BJS Open, 2025, zraf041
https://doi.org/10.1093/bjsopen/zraf041
Original Article

α-Fetoprotein model *versus* Milan criteria in predicting outcomes after hepatic resection for hepatocellular carcinoma: multicentre study

Chao Li[1,†], Yong-Kang Diao[1,†], Yi-Fan Li[2,†], Shao-Dong Lv[1,†], Xian-Ming Wang[3], Xue-Dong Wang[4], Qi-Xuan Zheng[5], Hong Wang[6], Han Liu[7], Kong-Ying Lin[8], Ying-Jian Liang[9], Ya-Hao Zhou[10], Wei-Min Gu[11], Ming-Da Wang[1], Lan-Qing Yao[1], Xin-Fei Xu[1], Jia-Hao Xu[1], Li-Hui Gu[1], Timothy M. Pawlik[12], Feng Shen[1,*] and Tian Yang[1,*]

[1]Department of Hepatobiliary Surgery, Eastern Hepatobiliary Surgery Hospital, Naval Medical University, Shanghai, China
[2]Department of Hepatobiliary Surgery, Affiliated Hospital of Nantong University, Nantong, China
[3]Department of General Surgery, First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Jinan, China
[4]Hepatopancreatobiliary Centre, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, Beijing, China
[5]Department of Hepatobiliary Surgery, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, China
[6]Department of General Surgery, Liuyang People's Hospital, Liuyang, China
[7]Department of Hepatobiliary and Pancreatic Surgery, General Surgery Centre, First Hospital of Jilin University, Changchun, China
[8]Department of Hepatobiliary Surgery, Mengchao Hepatobiliary Hospital, Fujian Medical University, Fujian, China
[9]Department of Hepatobiliary Surgery, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China
[10]Department of Hepatobiliary Surgery, Pu'er People's Hospital, Pu'er, China
[11]First Department of General Surgery, Fourth Hospital of Harbin, Harbin, Heilongjiang, China
[12]Department of Surgery, Ohio State University, Wexner Medical Centre, Columbus, Ohio, USA

*Correspondence to: Tian Yang, Department of Hepatobiliary Surgery, Eastern Hepatobiliary Surgery Hospital, Naval Medical University, 225 Changhai Road, Shanghai 200438, China (e-mail: yangtianehbh@smmu.edu.cn); Feng Shen, Department of Hepatobiliary Surgery, Eastern Hepatobiliary Surgery Hospital, Naval Medical University, 225 Changhai Road, Shanghai 200438, China (e-mail: fengshensmmu@gmail.com)

†Joint first authors.

Abstract

Background: The Milan criteria and the French α-fetoprotein (AFP) model have both been validated for predicting outcomes after liver transplantation for hepatocellular carcinoma, with the Milan criteria also used for predicting outcomes after hepatic resection. The aim of this study was to evaluate the AFP model's predictive value for recurrence and survival following hepatocellular carcinoma resection and compare its performance with that of the Milan criteria.

Methods: Data for patients who underwent hepatocellular carcinoma resection between 2002 and 2021 were analysed. For both the AFP model and Milan criteria, patients were divided into two groups: those with hepatocellular carcinoma within and beyond the AFP model (scores ≤ 2 and > 2 points, respectively) and the Milan criteria. Cumulative recurrence and overall survival rates were compared between patients within and beyond the AFP model. Predictions of recurrence and overall survival by the AFP model and Milan criteria were compared using net reclassification improvement and area under the receiver operating characteristic curve analyses.

Results: Among 1968 patients evaluated, 1058 (53.8%) and 940 (47.8%) were classified as beyond on the AFP model and Milan criteria, respectively. After controlling for competing factors on multivariable analyses, being beyond the AFP model was independently associated with recurrence and worse overall survival after resection of hepatocellular carcinoma. Time-dependent net reclassification improvement and area under the receiver operating characteristic curve analyses demonstrated that the AFP model was superior to the Milan criteria in predicting recurrence. Of note, patients who were classified as beyond both the Milan criteria and AFP model had an even higher risk of postoperative recurrence and mortality (hazard ratios 1.51 and 1.47, respectively).

Conclusion: The French AFP model demonstrated superior prognostic accuracy to the Milan criteria in predicting recurrence and survival after hepatocellular carcinoma resection. The AFP model not only effectively stratified patient risk but also identified a subgroup of high-risk patients among those beyond the Milan criteria.

Introduction

Hepatocellular carcinoma (HCC) is a leading cause of cancer-related deaths worldwide[1–4], with hepatic resection often serving as the primary curative treatment due to limited donor organs[5,6]. However, recurrence rates remain high (50–70%) within 5 years after surgery[6–10]. The Milan criteria have been widely accepted for the selection of patients for liver transplantation in HCC[11] because they predict postoperative recurrence following liver transplantation or hepatic resection for HCC[12–19]. However, in real-world clinical practice, using the Milan criteria could be challenging when treating patients with HCC[20,21], not least because of a lack of pretransplant α-fetoprotein (AFP) values[22–24]. Thus, other criteria to select HCC patients for liver transplantation have emerged, including the AFP model, the Hangzhou criteria, and the Metroticket 2.0 model[12–14,20,25,26].

Received: October 23, 2024. **Revised:** January 20, 2025. **Accepted:** February 21, 2025

AFP is the most commonly used HCC biomarker. AFP is used for diagnosis, surveillance, measuring the efficacy of treatment, and prognostic evaluation[27–29]. Among patients with HCC, high serum AFP concentrations have been associated with both an increased risk of recurrence and worse survival after liver transplantation[12–15] and hepatic resection[16–18]. In 2012, a French study group[20] proposed the French AFP model for HCC transplant candidates. This model incorporates tumour burden (the number and size of nodules) and preoperative AFP levels (*Table S1*), with scores over 2 points considered 'beyond' the model and associated with a greater risk of recurrence after liver transplantation[20,30–32]. The performance of the AFP model in predicting recurrence and survival after hepatic resection for HCC has not been thoroughly investigated.

Thus, the aim of the present study was to characterize the utility of the French AFP model in predicting recurrence and survival after hepatic resection for HCC, and to compare it to the Milan criteria.

Methods

Patient selection

This was a multicentre retrospective study conducted on patients who had undergone hepatic resection with curative intent for HCC between 2002 and 2021 at 11 Chinese hospitals. The study was approved by the institutional review board at each participating centre. The inclusion criteria were HCC diagnosis confirmed on resected specimens, curative hepatic resection (complete resection of all macroscopic tumours with microscopically clear resection margins (R0 resection) and no residual tumours on postoperative imaging within 4–6 weeks), no gross/macroscopic vascular invasion, and no previous anti-HCC treatment before resection. The exclusion criteria were age ≤ 18 years, recurrent HCC, ruptured HCC before surgery, palliative hepatic resection (R1) or R2 resection margins, residual tumours on imaging, and missing data for prognostic variables and/or follow-up data. Patients who died within 90 days after surgery or were lost to follow-up within 6 months were also excluded.

Informed consent was obtained from all patients. The study was performed according to the Declaration of Helsinki and the Ethical Guidelines for Clinical Studies of the participating centres.

Data collection

Data were collected on patient age, sex, American Society of Anesthesiologists grade, hepatitis B virus (HBV) status, preoperative Child–Pugh score, cirrhosis, preoperative serum AFP level within 1 week before surgery, largest tumour size, number of tumours, microvascular invasion, satellite nodules, tumour encapsulation, tumour differentiation, intraoperative blood loss, intraoperative blood transfusion, type of resection (anatomical or non-anatomical), extent of hepatectomy (major or minor), and resection margin status. A narrow resection margin was defined as a distance from the resection margin to the tumour edge of less than 1 cm. Tumour differentiation was defined according to the modified Edmondson criteria[33]. Anatomical resections were classified based on the Brisbane 2000 Nomenclature of Liver Anatomy and Resections, whereas non-anatomical resections included wedge resection or limited resection[34]. Major hepatectomy was defined as resection of three or more Couinaud liver segments, whereas minor hepatectomy was defined as resection of fewer than three liver segments[34].

Follow-up

Patients attended follow-up visits at 2-monthly intervals for the first 6 months, at 3-monthly intervals for the next 1.5 years, and then once every 6 months thereafter. Information regarding the patient's history, physical examination, serum AFP level, and imaging results (ultrasonography, contrast-enhanced computed tomography (CT), or magnetic resonance imaging (MRI)) was collected at each follow-up visit. Enhanced CT or MRI, hepatic angiography, bone scans, or positron emission tomography were performed when recurrence or distant metastasis was suspected. Tumour recurrence was defined as the new appearance of intrahepatic or extrahepatic tumour nodule(s) with radiological characteristics similar to those of the primary tumour. Treatment of recurrence included hepatic resection, liver transplantation, local ablation, transcatheter arterial chemoembolization, radiotherapy, systemic therapy, or palliative care, as appropriate. The dates of recurrence, last follow-up, and death were recorded.

Division of patients into groups beyond and within the AFP model and Milan criteria

AFP scores were calculated using the simplified version of the AFP model[20]. Patients with an AFP score of ≤ 2 points were classified as being 'within' the AFP model, and those with a score of > 2 points were classified as being 'beyond' the model. Patients were also classified as being 'within' and 'beyond' the Milan criteria.

Statistical analysis

The primary outcome of this study was cumulative recurrence, calculated from the date of surgery to the date when recurrence was diagnosed, or the date of the last follow-up for patients without recurrence. The secondary outcome was the overall survival (OS) rate, which was calculated from the date of surgery to a patient's death, or the date of the last follow-up for surviving patients.

Clinical characteristics and operative variables are reported as frequencies and percentages for categorical variables and as either mean(standard deviation) or median (interquartile range [IQR]) for continuous variables. Categorical and continuous variables were compared using Chi-square test and the Wilcoxon rank-sum test, respectively. The Kaplan–Meier method was used to estimate cumulative recurrence and OS rates, which were compared between groups using the log rank test. Hazard ratios (HRs) with 95% confidence intervals for recurrence and OS were calculated using univariate and Cox regression multivariable analyses.

The prediction of recurrence was assessed and compared between the AFP model and Milan criteria using the time-dependent net reclassification improvement (NRI) method[35]. The time-dependent area under the receiver operating characteristic curve (AUROC) was used to predict recurrence, as well as to compare the AFP model and Milan criteria[36]. Two-sided $P < 0.050$ was considered statistically significant. Statistical analyses were performed using SPSS® version 25.0 (IBM, Armonk, NY, USA) or R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria).

Results

A flow chart of the study is shown in *Fig. S1*. Of the 1968 patients undergoing HCC resection, 1684 (85.6%) were men and 284 (14.4%) were women. The median age at the time of surgery was 52 years (range 19–86 years). Most patients had chronic HBV

Table 1 Comparisons of clinical characteristics and operative variables between patients within and beyond the AFP model

	Overall ($n = 1968$)	Within AFP model ($n = 910$)	Beyond AFP model ($n = 1058$)	P*
Sex				0.065
Male	1684 (85.6%)	793 (87.1%)	891 (84.2%)	
Female	284 (14.4%)	117 (12.9%)	167 (15.8%)	
Age (years), mean(s.d.)	55 (13)	56 (12)	54 (13)	0.001†
ASA grade > II	236 (12.0%)	99 (10.9%)	137 (12.9%)	0.159
HBV-positive	1760 (89.4%)	808 (88.8%)	952 (90.0%)	0.252
Cirrhosis	1345 (68.3%)	652 (71.6%)	693 (65.5%)	0.003
Child–Pugh grade B	144 (7.3%)	69 (7.6%)	75 (7.1%)	0.675
Preoperative AFP level (ng/ml), median (i.q.r.)	152 (13–923)	16 (9–105)	726 (156–6529)	< 0.001†
≤ 100	885 (45.0%)	680 (74.7%)	205 (19.4%)	< 0.001
101–1000	600 (30.5%)	230 (25.3%)	370 (35.0%)	
> 1000	483 (24.5%)	0 (0%)	483 (45.7%)	
No. of nodules, median (i.q.r.)	1 (1–1)	1 (1–1)	1 (1–1)	< 0.001†
1–3	1916 (97.4%)	909 (99.9%)	1007 (95.2%)	< 0.001
> 3	52 (2.6%)	1 (0.1%)	51 (4.8%)	
Largest tumour size (cm), median (i.q.r.)	4.0 (3.0–6.5)	3.0 (2.0–4.0)	6.2 (4.0–9.2)	< 0.001†
≤ 3	646 (32.8%)	533 (58.6%)	113 (10.7%)	< 0.001
3–6	786 (39.9%)	377 (41.4%)	409 (38.7%)	
> 6	536 (27.2%)	0 (0%)	536 (50.7%)	
Tumour extent				< 0.001
Within Milan criteria	1028 (52.2%)	725 (79.7%)	303 (28.6%)	
Beyond Milan criteria	940 (47.8%)	185 (20.3%)	755 (71.4%)	
Microvascular invasion	628 (31.9%)	191 (21.0%)	437 (41.3%)	< 0.001
Satellite nodules	294 (14.9%)	80 (8.8%)	214 (20.2%)	< 0.001
Incomplete tumour encapsulation	1102 (56.0%)	433 (47.6%)	669 (63.2%)	< 0.001
Poor tumour differentiation	1134 (57.6%)	418 (45.9%)	716 (67.7%)	< 0.001
Intraoperative blood loss (ml), median (i.q.r.)	300 (150–500)	200 (100–400)	300 (200–550)	< 0.001†
Intraoperative blood transfusion	273 (13.9%)	77 (8.5%)	196 (18.5%)	< 0.001
Non-anatomical resection	706 (35.9%)	323 (35.5%)	383 (36.2%)	0.745
Major hepatectomy	350 (17.8%)	57 (6.3%)	293 (27.7%)	< 0.001
Narrow resection margin (<1 cm)	972 (49.4%)	364 (40.0%)	608 (57.5%)	< 0.001

AFP, α-fetoprotein; s.d., standard deviation; ASA, American Society of Anesthesiologists; HBV, hepatitis B virus; i.q.r., interquartile range; *Chi-square test, except †Wilcoxon rank-sum test.

infection (1760, 89.4%), and cirrhosis (1345, 68.3%). Clinical and operative variables differed significantly between patients beyond (score > 2 points) and within (score ≤ 2 points) the AFP model (*Table 1*). Patients beyond the AFP model were younger and had less cirrhosis, higher AFP levels, larger tumours, and more aggressive tumour characteristics, including multiple tumours, satellite nodules, poor differentiation, and incomplete encapsulation, than those within the AFP model (all $P < 0.001$). The group beyond the AFP model had greater blood loss, higher transfusion rates, and a greater percentage of patients undergoing major hepatectomy and with narrow resection margins (all $P < 0.001$).

At a median follow-up of 54.3 months, 964 patients (49.0%) had developed HCC recurrence (intrahepatic: 703, 35.7%; extrahepatic: 84, 4.3%; both: 177, 9.0%). Among the patients with recurrence, 902 died (45.8%), with 676 cancer-related deaths (34.3%). For the overall cohort, the 1-, 3-, and 5-year recurrence rates were 18.9, 37.5, and 46.3%, respectively, and OS rates at 1, 3, and 5 years were 93.4, 75.9, and 64.0%, respectively.

Recurrence and OS in patients beyond and within the AFP model

As indicated in *Table 2*, the overall incidence of recurrence and death was higher among patients beyond (score > 2 points) than within (score ≤ 2 points) the AFP model (57.8 *versus* 38.8% for recurrence; and 55.0 *versus* 35.2% for death; both $P < 0.001$). There was no significant difference in the incidence of non-cancer-specific deaths between patients beyond and within the AFP model (11.6 *versus* 11.3%, respectively; $P = 0.831$). The incidence of cancer-specific deaths was markedly higher in the group beyond than within the AFP model (43.4 *versus* 23.8%; $P < 0.001$).

Comparisons of cumulative recurrence and OS between patients beyond and within the AFP model are shown in *Fig. 1*. The 1-, 3-, and 5-year recurrence rates among patients beyond the AFP model scores were 26.1, 47.9, and 57.4%, respectively, much higher than corresponding rates for patients within the AFP model (10.6, 25.8, and 33.8%, respectively; HR 2.01, 95% confidence interval (c.i.) 1.77 to 2.30; $P < 0.001$). OS rates at 1, 3, and 5 years among patients beyond the AFP model were 90.3, 66.7, and 52.7%, respectively, which were worse than for patients within the AFP model (96.9, 86.4, and 76.9%, respectively; HR 2.02, 1.77 to 2.32; $P < 0.001$).

Recurrence and OS in patients beyond and within the Milan criteria

Clinical characteristics, operative variables, and long-term recurrence and survival between patients within and beyond the Milan criteria were compared (*Tables S2* and *S3*). The differences between the two groups were similar those seen for patients within and beyond the AFP model. The 1-, 3-, and 5-year recurrence rates among patients beyond the Milan criteria were 27.5, 49.6, and 58.7%, respectively, which were higher than those for patients within the Milan criteria (11.1, 26.7, and 35.2%, respectively; HR 2.00, 95% c.i. 1.75 to 2.27; $P < 0.001$). OS rates at 1, 3, and 5 years among patients beyond the Milan criteria were 89.1, 65.4, and 51.9%, respectively, and worse than corresponding values among patients within the Milan criteria (97.3, 85.4, and 75.0%, respectively; HR 1.97, 1.72 to 2.25; $P < 0.001$) (*Fig. 2*).

Table 2 Comparisons of long-term outcomes between patients within and beyond the AFP model

	Overall ($n = 1968$)	Within AFP model ($n = 910$)	Beyond AFP model ($n = 1058$)	P*
Recurrence during follow-up	964 (49.0%)	353 (38.8%)	611 (57.8%)	< 0.001
Intrahepatic only	703 (35.7%)	280 (30.8%)	423 (40.0%)	< 0.001
Extrahepatic only	84 (4.3%)	18 (2.0%)	66 (6.2%)	< 0.001
Intrahepatic and extrahepatic	177 (9.0%)	55 (6.0%)	122 (11.5%)	< 0.001
Death during follow-up	902 (45.8%)	320 (35.2%)	582 (55.0%)	< 0.001
Cancer-specific	676 (34.3%)	217 (23.8%)	459 (43.4%)	< 0.001
Non-cancer-specific	226 (11.5%)	103 (11.3%)	123 (11.6%)	0.831
Time to recurrence (months), median (95% c.i.)	74.3 (65.9, 82.7)	109.1 (89.6, 128.6)	39.3 (33.8, 44.8)	< 0.001†
1-year recurrence rate (%)	18.9	10.6	26.1	
3-year recurrence rate (%)	37.5	25.8	47.9	
5-year recurrence rate (%)	46.3	33.8	57.4	
OS (months), median (95% c.i.)	96.6 (89.4, 103.8)	119.2 (110.9, 127.5)	65.5 (58.2, 72.7)	< 0.001†
1-year OS rate (%)	93.4	96.9	90.3	
3-year OS rate (%)	75.9	86.4	66.7	
5-year OS rate (%)	64.0	76.9	52.7	

AFP, α-fetoprotein; c.i., confidence interval; OS, overall survival. *Chi-square test, except †log rank test.

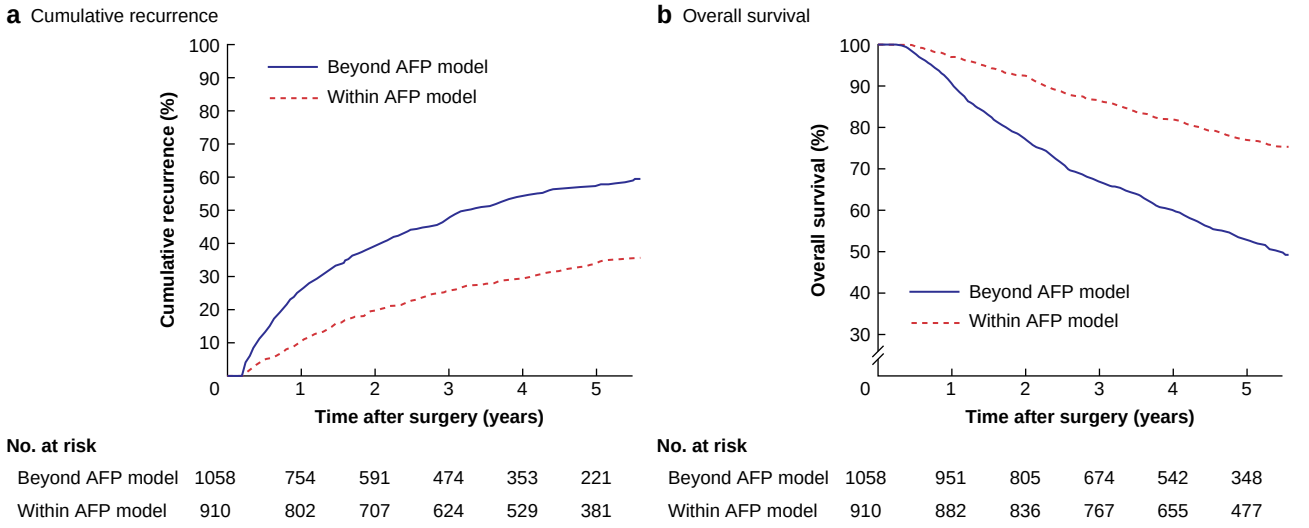

Fig. 1 Kaplan–Meier curves of cumulative recurrence and overall survival after hepatic resection for hepatocellular carcinoma after hepatic resection for hepatocellular carcinoma in patients within and beyond the AFP model

a Cumulative recurrence and **b** overall survival. AFP, α-fetoprotein. **a**,**b** $P < 0.001$ (log rank test).

Univariable and multivariable analyses in predicting recurrence and OS

Univariable and multivariable Cox regression analysis identified risk factors associated with postoperative recurrence after hepatic resection for HCC (*Table 3*). Multivariable regression models examining the AFP model and Milan criteria separately that included variables comprising each assessment tool revealed that both the AFP model and Milan criteria were independently associated with increased risks of recurrence (AFP model: HR 1.62, 95% c.i. 1.39 to 1.88; Milan criteria: HR 1.43, 1.23 to 1.67). Similar results were found regarding OS (AFP model: HR 1.54, 1.30 to 1.83; Milan criteria: HR 1.52, 1.28 to 1.81) (*Table 4*).

NRI and AUROC values between the AFP model and Milan criteria

Compared with the Milan criteria, the AFP model had improved predictive power to risk classify 1-, 2-, 3-, 4-, and 5-year recurrence (NRI 0.023–0.071; 1- and 2-year recurrence, $P < 0.050$; and 3-, 4-, and 5-year recurrence, $P > 0.050$). The plot for time-dependent NRI showed that the AFP model maintained better predictive power than the Milan criteria for the first 5 years of follow-up. Of note, the AFP model was particularly better than the Milan criteria in predicting HCC recurrence with the first 2 years after surgery (*Fig. 3a*). Time-dependent AUROC curves for the AFP model and Milan criteria showed that the AFP model maintained higher AUROC values than the Milan criteria at 1, 2, 3, 4, and 5 years (0.669 *versus* 0.644, 0.655 *versus* 0.638, 0.639 *versus* 0.626, 0.631 *versus* 0.624, and 0.629 *versus* 0.622, respectively) (*Fig. 3b*).

Subgroup analysis of patients with HCC beyond the Milan criteria

Of 940 patients who were beyond the Milan criteria, 185 (19.7%) and 755 (80.3%) were within and beyond the AFP model, respectively. *Figure S2* shows that patients who were beyond both the Milan criteria and the AFP model had significantly higher 1-, 3-, and 5-year recurrence rates than those who were beyond the Milan criteria but within the AFP model (29.7, 52.8, and 62.0% *versus* 18.7, 36.6, and 45.6%, respectively; HR 1.62,

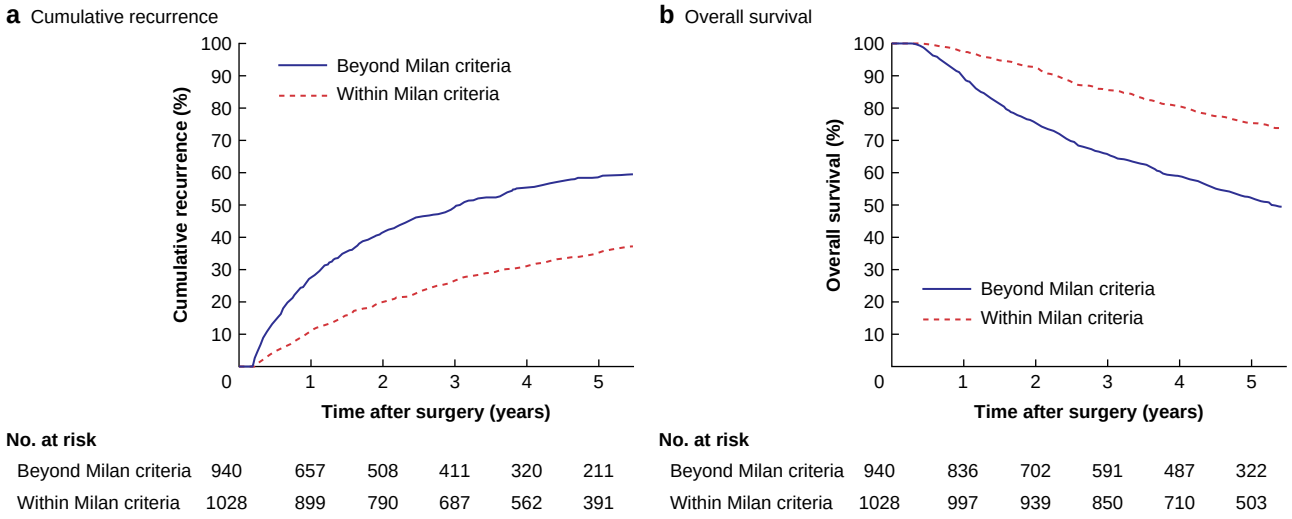

Fig. 2 Kaplan–Meier curves of cumulative recurrence and overall survival after hepatic resection for hepatocellular carcinoma in patients within and beyond the Milan criteria

a Cumulative recurrence and **b** overall survival. **a**,**b** $P < 0.001$ (log rank test).

Table 3 Univariable and multivariable Cox regression analysis of risk factors associated with recurrence

	Univariable analysis		Multivariable analysis*		Multivariable analysis†	
	HR	P	HR	P	HR	P
Male sex	1.04 (0.87, 1.25)	0.672				
Age > 60 years	0.76 (0.66, 0.88)	< 0.001	n.a.	0.848	n.a.	0.591
ASA grade > II	1.01 (0.83, 1.23)	0.940				
HBV-positive	1.53 (1.31, 1.79)	< 0.001	n.a.	0.791	n.a.	0.943
Cirrhosis	1.44 (1.25, 1.66)	< 0.001	1.23 (1.05 1.45)	0.013	n.a.	0.051
Child–Pugh grade B	1.34 (1.06, 1.69)	0.015	n.a.	0.095	n.a.	0.162
Preoperative AFP level (ng/ml)†						
≤ 100	1.00 (reference)				1.00 (reference)	
100–1000	1.95 (1.68, 2.27)	< 0.001			1.90 (1.62, 2.24)	< 0.001
> 1000	1.79 (1.45, 2.10)	< 0.001			1.52 (1.28, 1.82)	< 0.001
Beyond Milan criteria†	1.99 (1.75, 2.27)	< 0.001			1.43 (1.23, 1.67)	< 0.001
Beyond AFP model*	2.01 (1.77, 2.30)	< 0.001	1.62 (1.39, 1.88)	< 0.001		
Microvascular invasion	2.29 (2.02, 2.60)	< 0.001	1.48 (1.26, 1.73)	< 0.001	1.53 (1.31, 1.78)	< 0.001
Satellite nodules	2.02 (1.73, 2.36)	< 0.001	1.71 (1.43, 2.05)	< 0.001	1.59 (1.33, 1.91)	< 0.001
Incomplete tumour encapsulation	1.38 (1.22, 1.58)	< 0.001	1.25 (1.08, 1.46)	0.003	1.28 (1.10, 1.49)	0.002
Poor tumour differentiation	1.28 (1.13, 1.46)	< 0.001	n.a.	0.682	n.a.	0.345
Intraoperative blood loss > 600 ml	1.53 (1.28, 1.83)	< 0.001	n.a.	0.699	n.a.	0.875
Intraoperative blood transfusion	1.71 (1.45, 2.02)	< 0.001	1.41 (1.18, 1.69)	< 0.001	1.41 (1.18, 1.69)	< 0.001
Non-anatomical resection	1.45 (1.26, 1.67)	< 0.001	1.26 (1.06, 1.49)	0.007	1.26 (1.07, 1.50)	0.007
Major hepatectomy	1.37 (1.17, 1.61)	< 0.001	n.a.	0.917	n.a.	0.956
Narrow resection margin (< 1 cm)	1.56 (1.37, 1.77)	< 0.001	1.64 (1.43, 1.90)	< 0.001	1.58 (1.37, 1.83)	< 0.001

Values in parentheses are 95% confidence intervals. *'Beyond α-fetoprotein (AFP) model' variable was entered into the multivariable analysis. †'Preoperative AFP level' and 'beyond Milan criteria' variables were entered into the multivariate analysis. HR, hazard ratio; n.a., not available; ASA, American Society of Anesthesiologists; HBV, hepatitis B virus.

95% c.i. 1.29 to 2.04; $P < 0.001$). In addition, 1-, 3-, and 5-year OS rates were worse in the former than latter group (88.9, 62.3, and 48.1% *versus* 90.3, 77.8, and 67.0%, respectively; HR 1.60, 1.26 to 2.03; $P < 0.001$).

The results of univariate and multivariable Cox regression analyses identifying risk factors associated with recurrence and OS after hepatic resection for HCC beyond the Milan criteria are presented in *Tables* S4 and S5. After hepatic resection, patients who were beyond the Milan criteria and beyond the AFP model exhibited increased recurrence (HR 1.51, 1.17 to 1.94; $P = 0.002$) and worse OS (HR 1.47, 1.03 to 1.93; $P = 0.031$) than those beyond the Milan criteria but within the AFP model.

Discussion

This multicentre study demonstrated that the AFP model outperforms the Milan criteria in predicting recurrence and survival after hepatic resection for HCC. Key findings include superior accuracy of the AFP model, especially for early recurrence within 2 years, and effective risk stratification, identifying high-risk subgroups among patients beyond the Milan criteria and poorer outcomes for patients exceeding both the AFP model and Milan criteria. The AFP model's inclusion of serum levels of AFP, a biomarker of the aggressiveness of HCC, provides a more nuanced evaluation of tumour biology compared with the Milan criteria, which are based on tumour

Table 4 Univariable and multivariable Cox regression analysis of risk factors associated with overall survival

	Univariable		Multivariable*		Multivariable†	
	HR	P	HR	P	HR	P
Male sex	1.23 (1.03, 1.47)	0.021	n.a.	0.351	n.a.	0.478
Age > 60 years	1.30 (1.14, 1.49)	< 0.001	n.a.	0.066	n.a.	0.068
ASA grade > II	1.27 (1.05, 1.53)	0.016	n.a.	0.247	n.a.	0.314
HBV-positive	1.30 (1.13, 1.49)	< 0.001	n.a.	0.256	n.a.	0.284
Cirrhosis	1.14 (0.99, 1.31)	0.068	1.33 (1.10, 1.61)	0.003	1.31 (1.08, 1.58)	0.005
Child–Pugh grade B	1.94 (1.57, 2.39)	< 0.001	1.87 (1.47, 2.37)	< 0.001	1.82 (1.43, 2.31)	< 0.001
Preoperative AFP level (ng/ml)†						
≤ 100	1.00 (reference)				1.00 (reference)	
100–1000	1.96 (1.68, 2.29)	< 0.001			1.68 (1.40, 2.02)	< 0.001
> 1000	1.74 (1.47, 2.05)	< 0.001			1.36 (1.11, 1.66)	0.003
Beyond Milan criteria†	1.97 (1.72, 2.25)	< 0.001			1.52 (1.28, 1.81)	< 0.001
Beyond AFP model*	2.02 (1.77, 2.32)	< 0.001	1.54 (1.30, 1.83)	< 0.001		
Microvascular invasion	1.82 (1.59, 2.07)	< 0.001	1.43 (1.20, 1.70)	< 0.001	1.47 (1.24, 1.74)	< 0.001
Satellite nodules	1.95 (1.66, 2.28)	< 0.001	1.63 (1.34, 1.97)	< 0.001	1.51 (1.24, 1.83)	< 0.001
Incomplete tumour encapsulation	1.79 (1.56, 2.06)	< 0.001	1.36 (1.15, 1.61)	< 0.001	1.37 (1.16, 1.63)	< 0.001
Poor tumour differentiation	1.34 (1.17, 1.54)	< 0.001	n.a.	0.318	n.a.	0.137
Intraoperative blood loss > 600 ml	1.74 (1.44, 2.11)	< 0.001	n.a.	0.950	n.a.	0.759
Intraoperative blood transfusion	1.92 (1.63, 2.26)	< 0.001	1.54 (1.27, 1.86)	< 0.001	1.50 (1.24, 1.82)	1.499
Non-anatomical resection	1.01 (0.88, 1.15)	0.912				
Major hepatectomy	1.53 (1.31, 1.79)	< 0.001	n.a.	0.747	n.a.	0.506
Narrow resection margin (< 1 cm)	1.76 (1.53, 2.01)	< 0.001	1.97 (1.68, 2.31)	< 0.001	1.86 (1.58, 2.19)	< 0.001

Values in parentheses are 95% confidence intervals. *'Beyond α-fetoprotein (AFP) model' variable was entered into the multivariable analysis. †'Preoperative AFP level' and 'Beyond Milan criteria' variables were entered into the multivariable analysis. HR, hazard ratio; n.a., not available; ASA, American Society of Anesthesiologists; HBV, hepatitis B virus.

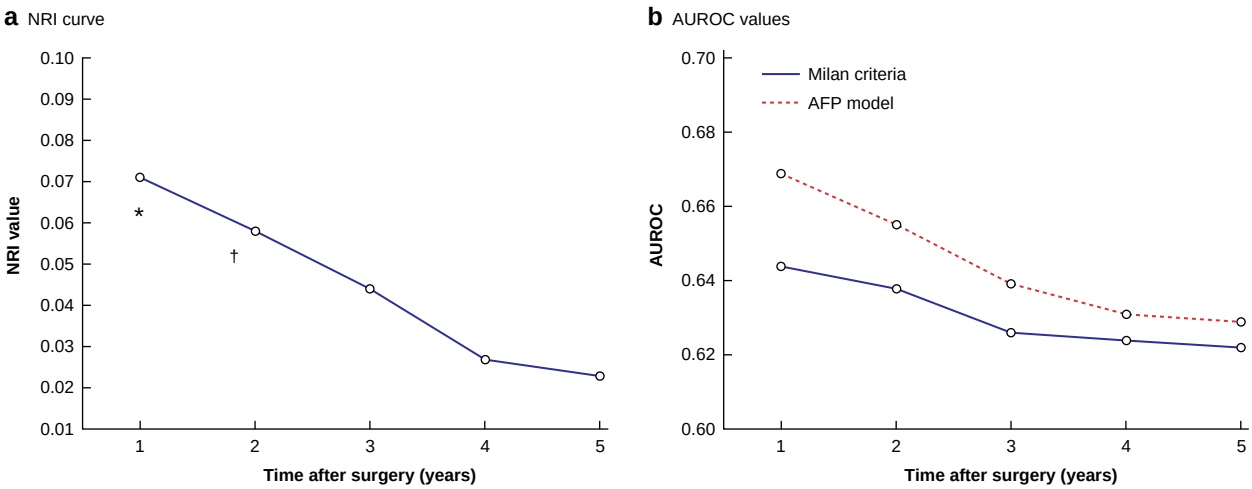

Fig. 3 Comparisons between the AFP model and Milan criteria

a Time-dependent net reclassification improvement (NRI) curve for predicting recurrence. Points were calculated by subtracting Milan criteria values from α-fetoprotein (AFP) model values. Patients censored before the endpoints for analysis were excluded. **b** Time-dependent area under the receiver operating characteristic curve (AUROC) values for predicting recurrence for the AFP model and Milan criteria. *$P = 0.003$, †$P = 0.035$ (log rank test).

burden. The results of the present study validate the utility of the AFP model in hepatic resection and align with previous evidence from liver transplantation studies[20,30–32].

The AFP model has the advantage of being a simple, widely accepted scoring system. Although novel tools incorporating artificial intelligence or nomograms are promising[37], they often lack sufficient validation for clinical application. This study emphasizes the importance of validating established models like the AFP model, rather than developing new ones, to enhance their adoption in practice. Time-dependent AUROC analysis showed that both the AFP model and Milan criteria were more predictive of early recurrence within 2 years after surgery than late recurrence, in which tumour biology differs. Early recurrence is more likely the result of occult metastasis from the primary tumour and is more likely to be aggressive, whereas late recurrence more likely represents different clonal origins[38–40]. Subgroup analysis highlighted the poor outcomes for patients exceeding both the AFP model and Milan criteria, with > 60% recurrence and < 50% survival rates at 5 years, compared with more favourable outcomes for patients within AFP model thresholds. These findings underscore the potential of the AFP model to refine risk stratification and guide adjuvant therapy trials for high-risk subgroups. Patients identified as being at high risk by both models may benefit from more intensive surveillance protocols and could be prioritized for adjuvant therapy trials[41–45]. The AFP model has several clinical applications, including aiding preoperative decision-making, personalizing postoperative surveillance, and selecting candidates for adjuvant therapies[43,44,46,47].

The present study has several limitations that warrant discussion. Its retrospective nature introduced potential selection bias and limited the control for all confounding factors. The predominance of HBV-related HCC in this cohort, reflecting the epidemiology of HCC in China, may limit the generalizability of these findings (such as hepatitis C virus, alcohol, non-alcoholic steatohepatitis)[48–50]. Further validation in diverse patient populations is necessary. Although the multicentre design of the enhances the generalizability of the results, the protocols across different centres could not be entirely standardized, which introduced potential variations in surgical techniques and perioperative management. Molecular markers or gene expression profiles were not included in this study, which could provide additional prognostic information beyond clinical and pathological factors[10]. Advancements in surgical techniques and adjuvant therapies may have influenced the results during the long follow-up in this study. In a subsequent study, comprehensive analysis will be conducted to explore temporal trends and potential variations in the clinical characteristics and management strategies for HCC. In addition to the Milan criteria, other predictive models and standards (such as the up-to-7 criteria) have significant clinical value. The Milan criteria remain the most widely recognized and used staging scoring system for liver transplantation worldwide. Although the present study demonstrated the prognostic value of the AFP model, it did not directly prove that its use improves patient outcomes; prospective studies evaluating the impact of AFP model-guided decision-making on long-term outcomes are needed.

Funding

This study was supported by grants from the National Natural Science Foundation of China (82425049, 82372813 and 82273074), Dawn Project Foundation of Shanghai (21SG36), Shanghai Health and Hygiene Discipline Leader Project (2022XD001), Shanghai Outstanding Academic Leader Programme (23XD1424900), the Natural Science Foundation of Shanghai (22ZR1477900), Shanghai Science and Technology Committee Rising-Star Programme (22QA1411600), Special Clinical Project of Shanghai Municipal Health Commission (20224Y0299), and National Natural Science Foundation Incubation Project (2022MS039 and 2022GZR005).

Author contributions

Chao Li (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Writing—original draft), Yong-Kang Diao (Conceptualization, Data curation, Formal analysis, Methodology, Software), Yi-Fan Li (Conceptualization, Formal analysis, Investigation, Validation, Writing—original draft), Shao-Dong Lv (Conceptualization, Writing—review & editing), Xian-Ming Wang (Conceptualization, Formal analysis, Project administration, Resources, Visualization, Writing—original draft), Xue-Dong Wang (Data curation, Project administration, Resources, Visualization), Qi-Xuan Zheng (Data curation, Investigation, Resources, Validation), Hong Wang (Data curation, Project administration, Validation), Han Liu (Data curation, Methodology, Resources, Validation), Kong-Ying Lin (Data curation, Methodology, Resources, Supervision), Ying-Jian Liang (Data curation, Project administration, Resources, Visualization), Ya-Hao Zhou (Data curation, Investigation, Resources, Software, Validation), Wei-Min Gu (Data curation, Investigation, Resources, Validation), Ming-Da Wang (Conceptualization, Data curation, Formal analysis, Supervision, Visualization), Lan-Qing Yao (Data curation, Project administration, Software, Validation), Xin-Fei Xu (Data curation, Investigation, Software, Validation), Jia-Hao Xu (Data curation, Investigation, Resources, Software, Validation), Li-Hui Gu (Data curation, Investigation, Methodology, Validation), Timothy M. Pawlik (Formal analysis, Supervision, Writing—review & editing), Feng Shen (Conceptualization, Formal analysis, Supervision, Writing—review & editing), and Tian Yang (Conceptualization, Data curation, Funding acquisition, Project administration, Supervision, Writing—review & editing)

Disclosure

The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS Open* online.

Data availability

The data set for this study is available upon request by contacting the corresponding author. The authors confirm that the data supporting the findings of this study are available within the article and its *Supplementary material*.

References

1. 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**:394–424
2. Akinyemiju T, Abera S, Ahmed M, Alam N, Alemayohu MA, Allen C *et al.* The burden of primary liver cancer and underlying etiologies from 1990 to 2015 at the global, regional, and national level: results from the global burden of disease study 2015. *JAMA Oncol* 2017;**3**:1683–1691
3. Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: new trends. *J Hepatol* 2020;**72**:250–261
4. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S *et al.* Hepatocellular carcinoma. *Nat Rev Dis Primers* 2021;**7**:6
5. Knox JJ, Cleary SP, Dawson LA. Localized and systemic approaches to treating hepatocellular carcinoma. *J Clin Oncol* 2015;**33**:1835–1844
6. Pinna AD, Yang T, Mazzaferro V, De Carlis L, Zhou J, Roayaie S *et al.* Liver transplantation and hepatic resection can achieve cure for hepatocellular carcinoma. *Ann Surg* 2018;**268**:868–875
7. Makuuchi M, Takayama T, Kubota K, Kimura W, Midorikawa Y, Miyagawa S *et al.* Hepatic resection for hepatocellular carcinoma—Japanese experience. *Hepatogastroenterology* 1998;**45**:1267–1274
8. Grazi GL, Ercolani G, Pierangeli F, Del Gaudio M, Cescon M, Cavallari A *et al.* Improved results of liver resection for hepatocellular carcinoma on cirrhosis give the procedure added value. *Ann Surg* 2001;**234**:71–78
9. Wang JH, Wang CC, Hung CH, Chen CL, Lu SN. Survival comparison between surgical resection and radiofrequency ablation for patients in BCLC very early/early stage hepatocellular carcinoma. *J Hepatol* 2012;**56**:412–418

10. Wang M, Diao Y, Yao L, Fan Z, Wang K, Wu H *et al.* Emerging role of molecular diagnosis and personalized therapy for hepatocellular carcinoma. *iLIVER* 2024;**3**:100083
11. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F *et al.* Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* 1996;**334**:693–699
12. Toso C, Meeberg G, Hernandez-Alejandro R, Dufour JF, Marotta P, Majno P *et al.* Total tumor volume and alpha-fetoprotein for selection of transplant candidates with hepatocellular carcinoma: a prospective validation. *Hepatology* 2015;**62**:158–165
13. Mazzaferro V, Sposito C, Zhou J, Pinna AD, De Carlis L, Fan J *et al.* Metroticket 2.0 model for analysis of competing risks of death after liver transplantation for hepatocellular carcinoma. *Gastroenterology* 2018;**154**:128–139
14. Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L *et al.* Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol* 2009;**10**:35–43
15. Vibert E, Azoulay D, Hoti E, Iacopinelli S, Samuel D, Salloum C *et al.* Progression of alphafetoprotein before liver transplantation for hepatocellular carcinoma in cirrhotic patients: a critical factor. *Am J Transplant* 2010;**10**:129–137
16. Yang SL, Liu LP, Yang S, Liu L, Ren JW, Fang X *et al.* Preoperative serum α-fetoprotein and prognosis after hepatectomy for hepatocellular carcinoma. *Br J Surg* 2016;**103**:716–724
17. Ma WJ, Wang HY, Teng LS. Correlation analysis of preoperative serum alpha-fetoprotein (AFP) level and prognosis of hepatocellular carcinoma (HCC) after hepatectomy. *World J Surg Oncol* 2013;**11**:212
18. Kim H, Lee SJ, Yoon M. Alpha-fetoprotein is correlated with intrahepatic recurrence of hepatocellular carcinoma after a hepatectomy. *Ann Surg Treat Res* 2020;**98**:168–176
19. Rahbari NN, Mehrabi A, Mollberg NM, Müller SA, Koch M, Büchler MW *et al.* Hepatocellular carcinoma: current management and perspectives for the future. *Ann Surg* 2011;**253**:453–469
20. Duvoux C, Roudot-Thoraval F, Decaens T, Pessione F, Badran H, Piardi T *et al.* Liver transplantation for hepatocellular carcinoma: a model including α-fetoprotein improves the performance of Milan criteria. *Gastroenterology* 2012;**143**:986–994.e3
21. Yao LQ, Fan ZQ, Wang MD, Diao YK, Chen TH, Zeng YY *et al.* Prognostic value of serum α-fetoprotein level as an important characteristic of tumor biology for patients undergoing liver resection of early-stage hepatocellular carcinoma (BCLC stage 0/A): a large multicenter analysis. *Ann Surg Oncol* 2024;**31**:1219–1231
22. Kim SJ, Kim JM. Prediction models of hepatocellular carcinoma recurrence after liver transplantation: a comprehensive review. *Clin Mol Hepatol* 2022;**28**:739–753
23. Mazzaferro V, Chun YS, Poon RT, Schwartz ME, Yao FY, Marsh JW *et al.* Liver transplantation for hepatocellular carcinoma. *Ann Surg Oncol* 2008;**15**:1001–1007
24. Pavel MC, Fuster J. Expansion of the hepatocellular carcinoma Milan criteria in liver transplantation: future directions. *World J Gastroenterol* 2018;**24**:3626–3636
25. Xu X, Lu D, Ling Q, Wei X, Wu J, Zhou L *et al.* Liver transplantation for hepatocellular carcinoma beyond the Milan criteria. *Gut* 2016;**65**:1035–1041
26. Vitale A, Cucchetti A, Qiao GL, Cescon M, Li J, Ramirez Morales R *et al.* Is resectable hepatocellular carcinoma a contraindication to liver transplantation? A novel decision model based on 'number of patients needed to transplant' as measure of transplant benefit. *J Hepatol* 2014;**60**:1165–1171
27. Trevisani F, Garuti F, Neri A. Alpha-fetoprotein for diagnosis, prognosis, and transplant selection. *Semin Liver Dis* 2019;**39**:163–177
28. Guan MC, Ouyang W, Wang MD, Liang L, Li N, Fu T-T *et al.* Biomarkers for hepatocellular carcinoma based on body fluids and feces. *World J Gastrointest Oncol* 2021;**13**:351–365
29. Mitsuhashi N, Kobayashi S, Doki T, Kimura F, Shimizu H, Yoshidome H *et al.* Clinical significance of alpha-fetoprotein: involvement in proliferation, angiogenesis, and apoptosis of hepatocellular carcinoma. *J Gastroenterol Hepatol* 2008;**23**:e189–e197
30. Piñero F, Tisi Baña M, de Ataide EC, Hoyos Duque S, Marciano S, Varón A *et al.* Liver transplantation for hepatocellular carcinoma: evaluation of the alpha-fetoprotein model in a multicenter cohort from Latin America. *Liver Int* 2016;**36**:1657–1667
31. Notarpaolo A, Layese R, Magistri P, Gambato M, Colledan M, Magini G *et al.* Validation of the AFP model as a predictor of HCC recurrence in patients with viral hepatitis-related cirrhosis who had received a liver transplant for HCC. *J Hepatol* 2017;**66**:552–559
32. Ren A, Li Z, Zhou X, Zhang X, Huang X, Deng R *et al.* Evaluation of the alpha-fetoprotein model for predicting recurrence and survival in patients with hepatitis B virus (HBV)-related cirrhosis who received liver transplantation for hepatocellular carcinoma. *Front Surg* 2020;**7**:52
33. Edmondson HA, Steiner PE. Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. *Cancer* 1954;**7**:462–503
34. Strasberg SM, Phillips C. Use and dissemination of the Brisbane 2000 Nomenclature of Liver Anatomy and Resections. *Ann Surg* 2013;**257**:377–382
35. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–172
36. Bandos AI, Rockette HE, Song T, Gur D. Area under the free-response ROC curve (FROC) and a related summary index. *Biometrics* 2009;**65**:247–256
37. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;**350**:g7594
38. Xu XF, Xing H, Han J, Li ZL, Lau WY, Zhou YH *et al.* Risk factors, patterns, and outcomes of late recurrence after liver resection for hepatocellular carcinoma: a multicenter study from China. *JAMA Surg* 2019;**154**:209–217
39. Yan WT, Li C, Yao LQ, Qiu HB, Wang MD, Xu XF *et al.* Predictors and long-term prognosis of early and late recurrence for patients undergoing hepatic resection of hepatocellular carcinoma: a large-scale multicenter study. *Hepatobiliary Surg Nutr* 2023;**12**:155–168
40. Wang MD, Li C, Liang L, Xing H, Sun LY, Quan B *et al.* Early and late recurrence of hepatitis B virus-associated hepatocellular carcinoma. *Oncologist* 2020;**25**:e1541–e1551
41. Zhu XD, Li KS, Sun HC. Adjuvant therapies after curative treatments for hepatocellular carcinoma: current status and prospects. *Genes Dis* 2020;**7**:359–369
42. Brown ZJ, Greten TF, Heinrich B. Adjuvant treatment of hepatocellular carcinoma: prospect of immunotherapy. *Hepatology* 2019;**70**:1437–1442
43. Pinato DJ, Fessas P, Sapisochin G, Marron TU. Perspectives on the neoadjuvant use of immunotherapy in hepatocellular carcinoma. *Hepatology* 2021;**74**:483–490
44. Akateh C, Black SM, Conteh L, Miller ED, Noonan A, Elliott E *et al.* Neoadjuvant and adjuvant treatment strategies for hepatocellular carcinoma. *World J Gastroenterol* 2019;**25**:3704–3721

45. Liang L, Li C, Wang MD, Wang H, Zhou Y-H, Zeng Y-Y *et al.* Development and validation of a novel online calculator for estimating survival benefit of adjuvant transcatheter arterial chemoembolization in patients undergoing surgery for hepatocellular carcinoma. *J Hematol Oncol* 2021;**14**:165
46. Xu X, Wang MD, Xu JH, Fan ZQ, Diao YK, Chen Z *et al.* Adjuvant immunotherapy improves recurrence-free and overall survival following surgical resection for intermediate/advanced hepatocellular carcinoma: a multicenter propensity matching analysis. *Front Immunol* 2023;**14**:1322233
47. Chen WY, Li C, Liu ZP, Kong QY, Sun LY, Zeng YY *et al.* Novel online calculator to predict reduced risk of early recurrence from adjuvant transarterial chemoembolisation for patients with hepatocellular carcinoma. *eGastroenterology* 2023;**1**: e100008
48. Sun LY, Wang NY, Diao YK, Yan CL, Fan ZP, Wei LH *et al.* Comparison between models for detecting hepatocellular carcinoma in patients with chronic liver diseases of various etiologies: ASAP score *versus* GALAD score. *Hepatobiliary Pancreat Dis Int* 2023; DOI: 10.1016/j.hbpd.2023.12.004 [Epub ahead of print]
49. Yang T, Wang MD, Xu XF, Li C, Wu H, Shen F. Management of hepatocellular carcinoma in China: seeking common grounds while reserving differences. *Clin Mol Hepatol* 2023;**29**:342–344
50. Yang T, Lu JH, Wu MC. Hepatocellular carcinoma in China. *BMJ* 2010;**340**:c1026