Internal Medicine
The Japanese Society of Internal Medicine
doi: 10.2169/internalmedicine.9064-17
Intern Med 57: 457-468, 2018
http://internmed.jp

【 ORIGINAL ARTICLE 】

Nomograms to Predict the Disease-free Survival and Overall Survival after Radiofrequency Ablation for Hepatocellular Carcinoma

Yoshitaka Takuma[1], Iwadou Shota[1], Hirokazu Miyatake[1], Shuji Uematsu[1], Ryouichi Okamoto[1], Yasuyuki Araki[1], Hiroyuki Takabatake[2], Youichi Morimoto[2] and Hiroshi Yamamoto[2]

Abstract:

Objective The purpose of this study was to construct nomograms for the disease-free survival (DFS) and overall survival (OS) of post-radiofrequency ablation (RFA) patients with hepatocellular carcinoma (HCC). Furthermore, we compared the prognostic predictive ability of these nomograms for estimating per-patient outcomes with that of traditional staging systems.

Methods We retrospectively enrolled 298 patients in the training set and 272 patients in the validation set who underwent RFA for HCC. The nomograms for the DFS and OS were constructed from the training set using the multivariate Cox proportional hazards model. The discriminatory accuracy of the models was compared with traditional staging systems by analyzing the Harrell's C-index.

Results The DFS nomogram was developed based on the tumor size, tumor number, aspartate aminotransferase (AST), albumin, age, and α-fetoprotein. The OS nomogram was developed based on the tumor size, the model for end-stage liver disease, AST, and albumin. Our DFS and OS nomograms had good calibration and discriminatory abilities in the training set, with C-indexes of 0.640 and 0.692, respectively, that were greater than those of traditional staging systems. The C-indexes of our DFS and OS nomograms were also greater than those of traditional staging systems in the validation set, with C-indexes of 0.614 and 0.657, respectively. RFA patients were stratified into low- and high-risk groups based on the median nomogram scores. High-risk patients receiving surgical resection (SR) were associated with a better DFS and OS than those undergoing RFA. However, the DFS and OS were similar between the low-risk RFA and SR groups.

Conclusion We constructed reliable and useful nomograms that accurately predict the DFS and OS after RFA for early-stage HCC patients. These graphical tools are easy to use and will assist physicians during the therapeutic decision-making process.

Key words: hepatocellular carcinoma, radiofrequency ablation, nomogram, disease-free survival, overall survival

(Intern Med 57: 457-468, 2018)
(DOI: 10.2169/internalmedicine.9064-17)

Introduction

Hepatocellular carcinoma (HCC) is the third-leading cause of cancer death globally (1). According to the European and American HCC management guidelines, the first-line managements for Barcelona Clinic Liver Cancer (BCLC) stage 0 and BCLC stage A patients are surgical resection (SR), radiofrequency ablation (RFA), and liver transplantation (2, 3). Among these, patients with a preserved liver function (Child-Pugh A and B) may also qualify for curative treatments, such as SR and local ablation, as liver

[1]Department of Internal Medicine, Hiroshima City Hospital, Japan and [2]Department of Gastroenterology, Kurashiki Central Hospital, Japan
Received: February 19, 2017; Accepted: May 28, 2017; Advance Publication by J-STAGE: November 20, 2017
Correspondence to Dr. Yoshitaka Takuma, takuma@enjoy.ne.jp

transplantation is restricted by organ donor shortages and high costs (4, 5). Although SR is considered the main curative treatment for early-stage HCC, SR increases the risk of postoperative liver failure compared with RFA. RFA is considered a viable alternative treatment to SR for early HCC (≤3 tumors, each ≤3 cm in diameter), especially in patients with an impaired liver function (6). Therefore, understanding the prognostic factors in patients with early HCC after RFA is important. However, research into the prognostic factors in HCC patients treated with RFA thus far has examined heterogeneous patient populations in terms of the tumor size, tumor number, and causes of HCC (7-9).

It is therefore necessary to construct clinically relevant and easy-to-use graphical indicators with a demonstrated ability to predict tumor recurrence and the patient survival, specifically in patients with early HCC after RFA. Such graphical tools are known to be easy to use and can assist physicians in the therapeutic decision-making process.

The purpose of this study was to construct nomograms and to independently validate these scoring systems for recurrence and death in early HCC patients after RFA. Furthermore, we compared the prognostic predictive ability of these nomograms for estimating per-patient outcomes with that of traditional staging systems.

Materials and Methods

Patients

Our study was a retrospective cohort study that reviewed the records of consecutive patients treated with RFA as the initial treatment for HCC from a prospectively-collected database at Hiroshima City Hospital from October 2001 to July 2014 (training set). Data from another independent cohort (validation set) of consecutive patients who underwent RFA for HCC at Kurashiki Central Hospital during the same study period were collected retrospectively. Informed consent was obtained from all patients regarding the use of their clinical data. The institutional review board approved this study. The diagnosis of HCC was corroborated by histology or was made according to the American Association for the Study of Liver Diseases practice guidelines (10) by combining a diagnostic α-fetoprotein (AFP) level increase (>200 ng/mL) with a typical vascular pattern for HCC on 1 dynamic imaging technique or a typical vascular pattern for HCC on 2 dynamic imaging techniques. The maximum diameter of the tumors was measured using axial computed tomography (CT) or magnetic resonance imaging (MRI).

The indications of initial treatment of RFA for HCC were (i) ≤3 tumors each ≤3 cm in diameter and (ii) a liver function of Child-Pugh class A or B. Patients with performance status ≥2 (11), simultaneous malignancies, or candidates for liver transplantation were excluded. In our study, 327 consecutive patients forming the training set and 301 consecutive patients forming the validation set underwent RFA for the initial treatment of HCC. Ultimately, 298 patients in the training set and 272 in the validation set were included in our study.

Treatment

RFA was performed percutaneously by senior hepatologists with more than 20 years' experience. Under ultrasonographic guidance, RFA was performed using the Cooltip Radiofrequency Ablation System (Radionics, Burlington, USA) under local anesthesia. One day post-procedure, the treatment response was evaluated by dynamic CT, and the technical success of RFA was defined as hypoattenuation with the surrounding liver parenchyma of the entire tumor in both the arterial and portal venous phases, appearing larger than the tumor itself on the CT images. Additional RFA was performed until complete ablation of the tumor, if needed.

Assessment and follow-up

Patients were followed in the clinic for 1-3 months and then every 3 months with serum AFP and des-gamma-carboxy prothrombin (DCP) concentrations as well as with ultrasonography, dynamic CT, or dynamic MRI.

Intrahepatic tumor recurrence was confirmed via contrast-enhanced CT, contrast-enhanced MRI, or angiography, along with an ultrasound-guided biopsy when necessary, using the same criteria for diagnosing the primary HCC. The study endpoints included the disease-free survival (DFS) and overall survival (OS) rates. Intrahepatic HCC recurrence was classified as recurrence either at a site distant from the primary tumor (distant intrahepatic recurrence) or adjacent to the treated site (local tumor progression).

The choice of treatment modalities for recurrent HCCs depended on the patient preferences and the clinical practices of surgeons and hepatologists. In general, when recurrence was detected, the patients were treated with SR, RFA, percutaneous ethanol injection, transcatheter arterial chemoembolization, systemic chemotherapy, or conservative treatment, depending on the site of the tumor, liver function, and general condition of the patient.

We also evaluated the usefulness of the nomogram for risk stratification and treatment choice by comparing the DFS and OS in similar-stage HCC (≤3 tumors each ≤3 cm in diameter) patients receiving SR or RFA. The median value of the nomogram scores was used to stratify RFA patients into two (low and high risk) groups for survival analyses. The log-rank test was also performed between patients receiving SR and RFA in the same risk group. Similar-stage HCC patients (82 patients in the training set and 95 in the validation set) receiving SR were divided into 2 (low and high risk) groups by the median value of nomogram scores derived from the patients receiving RFA in the training set.

Statistical analyses

Summary statistics for the study population are presented as percentages or as median values with the interquartile range (IQR). The Mann-Whitney U test and chi-squared test were used to analyze the differences in baseline demograph-

Table 1. Clinical Characteristics in the Training Set and the Validation Sets.

Variables	Training set (n=298)	Validation set (n=272)	p value*
Sex (male/female)	161/137 (54.0)	196/76 (72.0)	<0.001
Age (years)	71 (64-76)	72.0 (65-76)	0.276
Performance status (0/1)	283/15 (95.0)	257/15 (94.5)	0.797
BMI (kg/m^2)	22.9 (21.0-25.4)	23.0 (21.0-25.7)	0.809
INR	1.13 (1.06-1.23)	1.11 (1.04-1.19)	0.003
Albumin (g/dL)	3.7 (3.3-4.1)	3.7 (3.3-4.0)	0.550
Total bilirubin (mg/dL)	0.8 (0.6-1.1)	0.7 (0.5-1.0)	0.002
AST (IU/L)	52 (37-69)	50.5 (37-67)	0.711
ALT (IU/L)	41 (29-65)	44.5 (27-64.5))	0.936
Platelet counts (×10^4/μL)	9.6 (7.2-13.9)	9.6 (7.2-14.1)	0.922
Creatinine (mg/dL)	0.77 (0.62-0.90)	0.77 (0.67-0.89)	0.307
Child-Pugh score	5 (5-6)	5 (5-6)	0.816
MELD	8 (7-10)	8 (7-9)	0.040
Anti-HCV (+/-)	223/75 (74.8)	219/53 (80.5)	0.104
HBsAg (+/-)	28/270 (9.4)	23/249 (8.5)	0.694
AFP (>50/≤50 ng/mL)	70/228 (23.5)	62/210 (22.8)	0.844
DCP (>400/≤400 mAU/mL)	10/288 (3.4)	17/255 (6.3)	0.104
No. of nodules (1/2/3)	233/50/15	198/58/16	0.319
Size of largest tumor (mm)	17.0 (15.0-20.3)	16.0 (13.0-20.0)	0.125
CLIP score (0/1/2/3)	174/99/24/1	147/108/15/2	0.631
Okuda score (0/1/2/3)	248/44/5/1	218/42/11/1	0.187
JIS score (0/1/2/3)	128/125/43/2	121/107/38/6	0.864
TNM stage(I/II/III)	169/111/18	152/97/23	0.538
BCLC stage(0/A/C)	159/124/15	117/144/11	0.055

Data are expressed as the median (interquartile range) or n (%). *Mann-Whitney U test and chi-square test were used to analyze the differences in background and biochemical data between the two groups. p<0.05 indicated a significant difference. BMI: body mass index, INR: international normalized ratio, AST: aspartate aminotransferase, ALT: alanine aminotransferase, MELD: model for end-stage liver disease, Anti-HCV: antibody to hepatitis C virus, HBsAg: hepatitis B surface antigen, AFP: alpha-fetoprotein, DCP: des-gamma-carboxy prothrombin, CLIP: Cancer of the Liver Italian Program, JIS: Japan Integrated Staging, TNM: tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan, BCLC: Barcelona Clinic Liver Cancer

ics, clinical, and biochemical characteristics between the training and validation sets. The OS and DFS rates were analyzed using the Kaplan-Meier method. Intergroup differences were compared using the log-rank test. Factors potentially influencing the OS and DFS were identified through backward stepwise selection in a Cox regression model. Added variables that were significantly related to the survival in the univariate Cox models (p<0.05) were subsequently included in the multivariate model. Selected variables were then incorporated into the nomogram. Model performance was evaluated by assessing discrimination with Harrell's C-index (12), creating calibration plots using a 100 bootstrapped sample, and plotting Kaplan-Meier curves over the quartiles of prediction by the nomogram. The method suggested by Kang and Chen was used to compare the Harrell's C-index of various prognostic staging systems (13). During the external validation of the nomogram, the total points of each patient in the validation cohort were calculated according to the established nomogram. Statistical analyses were performed using the SPSS software program, ver. 16.0 for Windows (SPSS, Chicago, USA), and R version 3.0.2. Significance tests were two-sided. A p value <0.05 was considered to be statistically significant.

Results

Baseline characteristics of the training vs. validation sets

Table 1 shows the patient characteristics of the two groups. Patients in the training set had a higher rate of women (p<0.001), a higher international normalized ratio (INR) (p=0.003), a higher level of total bilirubin (p=0.002), and a higher model for end-stage liver disease (MELD) score (p=0.040) than those in the validation set.

Technical success and complications of RFA

Technical success was achieved in all 570 patients. In the

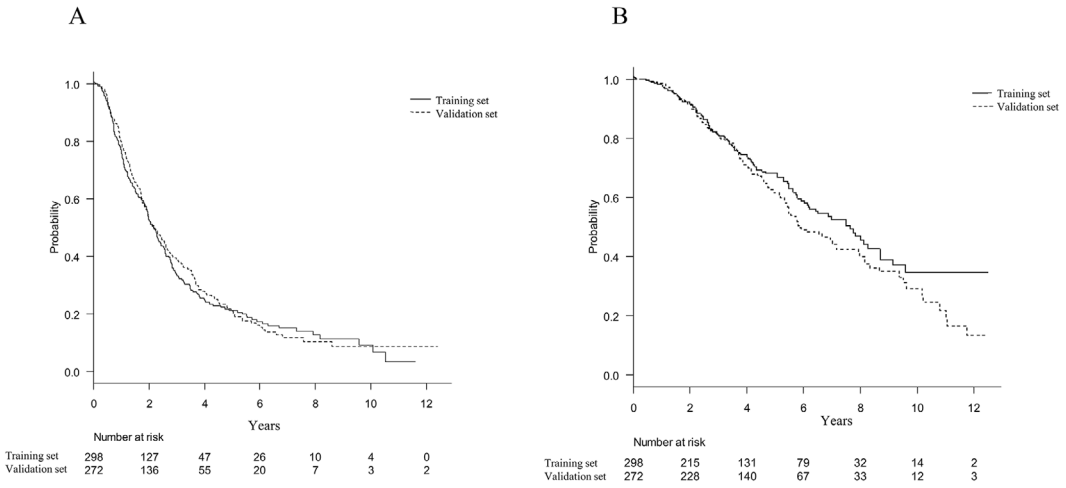

Figure 1. Cumulative DFS (A) and OS (B) curves of the two series of patients included in the study. The solid line represents the training set, and the dotted line represents the validation set. DFS: disease-free survival, OS: overall survival

training set, 1 RFA session was performed in 253 patients (84.8%), 2 sessions in 42 (14.1%), and 3 sessions in 3 (1.0%). In the validation set, 1 RFA session was performed in 221 patients (81.3%), 2 sessions in 47 (17.3%), and 3 sessions in 4 (1.5%). After RFA, complications occurred in 3 patients in the training set (1 diaphragmatic hernia, 1 minor burn, and 1 pleural effusions) (1.0%) and in 4 in the validation set (1 pneumothorax, 2 minor burns, and 1 pleural effusion) (1.5%).

Univariate and multivariate analyses and nomograms for the DFS and OS

The median observation period was 40 months (IQR, 23-74 months) for patients in the training set and 49 (IQR, 32-71) months for those in the validation set. Recurrence was noted in 62% (n=185) of HCCs in the training set and 68% (n=184) in the validation set. All recurrences following RFA in both cohorts were intrahepatic recurrences (151 distant intrahepatic and 34 local tumor progressions in the training set, and 148 distant intrahepatic and 36 local tumor progressions in the validation set). The 1-, 3-, and 5-year DFS rates were 77%, 34%, and 21%, respectively, in the training set and 80%, 39%, and 21%, respectively, in the validation set. The DFS rates did not differ significantly between the two cohorts (p=0.601) (Fig. 1A). Among the 298 patients in the training set, 190 were alive, 105 died, and 3 were lost to follow-up. Among the 272 patients in the validation set, 136 were alive, 131 died, and 5 were lost to follow-up. The 3-, 5-, and 7-year overall survival rates were 81%, 67%, and 52%, respectively, in the training set and 80%, 61%, and 44%, respectively, in the validation set. The OS rates did not differ significantly between the 2 cohorts (p=0.110) (Fig. 1B).

For the training set, the following were significantly associated with tumor recurrence in the univariate analyses (Table 2): a higher age (>60 years) (p=0.016), higher INR (>1.1) (p=0.012), lower serum albumin levels (≤3.5 g/dL) (p<0.001), higher total bilirubin levels (>1.0 mg/dL) (p<0.001), higher aspartate aminotransferase (AST) levels (>30 IU/L) (p<0.001), higher alanine aminotransferase (ALT) levels (>30 IU/L) (p=0.010), lower platelet counts ($\leq 15\times10^4/\mu L$) (p=0.040), higher MELD score (p=0.016), antibody to hepatitis C virus (anti-HCV) positivity (p=0.012), hepatitis B surface antigen (HBsAg) negativity (p=0.008), higher AFP levels (>50 ng/mL) (p=0.002), higher DCP levels (>400 mAU/mL) (p=0.023), number of nodules (2 and 3 nodules) (p=0.025 and p=0.004), and a larger tumor size (mm) (p=0.002). In the multivariate analyses, a higher age (>60 years) (p=0.037), lower serum albumin levels (≤3.5 g/dL) (p=0.013), higher AST levels (>30 IU/L) (p=0.001), higher AFP levels (>50 ng/mL) (p=0.036), number of nodules (2 and 3 nodules) (p=0.070 and p=0.035), and a larger tumor size (mm) (p=0.016) were independent risk factors associated with tumor recurrence (Table 2).

A nomogram for predicting the DFS of the training set is shown in Fig. 2A. The nomogram derived from the training set was developed based on the six independent prognostic markers (tumor size, tumor number, AST, albumin, age, and AFP). Each factor in the nomogram was assigned a weighted number of points, and the sum of points for each patient was associated with a specific predicted 1-, 3-, and 5-year DFS. For example, an 81-year-old man with a 22-mm single HCC nodule, an AST of 66 IU/L, a serum albumin of 3.5 g/dL, and an AFP of 19.3 ng/mL would have a total of 277 points (age=63 points, tumor size=66 points, tumor number=0 points, AST=100 points, and serum albumin =48 points, and AFp=0 points). For this patient, the predicted 1-, 3-, and 5-year DFS values were 70%, 20%, and 9%, respectively (Fig. 2A).

For the training, the following were significantly associated with a poor OS in the univariate analyses (Table 3): a higher age (>60 years) (p=0.011), performance status 1 (p=0.014), higher INR (>1.1) (p=0.012), lower serum albumin levels (≤3.5 g/dL) (p<0.001), higher total bilirubin lev-

Table 2. Risk of Tumor Recurrence in HCC Patients after RFA.

Variables	Univariate analysis		Multivariate analysis	
	HR (95%CI)	p value*	HR (95%CI)	p value*
Sex (male/female)	0.926 (0.706-1.214)	0.579		
Age (>60/≤60 years)	1.738 (1.107-2.728)	0.016	1.625 (1.030-2.565)	0.037
Performance status (1/0)	1.542 (0.813-2.925)	0.184		
BMI (>25/≤25 kg/m^2)	0.936 (0.692-1.265)	0.667		
INR (>1.1/≤1.1)	1.441 (1.082-1.919)	0.012		
Albumin (>3.5/≤3.5 g/dL)	0.560 (0.424-0.739)	<0.001	0.691 (0.515-0.926)	0.013
Total bilirubin (>1.0/≤1.0 mg/dL)	1.491 (1.109-2.003)	0.008		
AST (>30/≤30 IU/L)	2.587 (1.654-4.047)	<0.001	2.164 (1.361-3.441)	0.001
ALT (>30/≤30 IU/L)	1.529 (1.105-2.117)	0.010		
Platelet counts (>15×10^4/≤15×10^4/μL)	0.691 (0.486-0.984)	0.040		
Creatinine (>0.8/≤0.8 mg/dL)	1.039 (0.785-1.375)	0.790		
Child-Pugh classification (B/A)	1.287 (0.934-1.774)	0.123		
MELD	1.063 (1.011-1.118)	0.016		
Anti-HCV (+/-)	1.535 (1.097-2.148)	0.012		
HBsAg (+/-)	0.469 (0.266-0.824)	0.008		
AFP (>50/≤50 ng/mL)	1.628 (1.197-2.214)	0.002	1.400 (1.022-1.919)	0.036
DCP (>400/≤400 mAU/mL)	2.274 (1.117-4.626)	0.023		
No. of nodules (2/1)	1.491 (1.052-2.114)	0.025	1.384 (0.974-1.968)	0.070
No. of nodules (3/1)	2.398 (1.327-4.332)	0.004	1.900 (1.047-3.449)	0.035
Size of largest tumor (mm)	1.034 (1.007-1.061)	0.012	1.032 (1.006-1.059)	0.016

*p values were determined with Cox proportional hazards regression models. p<0.05 indicated a significant difference. HCC: hepatocellular carcinoma, HR: hazard ratio, CI: confidence intervals, RFA: radiofrequency ablation, BMI: body mass index, INR: international normalized ratio, AST: aspartate aminotransferase, ALT: alanine aminotransferase, MELD: model for end-stage liver disease, Anti-HCV: antibody to hepatitis C virus, HBsAg: hepatitis B surface antigen, AFP: alpha-fetoprotein, DCP: des-gamma-carboxy prothrombin

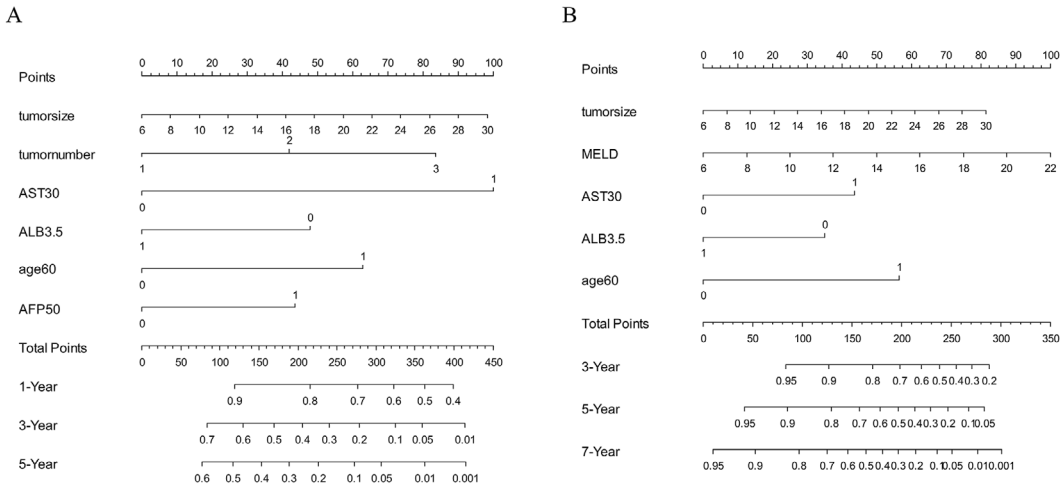

Figure 2. Nomogram for predicting (A) the disease-free survival and (B) the overall survival of HCC patients following RFA. HCC: hepatocellular carcinoma, RFA: radiofrequency ablation

els (>1.0 mg/dL) (p<0.001), higher AST levels (>30 IU/L) (p=0.015), lower platelet counts (≤15×10^4/μL) (p=0.029), Child-Pugh classification B (p=0.001), higher MELD score (p<0.001), anti-HCV positivity (p=0.016), HBsAg negativity (p=0.014), higher DCP levels (>400 mAU/mL) (p=0.021), and a larger tumor size (mm) (p=0.002). In the multivariate analyses, a higher age (>60 years) (p=0.017), lower serum albumin levels (≤3.5 g/dL) (p=0.008), higher AST levels (>30 IU/L) (p=0.041), higher MELD score (p=0.017), and a larger tumor size (mm) (p=0.003) were independent predictors of a poor OS (Table 3).

A nomogram for predicting the OS of the training set is shown in Fig. 2B. The nomogram derived from the training set was developed based on five independent prognostic markers: tumor size, MELD, AST, albumin, and age. Each factor in the nomogram was assigned a weighted number of points, and the sum of points for each patient was associated with a specific predicted 3-, 5-, and 7-year OS. Using the

Table 3. Risk of Death in HCC Patients after RFA.

Variables	Univariate analysis		Multivariate analysis	
	HR (95%CI)	p value*	HR (95%CI)	p value*
Sex (male/female)	0.795 (0.542-1.167)	0.241		
Age (>60/≤60 years)	2.716 (1.257-5.870)	0.011	2.582 (1.092-5.141)	0.017
Performance status (1/0)	1.738 (1.107-2.728)	0.014		
BMI (>25/≤25 kg/m^2)	0.849 (0.553-1.304)	0.456		
INR (>1.1/≤1.1)	1.713 (1.126-2.607)	0.012		
Albumin (>3.5/≤3.5 g/dL)	0.469 (0.317-0.694)	<0.001	0.556 (0.360-0.857)	0.008
Total bilirubin (>1.0/≤1.0 mg/dL)	2.050 (1.371-3.064)	<0.001		
AST (>30/≤30 IU/L)	2.348 (1.184-4.655)	0.015	2.079 (1.030-4.195)	0.041
ALT (>30/≤30 IU/L)	1.241 (0.775-1.986)	0.368		
Platelet counts (>15×10^4/≤15×10^4/μL)	0.544 (0.314-0.940)	0.029		
Creatinine (>0.8/≤0.8 mg/dL)	1.012 (0.676-1.515)	0.954		
Child-Pugh classification (B/A)	2.060 (1.333-3.183)	0.001		
MELD	1.140 (1.065-1.221)	<0.001	1.111 (1.019-1.211)	0.017
Anti-HCV (+/-)	1.898 (1.127-3.197)	0.016		
HBsAg (+/-)	0.237 (0.075-0.749)	0.014		
AFP (>50/≤50 ng/mL)	1.404 (0.903-2.183)	0.132		
DCP (>400/≤400 mAU/mL)	2.905 (1.174-7.187)	0.021		
No. of nodules (2/1)	1.488 (0.934-2.372)	0.095		
No. of nodules (3/1)	1.384 (0.559-3.425)	0.482		
Size of largest tumor (mm)	1.060 (1.022-1.100)	0.002	1.059 (1.020-1.100)	0.003

*p values were determined with Cox proportional hazards regression models. p<0.05 indicated a significant difference. HCC: hepatocellular carcinoma, HR: hazard ratio, CI: confidence intervals, RFA: radiofrequency ablation, BMI: body mass index, INR: international normalized ratio, AST: aspartate aminotransferase, ALT: alanine aminotransferase, MELD: model for end-stage liver disease, Anti-HCV: antibody to hepatitis C virus, HBsAg: hepatitis B surface antigen, AFP: alpha-fetoprotein, DCP: des-gamma-carboxy prothrombin

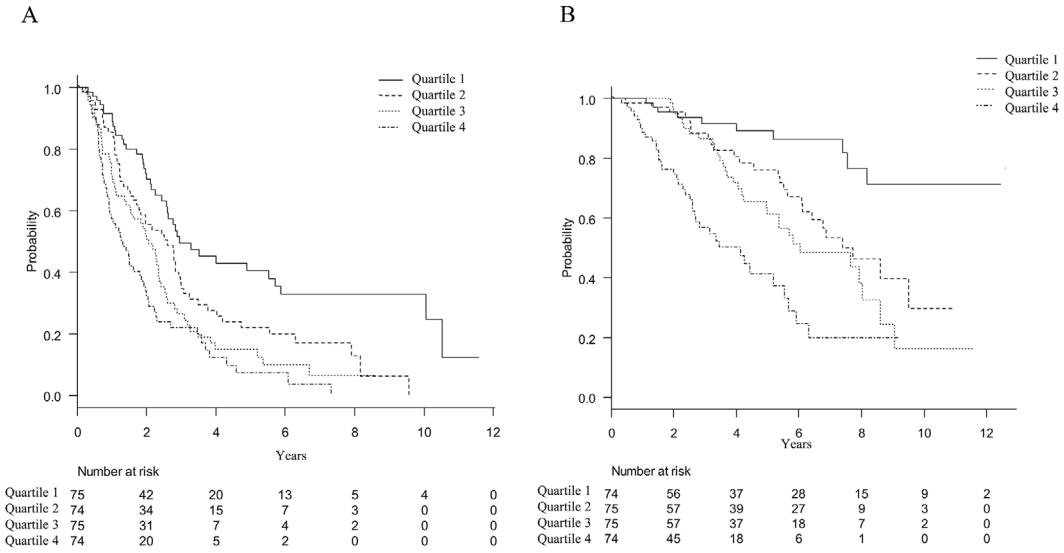

Figure 3. Kaplan-Meier curves demonstrating (A) the disease-free survival and (B) the overall survival for HCC patients following RFA according to the quartiles of the predicted disease-free or overall survival. HCC: hepatocellular carcinoma, RFA: radiofrequency ablation

nomogram, a higher score was associated with a worse prognosis. For example, a 71-year-old man with a 15-mm HCC nodule, a MELD of 11, an AST of 56 IU/L, and a serum albumin of 3.0 g/dL, would have a total of 197 points (age=56 points, tumor size=31 points, MELD=31 points, AST=44 points, and serum albumin=35 points). For this patient, the predicted 3-, 5-, and 7-year OS values were 70%, 50%, and 30%, respectively (Fig. 2B).

Model performance

The discrimination ability was assessed by dividing the predicted probability of the DFS and OS into quartiles in

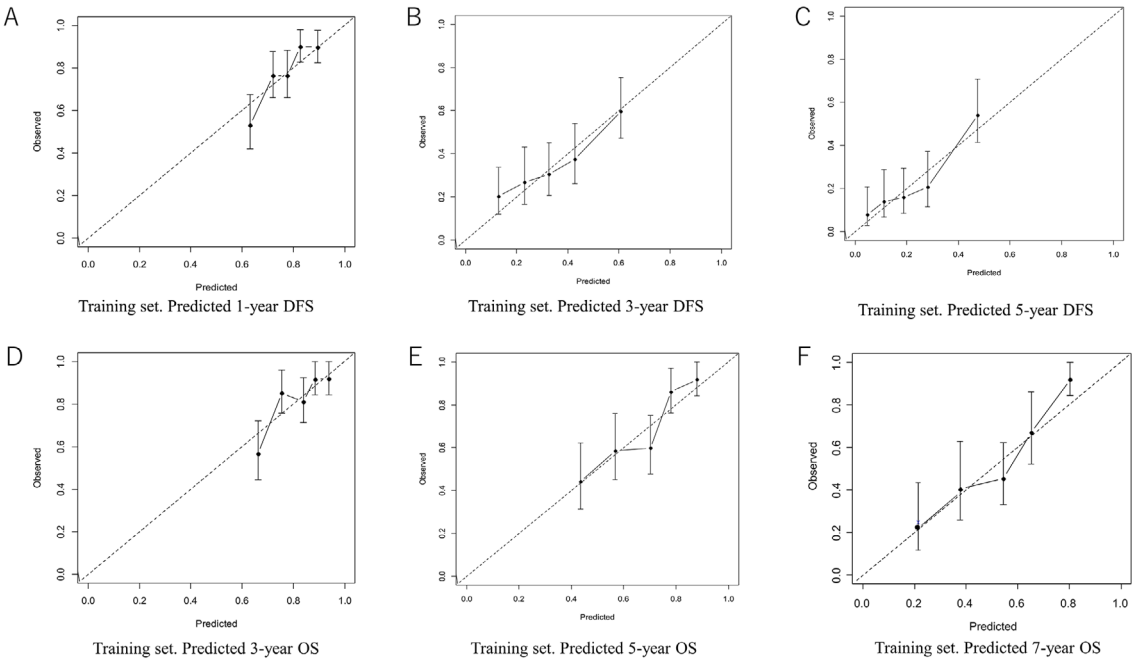

Figure 4. **Calibration curves for the nomogram-predicted probability of the 1- (A), 3- (B), and 5- (C) year DFS and the nomogram-predicted probability of the 3- (D), 5- (E), and 7- (F) year OS in the training set. Curves for a hypothetical ideal nomogram are represented by dashed lines, and those for the current nomogram are represented by solid lines. Vertical bars indicate 95% confidence intervals. DFS: disease-free survival, OS: overall survival**

Table 4. **Comparison of Staging Systems for Predicting Tumor Recurrence in Training Set.**

Variables	C-index (95%CI)
Nomogram	0.640 (0.597-0.683)
CLIP score	0.565 (0.527-0.602)
Okuda score	0.532 (0.503-0.562)
JIS score	0.556 (0.515-0.596)
TNM	0.540 (0.502-0.577)
BCLC	0.536 (0.499-0.574)

CLIP: Cancer of the Liver Italian Program, JIS: Japan Integrated Staging, TNM: tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan, BCLC: Barcelona Clinic Liver Cancer. In parentheses: percentages or 95% confidence interval

Table 5. **Comparison of Staging Systems for Predicting Mortality in Training Set.**

Variables	C-index (95%CI)
Nomogram	0.692 (0.634-0.750)
CLIP score	0.597 (0.539-0.655)
Okuda score	0.603 (0.545-0.661)
JIS score	0.590 (0.532-0.648)
TNM	0.538 (0.480-0.596)
BCLC	0.561 (0.503-0.619)

CLIP: Cancer of the Liver Italian Program, JIS: Japan Integrated Staging, TNM: tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan, BCLC: Barcelona Clinic Liver Cancer. In parentheses: percentages or 95% confidence interval

the training set. The DFS and OS stratified by quartile were then used to plot Kaplan-Meier curves (Fig. 3). The respective DFS rates for the 4 grades at 1, 3, and 5 years were as follows: 92%, 50%, and 41% in quartile 1; 86%, 37%, and 22% in quartile 2; 73%, 26%, and 15% in quartile 3; and 57%, 22%, and 7% in quartile 4 ($p<0.0001$) (Fig. 3A). Good stratification values are shown according to quartiles. The respective OS rates for the 4 grades at 3, 5, and 7 years were as follows: 92%, 89%, and 87% in quartile 1; 89%, 76%, and 54% in quartile 2; 87%, 61%, and 49% in quartile 3; and 57%, 42%, and 20% in quartile 4 ($p<0.0001$) (Fig. 3B). Good stratification values are shown according to quartiles.

The nomogram prediction was calibrated at the 1-, 3-, and 5-year DFS and the 3-, 5-, and 7-year OS using the training set (Fig. 4). Nomogram predictions seemed to be well calibrated with the actuarial survival.

The discrimination ability of the final model for the DFS and OS in the training set was assessed using the C-statistic. As shown in Table 4, the C-statistics for the DFS based on the nomogram, Cancer of the Liver Italian Program (CLIP) score, Okuda score, Japan Integrated Staging (JIS) score, tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan (TNM), and BCLC were 0.640, 0.565, 0.532, 0.556, 0.540, and 0.536, respectively. In

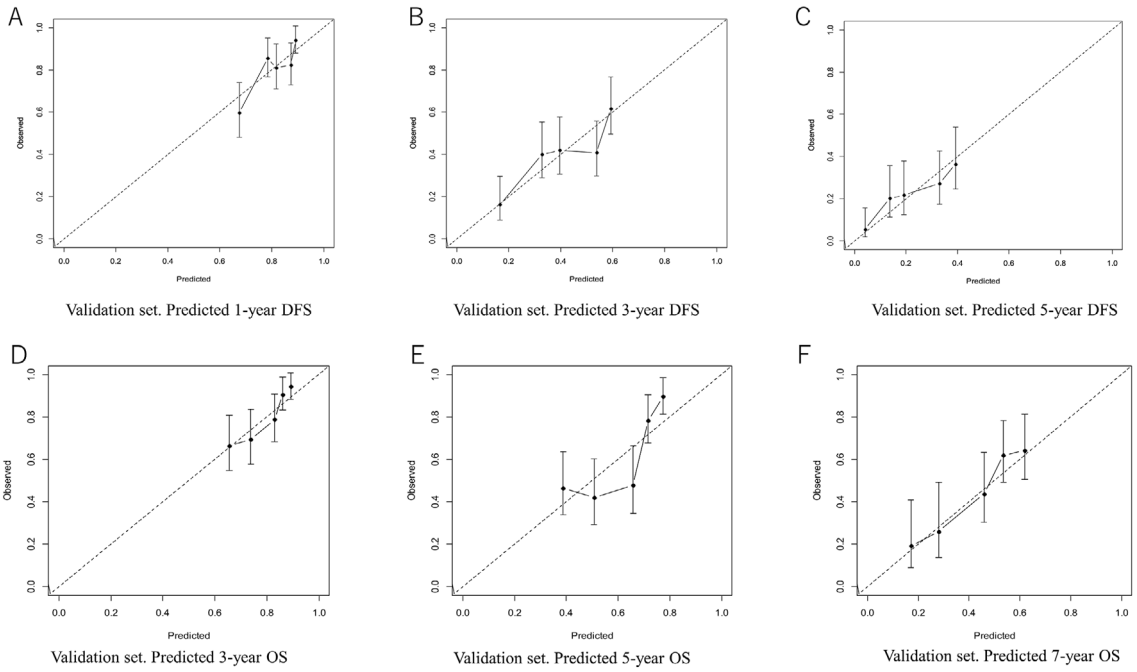

Figure 5. **Calibration curves for the nomogram-predicted probability of the 1- (A), 3- (B), and 5- (C) year DFS and the nomogram-predicted probability of the 3- (D), 5- (E), and 7- (F) year OS in the validation set. Curves for a hypothetical ideal nomogram are represented by dashed lines, and those for the current nomogram are represented by solid lines. Vertical bars indicate 95% confidence intervals. DFS: disease-free survival, OS: overall survival**

Table 6. **Comparison of Staging Systems for Predicting Tumor Recurrence in Validation Set.**

Variables	C-index (95%CI)
Nomogram	0.614 (0.575-0.654)
CLIP score	0.581 (0.545-0.617)
Okuda score	0.551 (0.522-0.580)
JIS score	0.583 (0.543-0.623)
TNM	0.554 (0.516-0.593)
BCLC	0.560 (0.523-0.597)

CLIP: Cancer of the Liver Italian Program, JIS: Japan Integrated Staging, TNM: tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan, BCLC: Barcelona Clinic Liver Cancer. In parentheses: percentages or 95% confidence interval

Table 7. **Comparison of Staging Systems for Predicting Mortality in Validation Set.**

Variables	C-index (95%CI)
Nomogram	0.657 (0.608-0.706)
CLIP score	0.604 (0.555-0.652)
Okuda score	0.598 (0.557-0.640)
JIS score	0.608 (0.556-0.661)
TNM	0.551 (0.498-0.604)
BCLC	0.586 (0.539-0.633)

CLIP: Cancer of the Liver Italian Program, JIS: Japan Integrated Staging, TNM: tumor-node-metastasis staging system developed by Liver Cancer Study Group of Japan, BCLC: Barcelona Clinic Liver Cancer. In parentheses: percentages or 95% confidence interval

addition, the nomogram showed the highest C-index for predicting tumor recurrence. The differences were statistically significant between the C-index of the nomogram and that of the other prognostic scores (all $p<0.05$).

As shown in Table 5, the C-statistics for the OS based on the nomogram in the training set, CLIP score, Okuda score, JIS score TNM, and BCLC were 0.692, 0.597, 0.603, 0.590, 0.538, and 0.561, respectively. In addition, the nomogram showed the highest C-index for predicting mortality. The differences were statistically significant between the C-index of the nomogram and that of the other prognostic scores (all $p<0.05$).

External validation

The scoring system built for the training set was externally validated in the validation set. The external validation analysis confirmed that these models performed well in terms of calibration (Fig. 5) and discrimination (C-indexes of 0.614 for the DFS and 0.657 for the OS) (Table 6, 7). The nomogram showed the highest C-index for predicting tumor recurrence, and the differences were statistically significant between the C-index of the nomogram and that of

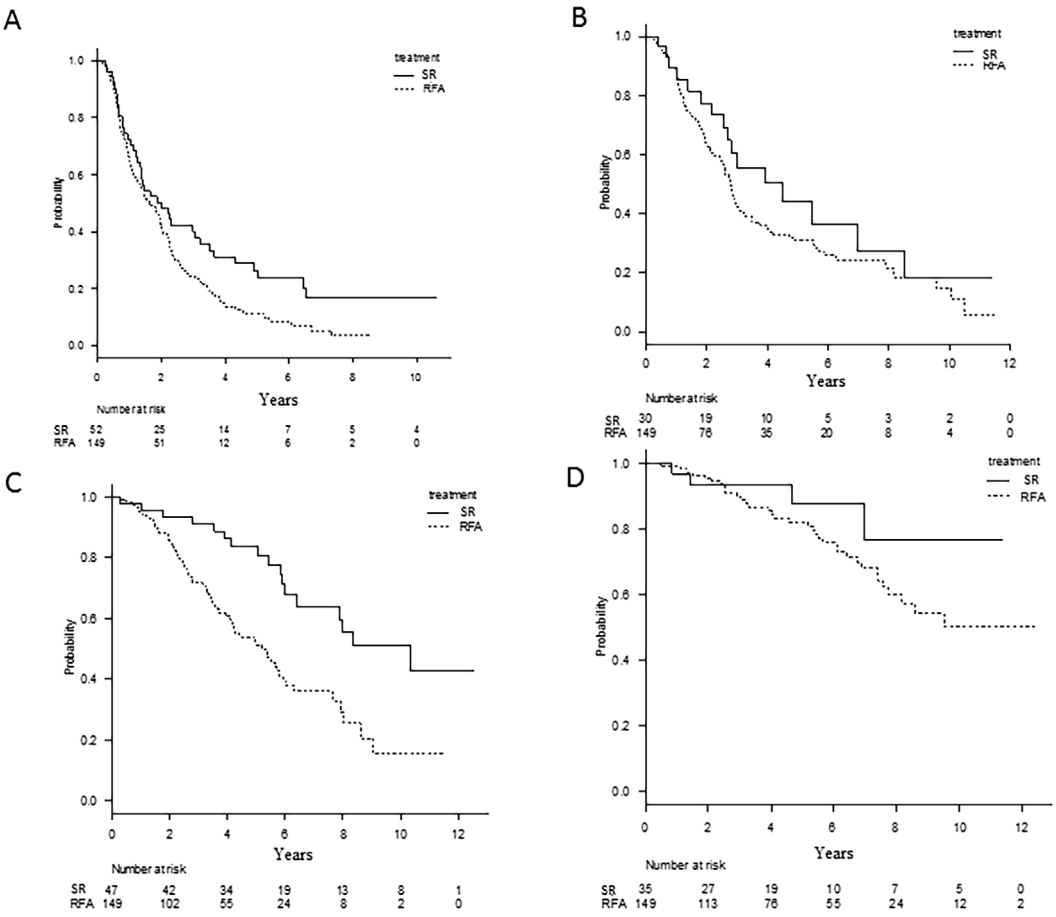

Figure 6. The DFS and OS in similar-stage HCC (≤ 3 tumors each ≤ 3 cm in diameter) patients receiving SR or RFA in the training set. (A) Patients undergoing SR had a better DFS than high-risk patients receiving RFA (p=0.0331). (B) The DFS was similar between the low-risk RFA and SR groups (p=0.184). (C) SR patients had a better OS than high-risk RFA patients (p=0.0004). (D) The OS was similar between the low-risk RFA and SR groups (p=0.165). DFS: disease-free survival, OS: overall survival, HCC: hepatocellular carcinoma, SR: surgical resection, RFA: radiofrequency ablation

the Okuda score (C-index 0.551, p=0.010), TNM (C-index 0.554, p=0.003), and BCLC (C-index 0.560, p=0.012; Table 6). However, there was no significant difference between the C-index of the nomogram and that of the CLIP score (C-index 0.581, p=0.114) or JIS score (C-index 0.583, p=0.151; Table 6).

The nomogram showed the highest C-index for predicting mortality, and the differences were statistically significant between the C-index of the nomogram and that of the Okuda score (C-index 0.598, p=0.039), TNM (C-index 0.551, p=0.002), and BCLC (C-index 0.586, p=0.014; Table 7). However, there was no significant difference between the C-index of the nomogram and that of the CLIP score (C-index 0.604, p=0.086) or JIS score (C-index 0.608, p=0.115; Table 7).

Risk group classification and the survival analysis between SR and RFA patients

The median of the nomogram scores in patients receiving RFA was 233 in the DFS and 158 in the OS, respectively. The DFS and OS were compared between SR and RFA patients in the respective high- and low-risk groups in the training set. High-risk patients receiving SR (n=52) were associated with a better DFS than those undergoing RFA (n=149) (p=0.0331; Fig. 6A). However, the DFS was similar between the low-risk RFA (n=149) and SR groups (n=30) (p=0.184; Fig. 6B). High-risk patients receiving SR (n=47) were associated with a better OS than those undergoing RFA (n=149) (p=0.0004; Fig. 6C). However, the OS was similar between the low-risk RFA (n=149) and SR groups (n=35) (p=0.165; Fig. 6D).

The DFS and OS were also compared between SR and RFA patients in the respective high- and low-risk groups in the validation set. Similar to the results in the training set, high-risk patients receiving SR (n=48) were associated with a better DFS than those undergoing RFA (n=145) (p=0.0473; Fig. 7A). However, the DFS was similar between the low-risk RFA (n=127) and SR groups (n=47) (p=0.141; Fig. 7B). High-risk patients receiving SR (n=47) were associated with a better OS than those undergoing RFA (n=125) (p=0.0148; Fig. 7C). However, the OS was similar between the low-risk RFA (n=147) and SR groups (n=48) (p=0.625;

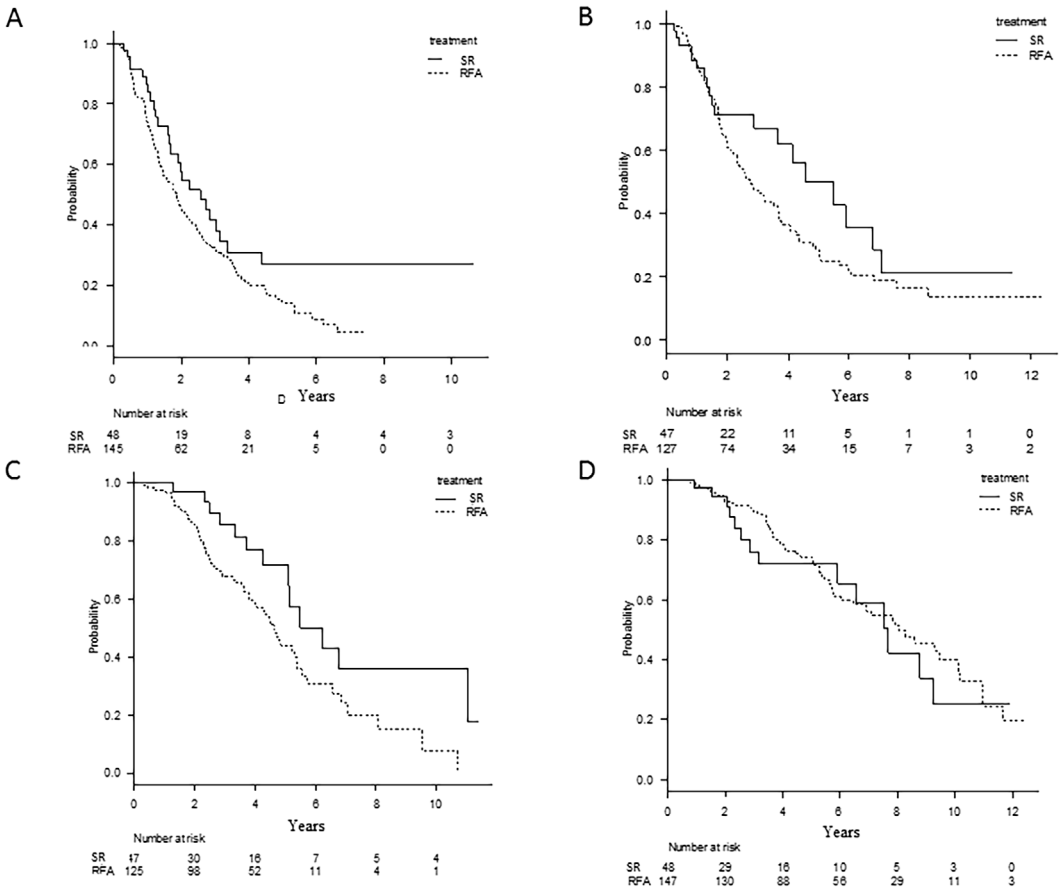

Figure 7. The DFS and OS in similar-stage HCC (≤ 3 tumors each ≤ 3 cm in diameter) patients receiving SR or RFA in the validation set. (A) Patients undergoing SR had a better DFS than high-risk patients receiving RFA (p=0.0473). (B) The DFS was similar between the low-risk RFA and SR groups (p=0.141). (C) SR patients had a better OS than high-risk RFA patients (p=0.0148). (D) The OS was similar between the low-risk RFA and SR groups (p=0.625). DFS: disease-free survival, OS: overall survival, HCC: hepatocellular carcinoma, SR: surgical resection, RFA: radiofrequency ablation

Fig. 7D).

Discussion

By examining the records of patients receiving RFA as the primary treatment for early-stage HCC, we constructed statistical predictive nomograms tailored to the individual patient with the ability to predict the DFS and OS after RFA. The nomograms derived from five or six clinical variables of pre-RFA patients were simple and easy-to-use graphical calculators. It is important to note that the nomograms did not include pathological findings.

Both the nomograms for the DFS and OS had C-indexes of >0.60, and they were more useful and reliable than the widely used traditional staging systems. Some staging systems have been developed to classify patients after RFA; however, none of these systems are specifically developed to predict outcomes after RFA. The TNM classification does not include the residual liver function, which influences the prognosis of HCC patients. In addition, the CLIP or Okuda scores may be less accurate scoring systems for tumor classification due to the subjective measure of tumor extension, and these systems may not be enough to stratify post-treatment risk in HCC patients (14, 15). The BCLC staging system, the scoring system with the most reliable ability to predict the survival and the most widely used (15, 16), was also less reliable than our prognostic models in predicting both the DFS and OS after RFA. Our models were also more reliable for predicting the DFS and OS than the JIS score in a Japanese cohort of HCC patients. The easy-to-use graphical models consist of clinical variables, including the tumor number, tumor size, AFP level, AST level, serum albumin level, MELD score, and age. These variables include tumor-related factors as well as those associated with the liver function. Other studies have shown that the serum albumin level was significantly associated with the DFS and OS (17-19). A low serum albumin level is thought to reflect advanced stages of liver disease, including enhanced liver carcinogenesis and liver failure (20). In this study, an increased age was a significant factor associated with both the DFS and OS. In recent studies of RFA for HCC, a multivariate analysis revealed that an increased age was significantly associated with a decreased OS (9, 21).

Similar to our study, the MELD score, representing the

residual liver function, was a significant factor associated with the OS in post-RFA patients (22). In addition to the Child-Pugh score, the MELD score is one of the most widespread scoring systems of liver cirrhotic patients. In previous studies (23, 24), a high serum AFP level, tumor size, and tumor number were predictive factors of distant HCC recurrence after RFA. The tumor size, tumor number, and serum AFP level are likely related to micrometastasis, and a high serum AFP level is usually observed in tumors of high-grade malignancy. Furthermore, a moderate increase in the AFP level, unrelated to tumors, is a well-known risk factor of HCC occurrence in cirrhotic liver (25). A high AST level was found to be an independent predictive factor for HCC recurrence after RFA (19), and active inflammation in the non-tumorous area and impairment of the liver function may be associated with metachronous multicentric carcinogenesis.

Successful antiviral treatment prior to the development of cirrhosis will prevent most of the morbidity and mortality associated with those infections. In the training set, 74.8% and 9.8% patients had HCC caused by hepatitis C and B virus, respectively, and some of them had a previous or ongoing use of an antiviral agent. However, we did not examine the effect of treatments such as successful antiviral therapy for hepatitis B or C, because viral infection as a cause of HCC was not a parameter in the nomograms.

These nomograms were developed to predict the DFS and OS after RFA for early-stage HCC. Importantly, these predictive nomograms may not only serve as individualized predictors for the DFS and OS but may also appropriately allocate treatment and follow-up strategies. High-risk patients receiving SR were associated with a better DFS and OS than those undergoing RFA. However, the DFS and OS were similar between the low-risk RFA and SR groups. For patients with higher risks of HCC recurrence, these graphical models may facilitate the appropriate allocation of patients to receive aggressive therapy, such as SR, to achieve better tumor control and a long-term survival. Furthermore, we conducted an internal validation with a resampling method by bootstrapping in both cohorts to calibrate the nomograms, and the utility of the nomogram derived from the training set was confirmed externally from different patient cohorts.

Several limitations associated with the present study warrant mention. First, this study was retrospective. Second, the choice of treatment modalities for primary and recurrent HCCs was dependent on the patient preferences and the clinical practices of surgeons and hepatologists. Third, the tumor localization, vascular pattern, and morphology as well as the technical success were not incorporated into the nomograms. Fourth, the age and serum albumin level were incorporated into the RFA-based nomograms. According to the nomograms, SR may be a better choice than RFA in high-risk patients with a high age and low serum albumin level. However, a high age and low serum albumin level are not favorable conditions for SR. A prospective study is required to validate the prognostic accuracy of these nomograms.

In conclusion, we constructed reliable and useful nomograms that accurately predict the DFS and OS after RFA for early-stage HCC patients. In clinical practice, these graphical tools are easy to use and will assist physicians in the therapeutic decision-making process because they are able to provide precise information on the per-patient prognosis and classify post-RFA patients into low- and high-risk groups.

The authors state that they have no Conflict of Interest (COI).

References

1. Gomaa AI, Khan SA, Toledano MB, et al. Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis. World J Gastroenterol **14**: 4300-4308, 2008.
2. European Association for the Study of the Liver; European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol **56**: 908-943, 2012.
3. Bruix J, Sherman M; American Association for the Study of Liver D. Management of hepatocellular carcinoma: an update. Hepatology **53**: 1020-1022, 2011.
4. Taura K, Ikai I, Hatano E, et al. Influence of coexisting cirrhosis on outcomes after partial hepatic resection for hepatocellular carcinoma fulfilling the Milan criteria: an analysis of 293 patients. Surgery **142**: 685-694, 2007.
5. Poon RT. Optimal initial treatment for early hepatocellular carcinoma in patients with preserved liver function: transplantation or resection? Ann Surg Oncol **14**: 541-547, 2007.
6. Lau WY, Lai EC. The current role of radiofrequency ablation in the management of hepatocellular carcinoma: a systematic review. Ann Surg **249**: 20-25, 2009.
7. Takahashi S, Kudo M, Chung H, et al. PIVKA-II is the best prognostic predictor in patients with hepatocellular carcinoma after radiofrequency ablation therapy. Oncology **75** (Suppl 1): 91-98, 2008.
8. Kobayashi M, Ikeda K, Kawamura Y, et al. High serum des-gamma-carboxy prothrombin level predicts poor prognosis after radiofrequency ablation of hepatocellular carcinoma. Cancer **115**: 571-580, 2009.
9. Waki K, Aikata H, Katamura Y, et al. Percutaneous radiofrequency ablation as first-line treatment for small hepatocellular carcinoma: results and prognostic factors on long-term follow up. J Gastroenterol Hepatol **25**: 597-604, 2010.
10. Bruix J, Sherman M; Practice Guidelines Committee AAftSoLD. Management of hepatocellular carcinoma. Hepatology **42**: 1208-1236, 2005.
11. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol **5**: 649-655, 1982.
12. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med **15**: 361-387, 1996.
13. Kang L, Chen W, Petrick NA, et al. Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach. Stat Med **34**: 685-703, 2015.
14. Grieco A, Pompili M, Caminiti G, et al. Prognostic factors for survival in patients with early-intermediate hepatocellular carcinoma undergoing non-surgical therapy: comparison of Okuda, CLIP, and BCLC staging systems in a single Italian centre. Gut **54**: 411-418, 2005.

15. Guglielmi A, Ruzzenente A, Pachera S, et al. Comparison of seven staging systems in cirrhotic patients with hepatocellular carcinoma in a cohort of patients who underwent radiofrequency ablation with complete response. Am J Gastroenterol **103**: 597-604, 2008.
16. Cillo U, Vitale A, Grigoletto F, et al. Prospective validation of the Barcelona Clinic Liver Cancer staging system. J Hepatol **44**: 723-731, 2006.
17. Yamanaka Y, Shiraki K, Miyashita K, et al. Risk factors for the recurrence of hepatocellular carcinoma after radiofrequency ablation of hepatocellular carcinoma in patients with hepatitis C. World J Gastroenterol **11**: 2174-2178, 2005.
18. Peng ZW, Zhang YJ, Chen MS, et al. Radiofrequency ablation as first-line treatment for small solitary hepatocellular carcinoma: long-term results. Eur J Surg Oncol **36**: 1054-1060, 2010.
19. Fuke H, Sugimoto K, Shiraki K, et al. Predictive factors for distant recurrence of HCV-related hepatocellular carcinoma after radiofrequency ablation combined with chemoembolization. Aliment Pharmacol Ther **27**: 1253-1260, 2008.
20. Huang YH, Wu JC, Chen CH, et al. Comparison of recurrence after hepatic resection in patients with hepatitis B vs. hepatitis C-related small hepatocellular carcinoma in hepatitis B virus endemic area. Liver Int **25**: 236-241, 2005.
21. Kao WY, Chiou YY, Hung HH, et al. Younger hepatocellular carcinoma patients have better prognosis after percutaneous radiofrequency ablation therapy. J Clin Gastroenterol **46**: 62-70, 2012.
22. Santambrogio R, Bruno S, Kluger MD, et al. Laparoscopic ablation therapies or hepatic resection in cirrhotic patients with small hepatocellular carcinoma. Dig Liver Dis **48**: 189-196, 2016.
23. N'Kontchou G, Mahamoudi A, Aout M, et al. Radiofrequency ablation of hepatocellular carcinoma: long-term results and prognostic factors in 235 Western patients with cirrhosis. Hepatology **50**: 1475-1483, 2009.
24. Shiina S, Tateishi R, Arano T, et al. Radiofrequency ablation for hepatocellular carcinoma: 10-year outcome and prognostic factors. Am J Gastroenterol **107**: 569-577; quiz 578, 2012.
25. Chen TM, Huang PT, Tsai MH, et al. Predictors of alpha-fetoprotein elevation in patients with chronic hepatitis C, but not hepatocellular carcinoma, and its normalization after pegylated interferon alfa 2a-ribavirin combination therapy. J Gastroenterol Hepatol **22**: 669-675, 2007.