

CT-Based Radiomics for Non-Invasive Prediction of Ki-67 Expression in Hepatocellular Carcinoma

Meilong Wu^{1,*}, Zhiyong Du^{1,*}, Ying Xiao², Yan Wang³, Jintao Yang¹, Zhike Li¹, Xini Liu¹, Shizhong Yang⁴, Tailai An¹

¹Division of Hepatobiliary and Pancreas Surgery, Department of General Surgery, Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University, The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, Guangdong, People's Republic of China;

²Department of Pathology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, People's Republic of China;

³Department of Radiology, Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University, The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, Guangdong, People's Republic of China; ⁴Hepatopancreatobiliary Center, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Institute for Precision Medicine, Tsinghua University, Beijing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Tailai An, Division of Hepatobiliary and Pancreas Surgery, Department of General Surgery, Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University, The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, Guangdong, 518002, People's Republic of China, Email antailai@mail2.sysu.edu.cn; Shizhong Yang, Hepato-Pancreato-Biliary Center, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, 102218, People's Republic of China, Email hepatobiliary2002@aliyun.com

Purpose: By far, non-invasive methods assessing Ki-67 are still scarce. This study was performed to evaluate the capability of radiomics based on contrast-enhanced CT in predicting expression of Ki-67 in hepatocellular carcinoma (HCC).

Patients and Methods: HCC patients who underwent curative hepatectomy were included. The optimal Ki-67 cutoff value for prognostic stratification was determined using maximum selection rank statistics. The Least Absolute Selection and Shrinkage Operator (LASSO) regression analysis was used for dimension reduction and data screening to obtain radiomics features. The radiomics model, clinical model and combined model integrating radiomics features and clinical indicators for predicting Ki-67 were constructed. The predictive efficacy among the models was compared using C-index, and further verified through DeLong test, decision curves and clinical impact curves.

Results: The optimal cutoff value for Ki-67 was 0.25. A total of 2553 radiomics features were obtained and after stability testing (by intraclass correlation coefficient ≥ 0.75) and feature selection (LASSO), four radiomics features that were highly correlated and stable with the expression of Ki-67 were included in the model construction. Multivariate analyses revealed that alpha-fetoprotein and intratumoral necrosis or radiomics features were independent predictors of Ki-67. The clinical model, radiomics model and combined model were constructed, respectively. The C-indices of Ki-67 for clinical model, radiomics model and the combined model were 0.75, 0.82 and 0.88. DeLong test, decision curves and clinical impact curves further confirmed that the inclusion of radiomics features improved the predictive efficacy of the model.

Conclusion: The comprehensive model based on contrast-enhanced CT radiomics could non-invasively and effectively predict the expression of Ki-67, suggesting its value in clinical decision-making.

Keywords: hepatocellular carcinoma, radiomics, Ki-67, prognosis

Introduction

Globally, hepatocellular carcinoma (HCC) is one of the most commonly diagnosed malignant tumors.¹ HCC is a heavy burden on the medical system and nearly a half of patients with HCC are from China.²⁻⁴ A series of clinicopathological parameters assessing outcomes of patients after treatments have been reported, of which, Ki-67 is a prognostic marker indicating proliferation activity.^{5,6} For patients with high risk of recurrence after curative-intent hepatectomy, transcatheter arterial chemoembolization (TACE) is recommended to reduce the risk of recurrence.⁷ Ki-67 was found to be up-regulated among unresponsive HCC and its high expression was significantly associated with

poor outcomes after TACE.^{8,9} In a study assessing the effects of Ki-67 on patients with liver metastasis of neuroendocrine tumor undergoing yttrium-90 (Y-90) or TACE treatment, it was demonstrated that patients with high expression of Ki-67 benefited more from Y-90 treatment while those with low expression of Ki-67 benefited more from TACE.¹⁰

Considering that Ki-67 was closely correlated with treatment, assessment of its expression is of huge significance. However, conventional method predicting (ie immunohistochemical staining) its expression is unsatisfactory. First, it is an invasive method, and HCC tissue must be obtained either via ultrasonography-guided fine needle biopsy or hepatectomy. Second, this method postpones formulating optimal treatment. Third, it costs more to assess Ki-67 expression in this way. However, these methods are not directly linked to patient prognosis. Therefore, a method that determines a cutoff value based on its association with survival outcomes is more clinically relevant, as Ki-67 is a known prognostic factor. Thus, it is necessary for us to develop a non-invasive and less expensive way assessing the expression of Ki-67. As a novel technology of precision medicine, radiomics could help us realize the transformation from qualitative analysis to quantitative analysis.^{11–13} By radiomics, we could perform more complex analysis of imaging features far beyond qualitative evaluation of tissue alterations, making it possible for us to perform quantitative imaging analysis to provide personalized diagnosis and treatment to patients with HCC.^{13–15} It is well known that CT has become a standard workhorse for diagnosis and staging of HCC, making a CT-based radiomics model a highly practical tool. The present study was performed to assess whether radiomics based on preoperative enhanced CT could accurately provide the expression of Ki-67.

Materials and Methods

Study Population

Clinicopathological data of patients with HCC undergoing curative-intent hepatectomy at Beijing Tsinghua Changgung Hospital, Tsinghua University between May 2015 and January 2022 were retrospectively reviewed. The inclusion criteria were as follows: pathologically diagnosed HCC; curative-intent hepatectomy as the initial treatment; complete CT information including three phases; complete clinicopathological and follow-up information. The exclusion criteria were as follows: incomplete clinicopathological data; incomplete imaging information or poor imaging quality; perioperative mortality.

Some inflammatory parameters were included, such as neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), lymphocyte-to-monocyte ratio (LMR), derived neutrophil-to-lymphocyte ratio (dNLR) and prognostic nutritional index (PNI). These inflammatory parameters were obtained adopting the same methods as in our previous study.¹⁶ The follow-ups took place on the phone or at the outpatient department. The case managers and hepatobiliary surgeons were responsible for follow-ups.

Staining and Counting of Ki-67

HCC tissues were cut into 4μm sections after being fixed in formalin and embedded in paraffin. These sections were then separated with xylene and rehydrated with various concentrations of alcohol (100%, 95%, 85%, and 75%). Endogenous peroxidase activity was then inhibited by adding 0.3% hydrogen peroxide. These sections were then boiled under pressure in a solution concentration containing 10mM citrate buffer solution (pH 6.0) to complete antigen retrieval. Subsequently, these sections were incubated with goat serum blocking solution for 30 min at 37°C. Ultimately, these sections were incubated using primary antibodies overnight at 4°C. Primary antibodies against Ki-67 were purchased from Invitrogen (14–5698-82). On the second day, after being washed three times, these sections were stained using 3,3'-Diaminobenzidine (DAB) and restained using hematoxylin. Then, the expression of Ki-67 was evaluated by observing the staining of nucleus. Brownish meant positive Ki-67 expression. There is no universally accepted method assessing Ki-67 expression in HCC, and we randomly chose three tumor regions to assess Ki-67 expression, which was consistent with a guideline on breast cancer published in 2022.¹⁷ The expression level of Ki-67 was calculated as the average of the ratio of positive cells in three randomly selected tumor regions.

Parameters for Contrast-Enhanced CT

All the CT imaging was performed by CT scanner at our center (16 or 64-slice spiral CT, GE Medical System, Siemens). The following parameters were acquired: the voltage was 120 kV; the electric current ranged from 100 to 500 mA; the rotation time was 0.6 second; the collimation of detector was 40mm; the scanning area was adjusted to include the whole liver; the scanning matrix was 512×512. After the conventional non-enhanced CT scan, the arterial phase, portal vein phase, and delayed phase enhanced CT scans were performed 32, 70, and 180 seconds later, respectively. 90 to 100 mL of Ultravist 370 (Bayer Schering Pharma, Berlin, Germany) was administered intravenously via a pump syringe (Ulrich CT Plus 150, Ulrich Medical, Ulm, Germany) at a rate of 3.0 mL/s. The enhanced CT images were reconstructed with a thickness of 5 mm. The contrast timing was controlled by computer software, suggesting that contrast timing variability was minimal.

Processing and Segmentation of CT Images

5-mm thick-layer enhanced CT image was exported as DICOM files, which were further segmented using 3D Slicer (4.10.2) to obtain region of interest (ROI). All the image segmentation and extraction were performed by a hepatobiliary surgeon with 5-year experience, who was blinded to the Ki-67 status and other clinical data. We performed the whole-tumor segmentation. Initially, these DICOM files were imported into 3D Slicer software and the standardized window width (350) and window position (40) of the abdomen window width and window position module recommended by the software were adopted. In the segmentation module, a semi-automatic method was used to segment the liver layer by layer to obtain 3D ROIs.

Image Segmentation and Extraction of Features

The Pyradiomics 2.2.0 software package was adopted to extract 3D radiomics features, and all the segmented images were resampled into $1 \times 1 \times 1 \text{ mm}^3$ size voxels to standardize voxel spacing. The recommended 25 HU box width discrete voxel intensity value was used to reduce image noises. In order to calculate the stability and repeatability of extracted features, 30 patients were randomly chosen to perform image segmentation and extraction of radiomics features by a hepatobiliary surgeon with 5-year experience and a radiologist with 10-year experience. The interval of radiomics extraction was at least 1 week. Intraclass correlation coefficients (ICCs) were used to calculate the stability of features extracted by the same researcher and the repeatability of features extracted by different researchers. $\text{ICC} < 0.4$ suggested poor correlation; $0.4 \leq \text{ICC} < 0.6$ suggested moderate correlation; $0.6 \leq \text{ICC} < 0.75$ suggested good correlation; $\text{ICC} \geq 0.75$ suggested excellent correlation.¹⁸ In this study, imaging features with $\text{ICC} \geq 0.75$ (both intra-investigators and inter-investigators) were included. Subsequently, the Least Absolute Selection and Shrinkage Operator (LASSO) regression analysis was performed to select radiomics features and the penalty parameters are adjusted by 10-fold cross verification. Ultimately, the radiomics score (Radscore) of each patient was obtained by weighting coefficient for each feature.

Statistical Analysis

The radiomics data of different patients were integrated using the Python 3.8. SPSS 26.0 (IBM, Armonk, NY) and R version 4.0.3 software (<http://www.r-project.org/>) was adopted for statistical analysis. R software was adopted for mapping. Z-score was used to standardize radiomics data. Heatmaps were plotted to demonstrate correlations between different radiomics features. ICCs were calculated using the “irr” package of the R software and features with $\text{ICC} \geq 0.75$ were selected. The “glmnet” of the R software was used to perform LASSO regression and 10-fold cross-validation to reduce the dimensionality of high-dimensional radiomics features. The logistics regression analysis was performed to identify independent predictive factors. Continuous variables were demonstrated as mean ($\pm$ standard deviation), while categorical ones were presented as numbers and percents.

The “nomogramFormula” package of the R software was used to calculate the Radscore of each patient. The optimal cutoff values for continuous variables were determined by maximum selection rank statistics using the “survminer” package of the R software. The survival curves were plotted using the Kaplan–Meier method and compared using the Log rank test. The receiver operating characteristic curves (ROCs) were plotted using the pROC software package and

were used to assess risks. The “rms” package in R software was used to develop the nomogram, and an evaluation of validation was conducted using the calibration curve. The C-index and the area under the ROC curve (AUC) were used to evaluate the predictive efficacy of the model. The DeLong test was adopted to compare area under curve of ROC. The decision curve and the clinical impact curve were used to assess predicted benefits. All the statistical analyses were two-sided in nature and $P < 0.05$ was considered statistically significant.

Results

Baseline Characteristics

A total of 88 patients were included, of which 65 ones were male. Sixty-five patients (73.9%) had positive HBsAg. Ten patients (11.4%) suffered from multiple tumors. The median expression of Ki-67 was 0.27 (0.01–0.80) (Table 1). The detailed process for patient selection was presented in Figure 1.

The Optimal Cutoff Value for Ki-67

It was demonstrated through the maximum selection rank statistics that the optimal cutoff value for Ki-67 was 0.25 (Figure 2). We then assessed the impacts of Ki-67 on disease-free survival, revealing that high expression of Ki-67 was significantly associated with poorer disease-free survival ($P=0.028$) (Figure 3). The 1-year disease-free survival for the

Table 1 Clinicopathological Data of Patients with Hepatocellular Carcinoma

Variables	Cohort
Gender (Male)	65(73.9%)
Age (>60 years)	46(52.3)
Age (years)	62 (33–84)
HBsAg (positive)	65(73.9%)
AFP (>20ng/mL)	58(65.9%)
Tumor diameter (cm)	4.50(0.70–14.00)
Tumor number (multiple)	10(11.4%)
Cirrhosis (present)	30(34.1%)
Tumor margin (smooth)	55(62.5%)
Tumor necrosis (present)	47(53.4%)
dNLR	1.40(–1.22–493.32)
PLR	103.85(29.05–666.67)
NLR	1.90(0.72–8.66)
LMR	3.41(0.05–8.55)
PNI	48.70(29.55–64.10)
Ki-67	0.27(0.01–0.80)

Abbreviations: AFP, alpha-fetoprotein; HBsAg, hepatitis B surface antigen; dNLR, derived neutrophil-lymphocyte ratio; NLR, neutrophil-lymphocyte ratio; PLR, platelet-lymphocyte ratio; PNI, prognostic nutritional index; LMR, lymphocyte-monocyte ratio.

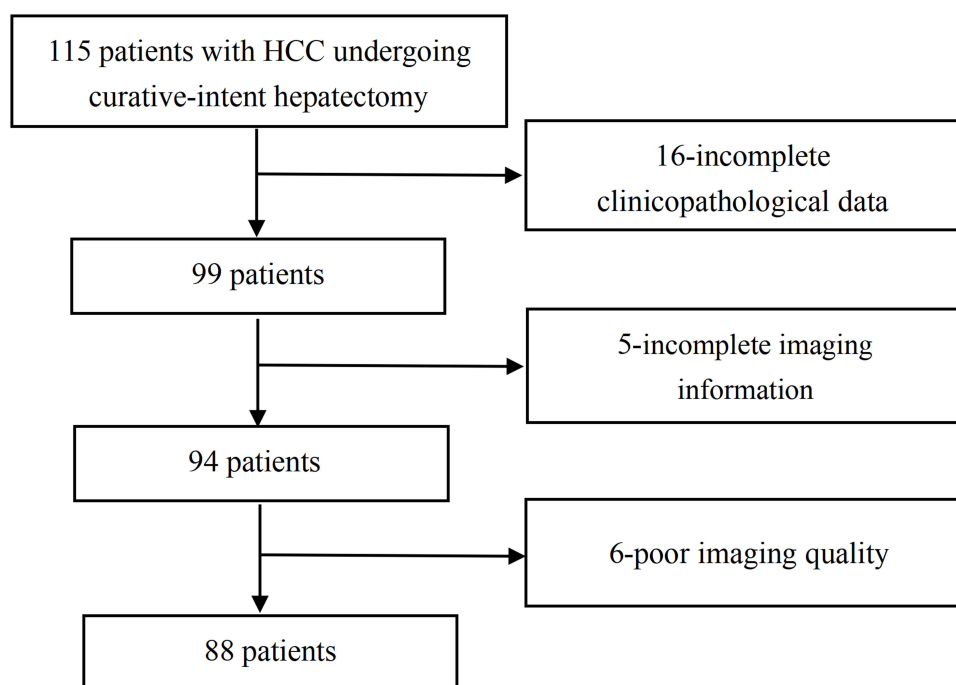


Figure 1 Patient inclusion flowchart.

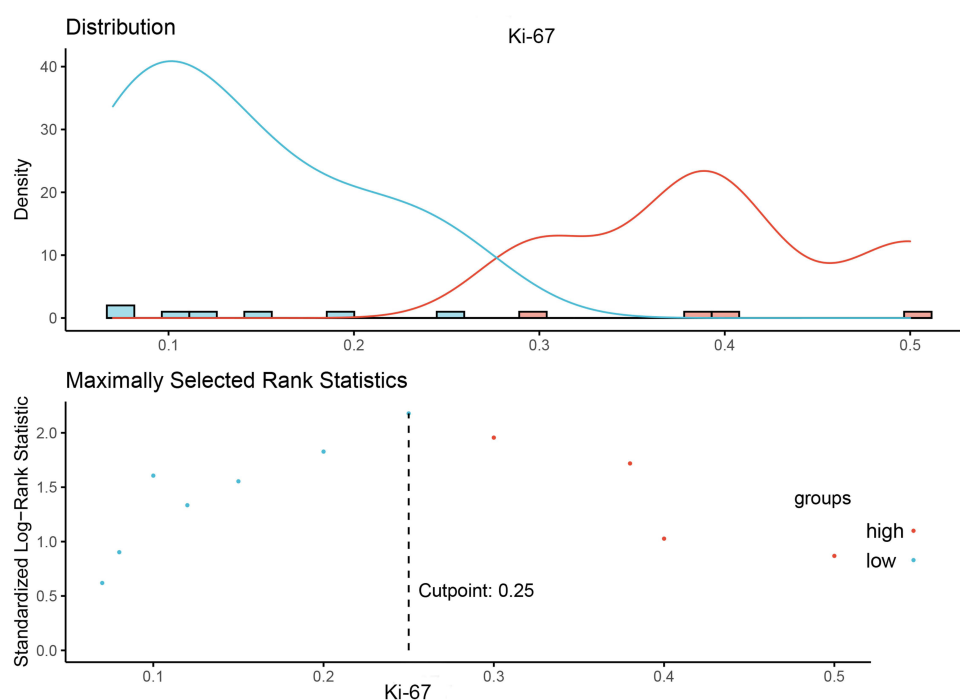


Figure 2 Optimal cutoff for Ki-67 identified by maximally selected rank statistics.

high Ki-67 group was 45.5% and the 3-year disease-free survival for the high Ki-67 group was 15.9%. The 1-year disease-free survival for the low Ki-67 group was 72.7% and the 3-year disease-free survival for the high Ki-67 group was 11.4%.

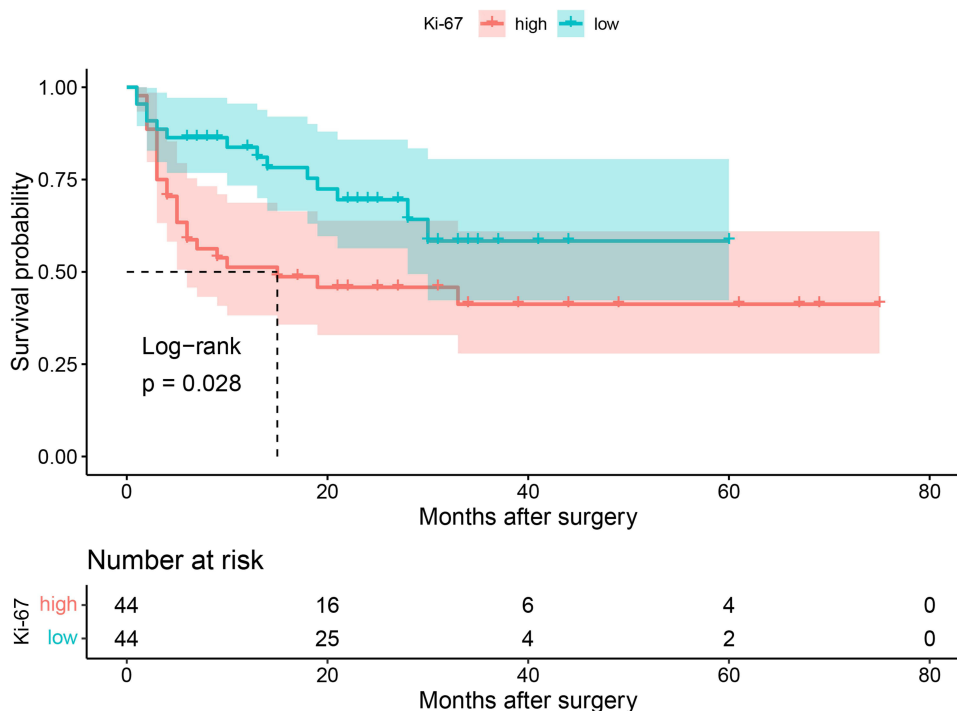


Figure 3 The impact of Ki-67 on disease-free survival by the Kaplan–Meier curve.

Radiomics Features and Foundation of Radiomics Signature

A total of 2553 features were obtained from three phases of contrast-enhanced CT and 1748 ones were with intra-class and inter-class ICCs ≥ 0.75 . The 1748 radiomics features are standardized using the Z-score function to avoid the impacts of too large a difference in data dimensions on the results. It was demonstrated through heatmaps that these radiomics feature were interrelated (Figure 4). Furthermore, LASSO regression analysis was performed to reduce the dimensionality of the high-dimensional data to avoid overfitting, results of which demonstrated that there were 12 non-0 coefficient features when the optimal solution was $\lambda=0.075$ and there were 17 non-0 coefficient features when the suboptimal solution was $\lambda=0.050$ (Figure 5).

The 12 optimal radiomics features were used for univariate and multivariate analyses to identify independent predictors of Ki-67 (Table 2). The process selecting the optimal features for the final radiomics was summarized as “The optimal λ value of 0.075 was selected via 10-fold cross-validation using the minimum criterion, resulting in 12 non-zero coefficients and these 12 features were then entered into a multivariate logistic regression to identify the most robust independent predictors for the final radiomics signature”. The multivariate analysis revealed that four indicators (wavelet-HHL-glszm-SmallAreaEmphasis-a, wavelet-LLL-firstorder-Median-a, original-shape-Elongation-d, and wavelet-LLH-gldm-DependenceNonUniformityNormalized-d) were independently associated with Ki-67 expression. These independently correlated radiomics features were further incorporated to construct the radiomics score (Radscore).

Independent Predictive Factors for Ki-67 Expression

Then, we performed logistics regression analysis to identify independent predictive factors for Ki-67 expression. Univariate analysis of clinical indicators revealed that alpha fetal protein (AFP) (OR (odds ratio) = 5.789, 95% CI (confidence interval): 2.127–15.757, $P=0.001$) and tumor necrosis (OR=2.793, 95% CI: 1.175–6.638, $P=0.020$) were predictors of ki67 (Table 3). When the Radscore was not included, AFP (OR=5.460, 95% CI: 1.963–15.189, $P=0.001$) and necrosis within the tumor (OR=2.552, 95% CI: 1.008–6.464, $P=0.048$) were independent predictive factors for Ki-67 expression (Table 4). However, when the Radscore was included, AFP (OR=7.078, 95% CI: 2.019–24.878, $P=0.002$) and

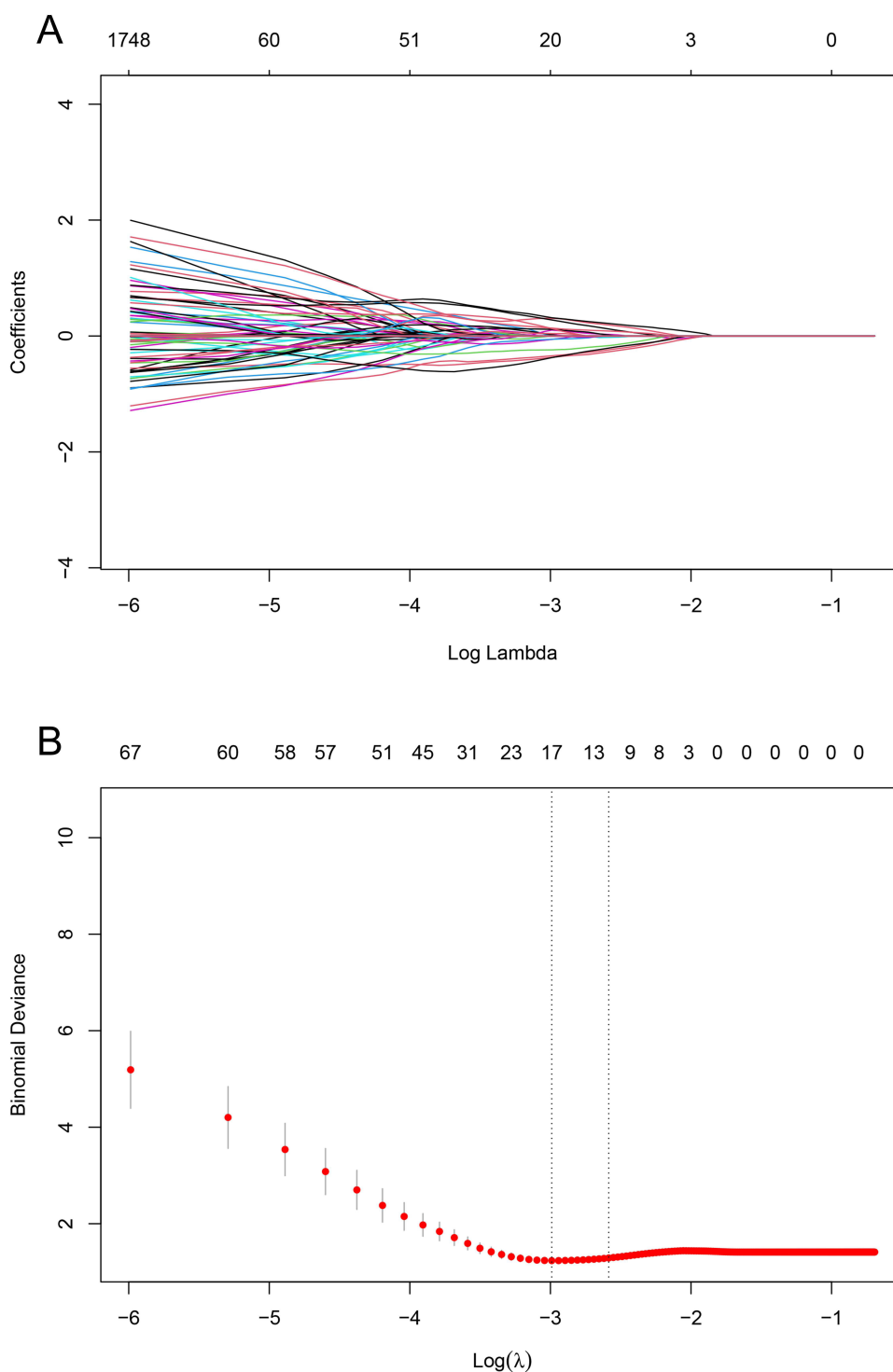


Figure 5 Extracting Radiomic Features Using LASSO Regression. **(A)** The LASSO coefficient distribution map of radiomics features. The Y-axis represents the value of the coefficient, the X-axis below represents $\log(\lambda)$, and the X-axis above represents the number of non-zero coefficients. **(B)** The tuning parameters (λ) of the LASSO model were selected by using 10 cross-validations. The relationship between binomial bias and $\log(\lambda)$ is plotted, and the numbers on the upper X-axis represent the number of predictors. Using the minimum criterion and one standard error of the minimum criterion, draw two vertical lines at the optimal value.

for AFP was 1.022 and VIF for tumor necrosis was 1.022. Given the fact that $VIF < 5$ means low collinearity, we may conclude that there is no remarkable collinearity between clinical indicators and radiomics features. Ultimately, the clinical practicability of different models in predicting Ki-67 expression was assessed by the clinical impact curve

Table 2 Univariate and Multivariate Analyses of Ki-67 Expression Predicted by Radiomics Features

Variables	Univariate		Multivariate	
	OR (95% CI)	P	OR (95% CI)	P
Wavelet-HHL-glszm-SmallAreaEmphasis-a	0.503(0.304–0.834)	0.008	0.380(0.185–0.781)	0.008
Wavelet-LLL-firstorder-Median-a	0.592(0.372–0.940)	0.026	0.369(0.178–0.764)	0.007
Original-firstorder-10Percentile-p	0.637(0.407–0.996)	0.049		
Wavelet-LHH-glszm-LargeAreaEmphasis-p	2.043(1.010–4.132)	0.047		
Wavelet-LLH-firstorder-Mean-p	2.359(1.243–4.476)	0.009		
Wavelet-LLL-gldm-MCC-p	1.492(0.942–2.362)	0.088		
Wavelet-LLL-firstorder-InterquartileRange-p	1.433(0.887–2.317)	0.142		
Original-shape-Elongation-d	0.552(0.317–0.962)	0.036	0.430(0.200–0.925)	0.031
Original-shape-Flatness-d	0.612(0.378–0.990)	0.046		
Wavelet-LHH-firstorder-Uniformity-d	2.478(0.918–6.687)	0.073		
Wavelet-LLH-gldm-DependenceNonUniformityNormalized-d	0.458(0.240–0.874)	0.018	0.318(0.118–0.858)	0.024
Wavelet-LLH-firstorder-Median-d	2.672(1.185–6.025)	0.018		

Abbreviations: a, arterial phase; CI, confidence interval; d, delay phase; p, portal phase; OR, odds ratio.

Table 3 Univariate Analysis of Clinical Indicators Predicting Ki-67 Expression

Variables	Univariate	
	OR (95% CI)	P
Gender (Male)	0.551(0.209–1.452)	0.228
Age (>60 year)	0.833(0.361–1.925)	0.670
HBsAg (positive)	1.815(0.689–4.783)	0.228
AFP (>20ng/mL)	5.789(2.127–15.757)	0.001
Tumor diameter (cm)	1.124(0.976–1.293)	0.105
Tumor number (multiple)	2.586(0.623–10.736)	0.191
Cirrhosis (present)	2.280(0.921–5.644)	0.075
Tumor margin (smooth)	0.503(0.209–1.211)	0.125
Tumor necrosis (present)	2.793(1.175–6.638)	0.020
dNLR	3.351(0.524–21.444)	0.202
PLR	0.997(0.990–1.003)	0.324
NLR	1.043(0.750–1.450)	0.802
LMR	0.838(0.617–1.138)	0.838
PNI	0.967(0.891–1.049)	0.417

Abbreviations: AFP, alpha-fetoprotein; HBsAg, hepatitis B surface antigen; CI, confidence interval; dNLR, derived neutrophil-lymphocyte ratio; LMR, lymphocyte-monocyte ratio; NLR, neutrophil-lymphocyte ratio; OR, odds ratio; PLR, platelet-lymphocyte ratio; PNI, prognostic nutritional index.

Table 4 Multivariate Analysis of Ki-67 Expression Prediction

Variables	Clinical Model		Combined Model	
	OR (95% CI)	P	OR (95% CI)	P
AFP (>20ng/mL)	5.460 (1.963–15.189)	0.001	7.807 (2.019–24.878)	0.002
Tumor necrosis (present)	2.552 (1.008–6.464)	0.048		
Radscore			1.098 (1.051–1.149)	0.000037

Abbreviations: AFP, alpha-fetoprotein; CI, confidence interval; OR, odds ratio; Radscore, radiomics score.

analysis, revealing that the combined model was associated with the highest clinical practicability while the clinical model owned the lowest clinical practicability (Figure 8).

Discussion

HCC is often diagnosed at an advanced stage. For advanced HCC, TACE, targeted therapy and immunotherapy were commonly adopted to treat advanced HCC.^{19,20} In clinical practice, it is observed that patients with the same stage HCC often have remarkably different prognosis, which might be explained by the fact that these patients own different molecular mutation patterns despite the same stage. Ki-67 is a vital indicating proliferation activity. As mentioned above, expression of Ki-67 would affect the responses of patients to different treatments.^{8–10} For breast cancer, Ki-67 is used as a predictive factor for cancer recurrence after surgery and changes in Ki-67 expression could be used to predict the preoperative “window of opportunity” in neoadjuvant endocrine therapy.^{21,22} Thus, it is of huge clinical significance for us to preoperatively predict Ki-67 expression since it would help us to select the optimal treatment for each patient. Those with high expression were more likely to experience recurrence and metastasis. Correspondingly, those with high expression of Ki-67 should undergo stricter follow-up and adjuvant therapy such as TACE, tyrosine kinase inhibitor and immunotherapy. Multiple methods identifying the optimal cutoff value for Ki-67 have been proposed, such as ROC curve, median and mean.^{23–27} However, none of these methods is associated with prognosis. Thus, methods associated with prognosis were more reasonable since Ki-67 was a known prognostic factor. It was revealed through the maximum selection rank statistics that the optimal cutoff value for Ki-67 was 0.25. The optimal cutoff value for Ki-67 was different from the median and mean of Ki-67 expression of this study. Survival analysis was further conducted based on the obtained cutoff. The results revealed that there was a significant difference in survival between the high-expression and low-expression groups of Ki-67. This further confirms that the cutoff values determined in this study have a good prognostic stratification ability and good clinical practicability.

Ki-67 has been proposed as a negative prognostic factor for HCC. In clinical practice, there is no consensus regarding the optimal cutoff value. For HCC, 0.2 is the proposed cutoff value for Ki-67. This difference may be due to different selected patients. More further studies are warranted to further identify the optimal cutoff value for Ki-67 in HCC.

In the present study, contrast-enhanced CT- based radiomics was adopted to assess Ki-67 in HCC with the aim of providing some guidance for future clinical practice. The results indicated that four radiomics features were independent predictive factors for Ki-67 expression. Clinical, radiomics and combined prediction models for predicting Ki-67 were constructed based on clinical indicators and radiomics characteristics. It was found that the combined prediction model had the best prediction efficiency and the clinical prediction model had the worst. The DeLong test found that the inclusion of radiomics features significantly improved the predictive efficacy of the Ki-67 prediction model. Consistent with this result, the analysis of the decision curve and the clinical impact curve revealed that the combined model had better clinical net benefits and clinical practicability. Radiomics features, which are quantitative imaging features obtained through artificial intelligence algorithms, are associated with key pathological and molecular characteristics, further supplementing prognostic information and enhancing decision-making capabilities.^{28,29} In our study, the inclusion of radiomics characteristics further consolidated the predictive efficacy. Notably, unlike prior studies that directly constructed radscore using data derived solely from LASSO regression,^{30,31} our approach further implemented

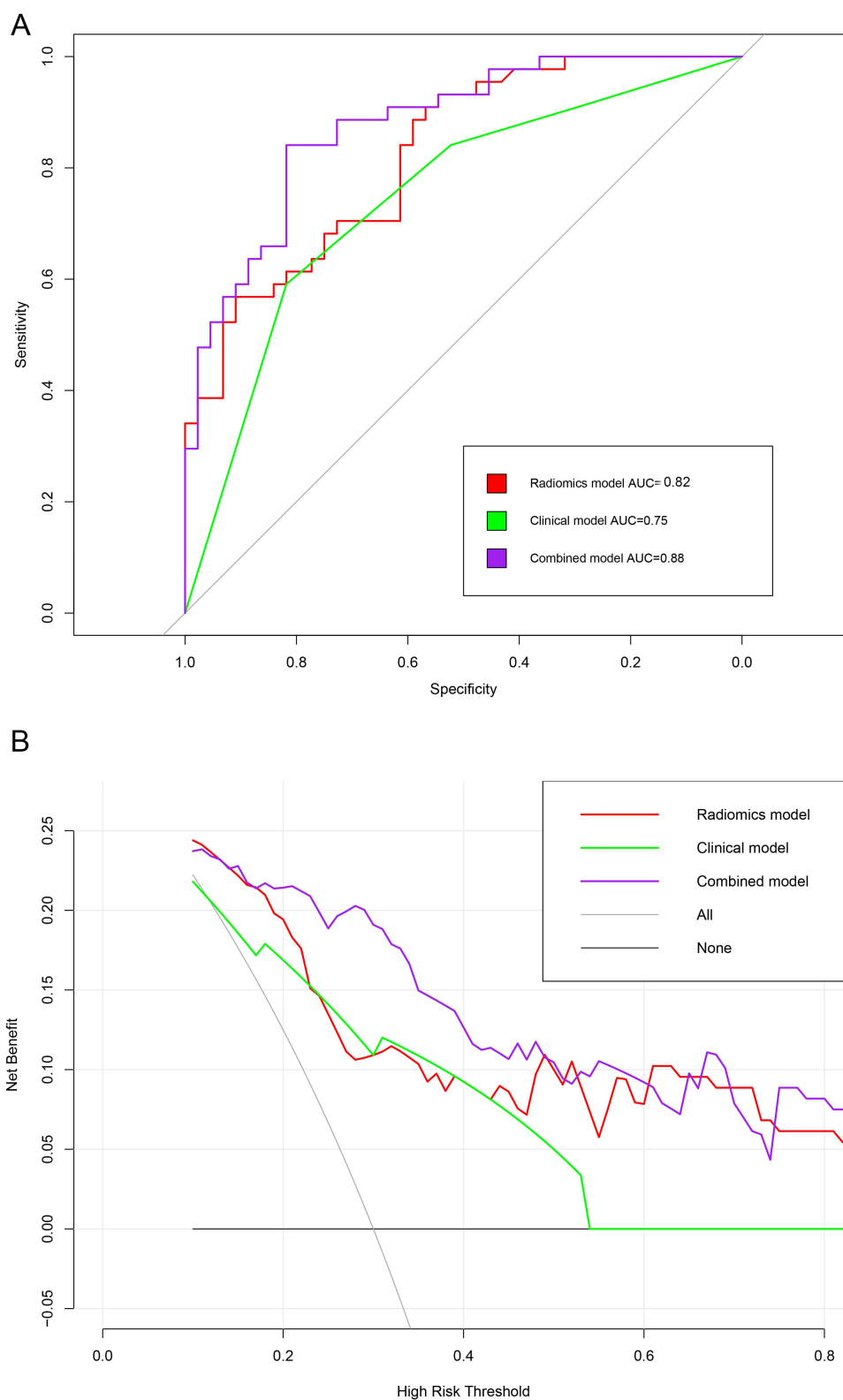


Figure 6 Evaluation of predictive efficacy of different models. **(A)** ROC curves of radiomics model, clinical models and combined model. **(B)** Decision curves of the radiomics model, clinical models and combined model.

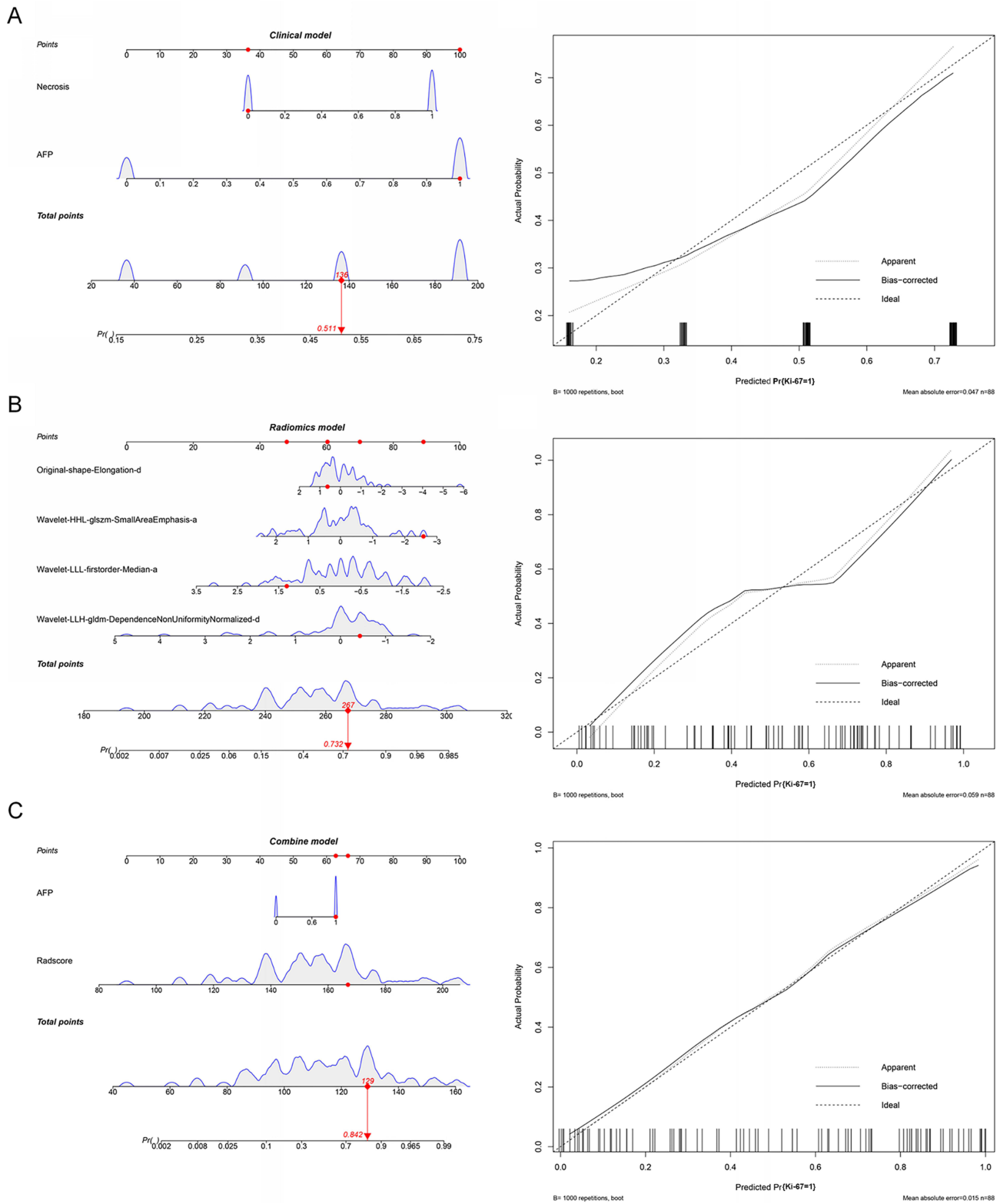


Figure 7 Nomogram and calibration curves of the predictive models for Ki-67. **(A)** radiomics model **(B)** clinical model **(C)** combined model. Red colored arrows: the probability corresponding to the total score obtained by adding up the scores of each red dot.

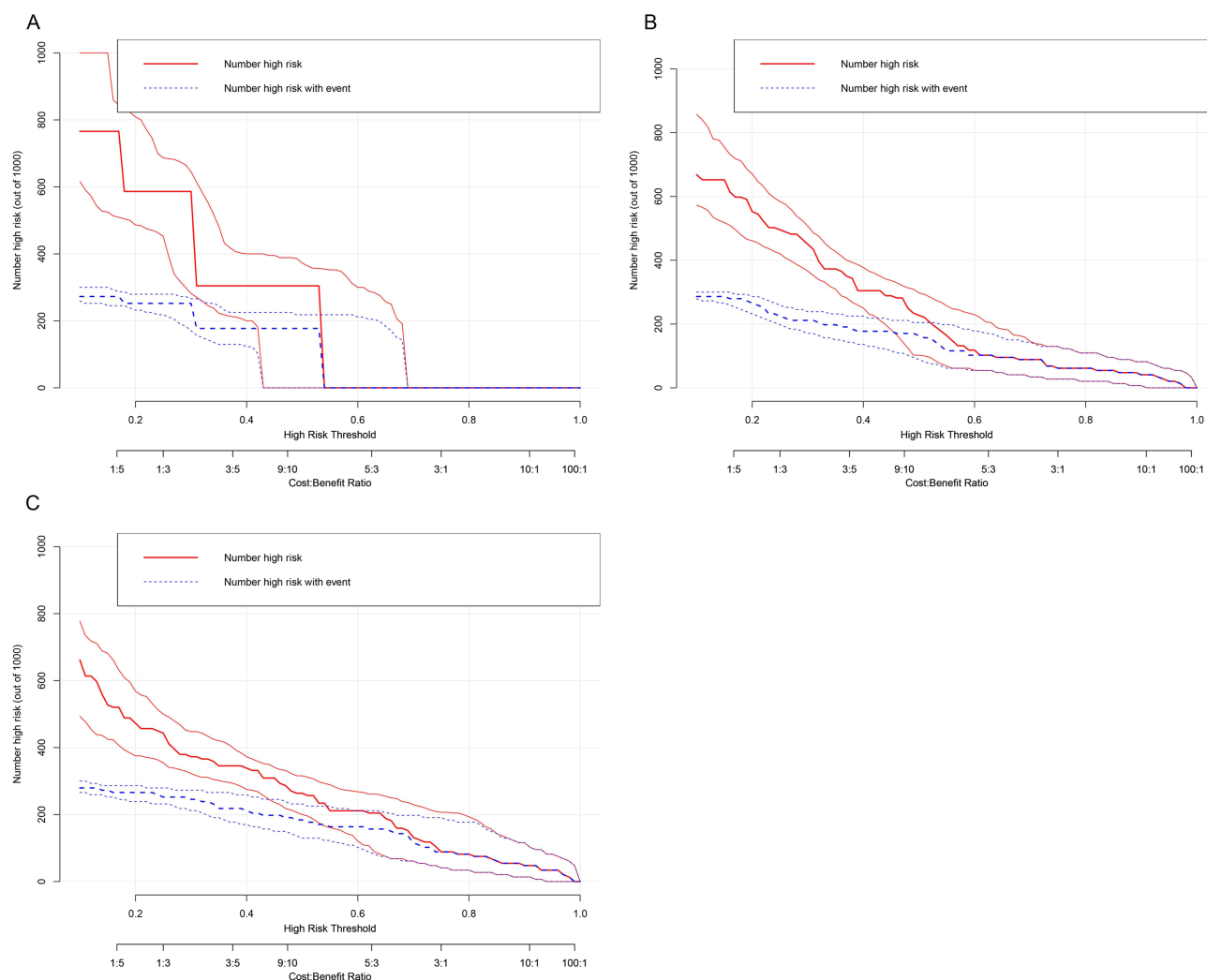


Figure 8 Clinical impact curves of different models. **(A)** radiomics model **(B)** clinical model **(C)** combined model.

multivariate analysis to identify Ki-67-independent radiomics features selected by LASSO regression, thereby reducing the risk of model overfitting. Considering all the results mentioned above, radiomics based on contrast-enhanced CT could well predict Ki-67 expression.

In this study, we combined qualitative assessment (intratumor necrosis) and quantitative assessment (the radiomics) in constructing the predictive model. Additionally, the model was a non-invasive method predicting Ki-67 expression. Furthermore, unlike previous studies, we adopted the maximum selection rank statistics to identify the optimal cutoff value for Ki-67 expression and it was further proved that high expression of Ki-67 was significantly associated with survival. The maximum selection rank statistics was more reasonable than conventional methods in identifying the optimal cutoff value. Additionally, the clinical practicability of different models in predicting Ki-67 expression was assessed by the clinical impact curve analysis. Thus, we assessed the predictive efficiency of this model adopting different statistical analyses, which was also an advantage of this study.

The limitations of this study are as follows. First, this study was a retrospective one in nature and selection biases were not absolutely avoidable. In the future, prospective studies are needed to further verify the conclusions of our study. Second, only patients from one center were included, and external validation could not be performed at present. In the future, we will perform a multi-centered study to verify conclusions of our study. Third, only 88 patients were included, and studies including a larger study population were needed. However, there are also some other studies including a relatively small study population. In the study by Philipp Schindler et al, the authors established an integrated model

that combines CT-based radiomics and imaging biomarkers with clinical variables to predict recurrence and recurrence-free survival in patients with HCC following liver transplantation (LT) and in this study, only 123 patients were included.²⁸ Additionally, Gulizaina Hapaer et al reported a radiomics-based model that could efficiently predict PD1 status and they only included 105 patients in this study.²⁷ Fourth, only CT imaging data were utilized without incorporating other imaging data. Multi-modal fusion may further enhance the efficacy of the analysis. Fifth, most patients in China were concurrently diagnosed with hepatitis B virus infection, which was less common among patients from western countries, suggesting that conclusions of this study should be further verified by studies carried out in western countries before its application among patients from these countries.

Overfitting is one of the problems related to a small sample size. In this study, we tried our best to minimize overfitting. LASSO analysis and subsequent multivariate regression analysis were performed to reduce overfitting. The causes of overfitting included the following three ones. The first one was a small sample size, and we may need to perform further studies including a larger sample size to reduce overfitting. The second one was an over complicated model. However, in our study, the predictive model was not over complicated. Thus, overfitting related to an over complicated model was not a concern. Third, the predictive model does not reflect the actual situation, which may lead to overfitting. In this study, we constructed a practical model predicting the Ki-67 expression and it was proven as a model with good predictive efficiency, a favorable agreement with the actual Ki-67 expression and a high clinical practicability. Thus, overfitting related to the third condition was not concern. Fourth, noisy data may also lead to overfitting. However, there are no absolutely efficient methods distinguishing noisy data from normal ones. Overfitting related to noisy data could not be absolutely eliminated. In the future, prospective multi-centered studies including a larger study population would solve most of the aforementioned limitations, especially overfitting related to a small sample size.

Conclusion

We adopted a prognostic Ki-67 cutoff value, which may be more clinically rational and better guide clinical decision-making. The present study demonstrates that incorporating radiomics features significantly improves the predictive performance of the Ki-67 prediction model. Compared with the clinical prediction model and radiomics prediction model, the integrated model combining clinical indicators and radiomics features exhibits superior predictive performance and clinical utility. However, due to the retrospective nature, the small sample size and lack of external validation, we still need to perform multi-centered prospective studies including a larger study population to further verify findings of our study.

Data Sharing Statement

The datasets are not publicly available due to patient privacy concerns but are available from Tailai An upon reasonable requests.

Ethics Approval and Informed Consent

This study was approved by the Ethics Committee (LL-KY-2024111-02) and all the patients included in this study had given his or her informed consent in the written form. Additionally, all methods were carried out in accordance with the Declaration of Helsinki and other relevant guidelines and regulations.

Acknowledgments

The authors appreciate Xue Tang from Shenzhen People's Hospital for her help in collecting data. An abstract of this work was presented as a poster at the 2024 IHPBA conference (abstract #419). The abstract did not include predictive model construction. This manuscript presents a significantly expanded analysis, including the development and validation of predictive models, and has not been previously published in a journal.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically

reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by Guangdong Basic and Applied Basic Research Foundation (2023A1515110688, 2024A1515220011), Guangdong Medical Science and Technology Research Foundation (A2024025), National Natural Science Foundation of China (82572375).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. *Nat Rev Clin Oncol*. 2023;20(12):864–884. doi:10.1038/s41571-023-00825-3
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
3. Liu L, Qin S, Lin K, et al. Development and comprehensive validation of a predictive prognosis model for very early HCC recurrence within one year after curative resection: a multicenter cohort study. *Int J Surg*. 2024;110(6):3401–3411. doi:10.1097/JS9.0000000000001467
4. Wei H, Fu F, Jiang H, et al. Development and validation of the OSASH score to predict overall survival of hepatocellular carcinoma after surgical resection: a dual-institutional study. *Eur Radiol*. 2023;33(11):7631–7645. doi:10.1007/s00330-023-09725-7
5. Menon SS, Guruvayoorappan C, Sakthivel KM, et al. Ki-67 protein as a tumour proliferation marker. *Clin Chim Acta*. 2019;491:39–45. doi:10.1016/j.cca.2019.01.011
6. Shi W, Hu J, Zhu S, et al. Expression of MTA2 and Ki-67 in hepatocellular carcinoma and their correlation with prognosis. *Int J Clin Exp Pathol*. 2015;8(10):13083–13089.
7. Wang H, Yu H, Qian YW, et al. Postoperative adjuvant transcatheter arterial chemoembolization improves the prognosis of patients with huge hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int*. 2021;20(3):232–239. doi:10.1016/j.hbpd.2020.12.018
8. Xu JX, Xing WT, Peng YC, et al. Outcomes of postoperative adjuvant transarterial chemoembolization for hepatocellular carcinoma according to the Ki67 index. *Future Oncol*. 2022;18(17):2113–2125. doi:10.2217/fon-2021-1443
9. Farris AB, Dursun N, Dhanasekaran R. Tumoral and angiogenesis factors in hepatocellular carcinoma after locoregional therapy. *Pathol Res Pract*. 2012;208(1):15–21. doi:10.1016/j.prp.2011.10.005
10. Singla S, LeVea CM, Pokuri VK, et al. Ki67 score as a potential predictor in the selection of liver-directed therapies for metastatic neuroendocrine tumors: a single institutional experience. *J Gastrointest Oncol*. 2016;7(3):441–448. doi:10.21037/jgo.2016.02.02
11. Liu Z, Wang S, Dong D, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics*. 2019;9(5):1303–1322. doi:10.7150/thno.30309
12. Bo Z, Song J, He Q, et al. Application of artificial intelligence radiomics in the diagnosis, treatment, and prognosis of hepatocellular carcinoma. *Comput Biol Med*. 2024;173:108337. doi:10.1016/j.compbiomed.2024.108337
13. Fang G, Fan J, Ding Z, et al. Application of biological big data and radiomics in hepatocellular carcinoma. *ILIVER*. 2023;2(1):41–49. doi:10.1016/j.iliver.2023.01.003
14. Xia TY, Zhou ZH, Meng XP, et al. Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model. *Radiology*. 2023;307(4):e222729. doi:10.1148/radiol.222729
15. Fu J, Cao SJ, Song L, et al. Radiomics/Radiogenomics in hepatocellular carcinoma: applications and challenges in interventional management. *ILIVER*. 2022;1(2):96–100. doi:10.1016/j.iliver.2022.07.001
16. Wu M, Yang S, Feng X, et al. Combining preoperative and postoperative inflammatory indicators can better predict the recurrence of hepatocellular carcinoma after partial hepatectomy. *J Inflamm Res*. 2021;14:3231–3245. doi:10.2147/JIR.S316177
17. Health Commission of the People's Republic of China N. National guidelines for diagnosis and treatment of breast cancer 2022 in China (English version). *Chin J Cancer Res*. 2022;34(3):151–175. doi:10.21147/j.issn.1000-9604.2022.03.02
18. Zheng BH, Liu LZ, Zhang ZZ, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. *BMC Cancer*. 2018;18(1):1148. doi:10.1186/s12885-018-5024-z
19. Ganesan P, Kulik LM. Hepatocellular carcinoma: new developments. *Clin Liver Dis*. 2023;27(1):85–102. doi:10.1016/j.cld.2022.08.004
20. Zhou H, Song T. Conversion therapy and maintenance therapy for primary hepatocellular carcinoma. *Biosci Trends*. 2021;15(3):155–160. doi:10.5582/bst.2021.01091
21. Zhang A, Wang X, Fan C, et al. The role of Ki67 in evaluating neoadjuvant endocrine therapy of hormone receptor-positive breast cancer. *Front Endocrinol*. 2021;12:687244. doi:10.3389/fendo.2021.687244
22. Chen J, Easwaralingam N, Warriar S, et al. Window of opportunity treatment in breast cancer. *ANZ J Surg*. 2020;90(1–2):34–40. doi:10.1111/ans.15487
23. Kaya R, Takanashi H, Nakajima A, et al. Prognostic significance of Ki67 during neoadjuvant chemotherapy in primary unresectable ovarian cancer. *J Obstet Gynaecol Res*. 2021;47(11):3979–3989. doi:10.1111/jog.14981
24. Liu Y, Gu Q, Xiao Y, et al. Prognostic value of Ki67 in epithelial ovarian cancer: post-neoadjuvant chemotherapy Ki67 combined with CA125 predicting recurrence. *Cancer Manag Res*. 2024;16:761–769. doi:10.2147/CMAR.S469132
25. Li HH, Qi LN, Ma L, et al. Effect of KI-67 positive cellular index on prognosis after hepatectomy in Barcelona clinic liver cancer stage A and B hepatocellular carcinoma with microvascular invasion. *Onco Targets Ther*. 2018;11:4747–4754. doi:10.2147/OTT.S165244

26. Hu X, Zhou J, Li Y, et al. Added value of viscoelasticity for MRI-based prediction of Ki-67 expression of hepatocellular carcinoma using a deep learning combined radiomics (DLCR) model. *Cancers*. 2022;14(11):2575. doi:10.3390/cancers14112575
27. Zhao F, Zhao J, Wei X, et al. Predicting Abiraterone efficacy in advanced prostate cancer: insights from marker of proliferation Ki67. *Prostate*. 2024;84(10):932–944. doi:10.1002/pros.24710
28. Wei J, Jiang H, Zhou Y, et al. Radiomics: a radiological evidence-based artificial intelligence technique to facilitate personalized precision medicine in hepatocellular carcinoma. *Dig Liver Dis*. 2023;55(7):833–847. doi:10.1016/j.dld.2022.12.015
29. Yao S, Ye Z, Wei Y, et al. Radiomics in hepatocellular carcinoma: a state-of-the-art review. *World J Gastrointest Oncol*. 2021;13(11):1599–1615. doi:10.4251/wjgo.v13.i11.1599
30. Hapaer G, Che F, Xu Q, et al. Radiomics-based biomarker for PD-1 status and prognosis analysis in patients with HCC. *Front Immunol*. 2025;16:1435668. doi:10.3389/fimmu.2025.1435668
31. Schindler P, von Beauvais P, Hoffmann E, et al. Combining radiomics and imaging biomarkers with clinical variables for the prediction of HCC recurrence after liver transplantation. *Liver Transpl*. 2025;31(10):1226–1237. doi:10.1097/LVT.000000000000060

Journal of Hepatocellular Carcinoma

Publish your work in this journal

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

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

Dovepress
Taylor & Francis Group