Original Article

Computed tomography radiomics signature via machine learning predicts *RRM2* and overall survival in hepatocellular carcinoma

Qian Li[1]^, Xiawei Long[1], Yan Lin[1]^, Rong Liang[1]^, Yongqiang Li[1], Lianying Ge[2]

[1]Department of Medical Oncology, Guangxi Medical University Cancer Hospital, Nanning, China; [2]Department of Endoscopy, Guangxi Medical University Cancer Hospital, Nanning, China

Contributions: (I) Conception and design: Q Li, L Ge; (II) Administrative support: Y Li; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Q Li, X Long; (V)Data analysis and interpretation: Q Li, X Long, Y Lin, R Liang; (VI)Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Lianying Ge, MD. Department of Endoscopy, Guangxi Medical University Cancer Hospital, Hedi Road No. 71, Nanning 530021, China. Email: gelianying2022@163.com.

Background: Radiomics can be used to noninvasively predict molecular markers to address the clinical dilemma that some patients cannot accept invasive procedures. This research evaluated the prognostic significance of the expression level of ribonucleotide reductase regulatory subunit M2 (*RRM2*) in individuals with hepatocellular carcinoma (HCC) and established a radiomics model for predicting the *RRM2* expression level.

Methods: Genomic data for HCC patients and corresponding computed tomography (CT) images were accessed at The Cancer Genome Atlas (TCGA) and The Cancer Imaging Archive (TCIA), which were utilized for prognosis analysis, radiomic feature extraction and model construction, respectively. The maximum relevance minimum redundancy algorithm (mRMR) and recursive feature elimination (RFE) were used for feature selection. Following feature extraction, a logistic regression algorithm was fitted to establish a dichotomous model that predicts *RRM2* gene expression. Establishment of the radiomics nomogram was carried out using the Cox regression model. Receiver operating characteristic (ROC) curve analysis was employed to assess the model performance. Clinical utility was determined by decision curve analysis (DCA).

Results: High *RRM2* expression acted as a risk factor for overall survival (OS) [hazard ratio (HR) =2.083, P<0.001] and was implicated in regulation of the immune response. Four optimal radiomics features were selected for prediction of *RRM2* expression. A predictive nomogram was established using the clinical variables and radiomics score (RS), and the areas under the ROC curve (AUCs) of the time-dependent ROC curve of the model were 0.836, 0.757, and 0.729 for the 1-, 3-, and 5-year periods, respectively. DCA confirmed that the nomogram had good clinical usefulness.

Conclusions: The *RRM2* expression level in HCC can considerably affect prognosis of these patients. Expression levels of *RRM2* and prognosis of HCC individuals can be predicted through radiomics features by utilizing CT scan data.

Keywords: Hepatocellular carcinoma (HCC); ribonucleotide reductase regulatory subunit M2 (*RRM2*); prediction model; radiomics

Submitted May 29, 2023. Accepted for publication Jun 19, 2023. Published online Jun 25, 2023.
doi: 10.21037/jgo-23-460
View this article at: https://dx.doi.org/10.21037/jgo-23-460

^ ORCID: Qian Li, 0000-0003-3332-0142; Yan Lin, 0000-0001-8034-5946; Rong Liang, 0000-0003-0343-3277.

Introduction

Globally, liver cancer is the sixth most prevalent malignancy and the third leading contributor to cancer death. Almost 75–85% of confirmed cases of primary liver cancer are hepatocellular carcinoma (HCC) (1). According to the Barcelona Clinic Liver Cancer (BCLC) staging and therapy approach, surgical resection is the most efficient option to improve prognoses in very early (0) and early-stage (a) HCC (2). However, HCC is highly aggressive, resulting in high postoperative recurrence frequency. Furthermore, the therapeutic options for advanced-stage patients are still limited. In September 2017, nivolumab, utilized for second-line treatment of advanced HCC, was approved by the US Food and Drug Administration (3), marking entry of systemic treatment for HCC into the era of immunotherapy. However, in the real world, the effect of immunotherapy is still modest. Screening patients who will benefit from immunotherapy through effective molecular markers is a breakthrough in treatment of advanced HCC. Hence, more effective predictive indicators are urgently required to improve the landscape of HCC diagnosis and management.

Ribonucleotide reductase regulatory subunit M2 (*RRM2*) is one of the two different subunits of ribonucleotide reductase. This enzyme is involved in catalysis of ribonucleotides to deoxyribonucleotides (4). *RRM2* is involved in providing precursors required for DNA synthesis and participates in inhibition of Wnt signaling. The signaling pathways in which *RRM2* is involved include *de novo* synthesis of adenosine DNA, the E2 factor (E2F) transcription factor network, activity of fluoropyrimidines, *de novo* biosynthesis of guanosine DNA, the cell cycle, and others (5). Data show that *RRM2* levels are substantially greater in HCC tissues than in non-HCC tissues and that *RRM2* small interfering RNA (siRNA) inhibits cell proliferation in HCC (6). Hepatitis B virus (HBV) infection is a pathogenic factor of HCC. HBV can activate *RRM2* expression to replicate viral DNA in host hepatocytes (7). Overexpression of *RRM2* enhances the proliferation and migration ability of Hep3B and Huh7 cells (8). These data indicate the potential of *RRM2* as an indicator for prognosis of HCC patients receiving cancer therapy.

Employing dynamic detection and quantitative response to tumor features, the high-throughput "image sequencing" technology radiomics is a noninvasive process predominantly utilized at the clinical level. Radiomics is an emerging field of research aimed at extracting quantitative imaging features from radiological images for expression of disease signatures. These radiological features may not be detectable by the radiologist through visual assessment. It has been reported that radiomics can predict the molecular classification of many disease outcomes. Quantitative radiomics has shown potential as a noninvasive imaging biomarker to characterize tumor diagnosis, progression/prognosis, and treatment response. Prior research has shown its effectiveness in early diagnosis and classification of HCC, as well as in assessment of the microenvironment and heterogeneity of a tumor (9,10).

This research focused on noninvasively predicting messenger RNA (mRNA) expression of *RRM2* in HCC tissues based on computed tomography (CT) radiomics and evaluating the association between our model and prognosis. We present this article in accordance with the TRIPOD reporting checklist (available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-460/rc).

Highlight box

Key findings

- Radiomics features derived from CT could efficiently predict the expression level of *RRM2*.

What is known and what is new?

- *RRM2* is a potential marker of prognosis in HCC. High expression level of *RRM2* was significantly correlated with a poor prognosis in patients with HCC.
- High expression of *RRM2* was associated with increased infiltration of M0 macrophages in HCC.

What is the implication, and what should change now?

- The radiomics prediction model of *RRM2* could benefit personalized clinical decision-making.

Methods

Data and image sources

Original clinicopathological and transcriptomics data were retrieved from The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov/); CT data were obtained from The Cancer Imaging Archive (TCIA; https://www.cancerimagingarchive.net/). The images and data retrieved from TCIA and TCGA are anonymous and available publicly. Each patient in TCIA was matched with those in TCGA via a unique ID number. Clinicopathological and transcriptomic data for 377 HCC patients from TCGA and CT images for 75 HCC patients from TCIA (accessed in October 2022) were retrieved. TCGA data were used for prognosis and immune infiltration analysis; TCIA images

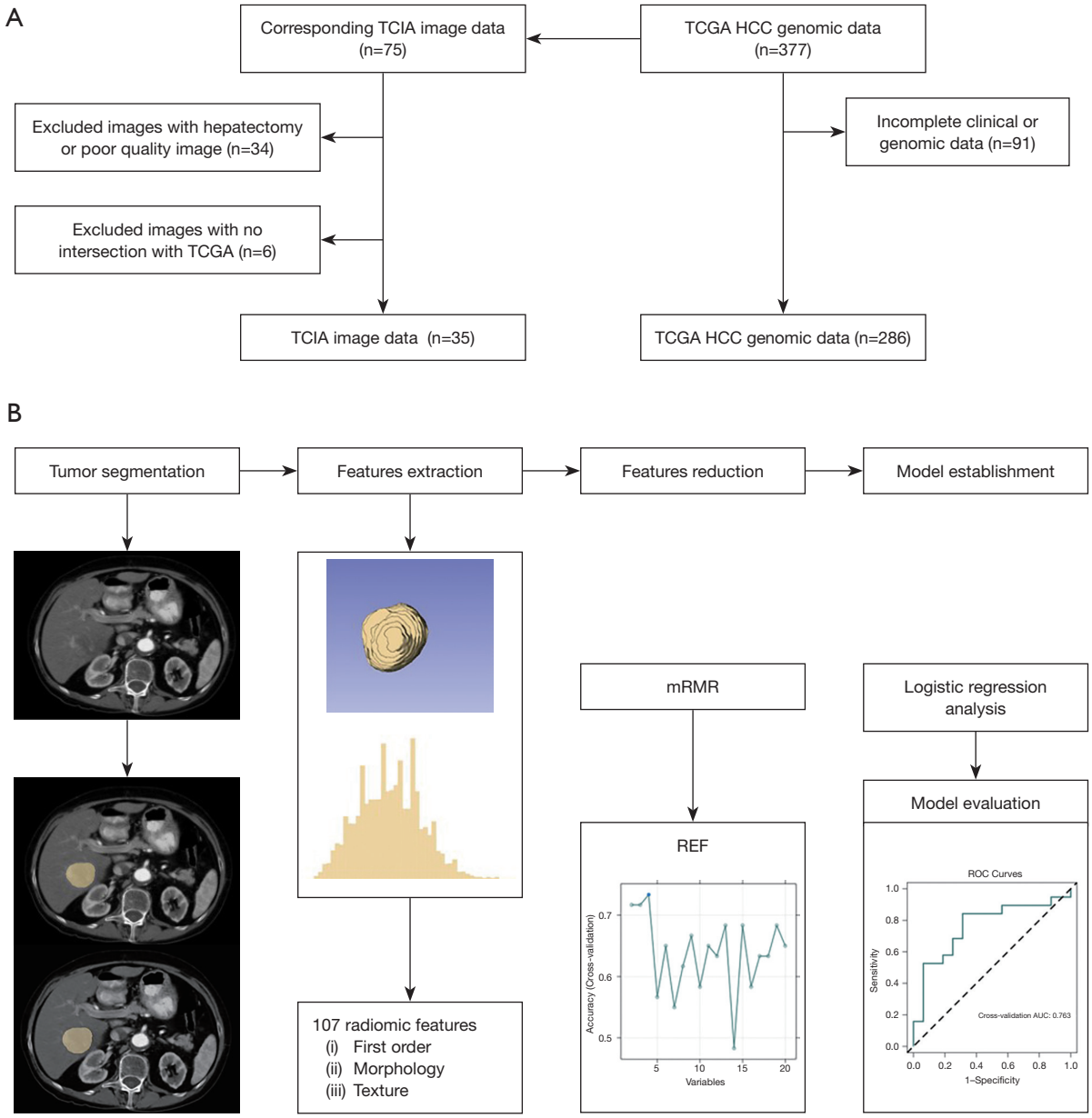

Figure 1 The brief flowchart of data collection and analysis. (A) Gene and image data screening process. (B) Brief flowchart of radiomic progression. TCIA, The Cancer Imaging Archive; TCGA, The Cancer Genome Atlas; HCC, hepatocellular carcinoma; mRMR, maximum relevance minimum redundancy algorithm; RFE, recursive feature elimination; ROC, receiver operating characteristic; AUC, area under the curve.

were utilized for extraction of radiologic features and establishment of a model. The criteria applied for exclusion of TCGA data included lack of survival data, survival time less than 30 days, HCC not as an initial diagnosis, lack of clinical data, and lack of RNA-sequencing (RNA-seq) data for the primary tumor. Exclusion criteria for TCIA data included no intersection with TCGA, CT images from tumor resection patients, and poor CT image quality. Finally, 286 patients from TCGA and 35 from TCIA were enrolled in this study. The brief flowchart is shown in *Figure 1A*. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Survival analysis and Cox regression analysis

The RNA-seq data files utilized in this analysis originated from UCSC Xena (https://xenabrowser.net/datapages/). Data in the format of fragments per kilobase of exon model per million mapped fragments (FPKM) were processed using the Xiantao tool (https://www.xiantao.love/products) through the Toil process, and differential expression analysis between samples was performed after $\log_2$ transformation. Statistical significance was established at P<0.05. Genetic data were combined with corresponding clinical data and processed through the 'survminer' package in R software to obtain the threshold value of *RRM2* expression, and the patients were classified into two expression groups (high and low). Kaplan-Meier analysis was employed to analyze overall survival (OS) in the two groups, draw a survival curve, and perform the log-rank test. Cox proportional hazard analyses (univariate and multivariate) were conducted to assess prognostic factors, including *RRM2* gene expression levels and clinical characteristics such as sex and age. Hazard ratios (HRs) and relevant 95% confidence intervals (CIs) were estimated. Interactions between *RRM2* expression and other variables included in the univariate Cox model were analyzed with the likelihood ratio test, and subgroup analyses were performed. Correlation analysis between the main variable *RRM2* and the clinical characteristics of the tumors was performed using Spearman's rank correlation coefficient.

Association between RRM2 and immune cell infiltration, immune genes, and enrichment analysis of differentially expressed genes

Gene expression matrices of HCC samples were uploaded to the CIBERSORTx database (https://cibersortx.stanford.edu/), and immune cell infiltration was calculated for each sample. Differential analysis of the immune cell infiltration level between the *RRM2* high- and low-expression groups was performed using the R package 'limma'. Correlation analysis between *RRM2* and immune genes was performed using Spearman's rank correlation coefficient. The gene expression profiles between the high- and low-risk cohorts were subjected to gene set enrichment analysis (GSEA) utilizing the gene sets Hallmark (h.all.v7.5.1.symbols.gmt) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (c2.cp.kegg.v7.5.1.symbols.gmt). GSEA was performed with the R package 'clusterprofiler'. Statistical significance was established at P<0.05.

Construction and evaluation of whole tumor and whole tumor & peritumoral models

A radiomic flowchart of this study is briefly shown in *Figure 1B*. By employing 3D slicer software (v4.10.2), the tumor area was manually outlined to obtain the whole tumor area; using the simpleITK package from Python (https://simpleitk.org/), 3 mm of automatic outward expansion on the basis of the whole tumor area was used to obtain the whole tumor and peritumoral areas. Extraction of radiomics features was achieved through the 'pyradiomics' package in Python and data normalization. The number of extracted radiomic features was 107, which included the following: first-order features (max, 10th percentile, etc.) morphological features (sphericity, elongation, etc.) and textural features [gray level zone matrix (GLSZM), neighborhood gray level difference matrix (NGTDM), gray level co-occurrence matrix (GLCM), and gray level run-length matrix (GLRLM), etc.]. Stabilized features were entered into the 'maximum relevance minimum redundancy algorithm (mRMR)' using the R package 'mRMRe', which aims to find a subset of correlated and complementary features to eliminate irrelevant and redundant features. Recursive feature elimination (RFE) feature screening was then implemented with the aid of the 'caret' R package, with the goal of finding a subset of predictors to be utilized for developing accurate models. The 20 leading features were selected through the mRMR method, and the best feature subset was further screened by RFE.

Intraclass correlation coefficients (ICCs) were calculated to evaluate the consistency of extracted radiomics features delineated separately by two physicians. After complete case delineation by one physician, ten samples were randomly chosen utilizing the 'random number table method' for delineation by another physician, along with subsequent extraction of their radiomics features. ICC ≥0.75 is generally regarded as good agreement, 0.51–0.74 as moderate, and lower than 0.50 as poor. Features with ICCs ≥0.75 were included.

A logistic regression model was developed according to the established optimal feature subsets. Logistic regression was conducted on the basis of linear regression. The sigmoid function was compounded with the following formula:

$$g(z) = y[x] = \frac{1}{1+e^{-\omega^T x}} \quad [1]$$

A logistic regression algorithm was fitted to the above filtered radiomic features using the R package 'stats' to build

a dichotomous model for predicting *RRM2* gene expression.

Receiver operating characteristic (ROC) curve analysis was employed to assess the model performance, which was revealed by calculating its specificity, sensitivity, and area under the ROC curve (AUC). The performance of the model was verified by AUC values from 5-fold cross-validation. Plotting of calibration curves and the Hosmer-Lemeshow goodness-of-fit test were conducted to assess calibration of the radiomics prediction models. Finally, the clinical utility of radiological evaluation was determined by decision curve analysis (DCA).

Differential analysis of RRM2 expression between whole tumor and whole tumor & peritumor model groups

The whole tumor and whole tumor & peritumoral radiomics models output a probabilistic radiomics score (RS) that predicted the expression level of *RRM2*. Afterward, the image omics marker RS was compared by the Wilcoxon test between the groups with high and low *RRM2* expression.

Construction and evaluation of a nomogram

After taking the intersection of TCIA-TCGA clinical samples, the cutoff value of RS was determined by the R package 'survminer', which was divided into low/high dichotomous variables. The associated clinical indicators and RS were utilized to establish univariate and multivariate Cox regression models. Variables were obtained by the minimal Akaike information criterion (AIC)-based stepwise selection method to construct nomograms for 1-, 2-, and 3-year survival probabilities from Cox regression.

The predictive ability of the factors at various time points was depicted by plotting the time-dependent ROC curve of the predictive model. A calibration curve was plotted with the abscissa as the predicted survival, ordinate as the actual survival, and diagonal as the predicted probability equal to the actual probability; deviations from the diagonal illustrated greater error in prediction. DCA of the prediction model was conducted to assess its clinical benefit.

Statistical analysis

All data were analyzed with the aid of R software v4.0.2 and python software v3.6.6. Quantitative data are presented as the mean ± standard deviation (SD) or median and interquartile range, and the groups were comparatively analyzed through Student's *t*-test or the Wilcoxon test, as appropriate. Categorical variables are presented as counts and percentages, and groups were comparatively analyzed using the chi-square test. Two-sided statistical tests were performed, and statistical significance was established at P<0.05. The 'survival' package of R software and Kaplan-Meier analysis were utilized to conduct survival analysis and to generate survival curves, respectively. The 'survival ROC', 'time ROC', 'RMS', and 'dcurves' packages were used to prepare the time-dependent ROC curve, calibration plot, DCA, and nomogram.

Results

Baseline characteristics of patients

The number of included HCC patients was 286, of which 196 (69%) were male and 90 (31%) were female. Baseline characteristics are shown in *Table 1*. Taking the optimal cutoff value of *RRM2* expression level (cutoff) =2.65938, the individuals were classified into a high-expression group (n=127) and a low-expression group (n=159). High *RRM2* expression correlated considerably with tumor grade (P<0.001) and hepatitis activity (P=0.028). There was no notable variation in age, sex, pathological stage, residual tumor, vascular invasion, and the distribution of alpha-fetoprotein (AFP) between the high- and low-expression groups.

Survival analysis and Cox regression analysis

In contrast with normal tissues, markedly elevated *RRM2* gene expression was detected in tumor tissues, with a median difference value of 2.039 between the two groups (1.747–2.330, P<0.001) (*Figure 2A*). The median OS was 2,486 days in the *RRM2* low-expression group and 1,685 days in the high-expression group. Kaplan-Meier curves revealed a strong association of high *RRM2* expression with worse OS (P<0.001) (*Figure 2B*).

Univariate Cox regression analysis showed that *RRM2* gene expression, pathological stage, residual tumor, vascular invasion, hepatitis activity, and AFP level were prognostic factors for OS (*Figure 2C*). High *RRM2* expression was a risk factor for OS (HR =2.083, P<0.001). In multivariate analysis, high *RRM2* expression (HR =2.038, P=0.002), pathological stage G3/G4 stage (HR =2.033, P=0.002), and residual tumor R1/R2/RX (HR =2.106, P=0.041) were risk factors for OS after multivariate adjustment. However, male sex (HR =0.588, P=0.018) acted as a protective factor for OS (*Figure 2C*).

Table 1 Characteristics of patients in TCGA

Variables	Total (n=286)	*RRM2*		P
		Low (n=159)	High (n=127)	
Age (years), n [%]				0.086
<60	138 [48]	69 [43]	69 [54]	
≥60	148 [52]	90 [57]	58 [46]	
Sex, n [%]				0.156
Female	90 [31]	44 [28]	46 [36]	
Male	196 [69]	115 [72]	81 [64]	
Tumor grade, n [%]				<0.001
G1/G2	176 [62]	112 [70]	64 [50]	
G3/G4	110 [38]	47 [30]	63 [50]	
Pathologic stage, n [%]				0.053
I/II	213 [74]	126 [79]	87 [69]	
III/IV	73 [26]	33 [21]	40 [31]	
Residual tumor, n [%]				0.744
R0	264 [92]	148 [93]	116 [91]	
R1/R2/RX	22 [8]	11 [7]	11 [9]	
Vascular invasion, n [%]				0.178
None	160 [56]	96 [60]	64 [50]	
Unknown	40 [14]	18 [11]	22 [17]	
Micro/macro	86 [30]	45 [28]	41 [32]	
Hepatic inflammation, n [%]				0.028
None	97 [34]	63 [40]	34 [27]	
Unknown	92 [32]	42 [26]	50 [39]	
Mild/severe	97 [34]	54 [34]	43 [34]	
AFP, n [%]				0.607
<400	153 [53]	88 [55]	65 [51]	
≥400	73 [26]	41 [26]	32 [25]	
Unknown	60 [21]	30 [19]	30 [24]	

TCGA, The Cancer Genome Atlas; *RRM2*, ribonucleotide reductase regulatory subunit M2; AFP, alpha-fetoprotein.

Subgroup analysis and interaction testing

Subgroup analysis of the population indicated that increased expression of *RRM2* was a risk factor for OS in the <60-year-old age group (HR =2.201, P=0.009) and ≥60-year-old age group (HR =2.074, P=0.009). The data of the interaction test (P=0.83) showed the absence of any notable interaction between the expression level of *RRM2* and age; that is, the effect of *RRM2* on OS was similar between the two age groups. Similarly, the influence of *RRM2* expression level on OS was similar among sex, tumor grade, pathological stage, and with or without vascular invasion (*Figure 3A*). Correlation heatmaps showed that the *RRM2* expression level was considerably linked to tumor grade (P=0.00049) and pathological stage (P=0.038) (*Figure 3B*).

Association between RRM2 and immune cell infiltration, immune genes, and enrichment analysis of differentially expressed genes

HCC samples were assessed concerning immune cell infiltration through cibersort, and the degree of M0 macrophage infiltration was considerably elevated in the high *RRM2* expression group in comparison to the low *RRM2* expression group (P<0.05). The infiltration degrees of $CD8^+$ T cells, M1 macrophages, and M2 macrophages were not significantly different between the two groups (P>0.05) (*Figure 4A*). A heatmap showed a significant relationship between expression of *RRM2* and several immune genes, such as *LAG3*, *CD276*, *CD86*, *CD48*, and *TNFSF4* (P<0.001) (*Figure 4B*). Differential genes between the two *RRM2* expression groups (high and low) were subjected to GSEA enrichment analysis, and the first 20 pathways were visualized. Concerning the hallmark gene sets, considerable enrichment of the G2M checkpoint, bile acid metabolism, and xenobiotic metabolism pathways, among others, was found (*Figure 4C*). For KEGG gene sets, differentially expressed genes were considerably enriched in the cytochrome P450 pathway utilized for the metabolism of xenobiotics, retinol metabolism, and drug metabolism cytochrome P450 pathway (*Figure 4D*).

Construction and evaluation of whole tumor and whole tumor & peritumoral models

Marasco *et al.* studied the effect of excision margins on the recurrence rate of patients undergoing hepatectomy and found that a wide margin (>1 cm) had a better prognosis than a narrow margin (<1 cm), because micrometastases may exist around the 1cm tumor (11). Therefore, we believe that the establishment of whole-tumor and whole-tumor plus peritumoral models can better predict the prognosis of patients and the expression of some molecular markers.

Using the mRMR algorithm for feature screening among 107 radiomics features, the first 20 features were selected,

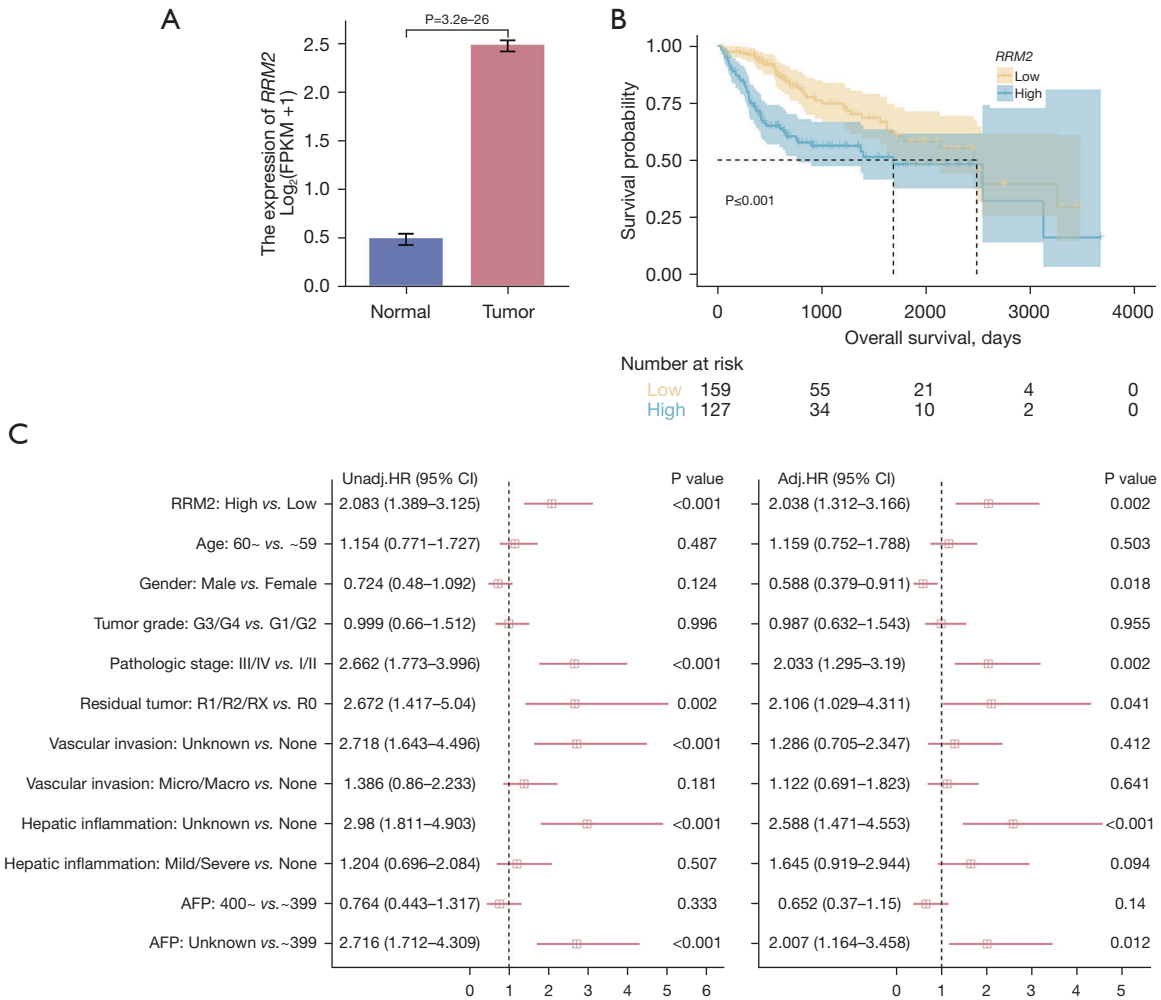

Figure 2 Comparison of clinical and survival data. (A) Comparison of *RRM2* expression level between healthy and tumor patients. (B) Kaplan-Meier curves of OS in patients with high/low *RRM2* expression levels. (C) Univariate and multivariate analyses for HCC patients. *RRM2*, ribonucleotide reductase regulatory subunit M2; FPKM, fragments per kilobase of exon model per million mapped fragments; AFP, alpha-fetoprotein; HR, hazard ratio; CI, confidence interval; OS, overall survival.

with four features being further screened out by RFE to construct the whole tumor model: radiomics formula = feature × corresponding coefficient (estimate) + intercept value (estimate). Specific characteristics and coefficients were utilized (*Figure 5A*). The median ICC value of features screened by the mRMR algorithm was 0.985, and there were 90 (84.1% of all features) radiomics features with ICC values ≥0.75. Four features of the whole tumor and peritumoral model were screened using the same method. Specific characteristics and coefficients were utilized (*Figure 5B*). The median ICC value of features was 0.983, with 100 radiomics signature ICC values ≥0.75 (93.5% of all signatures). The radiomics features selected by the whole tumor and whole tumor & peritumor models all had ICC values above 0.85.

Two logistic regression models were developed to predict the level of *RRM2* expression, and corresponding ROC curves were plotted and analyzed. The results indicated that the AUC of the whole tumor model was 0.776 (95% CI: 0.607–0.946); the AUC value of the 5-fold cross-validation was 0.763 (95% CI: 0.596–0.93) (*Figure 6A,6B*). The AUC of the whole tumor and peritumoral model was 0.803 (95% CI: 0.659–0.947), and its AUC value of the 5-fold cross-validation was 0.773 (95% CI: 0.609–0.937) (*Figure 6C,6D*). Based on the calibration curves and Hosmer-Lemeshow goodness of fit test of both models, the predicted probabilities of the models for *RRM2* expression were close

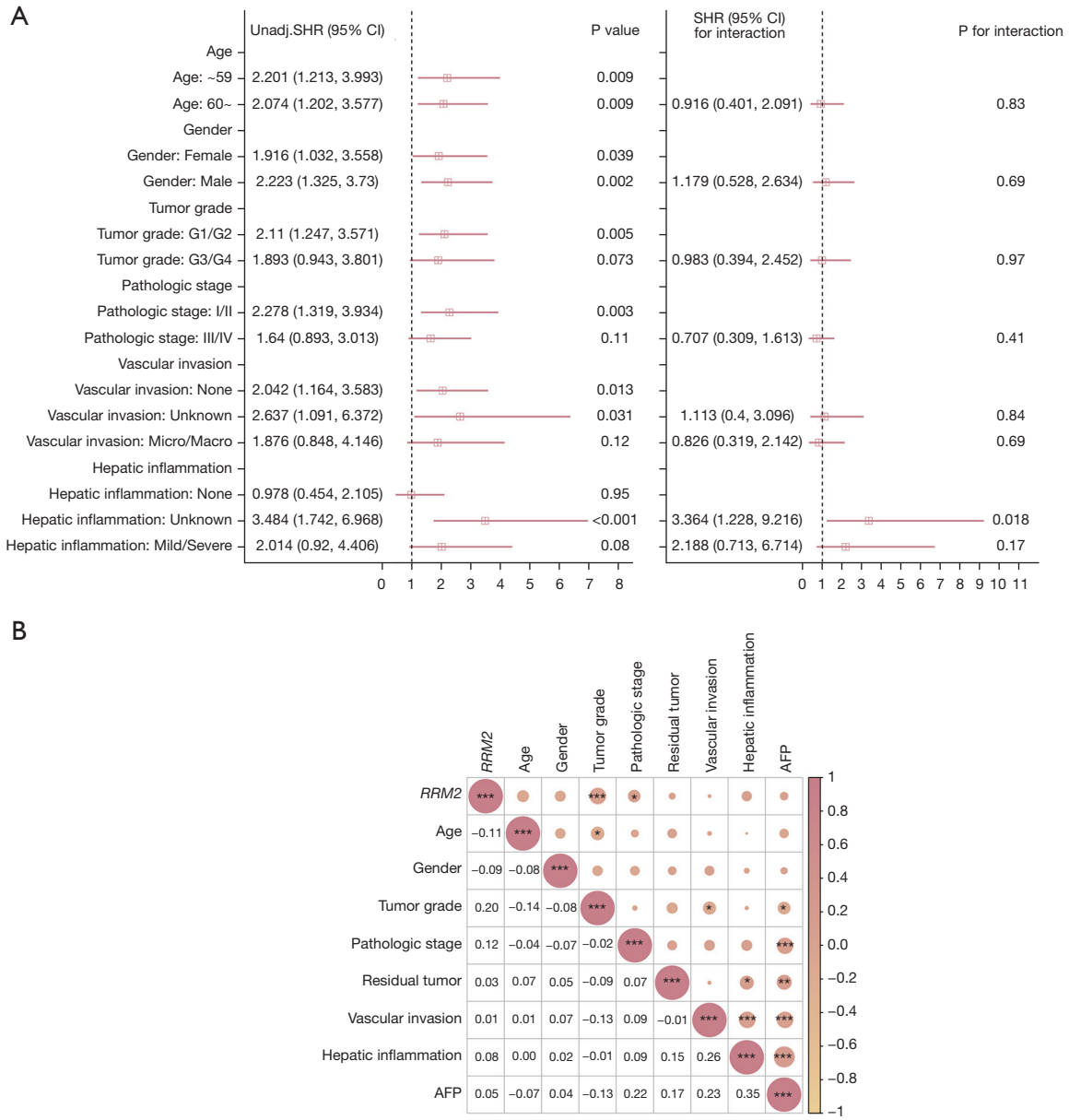

Figure 3 Subgroup analysis. (A) RRM2 expression level and the characteristics of the population. (B) *RRM2* expression level associated with clinical characteristics. *, P<0.05; **, P<0.01; ***, P<0.001. SHR, subdistribution hazard ratio; CI, confidence interval; *RRM2*, ribonucleotide reductase regulatory subunit M2; AFP, alpha-fetoprotein.

to the true values, indicating good agreement (P>0.05) (*Figure 6E,6F*). DCA showed high clinical utility for both models (*Figure 6G,6H*).

Following Delong's test, there was no significant variation in the whole tumor and whole tumor & peritumoral radiomics models. The respective P values in the training set and cross-validation set were 0.801 and 0.934. The AUC values of the whole tumor and peritumoral radiomics model were slightly increased in comparison to the whole tumor radiomics model. The latter was therefore chosen for subsequent analysis.

Differential analysis of RRM2 expression between whole tumor and whole tumor & peritumor model groups

In the whole tumor model, the probability RS to predict the

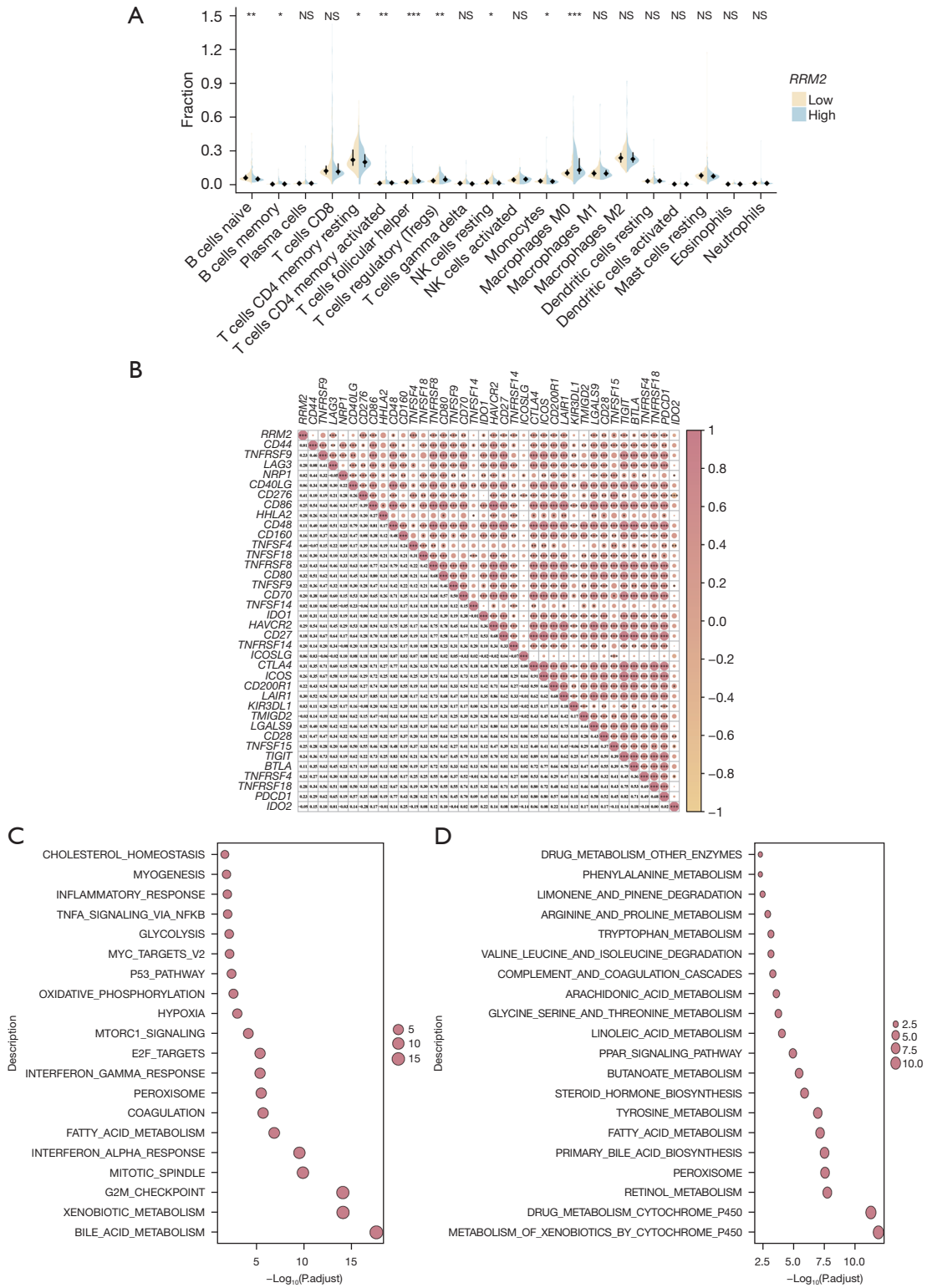

Figure 4 *RRM2* is closely related to immune regulation in the TME. (A) The infiltration level of various immune cells under different expression levels of RRM2 in HCC. (B) Expression of *RRM2* associated with immune-related genes. (C) Hallmark enrichment results show 20 paths. (D) KEGG pathways. *, P<0.05; **, P<0.01; ***, P<0.001; NS, not significant. *RRM2*, ribonucleotide reductase regulatory subunit M2; TME, tumor microenvironment; HCC, hepatocellular carcinoma; KEGG, Kyoto Encyclopedia of Genes and Genomes.

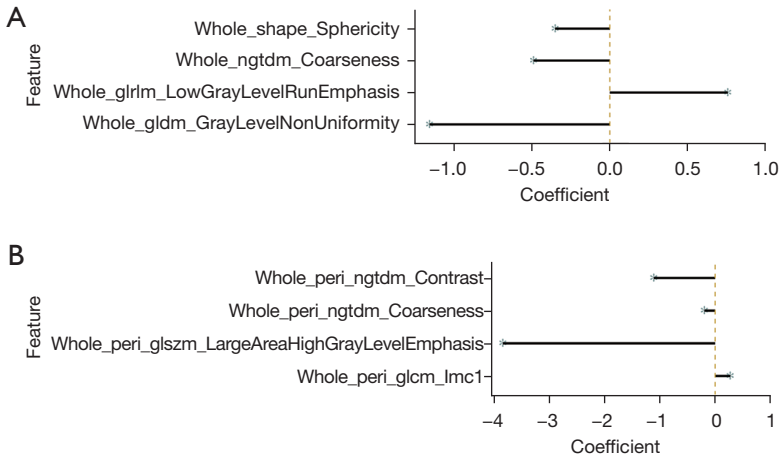

Figure 5 Summary of the model feature name and estimate. (A) The whole tumor model. (B) The whole tumor and peritumoral model.

RRM2 expression level varied considerably in distribution between the two *RRM2* expression groups (P=0.004), with increased values in the high *RRM2* expression group (*Figure 7A*). In the whole tumor and peritumor model, RS varied considerably between the two *RRM2* expression groups (P=0.001), also with increased RS values in the high *RRM2* expression group (*Figure 7B*).

Development and evaluation of the nomogram

The probability values (RS) predicted by the whole tumor and peritumoral model were merged after intersection with the full clinical dataset, resulting in information for 35 patients. The cutoff value of RS was taken as 0.6085, which was divided into low/high dichotomous variables, and baseline features of the patients were set (*Table 2*).

RS, sex, and degree of vascular invasion were selected based on multiple Cox regression, and a nomogram was developed to predict 1-, 2-, and 3-year OS (*Figure 8*). A time-dependent ROC curve, calibration, and clinical decision curves were employed to evaluate the predictive performance of the nomogram (*Figure 9*), and the accuracies of the OS predictions for the 1-, 2-, and 3-year periods were compared using ROC curves. The nomogram that predicted 1-, 2-, and 3-year OS rates had AUC values of 0.836, 0.757, and 0.729, respectively (*Figure 9A*). Calibration plots at 1, 2, and 3 years showed the curves at all times to be near the diagonal, indicating that the error of the predictions was small (*Figure 9B*). Regarding 1- and 2-year DCA curves, the model showed high clinical utility in the range of thresholds of 0.1 *vs.* 0.5. The 3-year DCA curve indicated high clinical utility in the range of thresholds of 0.3 *vs.* 0.7 (*Figure 9C-9E*). Thus, the prediction ability of the nomogram was verified.

Discussion

This research demonstrated that *RRM2* expression is a potential marker of prognosis in HCC patients. A radiomic model was developed and validated for prediction of the probability of *RRM2* expression in individuals with HCC. The data indicated that the model performed efficiently in predicting *RRM2* expression and OS in HCC patients.

RRM2 performs a vital function in DNA synthesis and repair, maintaining the stability of nucleotides by converting ribonucleotides to deoxyribonucleotides. It has been confirmed that *RRM2* is a potential biomarker in a variety of tumors (12-15). In HCC, studies have found that overexpression of *RRM2* is linked to low rates of survival (16), and expression of *RRM2* is strongly positively linked to infiltration of immune cells and immune checkpoint expression, the upregulation of *RRM2* mediated by ncRNAs correlates with poor prognosis and tumor immune infiltration of HCC (17). Through Kaplan-Meier analysis, this research showed an association between worsening prognosis of individuals with HCC with elevated expression levels of *RRM2* (P<0.01) by examining 286 TCGA samples. Univariate and multivariate Cox regression indicated the potential of increased *RRM2* expression to independently predict poor OS in individuals with HCC. Immune cell infiltration levels of liver cancer samples were also examined, with an increased degree of M0 macrophage

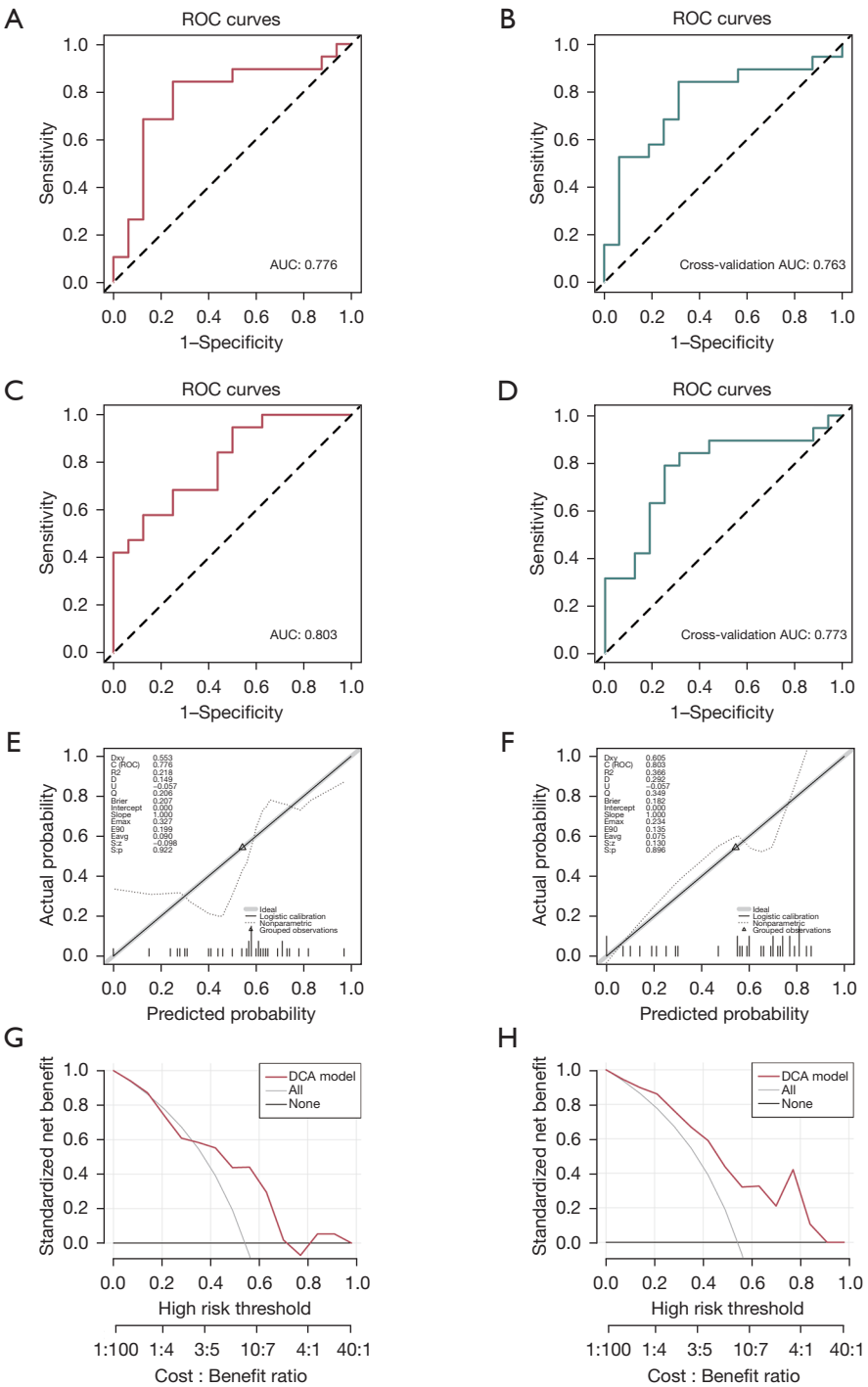

Figure 6 Evaluation of radiomic models. (A) The ROC of the whole tumor model. (B) The cross-validation ROC of the whole tumor model. (C) The ROC of the whole tumor & peritumoral model. (D) The cross-validation ROC of the whole tumor and peritumoral model. (E) Calibration curves of the whole tumor model. (F) Calibration curves of the whole tumor and peritumoral model. (G) DCA of the whole tumor model. (H) DCA of the whole tumor and peritumoral model. ROC, receiver operating characteristic; AUC, area under the ROC curve; DCA, decision curve analysis.

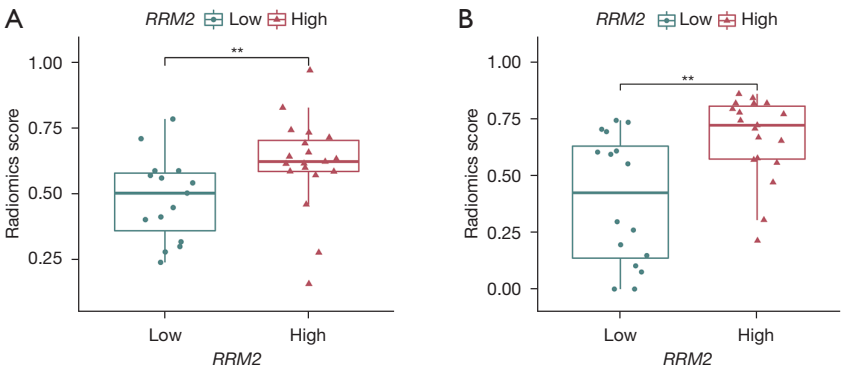

Figure 7 The expression level of *RRM2* in the two models. (A) The whole tumor model. (B) The whole tumor and peritumoral model. **, P<0.01. RRM2, ribonucleotide reductase regulatory subunit M2.

Table 2 Characteristics of patients in TCIA-TCGA

Variables	Total (n=35)	RS		P
		Low (n=18)	High (n=17)	
Age (years), n [%]				0.815
<60	12 [34]	7 [39]	5 [29]	
≥60	23 [66]	11 [61]	12 [71]	
Sex, n [%]				1
Female	13 [37]	7 [39]	6 [35]	
Male	22 [63]	11 [61]	11 [65]	
Tumor grade, n [%]				0.241
G1/G2	21 [60]	13 [72]	8 [47]	
G3/G4	14 [40]	5 [28]	9 [53]	
Pathologic stage, n [%]				0.909
I/II	24 [69]	13 [72]	11 [65]	
III/IV	11 [31]	5 [28]	6 [53]	
Residual tumor, n [%]				0.658
R0	30 [86]	16 [89]	14 [82]	
R1/R2/RX	5 [14]	2 [11]	3 [18]	
Vascular invasion, n [%]				1
None	25 [71]	13 [72]	12 [71]	
Micro/macro/unknown	10 [29]	5 [28]	5 [29]	
Hepatic inflammation, n [%]				1
None	15 [43]	8 [44]	7 [41]	
Mild/severe/unknown	20 [57]	10 [56]	10 [59]	
AFP, n [%]				0.815
<400	23 [66]	11 [61]	12 [71]	
≥400/unknown	12 [34]	7 [39]	5 [29]	

TCIA, The Cancer Imaging Archive; TCGA, The Cancer Genome Atlas; RS, radiomics score; AFP, alpha-fetoprotein.

infiltration in the *RRM2* high-expression group *vs.* the *RRM2* low-expression group (P<0.05). *RRM2* was notably linked to the immune genes *LAG3*, *CD276*, *CD86*, *CD48*, and *TNFSF4* (P<0.001). Therefore, noninvasive detection to predict the degree of *RRM2* expression is helpful for personalized clinical decision-making.

Radiomics provides an effective and noninvasive method for tumor diagnosis and efficacy prediction through extraction of multiple imaging features. With the development of molecular biology and artificial intelligence, an increasing number of molecular markers can be discovered and effectively predicted by radiomics. Che *et al.* (18) integrated clinicoradiological risk factors and the RS using the CRR model and predicted β-arrestin1 phosphorylation in HCC (AUC =0.898, 95% CI: 0.820–0.977). Wu *et al.* (19) developed and verified a radiomics nomogram based on CT radiomic features in individuals with HCC to predict Ki-67 expression (respective AUC values for training and validation groups: 0.884 and 0.819). MR-based radiomics features were closely linked to positive expression of glypican 3 (GPC3), an adverse prognostic factor for HCC, in the work of Gu *et al.* (20). Furthermore, the combination of AFP and radiomic characteristics can effectively predict GPC3 expression (respective AUC values for the training and validation groups: 0.926 and 0.914). In this research, four features were screened out from 107 imaging features to construct a whole-tumor model, and a whole-tumor and peritumoral model was constructed using the same method considering micrometastasis and microinfiltration of immune-associated cells of the surrounding area of HCC, which was used less frequently in previous studies (21,22). Two logistic regression models were built to predict the level of *RRM2* expression. The

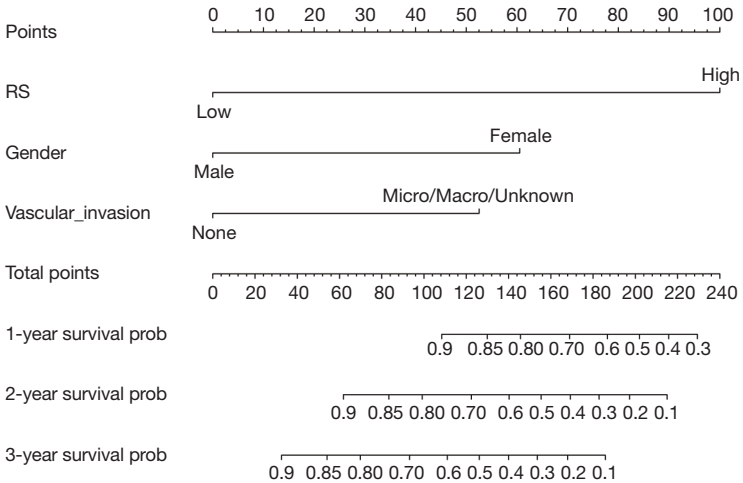

Figure 8 Nomogram of the prognostic model for OS in HCC. RS, radiomics score; OS, overall survival; HCC, hepatocellular carcinoma.

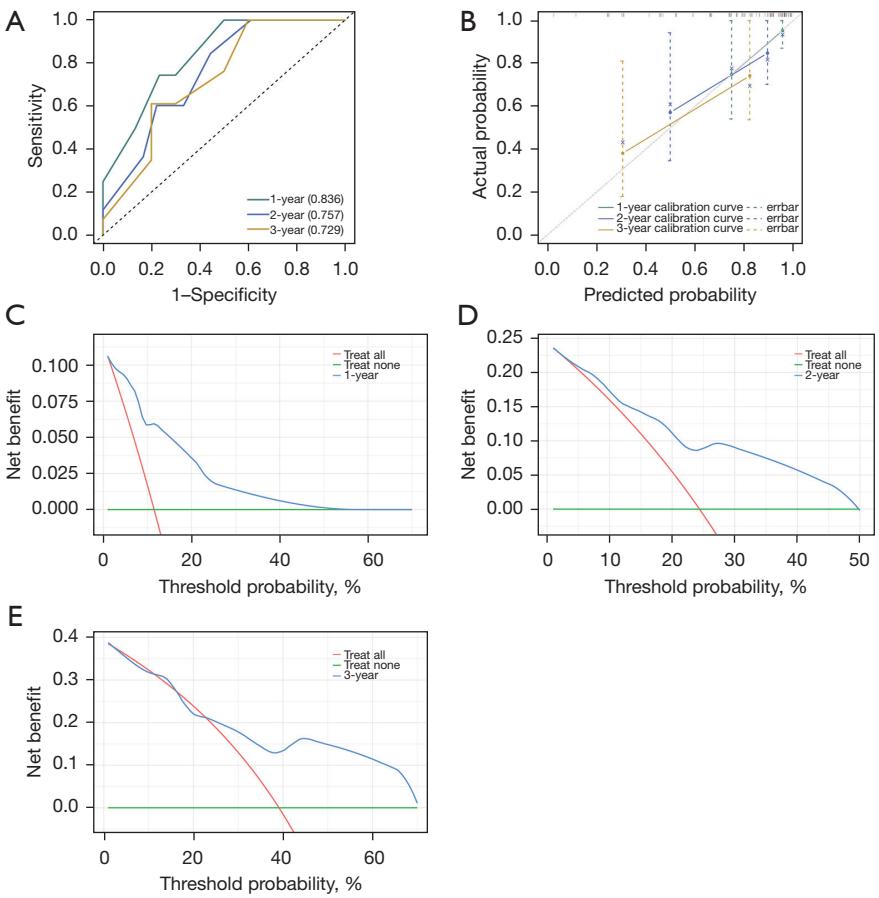

Figure 9 Evaluation of the nomogram. (A) The ROC curve of the nomogram. (B) Calibration curves of the nomogram. (C) The 1-year DCA of the nomogram. (D) The 2-year DCA of the nomogram. (E) The 3-year DCA of the nomogram. ROC, receiver operating characteristic; DCA, decision curve analysis.

calibration curves and Hosmer-Lemeshow goodness of fit test of the two models revealed considerable congruence between the prediction probability of high *RRM2* expression and the true value (P>0.05). The DeLong test revealed a lack of statistical variation in the two models, with respective P values in the training set and cross-validation set of 0.801 and 0.934. The AUC value of the whole-tumor & peritumoral imaging model was slightly higher than that the whole-tumor imaging model.

In the whole-peritumoral model, the RS of the group with high *RRM2* expression was considerably increased in comparison to the group with low *RRM2* expression (P<0.01), indicating that a higher RS value suggests a greater difference between images and greater tumor heterogeneity (23). Zheng *et al.* retrospectively extracted preoperative CT radiomic features of 319 solitary HCCs, utilized the least absolute shrinkage and selection operator (LASSO) logical model to generate the RS, and found that RS is an independent prognostic factor for HCC through multivariate analysis. Moreover, a nomogram based on RS has good prognostic performance (24). Prognostic models based on radiomic features combined with clinicopathological features have been recognized by an increasing number of studies. For example, Deng *et al.* (25) predicted the OS of individuals with HCC following radical hepatectomy by incorporating AFP, neutrophil-to-lymphocyte ratio (NLR), and radiomic features into a nomogram. The AUC of the ROC curve for 1-, 3-, and 5-year OS prediction was 0.850, 0.791, and 0.823 in the training cohort and 0.905, 0.884, and 0.911 in the validation cohort, respectively. A nomogram incorporating the RS and two clinical parameters [tumor embolus and albumin-bilirubin (ALBI) grade] was developed in another study evaluating CT imaging in patients with advanced HCC who were unresponsive to recommended first-line therapy and treated with programmed death 1 (PD-1) inhibitors to predict the probability of progression disease (PD) after PD-1 inhibitor therapy (AUC: training group, 0.894; validation group, 0.883) (26). In this research, Variables were obtained by the minimal AIC-based stepwise selection method. A nomogram was established for prediction of patient prognosis utilizing RS, sex, and vascular invasion according to multivariate Cox regression data; AUC values for the 1-, 2-, and 3-year OS prediction ability of the model were 0.836, 0.757, and 0.729, respectively. DCA showed that this model has high clinical practicability. Although widespread study of HCC radiomics has great potential for predicting tumor biomarkers, therapeutic response, and prognosis, several obstacles remain to be addressed before radiomic analysis can be widely used in the clinic. To date, most studies have been small sample single-center retrospective studies, with differences in the imaging providers, platforms, and protocols used (27). Therefore, a radiomic model based on a single center may not be generalizable. The radiomics data used in this study were from public databases, which avoids single-center data bias. However, the available images obtained from public databases in this study were still limited. On the other hand, there is a lack of consistent standards in radiomic studies concerning image acquisition protocols, segmentation processes, and radiomic tools used for analysis, which may lead to differences in radiomic feature measurements (28).

There are several limitations of this study. The research relied on public datasets, which may vary considerably in image quality, potentially impacting predictive analysis. Additionally, the images of a limited sample size of only 35 eligible patients in TCIA were accessed. Hence, validation of the predictive value through more data and clinical samples is needed.

Conclusions

The prognosis of individuals with HCC can be strongly impacted by *RRM2* expression levels. CT images based on radiomic characteristics can predict both the expression profile of *RRM2* and the prognosis of HCC patients.

Acknowledgments

Funding: This study was supported by the Guangxi Medical and Health Key Discipline Construction Project.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-460/rc

Peer Review File: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-460/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-460/coif). All authors report that this study was supported by the Guangxi Medical and Health Key Discipline Construction Project.

The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
2. Díaz-González Á, Reig M, Bruix J. Treatment of Hepatocellular Carcinoma. Dig Dis 2016;34:597-602.
3. El-Khoueiry AB, Sangro B, Yau T, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet 2017;389:2492-502.
4. Chabes A, Thelander L. Controlled protein degradation regulates ribonucleotide reductase activity in proliferating mammalian cells during the normal cell cycle and in response to DNA damage and replication blocks. J Biol Chem 2000;275:17747-53.
5. Wu L, Yin L, Ma L, et al. Comprehensive bioinformatics analysis of ribonucleoside diphosphate reductase subunit M2(RRM2) gene correlates with prognosis and tumor immunotherapy in pan-cancer. Aging (Albany NY) 2022;14:7890-905.
6. Gao J, Chen H, Yu Y, et al. Inhibition of hepatocellular carcinoma growth using immunoliposomes for co-delivery of adriamycin and ribonucleotide reductase M2 siRNA. Biomaterials 2013;34:10084-98.
7. Liu X, Xu Z, Hou C, et al. Inhibition of hepatitis B virus replication by targeting ribonucleotide reductase M2 protein. Biochem Pharmacol 2016;103:118-28.
8. Xu H, Li B. MicroRNA-582-3p targeting ribonucleotide reductase regulatory subunit M2 inhibits the tumorigenesis of hepatocellular carcinoma by regulating the Wnt/β-catenin signaling pathway. Bioengineered 2022;13:12876-87
9. Kim S, Shin J, Kim DY, et al. Radiomics on Gadoxetic Acid-Enhanced Magnetic Resonance Imaging for Prediction of Postoperative Early and Late Recurrence of Single Hepatocellular Carcinoma. Clin Cancer Res 2019;25:3847-55.
10. Gao L, Xiong M, Chen X, et al. Multi-Region Radiomic Analysis Based on Multi-Sequence MRI Can Preoperatively Predict Microvascular Invasion in Hepatocellular Carcinoma. Front Oncol 2022;12:818681.
11. Marasco G, Colecchia A, Colli A, et al. Role of liver and spleen stiffness in predicting the recurrence of hepatocellular carcinoma after resection. J Hepatol 2019;70:440-8.
12. Mazzu YZ, Armenia J, Chakraborty G, et al. A Novel Mechanism Driving Poor-Prognosis Prostate Cancer: Overexpression of the DNA Repair Gene, Ribonucleotide Reductase Small Subunit M2 (RRM2). Clin Cancer Res 2019;25:4480-92.
13. Abdel-Rahman MA, Mahfouz M, Habashy HO. RRM2 expression in different molecular subtypes of breast cancer and its prognostic significance. Diagn Pathol 2022;17:1.
14. Jin CY, Du L, Nuerlan AH, et al. High expression of RRM2 as an independent predictive factor of poor prognosis in patients with lung adenocarcinoma. Aging (Albany NY) 2020;13:3518-35.
15. Cao X, Xue F, Chen H, et al. MiR-202-3p inhibits the proliferation and metastasis of lung adenocarcinoma cells by targeting RRM2. Ann Transl Med 2022;10:1374.
16. Sun Q, Liu P, Long B, et al. Screening of significant biomarkers with poor prognosis in hepatocellular carcinoma via bioinformatics analysis. Medicine (Baltimore) 2020;99:e21702.
17. Mao G, Li L, Shan C, et al. High expression of RRM2 mediated by non-coding RNAs correlates with poor prognosis and tumor immune infiltration of hepatocellular carcinoma. Front Med (Lausanne) 2022;9:833301.
18. Che F, Xu Q, Li Q, et al. Radiomics signature: A potential biomarker for β-arrestin1 phosphorylation prediction in hepatocellular carcinoma. World J Gastroenterol 2022;28:1479-93.
19. Wu C, Chen J, Fan Y, et al. Nomogram Based on CT Radiomics Features Combined With Clinical Factors to

Predict Ki-67 Expression in Hepatocellular Carcinoma. Front Oncol 2022;12:943942.
20. Gu D, Xie Y, Wei J, et al. MRI-Based Radiomics Signature: A Potential Biomarker for Identifying Glypican 3-Positive Hepatocellular Carcinoma. J Magn Reson Imaging 2020;52:1679-87.
21. Li N, Wan X, Zhang H, et al. Tumor and peritumor radiomics analysis based on contrast-enhanced CT for predicting early and late recurrence of hepatocellular carcinoma after liver resection. BMC Cancer 2022;22:664.
22. Oura K, Morishita A, Tani J, et al. Tumor Immune Microenvironment and Immunosuppressive Therapy in Hepatocellular Carcinoma: A Review. Int J Mol Sci 2021;22:5801.
23. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441-6.
24. 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;18:1148.
25. Deng PZ, Zhao BG, Huang XH, et al. Preoperative contrast-enhanced computed tomography-based radiomics model for overall survival prediction in hepatocellular carcinoma. World J Gastroenterol 2022;28:4376-89.
26. Yuan G, Song Y, Li Q, et al. Development and Validation of a Contrast-Enhanced CT-Based Radiomics Nomogram for Prediction of Therapeutic Efficacy of Anti-PD-1 Antibodies in Advanced HCC Patients. Front Immunol 2020;11:613946.
27. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Med Phys 2017;44:1050-62.
28. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.

Cite this article as: Li Q, Long X, Lin Y, Liang R, Li Y, Ge L. Computed tomography radiomics signature via machine learning predicts *RRM2* and overall survival in hepatocellular carcinoma. J Gastrointest Oncol 2023;14(3):1462-1477. doi: 10.21037/jgo-23-460

