Clinic	0.845	0.759-0.931	0.495	0.688	0.599	0.935	0.935
	All	0.862	0.638-1.000	0.705	0.759	0.725	0.957	0.742
	Nomogram	0.948	0.832-1.000	0.815	0.897	0.801	0.963	0.871

Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model; AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value.

Table 6 Comparison of net reclassification improvement and integrated discrimination improvement among the clinical model, random forest radiomics model, and combined model

Comparisons between models	Training set		Testing set	
	NRI	IDI	NRI	IDI

ACCEPTED

Nomogram *vs* clinic	0.263	0.071	-0.034	-0.046
Nomogram *vs* ALL	0.032	0.048	0.397	0.166
ALL *vs* clinic	0.231	0.023	-0.431	-0.211

Clinic: Clinical model; ALL: Random forest radiomics model; Nomogram: Combined model; NRI: Net reclassification improvement; IDI: Integrated discrimination improvement.

ACCEPTED