



A nomogram to predict prognosis after surgery for young patients with hepatocellular carcinoma

Xingchen Li, Xinyu Bi, Jianjun Zhao, Zhiyu Li, Jianguo Zhou, Zhen Huang, Yefan Zhang, Hong Zhao, Jianqiang Cai

Department of Hepatobiliary Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Contributions: (I) Conception and design: J Cai, H Zhao; (II) Administrative support: J Cai, H Zhao; (III) Provision of study materials of patients: X Bi, J Zhao, Z Li; (IV) Collection and assembly of data: J Zhou, Z Huang, Y Zhang; (V) Data analysis and interpretation: X Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jianqiang Cai, MD; Hong Zhao, MD. Department of Hepatobiliary Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. Email: caijianqiang@cicams.ac.cn; zhaohong@cicams.ac.cn.

Background: Only few studies have been evaluated the clinical characteristics and prognosis of hepatocellular carcinoma (HCC) in young patients. The purpose of this study is to identify prognostic factors and develop an efficient and practical nomogram to predict cancer-specific survival (CSS) in young patients with HCC.

Methods: Four hundred and forty-one young patients with HCC who had undergone surgery from 2004–2015 were selected from the Surveillance, Epidemiology, and End Results (SEER) database. The competing risk model, Lasso and Cox regression were used to screen prognostic factors for CSS, and a prognostic nomogram was established using these factors. Thirty-nine young patients with HCC from the National Cancer Center, Cancer Hospital, Chinese Academy of Medical Science were used to validate our model. To further evaluate the predictive performance of our model, the concordance index was calculated and the calibration curves were drawn. The clinical usefulness was evaluated by decision curve analysis (DCA). Finally, all patients were grouped by our nomogram. The survival of different risk groups was analyzed using the Kaplan-Meier method, and the differences among survival curves were compared by the log-rank test.

Results: The median survival times of the SEER training group and the external National Cancer Center validation group were 41 and 52 months, respectively. Histological grade, tumor size, Alpha-fetoprotein (AFP), T stage, and M stage were selected as independent factors for CSS, and a prognostic nomogram was established. The concordance indices of the training and external validation groups were 0.76 (95% CI, 0.72 to 0.80) and 0.92 (SE=0.085), respectively. The calibration plots showed good agreement. DCA revealed that our nomogram resulted in a better clinical net benefit than the AJCC 7th edition and Barcelona Clinic Liver Cancer staging systems. Patients were divided into two risk groups according to the cut-off value of 125 of the total points from our nomogram. Kaplan-Meier plots for CSS were performed using the log-rank test, the P-value of which was <0.001.

Conclusions: The practical nomogram resulted in a more-accurate prognostic prediction for young HCC patients after curative liver resection.

Keywords: Hepatocellular carcinoma (HCC); young patients; nomogram; prognosis; prediction model

Submitted Dec 09, 2020. Accepted for publication Feb 07, 2021

doi: 10.21037/tcr-20-3411

View this article at: <http://dx.doi.org/10.21037/tcr-20-3411>

Introduction

Liver cancer is the fourth most common cause of cancer-related death, with the sixth-highest incidence rate worldwide (1). Hepatocellular carcinoma (HCC) accounts for the majority of primary liver cancer; however, the efficacy of therapies for HCC is poor. Although HCC is typically diagnosed in middle-aged and elderly populations, the peak incidence age differs substantially according to geographical regions and etiology. Hepatitis B virus (HBV)-associated HCC patients in East Asia and Africa are generally younger than those with hepatitis C virus (HCV)-associated HCC in Western countries (2). Moreover, in high-risk areas of East Asia, HCC is often diagnosed in individuals younger than 40 years of age, which is rare in European countries (3). Compared with middle-aged and elderly individuals, the liver cirrhosis rate of young patients is lower, and the liver function of these patients is better; however, HBV infection rates and serum alpha-fetoprotein (AFP) levels are higher among younger patients, and they generally have large tumor diameters, worse tumor differentiation, and stronger invasion (4,5). HCC patients often have a three-step development process of hepatitis-cirrhosis-liver cancer, but the process of cirrhosis in young patients is not apparent; therefore, researchers hypothesized that the tumorigenesis mechanisms of HCC are dissimilar at different ages (4). However, there are few studies on the risk factors for HCC in young patients. Several studies have found that the prognosis of HCC in young patients is better than that in older patients and that there is no significant difference in overall survival (OS) and disease-free survival (DFS) between the two age groups (5-7), while other studies have shown that the prognosis of young patients is worse (8-10). The disparities in these studies may be due to individual differences in the data of the patients; there is no strict matching between the clinical baseline data of the two groups of patients. In general, the clinicopathological characteristics of young patients with liver cancer are specific, and the prognosis is controversial.

The Surveillance, Epidemiology, and End Results (SEER) program is one of the largest and most representative cancer registration databases in North America and provides extensive data for different types of malignancies. It serves as an exhaustive source of information on malignancies in the United States, which affects approximately 35% of the U.S population (11).

Nomograms help to quantify various factors and use them for constructing a model to predict the outcome (12).

This model could be used to predict specific outcomes for any individual patient. The one in 10 rule is an empirical rule for how many predictor parameters can be estimated from data when conducting regression analysis while maintaining the risk of overfitting low in statistics. The rule states that one predictive variable can be studied for every ten events (13,14); due to the low incidence of HCC in young patients, it is not statistically effective to build the model with single-center data. In this study, the clinicopathological features and survival data of patients with HCC were obtained from the SEER database, which was used as a training cohort for constructing a nomogram to predict CSS. The prognostic evaluation models (PEM) constructed with the nomogram were then validated with an external cohort of HCC patients. We present the following article in accordance with the TRIPOD reporting checklist (available at <http://dx.doi.org/10.21037/tcr-20-3411>).

Methods

Ethics statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of the Cancer Hospital, Chinese Academy of Medical Sciences (ID: NCC2019C-016) and informed consent was taken from all the patients.

Data acquisition and the process of patient selection in the training group

The data were obtained from the SEER database [SEER 18 Regs Custom Data (with additional treatment fields), Nov 2018 sub (1973–2016 varying)] using the SEER*Stat software (version 8.3.6). In the SEER database, the surgical information about the primary site was only available after 1988. Generally, Collaborative Stage (CS) information must be acquired to obtain more detailed information on the American Joint Committee on Cancer (AJCC) staging. However, this information was only available between 2004 and 2015; hence, we included HCC patients from 2004 to 2015.

Inclusion criteria include (I) patients diagnosed with HCC from 2004 to 2015 (histologic type ICD-O-3=8170–8175), (II) a maximum age of 40 years, and (III) patients with available follow-up data. Exclusion criteria were listed as following: (I) patients with a history of other malignancies,

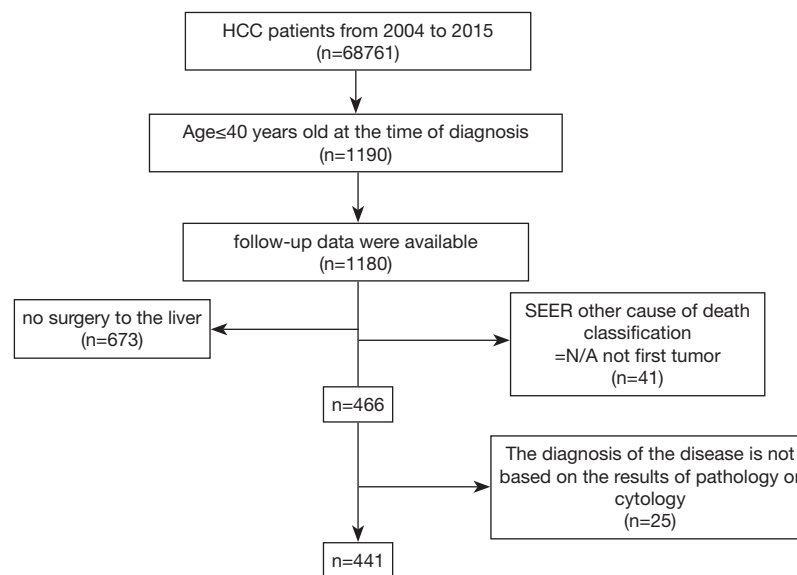


Figure 1 Flow chart of patients selection in the training group from the SEER database.

(II) recurrent HCC with surgical treatments, (III) the patient did not undergo surgical treatment, and (IV) the diagnosis of the disease was not based on the results of pathology or cytology. The selection process for enrolling patients is shown in *Figure 1*.

External validation group

Our study's independent external validation population was drawn from 39 consecutive young HCC patients from March 2012 to November 2017, who were treated with hepatic surgery at the National Cancer Center (NCC) Cancer Hospital, Chinese Academy of Medical Science. The inclusion and exclusion criteria were the same as for the SEER training group. The follow-up information was obtained via telephone, e-mail, letter, or outpatient service. Furthermore, similar variable information as for the SEER training group was included. All external validation group patients were staged with the following two major staging systems: the AJCC tumor-nodes-metastasis (TNM) staging system (7th edition 2010) and the Barcelona Clinic Liver Cancer (BCLC) staging system.

Definition of primary endpoint and staging system

In the statistical analysis process, we mainly applied the following two outcome indicators: cancer-specific survival (CSS) and survival time. CSS was determined as the time

interval from the diagnosis date until the date of tumor-related death. If a patient died due to a non-tumor cause, we classified it as another type of non-censored data. Patients who survived at the end of follow-up were classified as censored. In the SEER database, when the survival time is less than one month, it is defined as zero. According to standard epidemiological conventions, a survival of 0.5 months was assigned to those patients. HCC was staged using the 7th edition of the AJCC TNM staging system. Only the patients from 2010 to 2015 in the original SEER data had information regarding the 7th edition AJCC staging system. For other patients, the fields of CS Extension, CS Lymph Nodes, CS Mets at DX, and CS tumor size were used to determine the 7th edition AJCC staging.

Statistical analysis

First, risk factors related to CSS were screened based on a competing risk model. Gray's test was used to compare the cumulative incidence of death due to cancer and death due to other reasons, such as competing events (15). Factors with a P value of less than 0.2 entered the subsequent univariate Cox analysis. To further in-depth analyze the variables related to CSS, patients who did not die from HCC were excluded. Second, to minimize the possibility of overfitting of our model, variables with a P value less than 0.05 in the univariate Cox analysis were included in the

Lasso penalized cox model to filter the variables further. LASSO regression could set the coefficients of many irrelevant features exactly to zero based on the regulation weight λ . Optimal values of λ were determined by 10-fold cross-validation (16). Finally, the results of the LASSO regression were included in the multivariate cox regression analysis, and the final CSS-related variables were selected based on the P-value of less than 0.05. All P-values were reported to 3 decimal places. Subsequently, a nomogram of young liver cancer patients using these factors was constructed to predict 1-, 2-, and 3-year CSS.

The discriminative ability was evaluated using the c-index and receiver operating characteristics (ROC) curves of different times. The nomogram is further assessed for its predictive accuracy using calibration curves. The calibration curve can be used to observe the agreement between the predicted value and the actual value. In the calibration curve analysis, the prediction error was assessed using the nonparametric bootstrap procedure with 500 bootstrap replicates to minimize bias.

The decision curve analysis (DCA) was plotted to examine the usefulness and benefit of our nomogram. Our nomograms were subjected to 1,000 bootstrap resamples for internal validation. An independent external validation group from NCC was used to validate our model. The total score of each patient in the validation set was calculated according to the generated nomogram. Then, cox regression was performed using the total score as a factor. The C-index, calibration curve, and clinical decision curve were derived based on the cox regression in the validation group. The cut-off value of total score were analyzed using the X-tile program to distinguish the differences in young HCC patients' survival. The X-tile program is used to find the optimal value of continuous variables, and the program's application has been reported in numerous studies (17). We then stratified patients into different risk groups according to the cut-off value for the total points. The Survival curves of the different risk groups were plotted by the Kaplan-Meier method, and the log-rank test was used to determine significance.

Statistical analysis was performed using R (version 3.6.2) software. Statistical significance was set as above.

Results

Characteristics of training and validation groups

A total of 441 young HCC patients from 2004 to 2015 were included in our training group from the SEER database,

of whom 284 cases (6.6%) were male. The median survival time of the training group was 41.00 (17.00–86.00) months. Among them, the cause of death in 20 patients was non-tumor or unknown; therefore, in the process of screening for CSS risk factors, they were deleted. The validation group included 39 patients from the NCC who underwent surgical treatment in our hospital from March 2012 to November 2017. The last follow-up time was March 2020. Seven patients died, and the cause of death was HCC. The clinical and pathological features of the training and validation groups are listed in *Table 1*.

Prognostic factors related to CSS for young HCC patients

In Gray's test, marital status, histological grade, size group, AFP, T stage, N stage, and M stage were associated with CSS. The Cumulative incidence function (CIF) of young HCC patients with risk factors with a P-value <0.2 for CSS is displayed in *Figure 2*. After deleting the data of patients who died from non-tumor causes, the data of 421 patients entered the follow-up analysis. After primary Cox univariate analysis filtration, the remaining six variables were included in the Lasso regression. We applied 10× cross-validation to find the minimum of partial likelihood deviance which could make our model simpler. A simplified model helped to avoid overfitting as much as possible. Finally, the selected variables when the partial likelihood deviance was minimal ($\lambda = -4.44$) were grade, size group, AFP, T stage, and M stage. Additional details of the prognostic factor selection using the LASSO regression model are disclosed in *Figure 3*. In the multivariate Cox analysis, these five variables remained independently related to CSS. The variables screening process related to CSS details is outlined in *Table 2*. Collinearity diagnostics were examined for the potential presence of collinearity between the independent variables, and Variance inflation factors (VIF) ≤ 5 .

Construction and validation of the nomogram

Five variables, grade, size group, AFP, T stage, and M stage, were associated with CSS and were used to construct a nomogram, as shown in *Figure 4*. The C-index of nomogram in the SEER training group was 0.76 (95% CI, 0.72 to 0.80). Time-dependent ROC curves of CSS in the SEER group showed that the AUC at 1, 2 and 3 years was 0.81, 0.77, and 0.76, respectively. The C-index of the 7th AJCC system was 0.68 (95% CI, 0.63 to 0.72). The AUC for 1-, 2-, and 3- years of the 7th edition AJCC staging system

Table 1 Clinical and pathological features of the training and validation groups

| Variable | SEER training group | CICAMS validation group |
|--------------------------|---------------------|-------------------------|
| Number of patients | 441 | 39 |
| Sex | | |
| Male | 284 (64.40%) | 37 (94.87%) |
| Female | 157 (35.60%) | 2 (5.13%) |
| Age, median (Q1–Q3) | 32.00 (21.00–38.00) | 35.00 (31.50–39.00) |
| Race ^a | | |
| White | 265 (60.09%) | – |
| Black | 50 (11.34%) | – |
| Other | 122 (27.66%) | 39 (100.00%) |
| Unknown | 4 (0.91%) | – |
| Insurance recode | | |
| Insured and any Medicaid | 288 (65.31%) | 35 (89.74%) |
| Uninsured | 8 (1.81%) | 4 (10.26%) |
| Unknown | 145 (32.88%) | 0 (0.00%) |
| Marital status | | |
| Unmarried | 169 (38.32%) | 14 (35.90%) |
| Married | 260 (58.96%) | 25 (64.10%) |
| Unknown ^b | 12 (2.72%) | 0 (0.0%) |
| Grade | | |
| Grades I+II | 111 (25.17%) | 10 (25.64%) |
| Grades III+IV | 252 (57.14%) | 15 (38.46%) |
| Unknown | 111 (25.17%) | 14 (35.89%) |
| AJCC stage 7th | | |
| I | 190 (43.08%) | 30 (76.9%) |
| II | 93 (21.09%) | 3 (7.6%) |
| III | 74 (16.78%) | 5 (12.8%) |
| IV | 62 (14.06%) | 0 (0.00%) |
| Unknown | 22 (4.99%) | 1 (2.5%) |
| T stage | | |
| T1 | 210 (47.62%) | 30 (76.92%) |
| T2 | 106 (24.04%) | 3 (7.69%) |
| T3 | 89 (20.18%) | 5 (12.82%) |
| T4 | 21 (4.76%) | 0 (0.00%) |
| Tx | 15 (3.40%) | 1 (2.56%) |

Table 1 (continued)

Table 1 (continued)

| Variable | SEER training group | CICAMS validation group |
|------------------------------|---------------------|-------------------------|
| N stage | | |
| N0 | 388 (87.98%) | 39 (100.00%) |
| N1 | 40 (9.07%) | 0 (0.00%) |
| Nx | 13 (2.95%) | 0 (0.00%) |
| M stage | | |
| M0 | 402 (91.16%) | 39 (100.00%) |
| M1 | 35 (7.94%) | 0 (0.00%) |
| Mx | 4 (0.91%) | 0 (0.00%) |
| BCLC | | |
| 0 | – | 3 (7.6%) |
| A | – | 25 (64.1%) |
| B | – | 11 (28.2%) |
| C | – | 0 (0.00%) |
| Surgery to the liver-detail | | |
| Local tumor destruction | 32 (7.26%) | 0 (0.00%) |
| Wedge or segmental resection | 132 (29.93%) | 26 (66.6%) |
| Lobectomy | 128 (29.02%) | 4 (10.2%) |
| Extended lobectomy | 34 (7.71%) | 26 (66.6%) |
| Hepatectomy, NOS | 107 (24.26%) | 2 (5.1%) |
| Surgery, NOS | 8 (1.82%) | 7 (17.9%) |
| Surgery to LN | | |
| No | 302 (68.48%) | 35 (89.7%) |
| Yes | 136 (30.84%) | 4 (10.2%) |
| Unknown | 3 (0.68%) | 0 (0%) |
| AFP | | |
| Positive/elevated | 196 (44.44%) | 25 (64.1%) |
| Negative/normal | 144 (32.65%) | 13 (33.3%) |
| Unknown | 101 (22.90%) | 1 (2.5%) |
| Fibrosis score | | |
| 0–4 | 71 (16.10%) | 13 (34.21%) |
| 5–6 | 43 (9.75%) | 15 (39.47%) |
| Unknown | 327 (74.15%) | 10 (26.32%) |

Table 1 (continued)

Table 1 (continued)

| Variable | SEER training group | CICAMS validation group |
|---------------------------------|---------------------|-------------------------|
| Tumor size group (mm) | | |
| 0–20 | 55 (12.47%) | 1 (2.56%) |
| 21–50 | 114 (25.85%) | 21 (53.85%) |
| 51–100 | 109 (24.72%) | 10 (25.64%) |
| ≥101 | 140 (31.75%) | 6 (15.38%) |
| Unknown | 23 (5.22%) | 1 (2.56%) |
| Overall survival | | |
| Alive | 266 (60.32%) | 32 (82.05%) |
| Dead | 175 (39.68%) | 7 (17.95%) |
| Cancer-specific survival | | |
| Alive | 266 (60.32%) | 32 (82.05%) |
| Death due to cancer | 155 (35.15%) | 7 (17.95%) |
| Death due to other reasons | 20 (4.54%) | 0 (0.00%) |
| Survival months, median (Q1–Q3) | 41.00 (17.00–86.00) | 52.00 (29.00–65.00) |

^a, includes American Indian/AK Native, Asian/Pacific Islander; ^b, divorced/separated/single (never married)/unmarried or domestic partner/widowed.

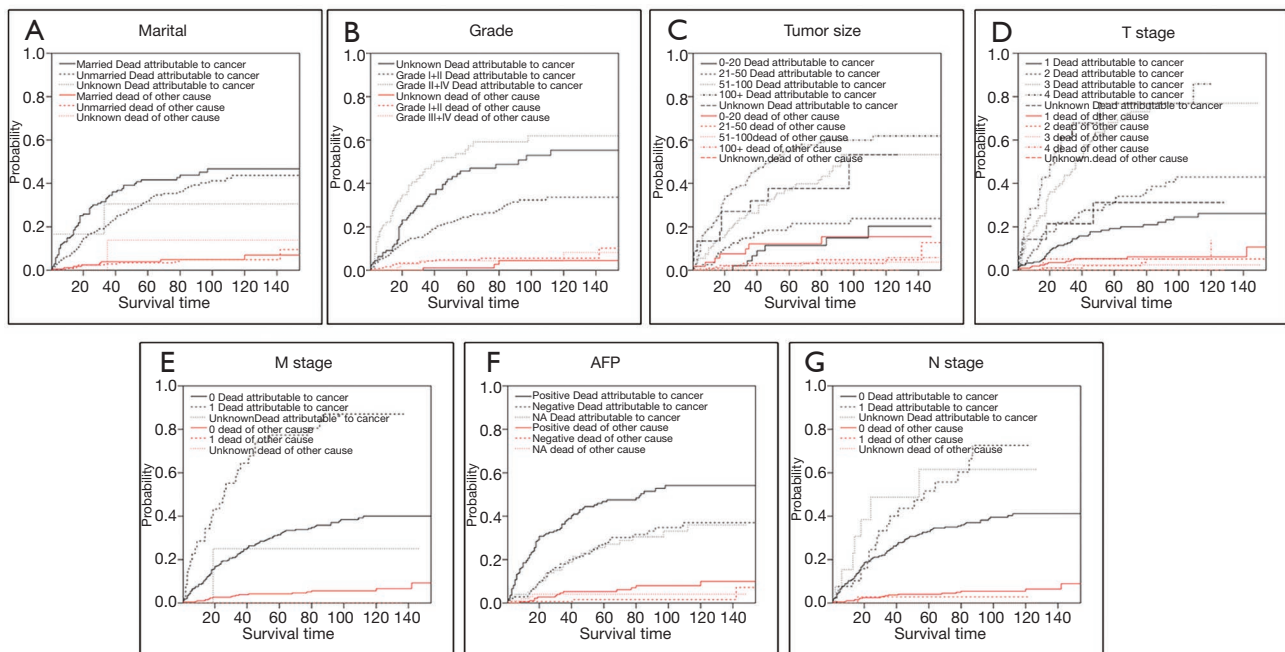


Figure 2 Cumulative incidence function (CIF) of SEER training patients with risk variables for cancer-specific death. (A) Marital status for cancer-specific death; (B) tumor grade for cancer-specific death; (C) tumor size for cancer-specific death; (D) T stage for cancer-specific death; (E) M stage for cancer-specific death; (F) Alpha-fetoprotein (AFP) status for cancer-specific death; (G) N stage for cancer-specific death.

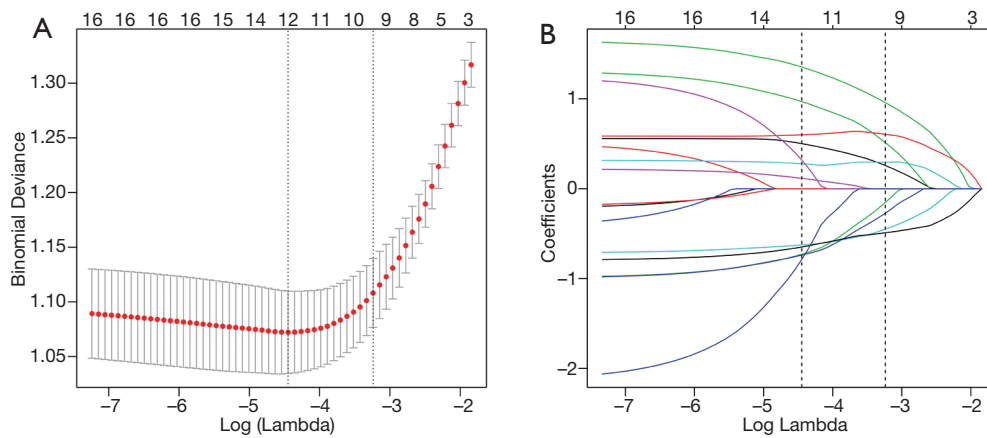


Figure 3 Prognostic factor selection using LASSO regression model. (A) Lasso regression search for the optimal coefficient when the Lambda was -4.44 ; (B) $10\times$ cross-validation was applied for searching the Lambda when partial likelihood deviance was the least.

Table 2 Variable screening process related to CSS

| | Gray's test | | CSS univariant COX | | CSS multivariant COX | | | |
|----------------|---------------|------------------------|--------------------|---------------|-------------------------|--------------------|---------------|-------------------------|
| | P value (CSS) | Filter results (P<0.2) | HR (95% CI) | P value (CSS) | Filter results (P<0.05) | HR (95% CI) | P value (CSS) | Filter results (P<0.05) |
| Sex | 0.302 | | | | | | | |
| Race | 0.213 | | | | | | | |
| Marital status | 0.178 | Chosen | | | Reject | | | |
| Married | | | 1 | reference | | | | |
| Unmarried | | | 0.73 (0.53, 1.01) | 0.057 | | | | |
| Unknown | | | 0.73 (0.23, 2.32) | 0.593 | | | | |
| Grade | <0.001 | Chosen | | | Chosen | | | Chosen |
| Unknown | | | 1 | reference | | 1 | reference | |
| I/II | | | 0.53 (0.37, 0.77) | <0.001 | | 0.53 (0.37, 0.77) | <0.001 | |
| III/IV | | | 1.51 (1.00, 2.28) | 0.051 | | 1.51 (1.00, 2.28) | 0.051 | |
| Surgery to LN | 0.645 | | | | | | | |
| No | | | | | | | | |
| Yes | | | | | | | | |
| Unknown | | | | | | | | |
| Size group | <0.001 | Chosen | | | Chosen | | | Chosen |
| 0-20 | | | 1 | reference | | 1 | reference | |
| 21-50 | | | 1.92 (1.17, 3.14) | 0.001 | | 1.51 (0.64, 3.55) | 0.346 | |
| 51-100 | | | 5.05 (3.21, 7.97) | <0.001 | | 3.46 (1.55, 7.71) | 0.002 | |
| >100 | | | 5.19 (3.25, 8.28) | <0.001 | | 5.46 (2.52, 11.86) | <0.001 | |
| Unknown | | | 1.96 (0.86, 4.43) | 0.108 | | 3.50 (1.30, 9.39) | 0.013 | |

Table 2 (continued)

Table 2 (continued)

| | Gray's test | | CSS univariant COX | | | CSS multivariant COX | | |
|--------------------|---------------|------------------------|--------------------|---------------|-------------------------|----------------------|---------------|-------------------------|
| | P value (CSS) | Filter results (P<0.2) | HR (95% CI) | P value (CSS) | Filter results (P<0.05) | HR (95% CI) | P value (CSS) | Filter results (P<0.05) |
| AFP | <0.001 | Chosen | | | Chosen | | | Chosen |
| Positive | | | 1 | reference | | 1 | reference | |
| Negative | | | 0.45 (0.31, 0.66) | <0.01 | | 0.45 (0.31, 0.66) | <0.001 | |
| Unknown | | | 0.44 (0.29, 0.68) | <0.01 | | 0.44 (0.29, 0.68) | <0.001 | |
| Fibrosis Score | 0.795 | | | | | | | |
| Fibrosis score 5–6 | | | | | | | | |
| Fibrosis score 0–4 | | | | | | | | |
| Unknown | | | | | | | | |
| Insurance | 0.234 | | | | | | | |
| T stage | <0.001 | Chosen | | | Chosen | | | Chosen |
| T1/0 | | | 1 | reference | | | | |
| T2 | | | 1.87 (1.19, 2.92) | <0.001 | | | | |
| T3 | | | 4.85 (3.23, 7.27) | <0.001 | | | | |
| T4 | | | 8.67 (4.75, 15.85) | <0.001 | | | | |
| TX | | | 1.55 (0.55, 4.34) | 0.402 | | | | |
| N stage | 0.010 | Chosen | | | Chosen | | | Chosen |
| N0 | | | 1 | reference | | | | |
| N1 | | | 1.76 (1.11, 2.79) | 0.017 | | | | |
| NX | | | 2.00 (0.93, 4.29) | 0.075 | | | | |
| M stage | <0.001 | Chosen | | | Chosen | | | Chosen |
| M0 | | | 1 | reference | | | | |
| M1 | | | 3.49 (2.33, 5.24) | <0.001 | | | | |
| MX | | | 0.67 (0.09, 4.81) | 0.693 | | | | |

was 0.70, 0.67, and 0.69, respectively. The C-index of NCC validation group was 0.92 (95% CI, 0.84 to 1.00).

Time-dependent ROC curves showed that the AUC at 1, 2 and 3 years was 0.95, 0.91, and 0.85, respectively. The C-index of the AJCC and BCLC staging systems were only 0.80 (95% CI, 0.66 to 0.95) and 0.61(95% CI, 0.42 to 0.80), respectively.

The calibration curves illustrated an optimal agreement in the probability of 1- and 3-year CSS between the prediction by nomogram and actual observation in both in the SEER training and NCC validation groups (see *Figure 5*

for details). Furthermore, the DCA was plotted and our nomogram showed better net benefits with a wider range of threshold probability than either the 7th edition AJCC or BCLC staging systems in the training and validation groups (see *Figure 6* for details).

Risk stratification according to the nomogram

According to the nomogram model, the total score of each training group patient were calculated. The total points ranged from 0 to 244. The best cut-off value of total score

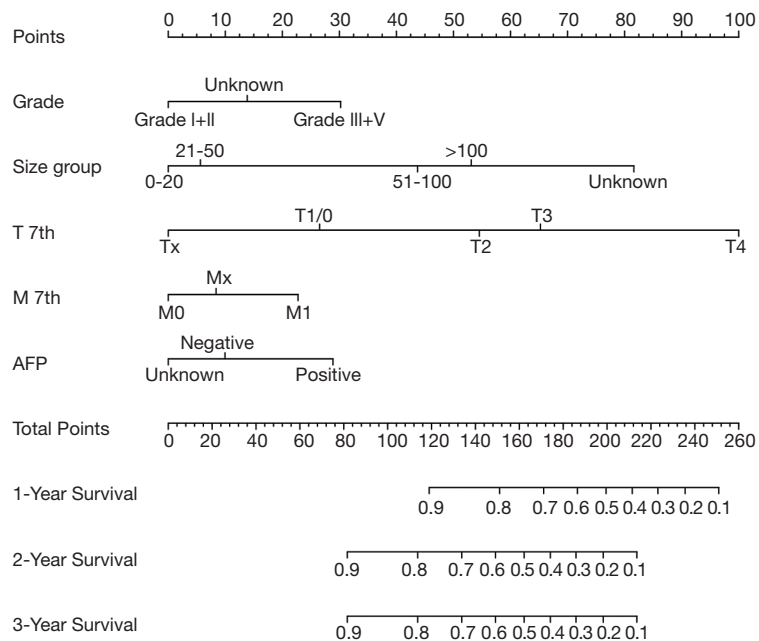


Figure 4 Nomogram predicting cancer specific-survival probability of young hepatocellular carcinoma (HCC) patients after surgery.

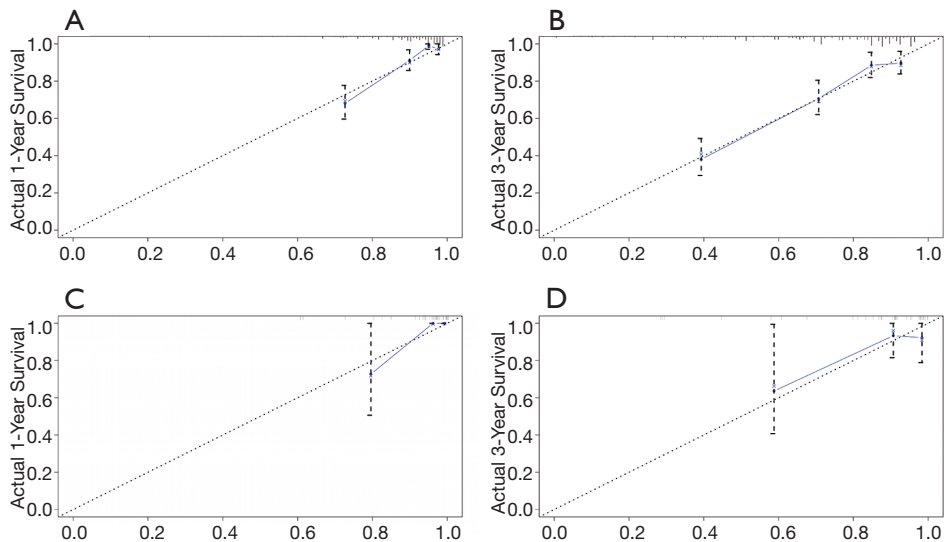


Figure 5 Calibration plot of the nomogram. (A) SEER training group nomogram-predicted probability of 1-year survival; (B) SEER training group nomogram-predicted probability of 3-year survival; (C) CICAMS validation group nomogram-predicted probability of 1-year survival; (D) CICAMS validation group nomogram-predicted probability of 3-year survival.

was 125 and the patients could be divided into low-risk (score ≤ 125) and high-risk group (score > 125). Kaplan–Meier survival curves of the training and validation groups were generated. The log-rank test was used in analysis and the P-value < 0.001 (see further details in *Figure 7*).

Discussion

This study included the largest cohort of young patients with HCC to date as we explored the characteristic of this tumor using the SEER database. The clinicopathological

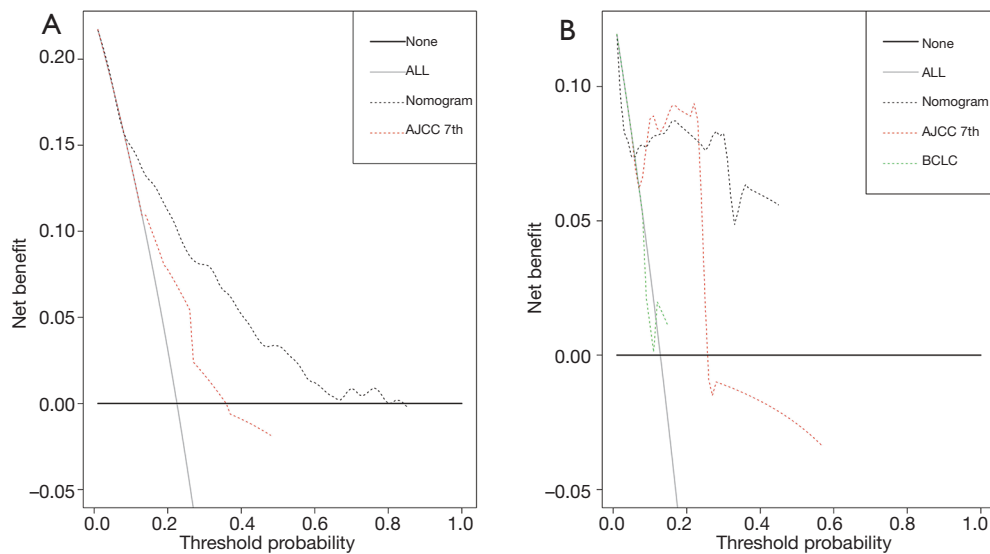


Figure 6 Decision curve analysis comparing the clinical utility of our nomogram model to the AJCC 7th staging system and BCLC staging system. (A) Decision curve analysis (DCA) of the training group (B) DCA of the validation group.

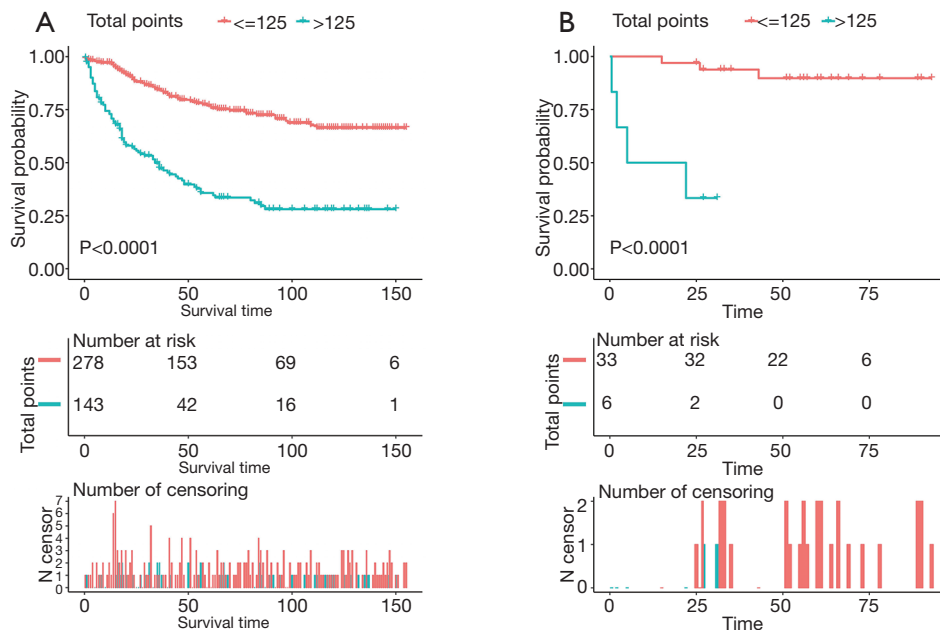


Figure 7 The survival curves of different risk groups according to the total points of 125. (A) Survival curves of the training group; (B) Survival curves of the validation group.

characteristics and prognosis of this population are significantly different from those of middle-aged and elderly patients with liver cancer. Our research focused on identifying prognostic risks and clinicopathological factors, constructing PEM, and validating with an external

validation cohort. We identified grade, size group, AFP, T stage, and M stage as independent prognostic factors for HCC in young patients. Moreover, a nomogram we constructed that contained the risk factors associated with the prognosis of HCC to predict the CSS of young

patients. The calibration curves for 1-, 2-, and 3-year CSS closely matched the ideal 45 degree line, and the *c*-indices of the training and external validation groups were 0.76 and 0.92, respectively, which was significantly higher than that of the conventional staging systems—the 7th edition AJCC and BCLC. These comprehensive analysis designs for screening prognostic factors contributed to the results with high credibility and good predictive effect. Through this model, doctors could estimate the post-operative survival of young patients more precisely and thereafter provide the guidance for the frequency of post-operative surveillance as well as adjuvant therapy in patients with poor prognosis.

In recent years, the number of young people with HCC has increased (18). There are no specific clinical manifestations of HCC in young patients, but the tumor typically grows insidious and fast. Young people seldom have regular physical examinations, and they often do not pay attention to it and are prone to delay the disease. Compared with middle-aged and elderly people, young people often show high AFP levels, large tumor diameter, low tumor differentiation, and high HBV infection rates, but a low degree of cirrhosis (19,20). A study demonstrated that a younger group of patients showed better liver function and had higher AFP levels than an older group; these outcomes are consistent with other previous reports. In addition, younger patients tend to have larger tumors than older patients (6). Overall, the prognosis of young patients with HCC is controversial (5-10). Several researchers have reported that young HCC patients tend to have poor prognoses, compared to middle-aged or older patients. This is most likely because young HCC patients are often diagnosed with advanced stages, as indicated by their advanced TNM stages and high serum AFP levels. In addition, certain patients with HCC may have developed the disease from HBV infection; the proportion is significantly higher than that in elderly patients (19). HBV can develop in non-cirrhotic livers and is also an underlying cause of HCC in young patients. This may be one of the reasons why there is a lesser degree of liver cirrhosis in young patients, and the prognosis of young patients with HCC is better. Other reports have indicated that there is no significant difference in OS and DFS between younger and older patients (5,6).

As a practical tool in clinical practice, the nomogram is more accurate and individualized in predicting the survival and prognosis of patients than traditional neoplasm staging

criteria and has increasingly attracted attention from clinicians. Since the results of the regression equation can be visualized by the nomogram, it has been used to predict the survival and recurrence of various tumors, such as breast cancer (21) and lung cancer (22). A nomogram model for survival or postoperative recurrence of liver cancer has been reported (23-26); the survival of patients with HCC is affected by numerous factors, but most of the predictive factors in these studies only involve the characteristics of tumors, and many are single-center studies without external verification. Since most of the patients with HCC are middle-aged and elderly, the previous prediction model is more suitable for these patients. Research has reported nomograms built on multivariate analyses, which showed that one of the independent predictors for HCC was cirrhosis (27), which is typically more serious in elderly patients compared to young HCC patients. We hope to establish a more comprehensive and practical nomogram model for the specific group of HCC in young patients, including tumor characteristics, pathological characteristics, and other factors. The incidence of HCC in young people is relatively low, which makes it more important to establish a prediction model by using public databases, to enlarge the advantage of sample size. In addition, the application of the nomogram model must be established on the premise that the model has been verified to ensure the reliability and accuracy of the prediction effect. Moreover, in the present study, the PEM developed were both internally and externally validated; the *c*-indices of the training and external validation groups were 0.76 and 0.92, respectively, and the calibration curves for the probability of survival showed good agreement.

This study had several limitations, including the lack of reports on hepatic virus infections, such as HBV/HCV, surgical procedures, and liver function, which was due to that fact that they were not completely recorded in the SEER database. Therefore, the optimal management of HCC in young patients could not be studied. In addition, the retrospective study design has several inherent and inevitable biases, which might be resolved by confirming the study results with large-scale prospective multicenter studies with a more comprehensive analysis of the risk factors. Regardless of these limitations, our study remains the largest population-based HCC study to date and provides a specific predictive model for HCC prognosis in young patients. Further validation with a larger HCC cohort is needed in the future.

Conclusions

Grade, tumor size, AFP, T stage, and M stage were independent prognostic factors associated with CSS in young HCC patients. The nomogram could be used in clinical practice to accurately predict the prognosis of young HCC patients.

Acknowledgments

Funding: This work was supported by the National Capital Health Research and Development of Special (No. 2018-1-4021) and the National Science and Technology Major Project (No. 2018ZX10723204).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <http://dx.doi.org/10.21037/tcr-20-3411>

Data Sharing Statement: Available at <http://dx.doi.org/10.21037/tcr-20-3411>.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr-20-3411>). The authors have no 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). The study was approved by the Institutional Review Board of the Cancer Hospital, Chinese Academy of Medical Sciences (ID: NCC2019C-016) and informed consent was taken from all the patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. International Agency for Research on Cancer WHO. Cancer Today. Available online: <https://gco.iarc.fr/today/home>
2. Klein WM, Molmenti EP, Colombani PM, et al. Primary liver carcinoma arising in people younger than 30 years. *Am J Clin Pathol* 2005;124:512-8.
3. Bosch FX, Ribes J, Diaz M, et al. Primary liver cancer: worldwide incidence and trends. *Gastroenterology* 2004;127:S5-16.
4. Namieno T, Kawata A, Sato N, et al. Age-related, different clinicopathologic features of hepatocellular carcinoma patients. *Ann Surg* 1995;221:308-14.
5. Lee CR, Lim JH, Kim SH, et al. A comparative analysis of hepatocellular carcinoma after hepatic resection in young versus elderly patients. *J Gastrointest Surg* 2012;16:1736-43.
6. Lee JS, Kim JM, Lee S, et al. The prognosis in cases of hepatocellular carcinoma after hepatectomy: young patients versus older patients. *Korean J Hepatobiliary Pancreat Surg* 2015;19:154-60.
7. Falkson G, Cnaan A, Schutt AJ, et al. Prognostic factors for survival in hepatocellular carcinoma. *Cancer Res* 1988;48:7314-8.
8. Ni YH, Chang MH, Hsu HY, et al. Hepatocellular carcinoma in childhood. Clinical manifestations and prognosis. *Cancer* 1991;68:1737-41.
9. Cho SJ, Yoon JH, Hwang SS, et al. Do young hepatocellular carcinoma patients with relatively good liver function have poorer outcomes than elderly patients? *J Gastroenterol Hepatol* 2007;22:1226-31.
10. Chen CH, Chang TT, Cheng KS, et al. Do young hepatocellular carcinoma patients have worse prognosis? The paradox of age as a prognostic factor in the survival of hepatocellular carcinoma patients. *Liver Int* 2006;26:766-73.
11. SEER. Software SID-SD. 2019. Available online: <https://seer.cancer.gov/data/index.html>
12. Kattan MW, Scardino PT. Evidence for the usefulness of nomograms. *Nat Clin Pract Urol* 2007;4:638-9.
13. 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* 1996;15:361-87.
14. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373-9.

15. Gray R. A Class of $\log\text{-rank}$ Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. *Ann Statist* 1988;16:1141-54.
16. Tibshirani R. Regression shrinkage selection via the LASSO. *J R Stat Soc Series B Stat Methodol* 2011;73:273-82.
17. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 2004;10:7252-9.
18. Zhang S, Zheng R, Zeng H, et al. The incidence differences among sex, geographical areas and mean age of diagnosis for liver cancer in China, 1989-2008. *Zhonghua Yu Fang Yi Xue Za Zhi* 2014;48:355-60.
19. Yamazaki Y, Kakizaki S, Sohara N, et al. Hepatocellular carcinoma in young adults: the clinical characteristics, prognosis, and findings of a patient survival analysis. *Dig Dis Sci* 2007;52:1103-7.
20. Kim JH, Choi MS, Lee H, et al. Clinical features and prognosis of hepatocellular carcinoma in young patients from a hepatitis B-endemic area. *J Gastroenterol Hepatol* 2006;21:588-94.
21. Katz A, Smith BL, Golshan M, et al. Nomogram for the prediction of having four or more involved nodes for sentinel lymph node-positive breast cancer. *J Clin Oncol* 2008;26:2093-8.
22. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol* 2015;33:861-9.
23. Li J, Liu Y, Yan Z, et al. A nomogram predicting pulmonary metastasis of hepatocellular carcinoma following partial hepatectomy. *Br J Cancer* 2014;110:1110-7.
24. Shim JH, Jun MJ, Han S, et al. Prognostic nomograms for prediction of recurrence and survival after curative liver resection for hepatocellular carcinoma. *Ann Surg* 2015;261:939-46.
25. Shen J, He L, Li C, et al. Prognostic nomograms for patients with resectable hepatocellular carcinoma incorporating systemic inflammation and tumor characteristics. *Oncotarget* 2016;7:80783-93.
26. Feng LH, Dong H, Lau WY, et al. Novel microvascular invasion-based prognostic nomograms to predict survival outcomes in patients after R0 resection for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2017;143:293-303.
27. Torzilli G, Donadon M, Belghiti J, et al. Predicting Individual Survival After Hepatectomy for Hepatocellular Carcinoma: a Novel Nomogram from the "HCC East & West Study Group". *J Gastrointest Surg* 2016;20:1154-62.

Cite this article as: Li X, Bi X, Zhao J, Li Z, Zhou J, Huang Z, Zhang Y, Zhao H, Cai J. A nomogram to predict prognosis after surgery for young patients with hepatocellular carcinoma. *Transl Cancer Res* 2021;10(4):1773-1786. doi: 10.21037/tcr-20-3411