

Shanghai Score: A Prognostic and Adjuvant Treatment-evaluating System Constructed for Chinese Patients with Hepatocellular Carcinoma after Curative Resection

Hui-Chuan Sun¹, Lu Xie², Xin-Rong Yang¹, Wei Li^{2,3,4}, Jian Yu², Xiao-Dong Zhu¹, Yong Xia⁵, Ti Zhang⁶, Yang Xu¹, Bo Hu¹, Li-Ping Du⁷, Ling-Yao Zeng², Jian Ouyang², Wei Zhang⁶, Tian-Qiang Song⁶, Qiang Li⁶, Ying-Hong Shi¹, Jian Zhou¹, Shuang-Jian Qiu¹, Qian Liu³, Yi-Xue Li^{2,4}, Zhao-You Tang¹, Yu Shyr⁷, Feng Shen⁵, Jia Fan¹

¹Department of Liver Surgery, Liver Cancer Institute, Zhongshan Hospital, Fudan University, Key Laboratory of Carcinogenesis and Cancer Invasion, Ministry of Education, Shanghai 200032, China

²Shanghai Center for Bioinformation Technology, Shanghai Academy of Science and Technology, Shanghai 201203, China

³Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

⁴Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

⁵Department of Hepatic Surgery, Eastern Hepatobiliary Surgery Hospital, The Second Military Medical University, Shanghai 200438, China

⁶Department of Hepatobiliary Surgery, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300040, China

⁷Center for Quantitative Sciences, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

Hui-Chuan Sun and Lu Xie contributed equally to this work.

Abstract

Background: For Chinese patients with hepatocellular carcinoma (HCC), surgical resection is the most important treatment to achieve long-term survival for patients with an early-stage tumor, and yet the prognosis after surgery is diverse. We aimed to construct a scoring system (Shanghai Score) for individualized prognosis estimation and adjuvant treatment evaluation.

Methods: A multivariate Cox proportional hazards model was constructed based on 4166 HCC patients undergoing resection during 2001–2008 at Zhongshan Hospital. Age, hepatitis B surface antigen, hepatitis B e antigen, partial thromboplastin time, total bilirubin, alkaline phosphatase, γ -glutamyltransferase, α -fetoprotein, tumor size, cirrhosis, vascular invasion, differentiation, encapsulation, and tumor number were finally retained by a backward step-down selection process with the Akaike information criterion. The Harrell's concordance index (C-index) was used to measure model performance. Shanghai Score is calculated by summing the products of the 14 variable values times each variable's corresponding regression coefficient. Totally 1978 patients from Zhongshan Hospital undergoing resection during 2009–2012, 808 patients from Eastern Hepatobiliary Surgery Hospital during 2008–2010, and 244 patients from Tianjin Medical University Cancer Hospital during 2010–2011 were enrolled as external validation cohorts. Shanghai Score was also implied in evaluating adjuvant treatment choices based on propensity score matching analysis.

Results: Shanghai Score showed good calibration and discrimination in postsurgical HCC patients. The bootstrap-corrected C-index (confidence interval [CI]) was 0.74 for overall survival (OS) and 0.68 for recurrence-free survival (RFS) in derivation cohort (4166 patients), and in the three independent validation cohorts, the CIs for OS ranged 0.70–0.72 and that for RFS ranged 0.63–0.68. Furthermore, Shanghai Score provided evaluation for adjuvant treatment choices (transcatheter arterial chemoembolization or interferon- α). The identified subset of patients at low risk could be ideal candidates for curative surgery, and subsets of patients at moderate or high risk could be

Address for correspondence: Dr. Jia Fan,
Liver Cancer Institute, Zhongshan Hospital, Fudan University,
Shanghai 200032, China
E-Mail: fan.jia@zs-hospital.sh.cn

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

© 2017 Chinese Medical Journal | Produced by Wolters Kluwer - Medknow

Access this article online

Quick Response Code:



Website:
www.cmj.org

DOI:
10.4103/0366-6999.218019

Received: 18-06-2017 **Edited by:** Yuan-Yuan Ji

How to cite this article: Sun HC, Xie L, Yang XR, Li W, Yu J, Zhu XD, Xia Y, Zhang T, Xu Y, Hu B, Du LP, Zeng LY, Ouyang J, Zhang W, Song TQ, Li Q, Shi YH, Zhou J, Qiu SJ, Liu Q, Li YX, Tang ZY, Shyr Y, Shen F, Fan J. Shanghai Score: A Prognostic and Adjuvant Treatment-evaluating System Constructed for Chinese Patients with Hepatocellular Carcinoma after Curative Resection. Chin Med J 2017;130:2650-60.

recommended with possible adjuvant therapies after surgery. Finally, a web server with individualized outcome prediction and treatment recommendation was constructed.

Conclusions: Based on the largest cohort up to date, we established Shanghai Score – an individualized outcome prediction system specifically designed for Chinese HCC patients after surgery. The Shanghai Score web server provides an easily accessible tool to stratify the prognosis of patients undergoing liver resection for HCC.

Key words: Adjuvant Treatment; Hepatocellular Carcinoma; Prognosis; Shanghai Score

INTRODUCTION

Hepatocellular carcinoma (HCC) is the third leading cause of cancer death worldwide.^[1] In China, HCC ranks the fourth most deadly cancer, but these cases make up nearly half of the HCC cases in the world.^[2] Currently, surgery remains one of the most effective treatments with curative potential, but long-term survival of HCC patients after surgery is mixed.^[3,4] Additionally, it is difficult to predict patient prognosis and outcome, and there is a lack of adjuvant treatment options for postsurgical HCC.

Since the first proposal of Okuda *et al.* in 1985,^[5] a dozen of staging systems have been developed and tested in different populations,^[4] including the Barcelona Clinic Liver Cancer (BCLC),^[6] Cancer of the Liver Italian Program score (CLIP),^[7] and the Japan Integrated Staging score (JIS).^[8] Recently, the Hong Kong Liver Cancer classification (HKLC) was also developed for guiding treatment.^[9] These systems are based on preoperative assessments, and pathological factors such as tumor differentiation, encapsulation, satellite lesions, presence of microvascular invasion, and liver cirrhosis were not considered. During the last 2 years, a few postoperative models were proposed to predict overall survival (OS),^[10] recurrence,^[11] or benefit from adjuvant treatment after surgery.^[12] These models were aimed at either early-stage^[11] or large^[13] HCCs, and the derivation cohorts for model training were relatively small, with most of them lacked multicenter validation.

Transcatheter arterial chemoembolization (TACE) can be used to treat high-risk HCC patients^[12,14,15] with large tumor size (>5 cm),^[16] tumor with microvascular invasion,^[17] or tumors with portal vein tumor thrombosis,^[18] but there are no standardized guidelines. Interferon- α (IFN) therapy is an antiviral treatment for patients with hepatitis B virus (HBV) or hepatitis C virus infection,^[19] but the effects of IFN on HCC are unclear. An objective evaluating system that incorporates multiple clinical indices with prognosis prediction should help identify which population of postsurgical HCC patients might benefit from certain adjuvant treatments.

The goal of this work was to establish a prognostic system for the large population of Chinese HCC patients undergoing curative liver resection to predict their OS and tumor recurrence, as well as evaluate the benefits from adjuvant therapies, including TACE and IFN. We termed this system the Shanghai Score and based it on a retrospective analysis of a large derivation cohort of 4166

postoperative HCC patients. Validation cohorts included three independent multicenter patient groups ($n = 1978$, $n = 808$, and $n = 244$). As a computational substitute to a nomogram, a web server for the Shanghai Score was developed and deployed at http://www.115.28.66.83/pro/theme/admin_4/shanghai.php.

METHODS

Ethical approval

Ethical approvals were obtained from the Zhongshan Hospital, Eastern Hepatobiliary Surgery Hospital, and Tianjin Medical University Cancer Hospital Research Ethics Committees (No. B2017-001), and informed consent was obtained from all patients.

Study population

The entrance criteria for all patients were as follows: (1) definitive pathological diagnosis of HCC based on the World Health Organization criteria; (2) no prior anticancer treatment; (3) curative resection, defined as complete resection of all tumor nodules, with the cut surface free of cancer by histological examination;^[20] and (4) greater than 1 month survival postsurgery and availability of complete clinicopathological and follow-up data. We retrospectively identified 4166 HCC patients who underwent curative resection from 2001 to 2008 at Zhongshan Hospital (Cohort 1). A same-center internal independent cohort of 1978 consecutive HCC patients undergoing resection from 2009 to 2012 was enrolled as Cohort 2. Two additional cohorts of eligible patients were enrolled as external validation cohorts: one from Eastern Hepatobiliary Surgery Hospital in Shanghai, who underwent resection from 2008 to 2010 (Cohort 3, $n = 808$), and another from Tianjin Medical University Cancer Hospital in Tianjin who underwent resection from 2010 to 2011 (Cohort 4, $n = 244$).

Postsurgical follow-up

Patients were followed up in the clinics every 2 months during the first postsurgical year, and at least every 3–4 months thereafter until March 30, 2011 (Cohort 1), May 30, 2015 (Cohort 2), May 30, 2012 (Cohort 3), and May 30, 2014 (Cohort 4). The end points for this study were OS and RFS.

Candidate variables

We selected candidate variables that are commonly assessed and clearly defined to enable model generalization and comparison among different institutions. Variables used to calculate the propensity match score included

age, hepatitis B surface antigen (HBsAg), hepatitis B e antigen (HBeAg), cirrhosis, α -fetoprotein, total bilirubin, partial thromboplastin time, alkaline phosphatase, γ -glutamyltransferase, differentiation, tumor size, vascular invasion, encapsulation, and tumor number.

Statistical analysis

Prognosis prediction modeling was performed as follows: Cox proportional hazards regression was implemented for Cohort 1 to predict OS. Missing data were imputed by approximate multiple imputation.^[21] To take both prediction accuracy and parsimony into consideration, we performed a backward step-down selection process for the final model, using the Akaike information criterion (AIC).^[22] For each patient, we derived an accurate risk score (termed the Shanghai Score) by summing the products of the values of selected variables and regression coefficients. Discrimination was evaluated using the concordance index (C-index)^[23] corrected 1000 times by bootstrapping.^[24] To compare the predictive performance with other staging systems, the C-index for each staging system was first evaluated, and then the homogeneity and monotonicity of gradient were evaluated by multivariate analysis, where the difference of likelihood ratio Chi-square and AIC was calculated. We externally validated the prediction power of the Shanghai Score using Cohort 2, Cohort 3, and Cohort 4.

To evaluate the adjuvant therapy benefit, HCC patients were classified into three groups (high, moderate, and low risk) based on the estimated Shanghai Score, and the effect of TACE or IFN in each group was further evaluated. A propensity score matching (PSM) analysis was performed to minimize selection bias and approximate a randomized trial, and 1:1 matching was accomplished using the nearest-neighbor matching method.^[3,25] All statistical analyses were performed using R 3.2.1 (<http://www.r-project.org/>) with package rms and customized R scripts.

Web server construction

Although nomograms are popular ways to predict patient outcome, we decided to construct our Cox regression model into a prognostic web server scoring system for HCC patients after resection. This approach provides the freedom for users to obtain individual patient risk ratios at different time points. The dynamic web interface was constructed with PHP: Hypertext Preprocessor technology (<http://www.php.net/>) with an Apache server (<http://www.apache.org/>). The analysis tool modules were developed using R, which allows collection of clinical features from uploaded patient data and easy return of results.

RESULTS

Patients' characteristics

The clinical characteristics in the four cohorts are listed in Table 1 and Supplementary Table 1. The mean follow-up time in Cohort 1 was 40.5 months (range, 1.0–120.7 months). At the end of follow-up, 1910 (45.8%)

patients had died and 2161 (51.9%) had recurrence. In Cohort 2, the mean follow-up time was 19.0 months (range, 1.0–43.9 months); during this time, 260 (12.6%) patients died and 609 (29.2%) had recurrence. In Cohort 3, the mean follow-up time was 24.4 months (range, 1–57 months), during which 168 (20.8%) patients died and 375 (46.4%) had recurrence. For Cohort 4, the mean follow-up time was 26.7 months (range, 1.0–47.5 months), during which 87 (35.6%) patients died and 147 (60.2%) had recurrence. TACE treatment occurred in all the four cohorts, and in Cohorts 1 and 2, some patients were treated with IFN based on the judgment of the doctor at the time.

Construction and internal validation of the Shanghai Score

The Shanghai Score was constructed with multivariate Cox regression modeling to predict OS and RFS. The clinical variables used are summarized in Table 2 and include age, HBsAg, HBeAg, partial thromboplastin time, total bilirubin, alkaline phosphatase, γ -glutamyltransferase, α -fetoprotein, tumor size, cirrhosis, vascular invasion, differentiation, encapsulation, and tumor number. The Shanghai Score is calculated by summing the products of the 14 variable values times each variable's corresponding regression coefficient. The performance of Shanghai Score was internally validated in Cohort 1, and the discrimination power, shown by C-index corrected with bootstrapping, was 0.74 (95% confidence interval: 0.73–0.75) for OS and 0.68 (0.67–0.69) for RFS [Supplementary Table 2]. The calibration curves showed good correlation between predicted and observed outcomes for OS prediction [Figure 1].

The performance of the Shanghai Score was further compared with 11 HCC staging systems, including the Okuda *et al.*,^[5] the Chinese University Prognostic Index,^[26] the Eastern staging system,^[27] BCLC,^[6] CLIP,^[7] JIS,^[8] Chinese Staging,^[28] American Joint Committee on Cancer-Tumor node metastasis (TNM),^[29] United Network for Organ Sharing-TNM,^[30] Japan-TNM,^[31] and HKLC^[9] system. Importantly, we found that the Shanghai Score showed the best discrimination (C-index) for both OS and RFS predictions [Supplementary Tables 2 and 3].

Independent validation of the Shanghai Score

In three independent HCC patient cohorts (Cohorts 2, 3, and 4), the Shanghai Score maintained similar predictive performances for OS and RFS. The C-indexes for OS and RFS were 0.72 and 0.66 in Cohort 2, 0.70 and 0.63 in Cohort 3, and 0.70 and 0.68 in Cohort 4, respectively [Supplementary Table 2]. The performance of the Shanghai Score was also best compared with the other 11 staging systems for Cohorts 2–4 [Supplementary Table 3].

Adjuvant treatment recommendations for hepatocellular carcinoma patients after surgery using the Shanghai Score

We first divided Cohort 1 patients into three subgroups (high, moderate, and low risk) based on the distribution of Shanghai Scores [at cutoff values of 4.922 and 4.129;

Table 1: Clinicopathological characteristics in training (Cohort 1), same-center validation (Cohort 2), and extra-center validation (Cohort 3, Cohort 4) datasets

Variables	Cohort 1 (n = 4166)	Cohort 2 (n = 1978)	Cohort 3 (n = 808)	Cohort 4 (n = 244)
Sex				
Male	3570 (86)	1693 (86)	685 (85)	208 (85)
Female	596 (14)	285 (14)	123 (15)	36 (15)
HBsAg				
Positive	3551 (85)	1704 (86)	714 (88)	177 (73)
Negative	615 (15)	274 (14)	94 (12)	67 (27)
HBcAb				
Positive	3795 (91)	1942 (98)	760 (95)	192 (79)
Negative	371 (9)	36 (2)	44 (5)	51 (21)
NA			4 (0)	1 (0)
HBeAg				
Positive	2844 (68)	504 (25)	254 (31)	24 (10)
Negative	1322 (32)	1474 (75)	553 (68)	219 (90)
NA			1 (0)	1 (0)
HCVAb				
Positive	76 (2)	25 (1)	28 (3)	17 (7)
Negative	4090 (98)	1953 (99)	765 (95)	225 (92)
NA			15 (2)	2 (1)
Liver cirrhosis				
None	1018 (24)	983 (50)	407 (50)	152 (62)
Yes	3148 (76)	995 (50)	401 (50)	92 (38)
Child-Pugh score				
A	4054 (97)	1920 (97)	799 (99)	241 (99)
B	111 (3)	58 (3)	9 (1)	3 (1)
C	1 (0)	0	0	0
Encapsulation				
None	1932 (46)	774 (39)	243 (30)	167 (68)
Yes	2234 (54)	1204 (61)	565 (70)	77 (32)
Differentiation				
I–II	3016 (72)	1369 (69)	427 (53)	137 (56)
III–IV	1150 (28)	609 (31)	240 (30)	58 (24)
NA			141 (17)	49 (20)
Tumor number				
Single	3220 (77)	1571 (79)	737 (91)	200 (82)
Two	501 (12)	244 (12)	64 (8)	21 (9)
Multiple	445 (11)	163 (8)	7 (1)	22 (9)
NA				1 (0)
Tumor size				
≤5 cm	2415 (58)	1304 (66)	715 (88)	146 (60)
>5 cm	1751 (42)	674 (34)	93 (12)	94 (39)
NA				4 (2)
Vascular invasion				
None	2675 (64)	1320 (67)	651 (81)	78 (32)
Micro	1036 (25)	536 (27)	157 (19)	138 (57)
Macro	455 (11)	122 (6)	0	27 (11)
NA				1 (0)
Hilar lymph node metastasis				
None	4102 (98)	1934 (98)	805 (100)	244 (100)
Yes	64 (2)	44 (2)	3 (0)	0 (0)
Age (years)	52 (44–60)	54 (46–61)	51 (44–59)	54 (47–60)
TB				
Level (g/L)	14 (11–19)	12 (9–16)	14 (11–18)	15 (12–21)
NA	1 (0)			

Contd...

Table 1: Contd...

Variables	Cohort 1 (n = 4166)	Cohort 2 (n = 1978)	Cohort 3 (n = 808)	Cohort 4 (n = 244)
ALB				
Level (g/L)	42 (39–45)	41 (38–43)	42 (39–44)	46 (42–49)
NA	27 (1)			2 (1)
ALT				
Level (U/L)	38 (26–57)	33 (23–49)	40 (27–57)	35 (22–52)
NA	4 (0)			2 (1)
PTT				
Level (s)	12 (11–13)	12 (11–13)	12 (12–13)	11 (10–12)
NA	4 (0)			2 (1)
AKP				
Level (U/L)	84 (66–109)	76 (62–95)	75 (73–96)	77 (62–101)
NA	22 (1)			1 (0)
GGT				
Level (U/L)	64 (38–115)	53 (34–93)	65 (39–77)	65 (39–114)
NA				6 (2)
AFP				
Level (ng/ml)	107 (8–1210)	33 (5–461)	31 (5–647)	52 (5–1150)
NA	1 (0)			3 (1)

Values were shown as *n* (%), or median (lower quartile–upper quartile). NA: Not applicable; HBsAg: Hepatitis B surface Antigen; HBcAb: Hepatitis B core antibody; HBeAg: Hepatitis B e Antigen; PTT: Partial thromboplastin time; TB: Total bilirubin; AKP: Alkaline phosphatase; GGT: γ -glutamyltransferase; AFP: α -fetoprotein; ALT: Alanine transaminase; ALB: Albumin; HCVAb: Hepatitis C virus antibody.

Table 2: Multivariate analysis of clinical indicators for prognostic prediction in training cohort (n = 4166)

Variables	Up:low*	Overall survival		
		Coefficient	P	HR (95% CI)
Age	60:44	0.0082	<0.001	1.01 (1.00–1.01)
HBsAg	Positive:negative	0.1703	0.020	1.19 (1.02–1.37)
HBeAg	Positive:negative	0.2954	<0.001	1.34 (1.22–1.48)
AFP	1210:8	0.0400	<0.001	1.04 (1.02–1.06)
TB	19:11	0.0023	0.020	1.00 (1.00–1.00)
PTT	13:11	0.0876	<0.001	1.09 (1.05–1.13)
AKP	109:66	0.1669	0.007	1.18 (1.05–1.33)
GGT	115:38	0.1775	<0.001	1.19 (1.12–1.27)
Size	7.5:3.0	0.0818	<0.001	1.09 (1.07–1.10)
Cirrhosis	Yes:no	0.1629	0.007	1.18 (1.05–1.32)
Vascular invasion	Micro:none	0.4057	<0.001	1.50 (1.35–1.67)
	Macro:none	0.8460	<0.001	2.33 (2.01–2.70)
Differentiation	III–IV:I–II	0.2981	<0.001	1.35 (1.22–1.48)
Encapsulation	Yes:no	–0.1743	<0.001	0.84 (0.76–0.93)
Number	Double:single	0.3740	<0.001	1.45 (1.27–1.66)
	Multiple:single	0.3989	<0.001	1.49 (1.30–1.70)

*Upper quartile vs. lower quartile. HR: Hazard ratio; HBsAg: Hepatitis B surface antigen; HBeAg: Hepatitis B e antigen; PTT: Partial thromboplastin time; TB: Total bilirubin; AKP: Alkaline phosphatase; GGT: γ -glutamyltransferase; AFP: α -fetoprotein; CI: Confidence interval.

Figure 2a and 2b]. The three subgroups showed distinct prognoses for both OS and RFS [Figure 2c and 2d]. To compare the effects of adjuvant TACE on the survival and recurrence of 420 low-risk HCC patients in Cohort 1, we used PSM^[25] to choose 420 patients without TACE treatment and found that there were no significant differences between the two groups [Figure 3a and 3b]. In the moderate-risk group, adjuvant TACE was associated with improved survival time in 358 HCC patients, compared with the corresponding PSM patients [Figure 3c and 3d]. For high-risk HCC patients, both OS and RFS of the 195 patients who received TACE were

improved significantly [Figure 3e and 3f]. In Cohort 2, adjuvant TACE had a damaging effect in the low-risk group [Supplementary Figure 1a and 1b], a slight beneficial effect (not statistically significant) in the moderate-risk group [Supplementary Figure 1c and 1d], and was associated with significantly improved OS in high-risk patients [Supplementary Figure 1e and 1f]. In Cohort 3, adjuvant TACE had no survival benefits in low-risk patients [Supplementary Figure 2a and 2b], a beneficial tendency in the moderate-risk group [Supplementary Figure 2c and 2d], and significantly improved RFS for high-risk patients [Supplementary Figure 2e and 2f]. Together, these results

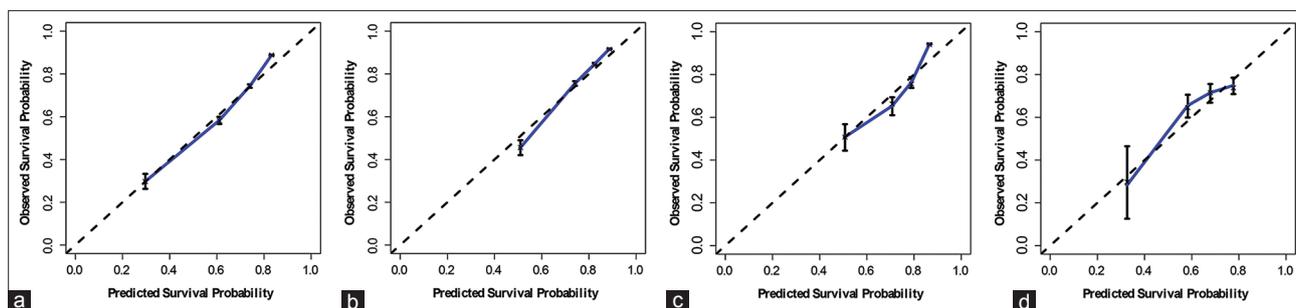


Figure 1: Calibration curves of the Shanghai Score in patient cohorts. The X-axis represents the estimated 3-year overall survival probability and the Y-axis represents the observed 3-year overall survival probability in Cohort 1 (a), Cohort 2 (b), Cohort 3 (c), and Cohort 4 (d).

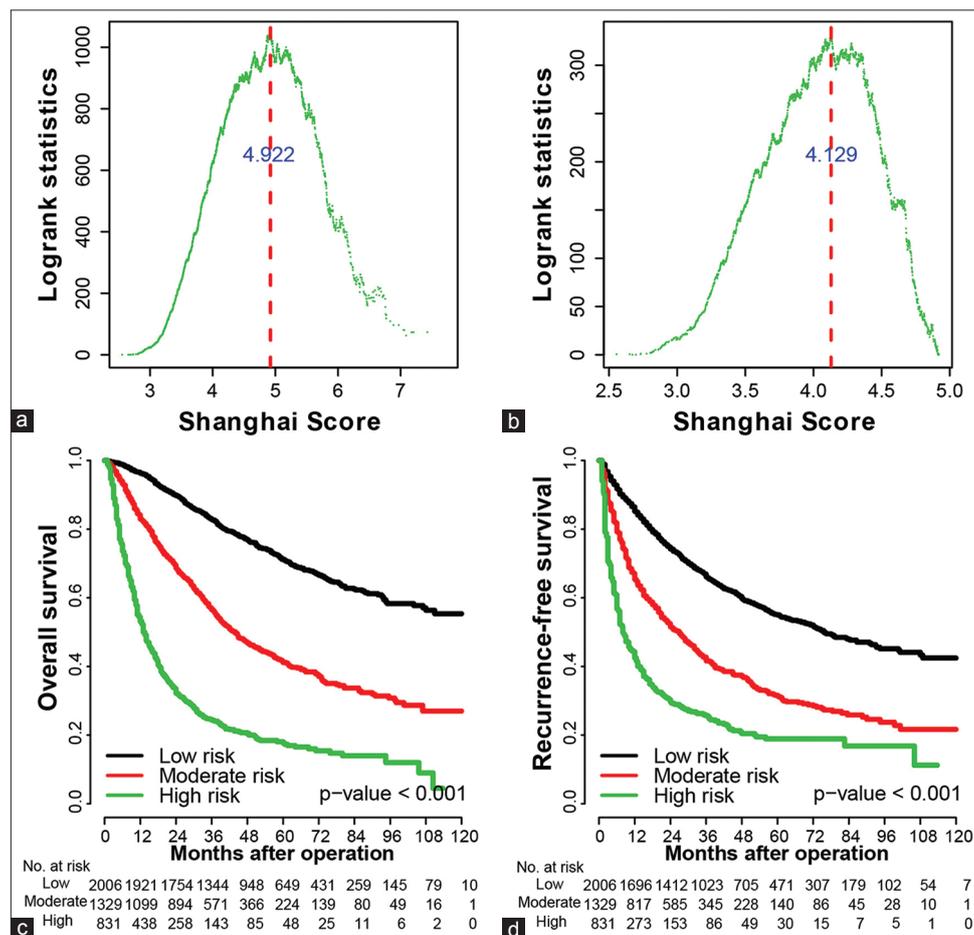


Figure 2: Identification of the cutoffs at which patients in Cohort 1 can be classified into different groups using the Shanghai Score. (a) The log-rank statistics distribution between groups based on different cutoffs. A Shanghai Score cutoff of 4.922 identified two groups (higher and lower) with the overall survival difference reaching the maximum. (b) The log-rank statistics distribution among lower risk patients. A Shanghai Score cutoff of 4.129 subtyped the lower risk patients into moderate- and low-risk patients, with the overall survival difference between the two groups reaching the maximum. (c and d) Kaplan–Meier curves for overall survival and recurrence-free survival according to the Shanghai Score.

suggest that a patient with a Shanghai Score prediction of high risk could be recommended for local TACE adjuvant therapy.

We next investigated the effect of IFN treatment on patient prognosis and found that it showed no survival benefit (either OS or RFS) for low-risk HCC patients in Cohort 1 after surgery compared with PSM patients [Figure 4a and 4b]. However, we found that IFN treatment was effective in prolonging OS and RFS in the moderate-risk

[Figure 4c and 4d] and high-risk groups [Figure 4e and 4f]. IFN treatment had no survival benefit in the low-risk group of Cohort 2 [Supplementary Figure 3a and 3b], but improved both OS and RFS in the moderate-risk group [Supplementary Figure 3c and 3d]. Therefore, a patient with a Shanghai Score falling into moderate- or high-risk groups might be recommended for systemic IFN adjuvant treatment.

In summary, using our newly developed Shanghai Score, we found that, in addition to routine treatment, more

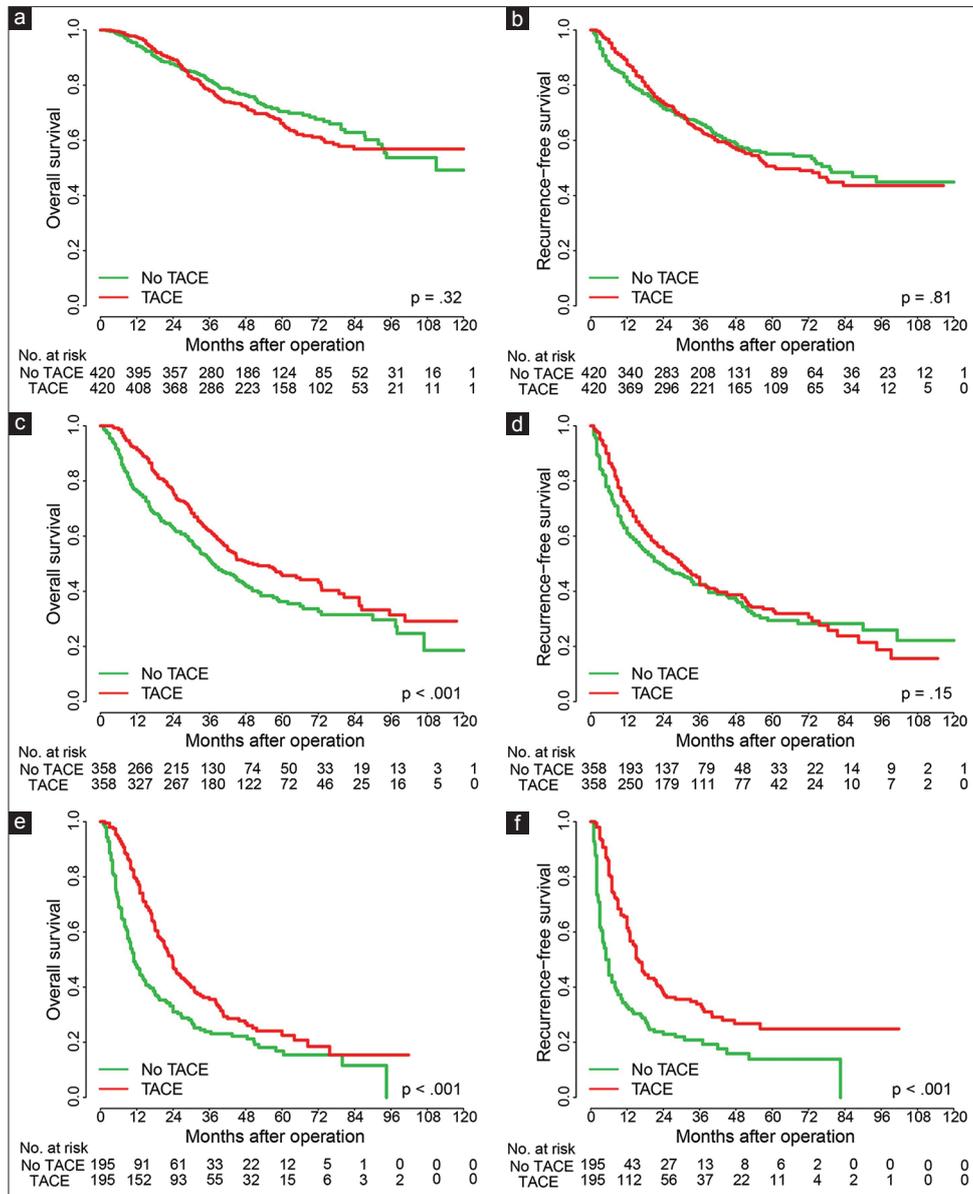


Figure 3: The survival benefit for postoperative transcatheter arterial chemoembolization (TACE) using the Shanghai Score for Cohort 1. Overall survival and recurrence-free survival in low-risk patients (a and b), moderate-risk patients (c and d), and high-risk patients (e and f); $n = 4166$ patients.

aggressive treatments should be advised for high-risk patients, such as increasing the frequency of surveillance and administration of TACE or IFN treatment, which could improve patient survival. For low-risk patients, a relatively conservative treatment may be advised instead of TACE or IFN, to avoid overtreatment. For moderate-risk patients, appropriate treatment should be recommended on an individual basis.

Web server

Using a similar method as nomogram, we implemented a software application as a simpler way for clinicians to use the Shanghai Score. Clinicians can use this software to evaluate prognosis and potential adjuvant therapy for HCC patients after hepatectomy by simply inputting various requested patient characteristics [Figure 5]. Compared to a nomogram, the web server has the advantages of quickly

assessing patient risk level and obtaining predicted OS and RFS at six different time points.

DISCUSSION

The European Association for the Study of the Liver/American Association for the Study of Liver Diseases guidelines recommend liver resection for patients with single nodules, fewer than three small nodules (≤ 3 cm), no clinically significant portal hypertension, and normal bilirubin, routed by BCLC. If judged by these criteria, 50% of patients undergoing liver resection in China would be considered unsuitable for hepatectomy. This restrictive approach was established more than 15 years ago, and has not evolved over time, despite significant improvement in surgical techniques and technologies.^[32] Recently, many Western and Eastern liver cancer centers advocated hepatectomy to treat HCC

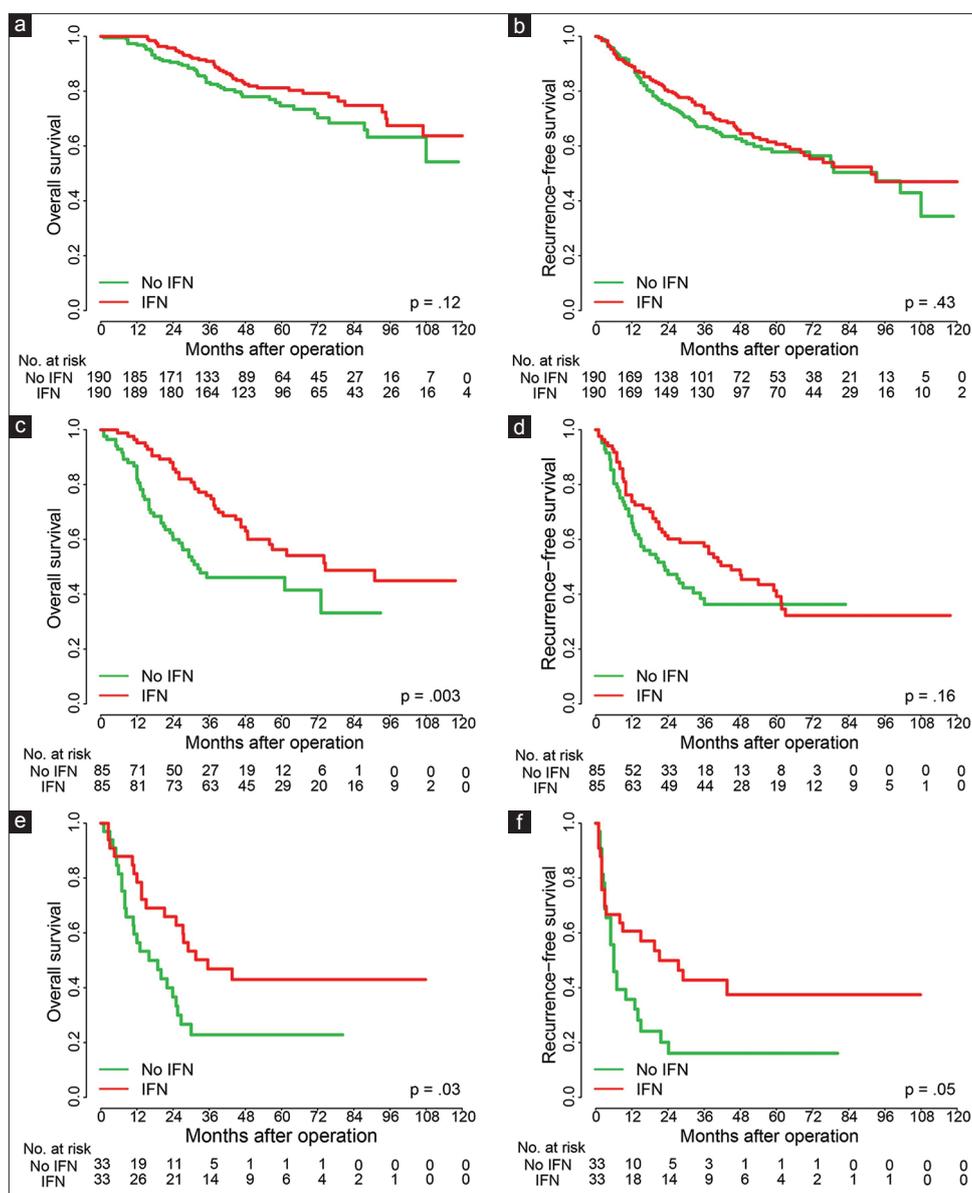


Figure 4: The survival benefit for postoperative interferon (IFN) using the Shanghai Score for Cohort 1. Overall survival and recurrence-free survival in low-risk patients (a and b), moderate-risk patients (c and d), and high-risk patients (e and f); $n = 4166$ patients.

outside of the BCLC criteria, and retrospective analyses demonstrated that these expanded patients had acceptable prognoses.^[4,32-36] Therefore, the National Liver Cancer Treatment Guideline of China recommended resection for selected patients with BCLC B or C stage. The extended indication for resection may be followed by a general increase in tumor recurrence and a decrease in survival, but may also save patients who may have curative potential. This is especially necessary for developing countries like China where most HCC patients are past early-stage disease at diagnosis, where liver resection still represents the cornerstone for any curability attempt.

The aim of the Shanghai Score is to construct a practical prognosis system to identify patients who may benefit from curative hepatectomy within the diverse population of Chinese HCC patients and provide an evaluation for possible

adjuvant therapy, including local TACE or systemic IFN treatment. Since there has been no standardized adjuvant therapy for HCC in the recent years, we only included TACE and IFN in the development of the Shanghai Score. More modern antiviral therapy or targeted drugs, like sorafenib, were not commonly used before 2012 when our retrospective patient cohorts were collected.

Up until now, there have been no standardized adjuvant treatments for HCC patients after surgery, and the practice guidelines of the American Association for the Study of Liver Diseases do not include TACE or IFN as suggested adjuvant therapy options. Although our previous studies demonstrated that TACE improves survival for patients at high risk for residual tumors^[15] and IFN treatment postpones recurrence and improves survival of HBV-infected patients,^[14] it is difficult to identify those who are most likely to benefit

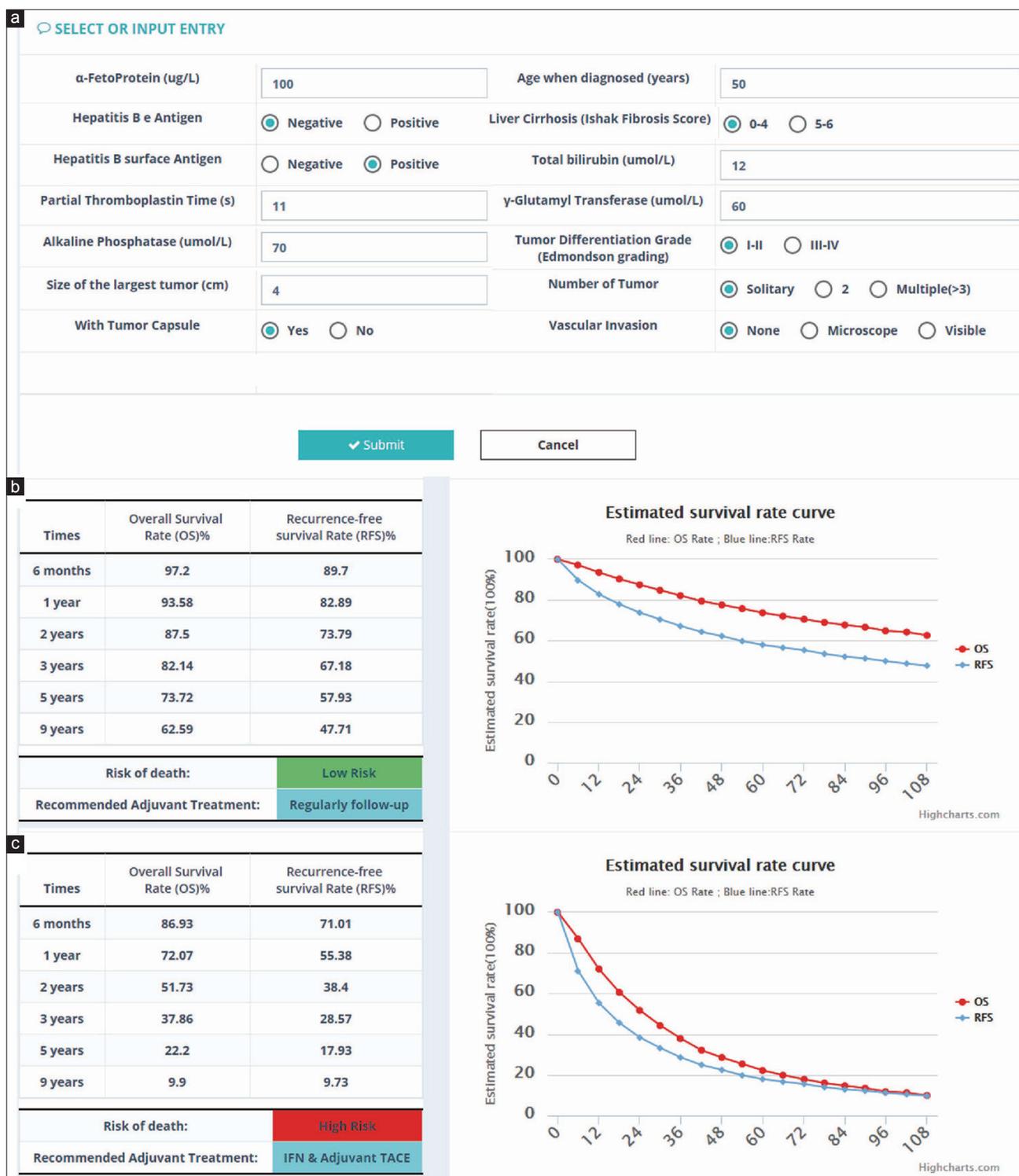


Figure 5: The Shanghai Score web server interface. (a) Clinical indexes included in the Shanghai Score are listed in the first page, and should either be selected or inputted, and submitted to back stage model calculation. (b and c) Two examples of results predicted by the Shanghai Score. An individualized estimate of survival probability and recurrence rate for each case is calculated for 6 months, and 1, 2, 3, 5, and 9 years after surgery, and personalized treatment recommendations are provided based on the survival benefit estimation.

from such adjuvant therapies. In clinical practice, a patient mainly relies on the clinician's limited personal experience to recommend adjuvant treatment after surgery, which may be subjective and biased. In this study, our patient-by-patient suggestive approach to Chinese HCC patients using the

Shanghai Score is based on a more powerful statistical assessment, by combining PSM with our score system; thus it should be more objective and reliable. Our results show that adjuvant TACE might be beneficial for patients with a high-risk score after surgery, while it might be redundant and

even harmful for patients with a low-risk score. Additionally, IFN treatment after surgery showed a survival benefit for HCC patients with a moderate- or high-risk score, but was not useful to patients with a low-risk score.

The Shanghai Score was developed using retrospective analyses of the largest number of HCC patients ever in China (a total of 7196 patients, distributed across one derivation cohort and three validation cohorts). In this scoring system, we included risk factors and demographic characteristics such as age, HBsAg, HBeAg, and cirrhosis; tumor-related factors such as tumor size, vascular invasion, differentiation, encapsulation, and tumor number; and the indexes reflecting liver impairment such as α -fetoprotein, total bilirubin, partial thromboplastin time, alkaline phosphatase, and γ -glutamyltransferase. Personal prognostic prediction is essential to guide postoperative treatment and counseling.^[32] The Shanghai Score system may be used as a clinical model for predicting prognosis and recommending adjuvant treatment for HCC patients after surgery, with web server-accessible continuous prognostic indexes to estimate risk for each patient. Our patient-by-patient approach is similar to nomogram, which is more practical and accurate than a staging system in predicting prognosis.^[37] In comparison with 11 other widely used staging systems, the Shanghai Score had the best predictive ability in our four HCC cohorts. Unfortunately, recently published postoperative prognosis models have also introduced different and new indexes, which we could not use for comparison. The emergence of new scoring and staging systems indicates that there are no widely acceptable standards for postoperative staging, as there are for preoperative staging.

There are a few limitations to our study. First, the work was based on retrospective analyses of multiple centers, and the follow-up intervals, postsurgery therapy, and postrecurrence therapy results were heterogeneous. However, the Shanghai Score was constructed based on patients who underwent resection from 2001 to 2008, and the performance was validated in three independent cohorts, in which patients all underwent resection after 2008. This indicates that the Shanghai Score can be applied prospectively. Second, the benefits of postsurgical TACE or IFN were estimated in patients using a PSM approach, in which other factors, such as treatment interval, were not be considered and may possibly introduce bias into the analysis. In addition, HBV DNA load was not routinely tested for at our institute until 2005, but the prognostic role of HBV DNA load had no significant advantage compared with other clinical factors (data not shown). Thus, we did not include HBV DNA load in this study. Additionally, more recent treatment options have not yet been included. Finally, our patients were all from China and most had a background of HBV infection; therefore, whether this prognostic system could be applicable to patients with a Western background is still unclear.

Supplementary information is linked to the online version of the paper on the Chinese Medical Journal website.

Financial support and sponsorship

This study was supported by grants from the National High Technology Research and Development Program (863 Program) of China (No. 2015AA020401), National Key Research and Development Program of China (No. 2016YFC0904101), State Key Program of National Natural Science Foundation of China (No. 81530077), National Natural Science Foundation of China (No. 81372317, 81472676, and 81572823), Projects from the Shanghai Science and Technology Commission (No. 14DZ1940300, 14411970200, 134119a1201, and 14140902301), and Specialized Research Fund for the Doctoral Program of Higher Education and Research Grants Council Earmarked Research Grants Joint Research Scheme (No. 20130071140008).

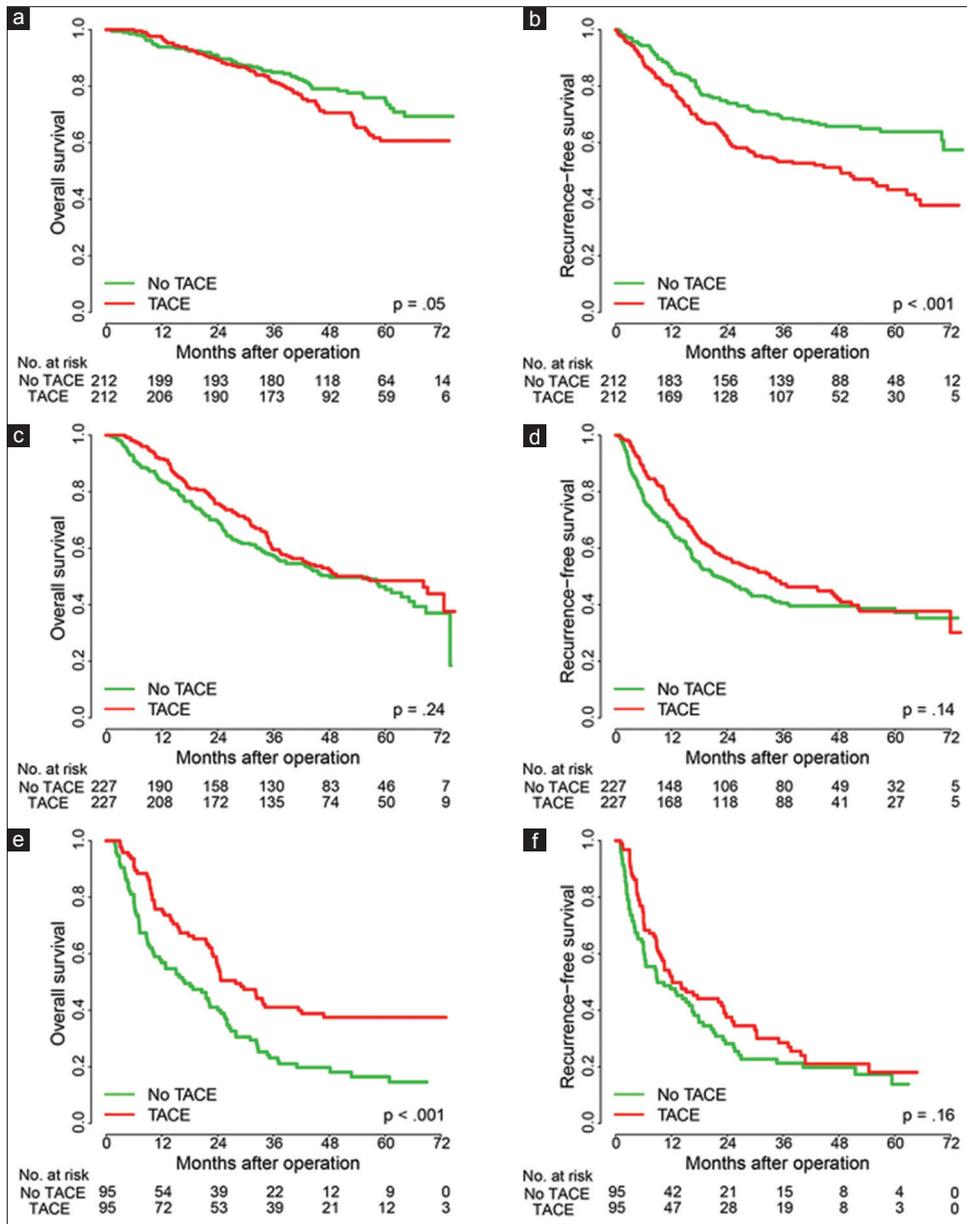
Conflicts of interest

There are no conflicts of interest.

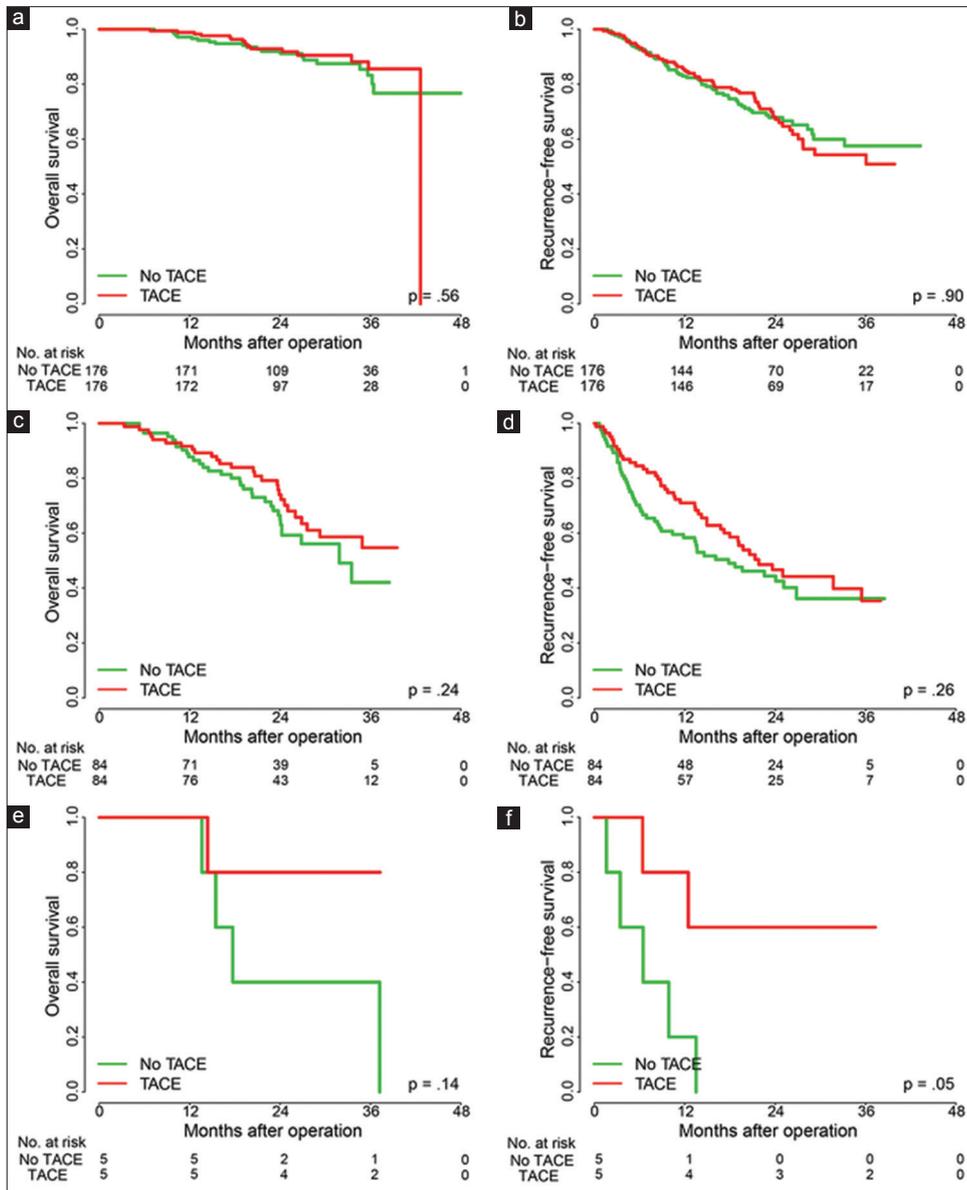
REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A, *et al*. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;65:87-108. doi: 10.3322/caac.21262.
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, *et al*. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;66:115-32. doi: 10.3322/caac.21338.
3. Bruix J, Takayama T, Mazzaferro V, Chau GY, Yang J, Kudo M, *et al*. Adjuvant sorafenib for hepatocellular carcinoma after resection or ablation (STORM): A phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Oncol* 2015;16:1344-54. doi: 10.1016/S1470-2045(15)00198-9.
4. Roayaie S, Jibara G, Tabrizian P, Park JW, Yang J, Yan L, *et al*. The role of hepatic resection in the treatment of hepatocellular cancer. *Hepatology* 2015;62:440-51. doi: 10.1002/hep.27745.
5. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, *et al*. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985;56:918-28. doi: 10.1002/1097-0142(19850815)56:4<918::AID-CNCR2820560437>3.0.CO;2-E.
6. Fedi P, Bafico A, Nieto Soria A, Burgess WH, Miki T, Bottaro DP, *et al*. Isolation and biochemical characterization of the human dkk-1 homologue, a novel inhibitor of mammalian Wnt signaling. *J Biol Chem* 1999;274:19465-72. doi: 10.1074/jbc.274.27.19465.
7. A new prognostic system for hepatocellular carcinoma: A retrospective study of 435 patients: The Cancer of the Liver Italian Program (CLIP) investigators. *Hepatology* 1998;28:751-5. doi: 10.1002/hep.510280322.
8. Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular carcinoma (CLIP score): Its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging Score (JIS score). *J Gastroenterol* 2003;38:207-15. doi: 10.1007/s005350300038.
9. Yau T, Tang VY, Yao TJ, Fan ST, Lo CM, Poon RT, *et al*. Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. *Gastroenterology* 2014;146:1691-7.e3. doi: 10.1053/j.gastro.2014.02.032.
10. Yang P, Qiu J, Li J, Wu D, Wan X, Lau WY, *et al*. Nomograms for pre- and postoperative prediction of long-term survival for patients who underwent hepatectomy for multiple hepatocellular carcinomas. *Ann Surg* 2016;263:778-86. doi: 10.1097/SLA.0000000000001339.
11. Shim JH, Jun MJ, Han S, Lee YJ, Lee SG, Kim KM, *et al*. Prognostic nomograms for prediction of recurrence and survival after curative liver resection for hepatocellular carcinoma. *Ann Surg* 2015;261:939-46. doi: 10.1097/SLA.0000000000000747.
12. Huang LF, Xing X, Wu D, Xia Y, Li J, Wang K, *et al*. A novel scoring system predicts adjuvant chemolipiodolization benefit for hepatocellular carcinoma patients after hepatectomy. *Oncotarget*

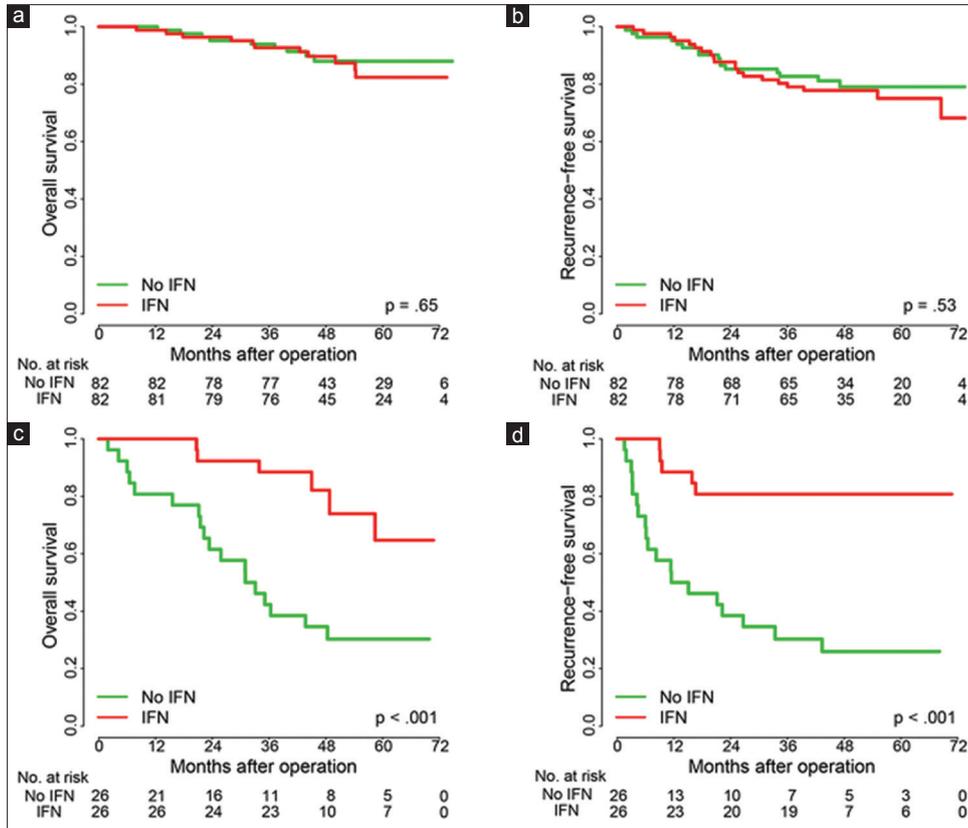
- 2016;7:25493-506. doi: 10.18632/oncotarget.8333.
13. Li Y, Xia Y, Li J, Wu D, Wan X, Wang K, *et al.* Prognostic nomograms for pre- and postoperative predictions of long-term survival for patients who underwent liver resection for Huge Hepatocellular carcinoma. *J Am Coll Surg* 2015;221:962-74. doi: 10.1016/j.jamcollsurg.2015.08.003.
 14. Sun HC, Tang ZY, Wang L, Qin LX, Ma ZC, Ye QH, *et al.* Postoperative interferon alpha treatment postponed recurrence and improved overall survival in patients after curative resection of HBV-related hepatocellular carcinoma: A randomized clinical trial. *J Cancer Res Clin Oncol* 2006;132:458-65. doi: 10.1007/s00432-006-0091-y.
 15. Ren ZG, Lin ZY, Xia JL, Ye SL, Ma ZC, Ye QH, *et al.* Postoperative adjuvant arterial chemoembolization improves survival of hepatocellular carcinoma patients with risk factors for residual tumor: A retrospective control study. *World J Gastroenterol* 2004;10:2791-4. doi: 10.3748/wjg.v10.i19.2791.
 16. Chen X, Zhang B, Yin X, Ren Z, Qiu S, Zhou J, *et al.* Lipiodolized transarterial chemoembolization in hepatocellular carcinoma patients after curative resection. *J Cancer Res Clin Oncol* 2013;139:773-81. doi: 10.1007/s00432-012-1343-7.
 17. Sun JJ, Wang K, Zhang CZ, Guo WX, Shi J, Cong WM, *et al.* Postoperative adjuvant transcatheter arterial chemoembolization after R0 hepatectomy improves outcomes of patients who have hepatocellular carcinoma with microvascular invasion. *Ann Surg Oncol* 2016;23:1344-51. doi: 10.1245/s10434-015-5008-z.
 18. Lu XJ, Dong J, Ji LJ, Luo JH, Cao HM, Xiao LX, *et al.* Safety and efficacy of TACE and gamma knife on hepatocellular carcinoma with portal vein invasion. *Gut* 2016;65:715-6. doi: 10.1136/gutjnl-2015-310292.
 19. Chen LT, Chen MF, Li LA, Lee PH, Jeng LB, Lin DY, *et al.* Long-term results of a randomized, observation-controlled, phase III trial of adjuvant interferon alfa-2b in hepatocellular carcinoma after curative resection. *Ann Surg* 2012;255:8-17. doi: 10.1097/SLA.0b013e3182363ff9.
 20. Bruix J, Han KH, Gores G, Llovet JM, Mazzaferro V. Liver cancer: Approaching a personalized care. *J Hepatol* 2015;62:S144-56. doi: 10.1016/j.jhep.2015.02.007.
 21. Alzola C, Harrell F. An Introduction to S and The Hmisc and Design Libraries; 2006. Available from: <https://cran.r-project.org/doc/contrib/Alzola+Harrell-Hmisc-Design-Intro.pdf>. [Last accessed on 2017 Jun 15].
 22. 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. doi: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.
 23. Harrell FE Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA* 1982;247:2543-6. doi: 10.1001/jama.1982.03320430047030.
 24. Efron B. Bootstrap methods: Another look at the jackknife. *Ann Stat* 1979;7:1-26. doi: 10.1214/aos/1176344552.
 25. Choi GH, Shim JH, Kim MJ, Ryu MH, Ryoo BY, Kang YK, *et al.* Sorafenib alone versus sorafenib combined with transarterial chemoembolization for advanced-stage hepatocellular carcinoma: Results of propensity score analyses. *Radiology* 2013;269:603-11. doi: 10.1148/radiol.13130150.
 26. Leung TW, Tang AM, Zee B, Lau WY, Lai PB, Leung KL, *et al.* Construction of the Chinese University Prognostic Index for hepatocellular carcinoma and comparison with the TNM staging system, the Okuda staging system, and the Cancer of the Liver Italian Program staging system: A study based on 926 patients. *Cancer* 2002;94:1760-9. doi: 10.1002/cncr.10384.
 27. Yang T, Zhang J, Lu JH, Yang LQ, Yang GS, Wu MC, *et al.* A new staging system for resectable hepatocellular carcinoma: Comparison with six existing staging systems in a large Chinese cohort. *J Cancer Res Clin Oncol* 2011;137:739-50. doi: 10.1007/s00432-010-0935-3.
 28. Sheng JM, Zhao WH, Wu FS, Ma ZM, Feng YZ, Zhou XR, *et al.* The Chinese classification system compared with TNM staging in prognosis of patients with primary hepatic carcinoma after resection. *Hepatobiliary Pancreat Dis Int* 2005;4:561-4.
 29. Chan AC, Fan ST, Poon RT, Cheung TT, Chok KS, Chan SC, *et al.* Evaluation of the seventh edition of the American Joint Committee on Cancer tumour-node-metastasis (TNM) staging system for patients undergoing curative resection of hepatocellular carcinoma: Implications for the development of a refined staging system. *HPB (Oxford)* 2013;15:439-48. doi: 10.1111/j.1477-2574.2012.00617.x.
 30. OPTN (The Organ Procurement and Transplantation Network) Policies. Sharing UNFO; 2007. Available from: <https://optn.transplant.hrsa.gov/>. [Last accessed on 2017 Jun 15].
 31. Vauthey JN, Lauwers GY, Esnaola NF, Do KA, Belghiti J, Mirza N, *et al.* Simplified staging for hepatocellular carcinoma. *J Clin Oncol* 2002;20:1527-36. doi: 10.1200/JCO.2002.20.6.1527.
 32. Sposito C, Di Sandro S, Brunero F, Buscemi V, Battiston C, Lauterio A, *et al.* Development of a prognostic scoring system for resectable hepatocellular carcinoma. *World J Gastroenterol* 2016;22:8194-202. doi: 10.3748/wjg.v22.i36.8194.
 33. Zhong JH, Ke Y, Gong WF, Xiang BD, Ma L, Ye XP, *et al.* Hepatic resection associated with good survival for selected patients with intermediate and advanced-stage hepatocellular carcinoma. *Ann Surg* 2014;260:329-40. doi: 10.1097/sla.0000000000000236.
 34. Yin L, Li H, Li AJ, Lau WY, Pan ZY, Lai EC, *et al.* Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan criteria: A RCT. *J Hepatol* 2014;61:82-8. doi: 10.1016/j.jhep.2014.03.012.
 35. Vitale A, Burra P, Frigo AC, Trevisani F, Farinati F, Spolverato G, *et al.* Survival benefit of liver resection for patients with hepatocellular carcinoma across different Barcelona Clinic Liver Cancer stages: A multicentre study. *J Hepatol* 2015;62:617-24. doi: 10.1016/j.jhep.2014.10.037.
 36. Shen J, He L, Li C, Wen T, Chen W, Lu C, *et al.* Prognostic nomograms for patients with resectable hepatocellular carcinoma incorporating systemic inflammation and tumor characteristics. *Oncotarget* 2016;7:80783-93. doi: 10.18632/oncotarget.13038.
 37. Cho CS, Gonen M, Shia J, Kattan MW, Klimstra DS, Jarnagin WR, *et al.* A novel prognostic nomogram is more accurate than conventional staging systems for predicting survival after resection of hepatocellular carcinoma. *J Am Coll Surg* 2008;206:281-91. doi: 10.1016/j.jamcollsurg.2007.07.031.



Supplementary Figure 1: The survival benefit for recommending postoperative TACE by Shanghai Score based on propensity score-matched approach in Cohort 2 ($n = 1978$). (a and b) OS and RFS in low-risk patients, TACE actually significantly made both overall survival and recurrence-free survival worse. (c and d) OS and RFS in moderate-risk patients. (e and f) OS and RFS in high-risk patients. TACE: Transcatheter arterial chemoembolization; OS: Overall survival; RFS: Recurrence-free survival.



Supplementary Figure 2: The survival benefit for recommending postoperative TACE by Shanghai Score based on propensity score-matched approach in Cohort 3 ($n = 808$). (a and b) OS and RFS in low-risk patients. (c and d) OS and RFS in moderate-risk patients. (e and f) OS and RFS in high-risk patients. TACE: Transcatheter arterial chemoembolization; OS: Overall survival; RFS: Recurrence-free survival.



Supplementary Figure 3: The survival benefit for recommending postoperative IFN by Shanghai Score based on propensity score-matched approach in Cohort 2 ($n = 1978$). (a and b) OS and RFS in low-risk patients. (c and d) OS and RFS in moderate-risk patients. IFN: Interferon; OS: Overall survival; RFS: Recurrence-free survival.

Supplementary Table 1: The distribution of 11 tumor stage systems in training and validation datasets

Tumor stage system	Cohort 1 (n = 4166), n (%)	Cohort 2 (n = 1978), n (%)	Cohort 3 (n = 808), n (%)	Cohort 4 (n = 244), n (%)
AJCC-TNM				
I	2273 (55)	1073 (54)	593 (73)	66 (27)
II	1175 (28)	546 (28)	202 (25)	140 (57)
III	654 (16)	315 (16)	10 (1)	17 (7)
IV	64 (2)	44 (2)	3 (0)	21 (9)
UNOS-TNM				
I	249 (6)	172 (9)	83 (10)	13 (5)
II	1879 (45)	984 (50)	622 (77)	110 (45)
III	1456 (35)	630 (32)	100 (12)	82 (34)
IV	582 (14)	192 (10)	3 (0)	39 (16)
Japan TNM				
I	249 (6)	171 (9)	83 (10)	13 (5)
II	2079 (50)	1272 (64)	663 (82)	170 (70)
III	1648 (40)	408 (21)	59 (7)	55 (23)
IV	190 (5)	127 (6)	3 (0)	6 (2)
Okuda				
I	3576 (86)	1724 (87)	766 (95)	193 (79)
II	561 (13)	254 (13)	41 (5)	51 (21)
III	29 (1)	0	1 (0)	0
CS stage				
I	1928 (46)	1159 (59)	698 (86)	134 (55)
II	1661 (40)	569 (29)	98 (12)	107 (44)
III	577 (14)	250 (13)	12 (1)	3 (1)
BCLC stage				
0–A	1827 (44)	1136 (57)	697 (86)	186 (76)
B	1839 (44)	720 (36)	111 (14)	30 (12)
C–D	500 (12)	122 (6)	0 (0)	28 (11)
JIS score				
0	242 (6)	168 (8)	82 (10)	13 (5)
1	2043 (49)	1242 (63)	656 (81)	169 (69)
2	1639 (39)	423 (21)	67 (8)	54 (22)
3	231 (6)	141 (7)	3 (0)	8 (3)
4	11 (0)	4 (0)	0	0
CLIP				
0	1924 (46)	1104 (56)	506 (63)	117 (48)
1–3	1902 (46)	866 (44)	302 (37)	114 (47)
4–6	340 (8)	8 (0)	0	13 (5)
CUPI				
<1	4135 (99)	1961 (99)	808 (100)	238 (98)
2–7	29 (1)	17 (1)	0	8 (2)
>8	2 (0)	0	0	0
Eastern stage				
0–1	1208 (29)	956 (48)	357 (44)	53 (22)
2–3	2289 (55)	852 (43)	411 (51)	190 (78)
4–8	669 (16)	170 (9)	40 (5)	1 (0)
HKLC				
I	1741 (42)	910 (46)	6 (1)	50 (20)
II	2246 (54)	958 (48)	579 (72)	112 (46)
III	179 (4)	110 (6)	223 (28)	82 (34)

BCLC: Barcelona Clinic Liver Cancer Staging; AJCC: American Joint Committee on Cancer; TNM: Tumor node metastasis; CLIP: Cancer of the Liver Italian Program score; CS: Chinese Staging; UNOS: United Network for Organ Sharing; JIS: Japan Integrated Staging score; CUPI: Chinese University Prognostic Index staging; HKLC: Hong Kong Liver Cancer Staging System.

Supplementary Table 2: Performance of different staging systems in training cohort and validation cohorts by univariate analysis

Univariate analysis	C-index (95% CI)							
	Cohort 1 (n = 4166)		Cohort 2 (n = 1978)		Cohort 3 (n = 808)		Cohort 4 (n = 244)	
	OS	RFS	OS	RFS	OS	RFS	OS	RFS
Shanghai score	0.74 (0.73–0.75)	0.68 (0.67–0.69)	0.72 (0.70–0.74)	0.66 (0.64–0.68)	0.70 (0.66–0.74)	0.63 (0.60–0.66)	0.70 (0.65–0.76)	0.68 (0.63–0.72)
BCLC	0.66 (0.65–0.68)	0.62 (0.60–0.63)	0.65 (0.63–0.67)	0.61 (0.59–0.63)	0.57 (0.53–0.61)	0.56 (0.53–0.59)	0.61 (0.56–0.66)	0.60 (0.56–0.64)
AJCC-TNM	0.66 (0.65–0.67)	0.61 (0.60–0.62)	0.66 (0.64–0.68)	0.62 (0.60–0.63)	0.60 (0.56–0.64)	0.57 (0.55–0.60)	0.60 (0.54–0.65)	0.58 (0.54–0.62)
CLIP	0.64 (0.62–0.65)	0.60 (0.58–0.61)	0.62 (0.60–0.64)	0.58 (0.57–0.60)	0.59 (0.55–0.63)	0.56 (0.53–0.59)	0.63 (0.57–0.70)	0.65 (0.61–0.70)
Okuda	0.57 (0.56–0.58)	0.55 (0.54–0.55)	0.56 (0.54–0.57)	0.54 (0.53–0.55)	0.53 (0.50–0.55)	0.52 (0.51–0.54)	0.59 (0.54–0.64)	0.57 (0.53–0.61)
CS stage	0.66 (0.65–0.68)	0.61 (0.60–0.62)	0.66 (0.64–0.68)	0.62 (0.60–0.63)	0.55 (0.52–0.58)	0.55 (0.53–0.57)	0.61 (0.55–0.66)	0.60 (0.56–0.63)
UNOS-TNM	0.66 (0.65–0.67)	0.62 (0.60–0.63)	0.67 (0.65–0.69)	0.62 (0.61–0.64)	0.57 (0.53–0.60)	0.56 (0.53–0.58)	0.64 (0.59–0.69)	0.64 (0.59–0.68)
JIS score	0.65 (0.64–0.66)	0.60 (0.59–0.61)	0.66 (0.64–0.67)	0.61 (0.60–0.63)	0.54 (0.52–0.57)	0.52 (0.50–0.54)	0.62 (0.57–0.67)	0.61 (0.57–0.65)
Japan-TNM	0.64 (0.63–0.65)	0.60 (0.59–0.61)	0.65 (0.63–0.67)	0.61 (0.59–0.62)	0.54 (0.51–0.57)	0.53 (0.51–0.55)	0.62 (0.56–0.67)	0.61 (0.57–0.65)
CUPI	0.64 (0.63–0.65)	0.59 (0.58–0.60)	0.51 (0.50–0.51)	0.50 (0.50–0.51)	0.58 (0.54–0.63)	0.57 (0.54–0.59)	0.61 (0.55–0.67)	0.61 (0.57–0.66)
Eastern stage	0.66 (0.65–0.68)	0.61 (0.60–0.63)	0.63 (0.61–0.65)	0.60 (0.59–0.62)	0.58 (0.53–0.63)	0.56 (0.53–0.59)	0.64 (0.59–0.69)	0.61 (0.57–0.65)
HKLC	0.65 (0.64–0.66)	0.61 (0.60–0.62)	0.65 (0.63–0.67)	0.61 (0.59–0.62)	0.61 (0.57–0.65)	0.58 (0.56–0.61)	0.61 (0.56–0.67)	0.60 (0.55–0.64)

OS: Overall survival; RFS: Recurrence-free survival; BCLC: Barcelona Clinic Liver Cancer staging; AJCC: American Joint Committee on Cancer; TNM: Tumor node metastasis; CLIP: Cancer of the Liver Italian Program score; CS: Chinese Staging; UNOS: United Network for Organ Sharing; JIS: Japan Integrated Staging score; CUPI: Chinese University Prognostic Index staging; HKLC: Hong Kong Liver Cancer Staging System; C-index: Concordance index; CI: Confidence interval.

Supplementary Table 3: Performance of different staging systems in training and validation datasets*

Items	Cohort 1 (n = 4166)			Cohort 2 (n = 1978)			Cohort 3 (n = 808)			Cohort 4 (n = 244)		
	ΔLHR	P	ΔAIC	ΔLHR	P	ΔAIC	ΔLHR	P	ΔAIC	ΔLHR	P	ΔAIC
Removing Shanghai Score	-272.359	<0.001	270.359	-78.960	<0.001	76.960	-32.400	<0.001	30.400	-9.937	0.001	7.937
Removing BCLC	-0.981	0.320	-1.019	-28.185	<0.001	26.185	-0.151	0.698	-1.849	-1.454	0.224	-0.546
Removing TNM	-1.277	0.257	-0.723	-0.855	0.356	-1.145	-0.127	0.724	-1.873	-0.172	0.679	-1.828
Removing CLIP	-0.371	0.542	-1.629	-0.125	0.723	-1.875	-0.363	0.548	-1.637	-1.116	0.291	-0.884
Removing OKUDA	-0.966	0.327	-1.034	-2.071	0.145	0.071	-0.833	0.374	-1.167	-0.329	0.566	-1.671
Removing C stage	0	0.989	-2.000	-20.184	<0.001	18.184	-0.202	0.657	-1.798	-0.001	0.978	-1.999
Removing UNOS	-0.052	0.820	-1.948	-22.213	<0.001	20.213	-0.169	0.685	-1.831	-1.790	0.177	-0.210
Removing JIS score	-1.452	0.223	-0.548	-5.134	0.027	3.134	-0.002	0.961	-1.998	-0.233	0.639	-1.767
Removing Japan TNM	-0.906	0.337	-1.094	-4.531	0.036	2.531	-0.005	0.941	-1.995	-1.360	0.266	-0.640
Removing CUPI	-0.103	0.749	-1.897	-0.117	0.755	-1.883	-0.008	0.929	-1.992	-0.002	0.966	-1.998
Removing Eastern System	-0.110	0.741	-1.890	-2.883	0.091	0.883	-2.869	0.092	0.869	-4.441	0.036	2.441
Removing HKLC	-0.061	0.805	-1.939	-0.127	0.721	-1.873	-1.797	0.182	-0.203	-1.617	0.201	-0.383

*In multivariate analysis, the independent contribution of each staging system in the full model was assessed by removing the concerned system. ΔLHR was calculated by first fitting full model and then removing one concerned system, at last $\Delta\text{LHR} = \text{LHR}_{\text{removing model}} - \text{LHR}_{\text{full model}}$; the same to ΔAIC . Larger drop of ΔLHR and increase of AIC indicate more information of the removed system. LHR: Likelihood ratio; AIC: Akaike information criterion; BCLC: Barcelona Clinic Liver Cancer staging; AJCC: American Joint Committee on Cancer; TNM: Tumor node metastasis; CLIP: Cancer of the Liver Italian Program score; CS: Chinese Staging; UNOS: United Network for Organ Sharing; JIS: Japan Integrated Staging score; CUPI: Chinese University Prognostic Index Staging; HKLC: Hong Kong Liver Cancer Staging System.