A practical nomogram predicting the prognosis of hepatocellular carcinoma patients with lymph node metastasis: A population-based analysis

CURRENT STATUS: Posted

Kai Zhang, Changcheng Tao, Jianxiong Wu, Weiqi Rong

Kai Zhang
Chinese Academy of Medical Sciences & Peking Union Medical College Hospital

Changcheng Tao
Chinese Academy of Medical Sciences & Peking Union Medical College Hospital

Jianxiong Wu
Chinese Academy of Medical Sciences and Peking Union Medical College
✉ wujianxiong@cicams.pumc.edu.cn Corresponding Author

Weiqi Rong
Chinese Academy of Medical Sciences & Peking Union Medical College
✉ rongweiqi@cicams.pumc.edu.cn Corresponding Author

Prescreen

10.21203/rs.3.rs-27749/v1

Subject Areas

Oncology Surgery
Keywords

*Hepatocellular carcinoma, prognosis, lymph nodes metastasis, prediction model, nomogram*
Abstract

Background: Lymph node (LN) metastasis is associated with poor survival outcomes in patients with hepatocellular carcinoma (HCC) patients and because of the reported low probability of lymph node metastasis, research into the Anchorprognoses of such patients is difficult to conduct. In this study, we aimed to develop a nomogram model to predict the prognosis of HCC patients with lymph node metastasis.

Methods: HCC patients diagnosed with LN metastasis from 2010 to 2015 were enrolled from the Surveillance, Epidemiology, and End Results (SEER) database. Univariate Cox regression and lasso regression were used to screen prognostic factors. Cox multiple-factor analysis was used to investigate the independent risk factors for survival. We developed a prognostic nomograms using these independent risk factors. The predictive performance of our nomogram model was evaluated according to the concordance index (C-index) and calibration curve. The net clinical benefit was assessed via decision curve analysis (DCA). Patients were divided into different risk groups according to the model. A survival curve was drawn using the Kaplan-Meier method and the difference was compared by the log-rank test.

Results: There were 944 patients in the training cohort and 402 patients in the validation cohort. Grade, T stage, surgery to the liver, chemotherapy, radiation recode, AFP, fibrosis score, tumor size group, M stage were selected as independent prognostic factors, and we developed nomograms using these variables. The calibration curves for probability of survival showed good agreement between the prediction by our model and actual observation in both the training and validation groups. DCA indicated that the nomogram had positive net benefits.

Conclusions: The nomogram can accurately predict the prognosis of HCC patients with lymph nodes metastasis and provide a reasonable basis for treatment.

Keywords: Hepatocellular carcinoma, prognosis, lymph nodes metastasis, prediction model, nomogram

Introduction

Hepatocellular carcinoma (HCC) is the seventh most prevalent tumor worldwide with 841,080 new cases occurring every year[1, 2]. The dominant pathogenic factors that contribute to the incidence rate of HCC vary according to countries and regions, including hepatitis B infection in China[3], hepatitis C infection in Africa and alcohol intake in Western countries[4]. Extrahepatic metastasis occurs in almost 30%-50% of patients during the course of the disease[5]. The lymph nodes are the second most common site of extrahepatic metastases in HCC[6]. The literature reports the incidence of LN metastasis range from 1.23% to 7.5% in some studies with large sample sizes[7-9]. Other research has shown that the incidence might reached approximately 30% of the average rate[10, 11]. Although a large proportion of the data are derived from autopsies, they might reflect that the occurrence of LN metastasis is underestimated, and more patients have lymph node metastasis.

According to the Barcelona staging, patients with lymph nodes metastasis are assigned to the C phase[12] and systemic therapy is seen as the primary treatment. The same situation has also been observed in other staging systems, such as the American Joint Committee on Cancer (AJCC) staging system and the NCCN guidelines[13]. The main reason for this phenomenon is that HCC patients with lymph nodes metastasis have a poor prognosis. A recent study showed that the median progression-free survival (PFS) time after surgery is 16.3 months for HCC patients without nodal involvement, but only 5.8 months for the group with lymph nodes metastasis[7]. It is undeniable that lymph node metastasis is a poor prognostic factor for hepatocellular carcinoma[9]. However, with the development of various treatments and drugs in recent years, the prognosis of HCC patients with lymph node metasteses has been improving[14-16]. Previous studies have shown that patients who were diagnosed with stage IV demonstrated a different prognosis, indicating that hereditary factors could contribute to the prognosis
of such patients\textsuperscript{[16]}. The prognosis of HCC patients with lymph node metastases is different, although they were treated with similar external beam radiotherapy in a study\textsuperscript{[17]}. The selection of the appropriate treatment should be based on accurate identification of different prognosis groups. Therefore, it is important to distinguish patients’ different prognoses. Because of its reported low probability, the grouping of patients requires a large sample size. This situation increases the difficulty of such studies. As far as we know, a study which constructing a prognostic model for the risk assessment of HCC patients with LN metastasis has not yet been reported. Therefore, the purpose of our study was to distinguish the different prognostic groups of HCC patients with LN metastasis and to assist and guide clinicians in making treatment decisions.

The nomogram is an effective statistical tool that represents a graphical solution to a predictive model, and it can accurately predict the outcomes of individual patients\textsuperscript{[18]}. As previously mentioned, this type of study is difficult because of its reported low probability. To expand the sample size and fully identify the factors affecting the prognosis of HCC patients with LN metastasis, we analyzed medical records from the Surveillance, Epidemiology and End Result (SEER) database. The SEER database collects several types of cancer patient data from electronic pathology reports and is an authoritative source of information on cancer, covering approximately 34.6% of the U.S. population (http://seer.cancer.gov/). In the present study, we downloaded data on HCC patients with LNs metastasis from the SEER registry between 2010 and 2015. Then, we divided these patients into a training group and validation group. A nomogram was constructed using the training group, and the validation group was used to evaluate its ability to predict patient survival.

Material And Methods

Patient selection

The data were acquired from the SEER database of the US National Cancer Institute (SEER 18 Regs Custom Data (with additional treatment fields), Nov 2018 sub (1973-2016 varying)). The data were obtained via the SEER*Stat software (version 8.3.6; http://seer.cancer.gov/seerstat/). Because some important prognostic factors were not available before 2010, patients diagnosed with HCC with LN metastases between 2010 and 2015 were finally included in our research.

The inclusion criteria were as follows: (1) hepatocellular carcinoma patients from 2010 to 2015, for whom the site code ICD-O-3 was liver; (2) according to the 7th edition of AJCC TNM staging, lymph-node metastasis patients were enrolled; (3) the patients was older than 18 years old; and (4) follow-up data were available. The exclusion criteria were as follows: (1) samples without follow-up time information data; (2) demographic and treatment Information was not complete; (3) data on vital prognostic factors and tumor staging information were missing; (4) and other cause of death classification was not the first tumor. The patients were randomized into the training group (accounting for 70%) and the validation group (accounting for 30%) using R software. The caret package was used, and the seed was 1988. To acquire more information, the selection process is shown in Figure 1.

Statistical analysis

The data regarding patients from the SEER database comprised sex, age group, race, grade, T stage, diagnostic confirmation information, surgery to the liver, surgery to LN, bone metastasis, brain metastasis, liver metastasis, pulmonary metastasis, chemotherapy, radiation, insurance, marital status, AFP, fibrosis score, tumor size group and M stage. The patients were censored as alive or dead of other causes. Categorical variables were compared using the chi-square test and Fisher’s exact test. Numerical variables were compared using the Mann-Whitney U-test. Kaplan-Meier survival curves were generated, and survival distributions were estimated by the log-rank test. Univariate and multivariate Cox proportional hazard models were used to identify factors associated with survival in the training group. According to the results of multivariate analysis, we constructed a nomogram using R software (version 3.4.3, https://www.r-project.org/), which was internally validated by bootstrapping in 1000 bootstrap samples. The c-index was used to compare the discriminative ability of the nomogram and AJCC 7th system (IVA/IVB) in the training and validation groups. Calibration curves were created to assess the
predictive accuracy in the two groups \[19\]. Then, we calculated the risk score of each training group patient according to the Cox regression model. The X-title program (versions 3.6.1) was used to select the optimal cut-off for the risk score to distinguish the differences in patient survival \[20, 21\]. Then, the training and validation groups of patients were divided into high and low risk groups according to risk score. Survival curves were plotted according to the Kaplan-Meier method, and the log-rank test was employed for survival analysis. Statistical analyses were performed with SPSS software, version 25 (IBM) and R software (version 3.3.4). All of the tests were two-sided, and \(p<0.05\) was considered statistically significant.

**Results**

**Training and validation group patients’ characteristics**

As shown in Figure 1, a total of 40,173 patients diagnosed with HCC from 2010 to 2015 were included in our research, of which 2,662 cases (6.6%) had LN metastasis. According to the above exclusion criteria, 1346 patients were finally enrolled. We allocated 944 patients into the training group and the others into the validation group. The clinical details of the patients are shown in Table 1.

Of all the patients, 81.7% were male, and 69.5% were white in race. Patients that older than 60 accounted for 54.3%. The AFP level was elevated in 74.82% of patients, and 14.0% of patients were grade III/IV for the Edmondson-Steiner classification. More than half (86.3%) of the patients did not receive radiation treatment. The same trend was observed for surgery to the liver/LN. The median (Q1-Q3) follow-up time of the patients was 5.00 (2.00-12.00) months. Finally, we found that there was no significant difference in on most demographic or baseline features between the training group and validation group. The incidence of pulmonary metastasis was different between the two groups. It was higher in the training group than in the validation group, but there was no significant difference between the training group and the overall population.

**Prognostic factors for HCC patients with lymph node metastasis**

As shown in Table 2, in the univariate analysis, grade, T stage, surgery to the liver, surgery to LN, bone metastasis, brain metastases, pulmonary metastasis, chemotherapy, radiation recode, Insurance recode, AFP, tumor size group, fibrosis score, M stage were associated with overall survival (OS).

To reduce the risk of over-fitting our model, we applied the Lasso regression method, which can compress partial factorial regression coefficients to zero \[22\]. The glmnet package was used in the R software. 10x cross validation was applied to search for the least partial likelihood deviance which can represent the complexity of the model. Finally, the variables that we chose when the partial likelihood deviance is least (lambda=-4.37) were age group, grade, T stage, surgery to the liver, surgery to LNs, bone metastasis, brain metastasis, pulmonary metastasis, Intrahepatic metastasis, chemotherapy, radiation recode, Insurance recode, AFP, tumor size group, fibrosis score, M stage. Combined with the results of Cox univariate analysis, we removed the variables of age group and intrahepatic metastasis and 14 variables were included in multivariate analysis. For more details see Figure2a-b.

In the multivariate analysis, grade, T Stage, surgery to the liver, chemotherapy, radiation recode, AFP, fibrosis score, tumor size group, M stage remained independently related to OS. The details are summarized in Table 3.

**Construction and validation of the nomogram**

A nomogram was formulated based on the results of the multivariate Cox regression analysis, as shown in Figure 3. In the training group, the Harrell’s C-index for OS prediction was 0.70 (95% CI, 0.68 to 0.72), and the area under ROC curve (AUC) for 1 and 2 years was 0.76 and 0.80, respectively. In the validation group, the Harrell’s C index for OS prediction was 0.73 (95% CI, 0.70 to 0.76), and the area under ROC curve (AUC) for 1 and 2 year was 0.79 and 0.75. However, in the training and validation groups, the C-index of the AJCC staging system was only 0.58 (95% CI, 0.56 to 0.60) and 0.59 (95% CI, 0.56 to 0.62), respectively. The nomogram model showed better discrimination.
We further assessed the accuracy of our model predictions by calibration plot. The calibration plot for the probability of survival at 1 and 2 years showed good agreement between the prediction by nomogram and actual observation. See further details in Figure 4a-d.

Furthermore, the decision curve analysis (DCA) was plotted to observe the clinical benefits to the patient. The DCA indicated that our nomogram had a positive net benefit with a wide scale of threshold probabilities in the training group and validation group. See further details in Figure 5a-d.

The establishment of different risk groups according to the model

According to Cox regression model, the risk score of each training group patient in the model was calculated. The X-title program was used to select optimal cut-off for risk score to distinguish the difference in patient survival. The point of grouping was 1.12 when the training group was divided into a low-risk group (score ≤ 1.12) and a high-risk group (score > 1.12). The corresponding total points were 261. The survival curves of the training group and the validation group were drawn, and the log-rank test was performed, with the p-value < 0.001. See further details in Figure 6a-b.

Discussion

According to existing reports, the incidence of lymph node metastasis during the treatment of liver cancer is 1.6% to 5.9%, while it is 25.5% on autopsy, indicating that lymph node metastasis might be neglected[11]. In our research, 6.6% of all HCC patients had LN metastasis, which was consistent with the results reported in the previous study. Regarding the prognosis of such patients, the emergence of novel treatments, including radiation, ablation, interventional therapy, and sorafenib, has improved the prognosis[15, 23-25]. Our study also showed that the HCC patients with LN metastasis can benefit from radiation and chemotherapy. In terms of surgical treatment, our study showed that patients with LN metastasis had no benefit from lymphadenectomy, and previous studies have shown similar results[26-28]. However, our research analysis showed a benefit for surgery at the primary site: the liver. We consider three possible reasons for the finding. First of all, more than 38% of patients included in our study were diagnosed without histology or cytology. Whereas the diagnosis of primary liver cancer can be made by clinical diagnosis, imaging and laboratory examinations[29-31], we included patients whose diagnoses were established according to clinical diagnostic criteria. In this case, the diagnosis of lymph node metastasis in many patients was based on clinical and radiographic findings and many patients might not actually have had cancerous lymph node metastasis. HCC patients often have chronic inflammation of the liver, such as hepatitis B, hepatitis C and non-alcoholic fatty liver disease. Inflammation of the liver can also cause enlarged lymph nodes and a study showed that the proportion of enlarged lymph nodes in hepatitis B virus-infected patients reached 9.4%[32]. According to the above discussion, in the absence of a pathological diagnosis of the lymph node, HCC patients with LN metastasis can be classified with cancerous metastasis and benign perihepatic lymph node enlargement (PLNE). A study showed that PLNE was an independent positive prognostic factor that might improve the prognosis of HCC patients[33]. In this case, some patients might benefit from surgery on the primary site. Second, only a few patients had their livers operated on, which might have affected the results of our statistical analysis. Third, patients undergoing surgery might have better basic indicators, such as performance status and liver function, than non-surgical patients. These facts might affect the outcome. In summary, cytological or histological confirmation is recommended to determine whether lymph node metastasis is truly present, especially as to the patients with hepatitis, and we should choose the treatment more carefully for HCC patients without pathological diagnosis of LN metastasis.

The Edmondson-Steiner classification indicates a pathological grade. A higher grade incidates a lower degree of differentiation and a higher degree of malignancy. Some reports in the literature have shown a relationship between Edmondson-Steiner classification and HCC patients' prognosis, and a higher grade indicates that the patients' prognosis is likely to be worse[34, 35]. Zhang et al reviewed the degree of cirrhosis affecting the prognosis of patients and found that the histological severity of cirrhosis is a vital adverse factor that affects the long-term outcomes of HCC patients undergoing liver surgery[36]. Similar results have been found in many
reports\textsuperscript{[37-39]}. The effect of AFP levels on patient prognosis is controversial. Some studies have found that elevated AFP levels could worsen HCC patients' prognosis\textsuperscript{[40-42]}. Other scholars have found that AFP has no significant effect on patient survival\textsuperscript{[43]}, perhaps because the studies are from different populations. In our study, we found that an elevated AFP level is an adverse prognostic factor.

With respect to pathological stage and tumor characteristics, T Stage, M Stage and tumor size group were associated with the prognosis of HCC patients with LN metastasis. Wu et al found that tumor size could be used as an independent risk predictor associated with survival in HCC\textsuperscript{[44]}. In combination with T stage, we grouped the patients according to different tumor sizes and obtained similar conclusions. The effect of M stage on patient prognosis is not in doubt. The previous literature has shown that the prognosis of HCC patients with different metastatic sites is different\textsuperscript{[45, 46]}. Therefore, we included the following prognostic indicators: bone metastasis, brain metastasis, intrahepatic metastasis, pulmonary metastasis and M Stage. Finally, we found that only M Stage of the above factors was an independent risk factor for HCC patients with LN metastasis. This outcome suggests that differences in metastatic site might not be as important in such patients as in those with HCC without LN metastasis. T Stage was also an independent risk factor for worse prognosis. In general, with the increase in T stage, the prognosis of patients became worse. However, in our analysis, patients with stage T3 had a higher risk of death than patients with stage T4. The 7th edition AJCC stage was used by SEER in the original data. According to the 7th edition of AJCC staging\textsuperscript{[47]}, T stage includes not only information about tumor size but also about vascular invasion and number of hepatic tumors. T3a indicates multiple tumors, at least one of which is \(>5\) cm. T3b indicates that the tumor has invaded the main trunk of the portal vein or/and hepatic vein, which would lead to worse prognosis\textsuperscript{[48, 49]}, and it has been included in stage T4 by the 8th AJCC cancer staging manual. In the 7th edition AJCC staging manual, T4 is defined as the tumor invading adjacent organs except for the gallbladder or penetrating the serous membrane. Therefore, for patients with liver cancer with lymph node metastasis, the prognostic significance of the number of liver tumors and vascular invasion might be greater than that of the invasion of adjacent organs.

As previously mentioned, the prognosis of HCC patients with LN metastasis is improving, and some studies have shown that stage IV HCC patients demonstrated a different prognosis\textsuperscript{[14-16]}. Therefore, it is important to distinguish the difference in prognosis in such patients.

On the basis of identifying the risk factors, we built the nomogram model and verified it. The model could well distinguish the difference of prognoses of HCC patients with LN metastasis. It could provide a basis for the choice of treatment for such patients. Only under the circumstance of reasonable differentiation of patients with different prognoses can a reasonable treatment plan be put forward. The establishment of the model could help to better distinguish the different prognoses of HCC patients with LN metastasis and provide a basis for follow-up treatment. To the best of our knowledge, no similar studies have yet been performed. Comparing to other models which need to use some technologies, our model applies a number of clinically accessible indicators. We evaluated predictors that were clinically relevant, so that the model can be easily applied in clinical practice.

However, our research still has some shortcomings. First, bias is inevitable in this type of retrospective study. For example, we removed many patients with lymph node metastasis whose important clinical data were unknown. A number of important prognostic factors were also missing from the enrolled patients. Some significant prognostic values were not recorded in the SEER database, such as liver function tests, hepatitis B or hepatitis C infections, and details of chemotherapy, radiation therapy and surgery. Second, internal validation was used for the model which might affect the accuracy of the models in general HCC patients with LNs metastasis. Next, we should further validate this model with our own clinical data. Prospective, randomized, controlled studies must be further implemented.

\textbf{Conclusion}

In conclusion, we showed that grade, T stage, surgery to the liver, chemotherapy, radiation recode, AFP, fibrosis score, tumor size group and M stage are independent risk factors for survival of HCC patients with LN
metastases. We established a nomogram to distinguish between patients with a good prognosis and those with a poor prognosis. Internal validation demonstrated the accuracy and usefulness of the nomogram. A reasonable treatment could be devised according to different risk scores. Further studies are warranted.

**Abbreviations**

HCC, Hepatocellular carcinoma; LN, Lymph node; SEER, Surveillance, Epidemiology, and End Results; C-index, concordance index; DCA, decision curve analysis; AFP, alpha-fetoprotein; AJCC, American Joint Committee on Cancer; PFS, progression-free survival; OS, overall survival; DCA, decision curve analysis; PLNE, perihepatic lymph node enlargement.

**Declarations**

**Acknowledgments**

Not applicable

**Author Contributions**

All authors contributed significantly to this work. Kai Zhang, Changcheng Tao, Weiqi Rong performed the research study and collected the data; Kai Zhang and Jianxiong Wu analyzed the data; Kai Zhang, Jianxiong Wu and Weiqi Rong designed the research study; Kai Zhang wrote the paper; Weiqi Rong prepared all the Tables. All authors reviewed the manuscript. In addition, all authors have read and approved the manuscript.

**Funding**

This work was supported by grants from the National Key Research and Development Program of China (No. 2016YFD0400604), CAMS Innovation Fund for Medical Science (CIFMS) (CAMS-2016-I2M-3-025).

**Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Ethics approval and consent to participate**

This research was exempted by the Ethics Committee of Cancer Hospital, Chinese Academy of Medical Sciences (Beijing, China).

**Consent for publication**

Not applicable.

**Conflicts of Interest**

The authors have no conflicts of interest to declare.

**Author details**

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
[1] McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. Clin Liver Dis. 2015. 19(2): 223-38.

[2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018. 68(6): 394-424.

[3] Fan JH, Wang JB, Jiang Y, et al. Attributable causes of liver cancer mortality and incidence in china. Asian Pac J Cancer Prev. 2013. 14(12): 7251-6.

[4] Augustine MM, Fong Y. Epidemiology and risk factors of biliary tract and primary liver tumors. Surg Oncol Clin N Am. 2014. 23(2): 171-88.

[5] Natsuiaka M, Omura T, Akaike T, et al. Clinical features of hepatocellular carcinoma with extrahepatic metastases. J Gastroenterol Hepatol. 2005. 20(11): 1781-7.

[6] Uka K, Aikata H, Takaki S, et al. Clinical features and prognosis of patients with extrahepatic metastases from hepatocellular carcinoma. World J Gastroenterol. 2007. 13(3): 414-20.

[7] Lee CW, Chan KM, Lee CF, et al. Hepatic resection for hepatocellular carcinoma with lymph node metastasis: clinicopathological analysis and survival outcome. Asian J Surg. 2011. 34(2): 53-62.

[8] Changchien CS, Chen CL, Yen YH, et al. Analysis of 6381 hepatocellular carcinoma patients in southern Taiwan: prognostic features, treatment outcome, and survival. J Gastroenterol. 2008. 43(2): 159-70.

[9] Xiaohong S, Huikai L, Feng W, Ti Z, Yunlong C, Qiang L. Clinical significance of lymph node metastasis in patients undergoing partial hepatectomy for hepatocellular carcinoma. World J Surg. 2010. 34(5): 1028-33.

[10] Liver Cancer Study Group of Japan. Primary liver cancer in Japan. Clinicopathologic features and results of surgical treatment. Ann Surg. 1990. 211(3): 277-87.

[11] Watanabe J, Nakashima O, Kojiro M. Clinicopathologic study on lymph node metastasis of hepatocellular carcinoma: a retrospective study of 660 consecutive autopsy cases. Jpn J Clin Oncol. 1994. 24(1): 37-41.

[12] Llovet JM, Brú C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis. 1999. 19(3): 329-38.

[13] Tellapuri S, Sutphin PD, Beg MS, Singal AG, Kalva SP. Staging systems of hepatocellular carcinoma: A review. Indian J Gastroenterol. 2018. 37(6): 481-491.

[14] Zeng ZC, Tang ZY, Fan J, et al. Consideration of role of radiotherapy for lymph node metastases in patients with HCC: retrospective analysis for prognostic factors from 125 patients. Int J Radiat Oncol Biol Phys. 2005. 63(4): 1067-76.

[15] Yuan Z, Xing A, Zheng J, Li W. Safety and technical feasibility of percutaneous ablation for lymph node metastases of hepatocellular carcinoma. Int J Hyperthermia. 2019. 36(1): 160-168.

[16] Berger Y, Spivack JH, Heskel M, Aycart SN, Labow DM, Sarpel U. Extrahepatic metastasectomy for hepatocellular carcinoma_ Predictors of long-term survival. J Surg Oncol. 2016. 114(4): 469-74.

[17] Wee CW, Kim K, Chie EK, Yu SJ, Kim YJ, Yoon JH. Prognostic stratification and nomogram for survival prediction in hepatocellular carcinoma patients treated with radiotherapy for lymph node metastasis. Br J Radiol. 2016. 89(1065): 20160383.

[18] Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015. 16(4): e173-80.

[19] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework
for traditional and novel measures. Epidemiology. 2010. 21(1): 128-38.

[20] Yang L, Chen H, Zhao M, Peng P. Prognostic value of circulating vitamin D binding protein, total, free and bioavailable 25-hydroxy vitamin D in patients with colorectal cancer. Oncotarget. 2017. 8(25): 40214-40221.

[21] Mao A, Zhou X, Liu Y, Ding J, Miao A, Pan G. KLF8 is associated with poor prognosis and regulates glycolysis by targeting GLUT4 in gastric cancer. J Cell Mol Med. 2019. 23(8): 5087-5097.

[22] Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997. 16(4): 385-95.

[23] Zhang H, Chen Y, Hu Y, et al. Image-guided intensity-modulated radiotherapy improves short-term survival for abdominal lymph node metastases from hepatocellular carcinoma. Ann Palliat Med. 2019. 8(5): 717-727.

[24] Chow P, Gandhi M, Tan SB, et al. SIRveNIB: Selective Internal Radiation Therapy Versus Sorafenib in Asia-Pacific Patients With Hepatocellular Carcinoma. J Clin Oncol. 2018. 36(19): 1913-1921.

[25] Cheng AL, Guan Z, Chen Z, et al. Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma according to baseline status: subset analyses of the phase III Sorafenib Asia-Pacific trial. Eur J Cancer. 2012. 48(10): 1452-65.

[26] Yang A, Xiao W, Ju W, et al. Prevalence and clinical significance of regional lymphadenectomy in patients with hepatocellular carcinoma. ANZ J Surg. 2019. 89(4): 393-398.

[27] Wu X, Li B, Qiu J, et al. Hepatectomy Versus Hepatectomy With Lymphadenectomy in Hepatocellular Carcinoma: A Prospective, Randomized Controlled Clinical Trial. J Clin Gastroenterol. 2015. 49(6): 520-8.

[28] Amini N, Ejaz A, Spolverato G, Maithel SK, Kim Y, Pawlik TM. Management of lymph nodes during resection of hepatocellular carcinoma and intrahepatic cholangiocarcinoma: a systematic review. J Gastrointest Surg. 2014. 18(12): 2136-48.

[29] Vogel A, Cervantes A, Chau I, et al. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019. 30(5): 871-873.

[30] Chou R, Cuevas C, Fu R, et al. Imaging Techniques for the Diagnosis of Hepatocellular Carcinoma: A Systematic Review and Meta-analysis. Ann Intern Med. 2015. 162(10): 697-711.

[31] Song P, Tang Q, Feng X, Tang W. Biomarkers: evaluation of clinical utility in surveillance and early diagnosis for hepatocellular carcinoma. Scand J Clin Lab Invest Suppl. 2016. 245: S70-6.

[32] Sato M, Hikita H, Hagiwara S, et al. Potential associations between perihepatic lymph node enlargement and liver fibrosis, hepatocellular injury or hepatocarcinogenesis in chronic hepatitis B virus infection. Hepatol Res. 2015. 45(4): 397-404.

[33] Tian F, Wu JX, Yu WB. Prognostic Factors and Clinical Characteristics for Hepatocellular Carcinoma Patients with Benign Enlarged Perihepatic Lymph Nodes: a Single-Center Experience from China. J Gastrointest Surg. 2015. 19(10): 1822-32.

[34] Zhou L, Rui JA, Zhou WX, Wang SB, Chen SG, Qu Q. Edmondson-Steiner grade: A crucial predictor of recurrence and survival in hepatocellular carcinoma without microvascular invasio. Pathol Res Pract. 2017. 213(7): 824-830.

[35] Zhou L, Rui JA, Wang SB, Chen SG, Qu Q. Prognostic factors of solitary large hepatocellular carcinoma: the importance of differentiation grade. Eur J Surg Oncol. 2011. 37(6): 521-5.

[36] Zhang EL, Liang BY, Chen XP, Huang ZY. Severity of liver cirrhosis: a key role in the selection of surgical modality for Child-Pugh A hepatocellular carcinoma. World J Surg Oncol. 2015. 13: 148.
[37] Sasaki Y, Imaoka S, Masutani S, et al. Influence of coexisting cirrhosis on long-term prognosis after surgery in patients with hepatocellular carcinoma. Surgery. 1992. 112(3): 515-21.

[38] Liang BY, Huang ZY, Liu YA, Hou P, Ji GB, Chen XP. [Influence of cirrhosis on long-term outcomes after liver resection in patients with a single small hepatocellular carcinoma]. Zhonghua Wai Ke Za Zhi. 2012. 50(10): 865-9.

[39] Choi WM, Lee JH, Ahn H, et al. Forns index predicts recurrence and death in patients with hepatitis B-related hepatocellular carcinoma after curative resection. Liver Int. 2015. 35(8): 1992-2000.

[40] Becker-Assmann J, Fard-Aghaie MH, Kantas A, et al. [Diagnostic and prognostic significance of α-fetoprotein in hepatocellular carcinoma]. Chirurg. 2020.

[41] She WH, Chan A, Cheung TT, Lo CM, Chok K. Survival outcomes of liver transplantation for hepatocellular carcinoma in patients with normal, high and very high preoperative alpha-fetoprotein levels. World J Hepatol. 2018. 10(2): 308-318.

[42] Zhang XF, Yin ZF, Wang K, Zhang ZQ, Qian HH, Shi LH. Changes of serum alpha-fetoprotein and alpha-fetoprotein-L3 after hepatectomy for hepatocellular carcinoma: prognostic significance. Hepatobiliary Pancreat Dis Int. 2012. 11(6): 618-23.

[43] Giannini EG, Marenco S, Borgonovo G, et al. Alpha-fetoprotein has no prognostic role in small hepatocellular carcinoma identified during surveillance in compensated cirrhosis. Hepatology. 2012. 56(4): 1371-9.

[44] Wu G, Wu J, Wang B, Zhu X, Shi X, Ding Y. Importance of tumor size at diagnosis as a prognostic factor for hepatocellular carcinoma survival: a population-based study. Cancer Manag Res. 2018. 10: 4401-4410.

[45] Oweira H, Petrausch U, Helbling D, et al. Prognostic value of site-specific extra-hepatic disease in hepatocellular carcinoma: a SEER database analysis. Expert Rev Gastroenterol Hepatol. 2017. 11(7): 695-701.

[46] Chen J. The prognostic analysis of different metastatic patterns in advanced liver cancer patients: A population based analysis. PLoS One. 2018. 13(8): e0200909.

[47] Edge SB. American Joint Committee on Cancer. AJCC cancer staging manual. 7th ed. New York: Springer. 2010.

[48] Vauthey JN, Lauwers GY, Esnaola NF, et al. Simplified staging for hepatocellular carcinoma. J Clin Oncol. 2002. 20(6): 1527-36.

[49] Kamarajah SK, Frankel TL, Sonnenday C, Cho CS, Nathan H. Critical evaluation of the American Joint Commission on Cancer (AJCC) 8th edition staging system for patients with Hepatocellular Carcinoma (HCC): A Surveillance, Epidemiology, End Results (SEER) analysis. J Surg Oncol. 2018. 117(4): 644-650.

### Tables

Table 1. Comparison of the demographics of the training and validation groups

| Variable       | All patients | Training group number (%) | Validation group number (%) | p     |
|----------------|--------------|----------------------------|----------------------------|-------|
| Number of patients | 1346 | 944 | 402 |       |
| Sex            |              |                            |                            |       |
| Male           | 1099 (81.65%)| 773 (81.89%)               | 326 (81.09%)               |       |
| Female         | 247 (18.35%) | 171 (18.11%)               | 76 (18.91%)                |       |
| Age group      |              |                            |                            |       |
| ≥61            | 731 (54.31%) | 513 (54.34%)               | 218 (54.23%)               |       |
| 18-60          | 615 (45.69%) | 431 (45.66%)               | 184 (45.77%)               |       |
| Race           |              |                            |                            |       |

11
|                          | White | Black | Other |
|--------------------------|-------|-------|-------|
|                          | 935 (69.47%) | 653 (69.17%) | 282 (70.15%) |
|                          | 220 (16.34%) | 163 (17.27%) | 57 (14.18%) |
|                          | 191 (14.19%) | 128 (13.56%) | 63 (15.67%) |
| **Grade**                |       |       |       |
| Grades I+II              | 269 (19.99%) | 189 (20.02%) | 80 (19.90%) |
| Grades III+IV            | 188 (13.97%) | 130 (13.77%) | 58 (14.43%) |
| Unknown                  | 889 (66.05%) | 625 (66.21%) | 264 (65.67%) |
| **T stage**              |       |       |       |
| T1                       | 245 (18.20%) | 177 (18.75%) | 68 (16.92%) |
| T2                       | 233 (17.31%) | 160 (16.95%) | 73 (18.16%) |
| T3                       | 748 (55.57%) | 518 (54.87%) | 230 (57.21%) |
| T4                       | 120 (8.92%) | 89 (9.43%) | 31 (7.71%) |
| **Diagnostic confirmation** |   |   |   |
| Histology/cytology\(^b\) | 829 (61.59%) | 594 (62.92%) | 235 (58.46%) |
| Clinical/radiography/text\(^c\) | 517 (38.41%) | 350 (37.08%) | 167 (41.54%) |
| **Surgery to the liver** |       |       |       |
| No surgery               | 1254 (93.16%) | 880 (93.22%) | 374 (93.03%) |
| Surgery                  | 92 (6.84%) | 64 (6.78%) | 28 (6.97%) |
| **Surgery to LNs**       |       |       |       |
| None                     | 1283 (95.32%) | 898 (95.13%) | 385 (95.77%) |
| Yes                      | 63 (4.68%) | 46 (4.87%) | 17 (4.23%) |
| **Bone metastasis**      |       |       |       |
| No                       | 1221 (90.71%) | 856 (90.68%) | 365 (90.80%) |
| Yes                      | 125 (9.29%) | 89 (9.32%) | 37 (9.20%) |
| **Brain metastasis**     |       |       |       |
| No                       | 1339 (99.48%) | 937 (99.26%) | 402 (100.00%) |
| Yes                      | 7 (0.52%) | 7 (0.74%) | 0 (0.00%) |
| **Intrahepatic metastasis** | | | |
| No                       | 1302 (96.73%) | 917 (97.14%) | 385 (95.77%) |
| Yes                      | 44 (3.27%) | 27 (2.86%) | 17 (4.23%) |
| **Pulmonary metastasis** |       |       |       |
| No                       | 1182 (87.82%) | 818 (86.65%) | 364 (90.55%) |
| Yes                      | 164 (12.18%) | 126 (13.35%) | 38 (9.45%) |
| **Chemotherapy**         |       |       |       |
| No                       | 597 (44.35%) | 414 (43.86%) | 183 (45.52%) |
| Yes                      | 749 (55.65%) | 530 (56.14%) | 219 (54.48%) |
| **Radiation recode**     |       |       |       |
| No                       | 1161 (86.26%) | 823 (87.18%) | 338 (84.08%) |
| Yes                      | 185 (13.74%) | 121 (12.82%) | 64 (15.92%) |
| **Insurance recode**     |       |       |       |
| Insured and any medicaid | 1257 (93.39%) | 887 (93.96%) | 370 (92.04%) |
| Uninsured                | 89 (6.61%) | 57 (6.04%) | 32 (7.96%) |
| **Marital status**       |       |       |       |
| Unmarried\(^d\)          | 662 (49.18%) | 478 (50.64%) | 184 (45.77%) |
| Married                  | 684 (50.82%) | 466 (49.36%) | 218 (54.23%) |
| **AFP**                  |       |       |       |
| Positive/elevated        | 999 (74.22%) | 707 (74.89%) | 292 (72.64%) |
| Negative/normal          | 198 (14.71%) | 139 (14.72%) | 59 (14.68%) |
| Unknown                  | 149 (11.07%) | 98 (10.38%) | 51 (12.69%) |
| **Fibrosis score**       |       |       |       |
| 0-4                      | 1008 (74.89%) | 713 (75.53%) | 295 (73.38%) |
| 5-6                      | 75 (5.57%) | 52 (5.51%) | 23 (5.72%) |
| Unknown\(^a\)            | 263 (19.54%) | 179 (18.96%) | 84 (20.90%) |
| **Tumor size group(mm)** |       |       |       |
| 0-20                     | 60 (4.46%) | 41 (4.34%) | 19 (4.73%) |
| 21-50                    | 334 (24.81%) | 238 (25.21%) | 96 (23.88%) |
| 51-100                   | 601 (44.65%) | 431 (45.66%) | 170 (42.29%) |
| ≥101                     | 351 (26.08%) | 234 (24.79%) | 117 (29.10%) |
| **M stage**              |       |       |       |
| M0                       | 784 (58.25%) | 539 (57.10%) | 245 (60.95%) |
| M1                       | 562 (41.75%) | 405 (42.90%) | 157 (39.05%) |
| **Survival months**      | 5.00 (2.00-12.00) | 5.00 (2.00-12.00) | 5.00 (2.00-12.00) |

\(^a\): includes American Indian/AK Native, Asian/Pacific Islander  
\(^b\): positive histology/positive exfoliative cytology
### Variable - Univariate analysis of overall survival for the primary group

| Variables                      | HR  | Univariate analysis | P-value |
|--------------------------------|-----|---------------------|---------|
| **Sex**                        |     |                     |         |
| Male                           | 1   | reference          |         |
| Female                         | 0.92| 1.1-0.76            | 0.3518  |
| **Age group**                  |     |                     |         |
| 18-60                          | 1   | reference          |         |
| ≥61                            | 0.91| 1.04-0.79           | 0.1714  |
| **Race**                       |     |                     |         |
| White                          | 1   | reference          |         |
| Black                          | 1.01| 1.22-0.84           | 0.8917  |
| Other                          | 0.95| 1.18-0.77           | 0.66    |
| **Grade**                      |     |                     |         |
| Grades I+II                    | 1   | reference          |         |
| Grades III+IV                  | 1.66| 2.13-1.3            | <0.0001 |
| Unknown                        | 1.22| 1.46-1.01           | 0.0366  |
| **T Stage**                    |     |                     |         |
| T1                             | 1   | reference          |         |
| T2                             | 0.98| 1.24-0.76           | 0.8385  |
| T3                             | 1.46| 1.77-1.21           | 0.0001  |
| T4                             | 1.73| 2.28-1.32           | <0.0001 |
| **Diagnostic confirmation**    |     |                     |         |
| Positive histology             | 1   | reference          |         |
| Clinical diagnosis only        | 0.92| 1.06-0.79           | 0.2353  |
| **Surgery to the liver**       |     |                     |         |
| No surgery                     | 1   | reference          |         |
| Surgery                        | 0.25| 0.36-0.17           | <0.0001 |
| **Surgery to LN**              |     |                     |         |
| None                           | 1   | reference          |         |
| LN removed                     | 0.45| 0.65-0.31           | <0.0001 |
| **Bone metastasis**            |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 1.5 | 1.9-1.19            | 0.0006  |
| **Brain metastasis**           |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 3.4 | 7.18-1.61           | 0.0013  |
| **Intrahepatic metastasis**    |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 0.87| 1.33-0.57           | 0.5116  |
| **Pulmonary metastasis**       |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 1.82| 2.22-1.49           | <0.0001 |
| **Chemotherapy**               |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 0.64| 0.74-0.56           | <0.0001 |
| **Radiation record**           |     |                     |         |
| No                             | 1   | reference          |         |
| Yes                            | 0.79| 0.97-0.64           | 0.0271  |
| **Insurance Recode**           |     |                     |         |
| Insured                        | 1   | reference          |         |
| Uninsured                      | 1.43| 1.91-1.07           | 0.0146  |
| **Marital status**             |     |                     |         |
| Married                        | 1   | reference          |         |
| Single (never married)         | 0.9 | 1.03-0.78           | 0.1393  |
| **AFP**                        |     |                     |         |
| Positive/elevated              | 1   | reference          |         |
| Negative/normal                | 0.62| 0.76-0.5            | <0.0001 |
| Unknown                        | 0.93| 1.17-0.73           | 0.5328  |
| **Fibrosis score**             |     |                     |         |
| Tumor size group (mm) | 1 reference | 0.94-0.49 | 0.0185 |
|----------------------|-------------|-----------|---------|
| 0-4                  | 0.67        | 0.49      |         |
| 5-6                  | 0.9         | 1.08-0.75 | 0.2515  |

| M Stage   | 1 reference | 1.59-0.75 | 0.6545 |
|-----------|-------------|-----------|---------|
| M0        | 1.09        | 0.75      |         |
| M1        | 1.49        | 1.59-0.75 | 0.03    |

| M Stage   | 2.15-1.04   | 2.53-1.2  | 0.0036  |
|-----------|-------------|-----------|---------|
| ≥101      | 1.74        | 2.15-1.04 |         |

Table 3. Multivariate analysis of overall survival for the primary group.
| Variables                                | HR     | multivariate analysis 95% CI | p-value |
|------------------------------------------|--------|-----------------------------|---------|
| **Grade**                                |        |                             |         |
| Grades I+II                              | 1 reference |                             |         |
| Grades III+IV                            | 1.56   | 1.21- 2.00                  | **0.0006** |
| Unknown                                  | 1.15   | 0.95- 1.38                  | 0.1513  |
| **T Stage**                              |        |                             |         |
| T1                                       | 1 reference |                             |         |
| T2                                       | 1.1    | 0.85- 1.43                  | 0.4718  |
| T3                                       | 1.25   | 1.02- 1.54                  | **0.0328** |
| T4                                       | 1.2    | 0.89- 1.60                  | 0.2309  |
| **Surgery to the liver**                 |        |                             |         |
| No surgery                               | 1 reference |                             |         |
| Surgery                                 | 0.3    | 0.20- 0.44                  | **<0.0001** |
| **Surgery to LN**                        |        |                             |         |
| None                                     | 1 reference |                             |         |
| LN removed                               | 0.76   | 0.51- 1.14                  | 0.1815  |
| **Bone metastasis**                      |        |                             |         |
| No                                       | 1 reference |                             |         |
| Yes                                      | 1.26   | 0.97- 1.65                  | 0.0872  |
| **Brain metastases**                     |        |                             |         |
| No                                       | 1 reference |                             |         |
| Yes                                      | 1.96   | 0.92- 4.19                  | 0.0823  |
| **Pulmonary metastasis**                 |        |                             |         |
| No                                       | 1 reference |                             |         |
| Yes                                      | 1.14   | 0.91- 1.43                  | 0.2593  |
| **Chemotherapy**                         |        |                             |         |
| No                                       | 1 reference |                             |         |
| Yes                                      | 0.57   | 0.50- 0.66                  | **<0.0001** |
| **Radiation recode**                     |        |                             |         |
| No                                       | 1 reference |                             |         |
| Yes                                      | 0.63   | 0.50- 0.79                  | **<0.0001** |
| **Insurance recode**                     |        |                             |         |
| Insured                                  | 1 reference |                             |         |
| Uninsured                                | 1.13   | 0.84- 1.51                  | 0.4294  |
| **AFP**                                  |        |                             |         |
| Positive/elevated                        | 1 reference |                             |         |
| Negative/normal                          | 0.66   | 0.53- 0.82                  | **0.0002** |
| Unknown                                  | 0.84   | 0.66- 1.07                  | 0.1618  |
| **Fibrosis score**                       |        |                             |         |
| Unknown                                  | 1 reference |                             |         |
| 0-4                                      | 0.63   | 0.45- 0.89                  | **0.0078** |
| 5-6                                      | 0.94   | 0.79- 1.14                  | 0.5484  |
| **Tumor size group(mm)**                |        |                             |         |
| 0-20                                     | 1 reference |                             |         |
| 21-50                                    | 1.21   | 0.82- 1.77                  | 0.3339  |
| 51-100                                   | 1.45   | 0.98- 2.13                  | 0.0616  |
| ≥101                                     | 1.81   | 1.21- 2.71                  | **0.0037** |
| **M stage**                              |        |                             |         |
| M0                                       | 1 reference |                             |         |
| M1                                       | 1.4    | 1.18- 1.67                  | **0.0001** |

Figures

[Site recode ICD-O-3/WHO 2008=liver from 2010 to 2015 (n=46345)](image)
Histologic Type ICD-O-3=8170-8175
(n=40241)

\[\text{Derived AJCC } N, 7th \text{ ed } N=N0,NX\]
(n=37511)

\[\text{Derived AJCC } N, 7th \text{ ed } N=N1\]
(n=2662)

\[\text{Race recode = unknown} (n=7)\]
\[\text{Marital status at diagnosis = unknown} (n=87)\]

n=1873

n=1779
Figure 1
The flow diagram of the patients' selection.
Lasso regression search for the optimal coefficient. (a) Lasso regression search for the optimal coefficient when the Lambda was -4.37. (b) 10x cross-validation was applied for search the Lambda when partial likelihood deviance was the least.
Figure 3
Nomogram predicting 1-and 2-year overall survival probability.
Figure 4
Calibration curve for 1 and 2 years in the training and validation group. (a) Calibration curve for 1 year in the training group. (b) Calibration curve for 2 years in the training group. (c) Calibration curve for 1 year in the validation group. (d) Calibration curve for 2 year in the validation group.
DCA of validation group in terms of 1 year
Clinical decision curve analysis (DCA) of the training and validation group in terms of 1 and 2 years. (a) Clinical decision curve analysis (DCA) of the training group in terms of 1 year. (b) Clinical decision curve analysis (DCA) of the training group in terms of 2 years. (c) Decision curve analysis (DCA) of the validation group in terms of 1 year. (d) Decision curve analysis (DCA) of the validation group in terms of 2 years.
Figure 6

Survival curves of the training and validation of different groups according to risk score. (a) Survival curves of the training set of different groups according to risk score. (b) Survival curves of validation set of different groups according to risk score.