

Keywords: hepatocellular carcinoma; HCC; intermediate-stage; hepatic resection; NSP score; Child–Pugh A

Intermediate-stage hepatocellular carcinoma treated with hepatic resection: the NSP score as an aid to decision-making

Yong-Fa Zhang^{1,2,3,7}, Jie Zhou^{4,7}, Wei Wei^{1,2,3}, Ru-Hai Zou^{2,3,5}, Min-Shan Chen^{1,2,3}, Wan Yee Lau^{1,6}, Ming Shi^{*1,2,3} and Rong-Ping Guo^{*1,2,3}

¹The Department of Hepatobiliary Oncology of Sun Yat-sen University Cancer Center, Guangzhou 510060, China; ²State Key Laboratory of Oncology in South China, Guangzhou 510060, China; ³Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; ⁴Department of Hepatobiliary Surgery, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China; ⁵Department of Ultrasonography of Sun Yat-sen University Cancer Center, Guangzhou 510515, China and ⁶Department of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong, China

Background: The subgroups of patients with intermediate-stage (BCLC-B) hepatocellular carcinoma (HCC) who would truly benefit from hepatic resection (HR) are unknown. An objective point score was established to guide the selection of these patients for HR.

Methods: In all, 255 consecutive patients with intermediate-stage HCC treated with HR were evaluated retrospectively and included in this study (the training cohort). The variables on overall survival (OS, log-rank test) were investigated and a point score (the NSP score) was developed by using a Cox-regression model and validated in an independent external cohort from another institution ($n = 169$).

Results: The NSP score differentiated two groups of patients (≤ 1 , > 1 point) with distinct prognoses (median OS, 61.3 vs 19.3 months; $P < 0.001$). A high NSP score was associated with increased major adverse events after HR (5.6 vs 13.8%, $P = 0.027$). Its predictive accuracy as determined by the area under the receiver operating characteristic curve (AUC) at 1, 3, and 5 years (AUCs 0.688, 0.739, and 0.732) was greater than the other six staging systems for HCC (0.513–0.677). The findings were supported by the validation cohort.

Conclusions: The NSP scoring system is more accurate in selecting patients with intermediate-stage HCC for HR.

Hepatocellular carcinoma (HCC) is one of the most common malignancies in Asia and Africa, and its incidence in the Western world is increasing (Torre *et al*, 2015). The Barcelona Clinic Liver Cancer (BCLC) classification (Llovet *et al*, 1999) has been endorsed as the best staging system and treatment algorithm for HCC by the European Association for the Study of Liver Disease (EASL) and the American Association for the Study of Liver Disease (Bruix *et al*, 2005; European Association For The Study Of The Liver and

European Organisation For Research And Treatment Of Cancer, 2012). In addition to estimating prognosis, the main advantage of the BCLC staging system is that it establishes a link between staging and treatment indications. For patients with intermediate-stage (BCLC-B) HCC, transarterial chemoembolization (TACE) is recommended as the first-line therapy with improved 2-year survival compared with conservative treatment. Hepatic resection (HR) is indicated only for patients with early-stage HCC

*Correspondence: Dr M Shi; E-mail: shiming@sysu.edu.cn or Dr R-P Guo; E-mail: guorongp@mail.sysu.edu.cn

⁷These authors contributed equally to this study.

Revised 5 July 2016; accepted 29 August 2016; published online 4 October 2016

© 2016 Cancer Research UK. All rights reserved 0007–0920/16

(BCLC-A) with satisfactory liver function (BCLC-A; Bruix *et al*, 2005; European Association For The Study Of The Liver and European Organisation For Research And Treatment Of Cancer, 2012). Furthermore, HR for intermediate-stage HCC is considered to be a poor option which is associated with an unfavourable prognosis (Bruix *et al*, 2004).

However, many large centres with extensive experience in treating HCC, particularly those in Asia, do not subscribe to these guidelines (Chow, 2012; Yamakado and Kudo, 2014; Colombo and Sangiovanni, 2015; Zhong *et al*, 2015). In real-world practice, a large proportion of patients with intermediate-stage HCC receive surgical resection (Torzilli *et al*, 2013; Roayaie *et al*, 2015), particularly in regions where donor lives are scarce. Numerous recent studies have demonstrated that HR for intermediate-stage HCC appears to benefit patients in terms of the 5-year survival rates, which ranges from 23.9 to 61.2%, and median survival time, which ranges from 22.5 to 60.4 months (Wang *et al*, 2008; Delis *et al*, 2010; Luo *et al*, 2011; Zhong *et al*, 2013; Jianyong *et al*, 2014; Yin *et al*, 2014; Cucchetti *et al*, 2015). As patients with intermediate-stage HCC encompasses a variety of tumour burden, hepatic function (Child–Pugh A or B) and disease aetiology, aggressive surgical resection may not be the optimal therapy for all these patients. Therefore, a suitable subset of patients who would truly benefit from HR remains to be identified.

In this study, a simple scoring system was established based on preoperative data, with the goal of accurately predicting the postoperative long-term survival of patients with intermediate HCC. The aim was to identify the subgroup of patients with intermediate-stage HCC who would benefit most from HR.

MATERIALS AND METHODS

Eligibility. Intermediate-stage HCC (BCLC-B) was defined on the basis of the BCLC classification as follows: 2 to 3 tumours, of which at least 1 was >3 cm in diameter; >3 tumours of any diameter; and the absence of extrahepatic metastasis, the absence of tumour invasion into the portal or hepatic veins and performance status 0 (Forner *et al*, 2010). All patients initially diagnosed with intermediate-stage HCC by histology or dynamic computed tomography (CT)/magnetic resonance imaging scans according to EASL diagnostic criteria (European Association For The Study Of The Liver and European Organisation For Research And Treatment Of Cancer, 2012) who were treated with HR and pathologically proven multiple HCCs were retrospectively screened for eligibility.

Inclusion and Exclusion Criteria. Only patients who met all of the following criteria were enrolled in the study. (a) Good surgical risk patients >18 years and ≤75 years of age; (b) patients diagnosed with intermediate-stage HCC; (c) well-preserved liver function (Child–Pugh class A); (d) no anticancer treatment before surgery; (e) resectable disease according to previously reported criteria (Luo *et al*, 2011), which was defined as the possibility to completely remove all tumours while retaining a sufficient liver remnant to maintain postoperative liver function, as assessed by our surgical team.

Patients were excluded from the study if they had one or more of the following: (a) a platelet count of $<50 \times 10^9 l^{-1}$; (b) Child–Pugh class B or C; (c) palliative tumour resection; and (d) incomplete data.

The results of the training cohort were then confirmed in an independent external validation cohort using the same inclusion/exclusion criteria.

Radiology Interpretation. All patients underwent a standardised liver imaging protocol (Pomfret *et al*, 2010). The image data were directly interfaced with a Picture Archiving and Communication

System (PACS; Centricity, GE Healthcare, Milwaukee, WI, USA), which was used to display all image data on monitors with adjustment of the optimal window setting for each case, and viewed by two experienced abdominal radiologists. For the measurement of HCC lesions, the largest diameter was recorded using electronic calipers on the PACS monitor. Any disagreements were resolved by consensus.

Data Collection. This retrospective study was conducted in accordance with the ethical guidelines of the 1975 Declaration of Helsinki, and written informed consent was obtained from each participant. The study procedure was approved by the institutional review board of the Sun Yat-sen University Cancer Center and the Nanfang Hospital of Southern Medical University. The study was censored on 31 May 2015.

In all cases, data (demographics, clinical, biological, radiological, treatment outcomes, and adverse events) were prospectively collected. HCC was staged according to the American Joint Committee on Cancer seventh edition (Edge and Compton, 2010; Okuda *et al*, 1985), Cancer of the Liver Italian Program (CLIP; Kudo *et al*, 2003), and the Chinese University Prognostic Index (CUPI; Leung *et al*, 2002). Furthermore, information on the recently proposed BCLC-B sub-classification (Bolondi *et al*, 2012) and the NDR scoring system for multiple HCC (Yang *et al*, 2015) was provided.

Adverse events that occurred in the hospital during admission for surgery or death within 90 days after HR were documented according to the Common Terminology Criteria for Adverse Events version 3.0 (CTCAE; Trotti *et al*, 2003).

Hepatic Resection. HR was performed using techniques as described previously (Shi *et al*, 2007; Luo *et al*, 2011). Intraoperative ultrasonography was performed routinely to assess the number and size of lesions and the relationship of the tumours to vascular structures. Pringle's manoeuvre was routinely used with a clamp/unclamp time of 10 min/5 min. Anatomic resection was our preferred surgical method for multiple nodules in one segment or in neighbouring segments. For multiple bilobar nodules, anatomic resection was preferred for the main tumour, whereas satellite nodules were resected nonanatomically with a negative resection margin. When an inadequate liver remnant was found, nonanatomic resection was performed with a negative resection margin. A negative resection margin was defined as a lack of visible tumour cells at the margins of the remnant liver nearest to the gross edge of the tumour.

Follow-up. Follow-up examinations were conducted using laboratory findings (including serum AFP, liver function, and blood tests), abdominal ultrasonography, and contrast-enhanced CT every 3 months for the first year and every 6 months thereafter for a total of 60 months after treatment. All patients with HBV-related HCC who were prepared for resection for their HCC in our hospital were counselled by a hepatologist for antiviral therapy regardless of the serum HBV DNA result (Wong *et al*, 2011). For patients in whom tumour recurrence was detected after undergoing tumour resection, the treatment choice was determined by the characteristics of the recurrent tumour, the patient's request, and the results of discussion by our multidisciplinary team (Luo *et al*, 2011).

Statistical Analysis. The main end point of the study was overall survival (OS), which was calculated from the date of resection until death or the end of the follow-up period. Disease-free survival (DFS) was defined as the interval between the operation and the date of diagnosis of the first recurrence or the last follow-up. Survival curves were calculated using the Kaplan–Meier method. Median survival times and their 95% confidence intervals (CI) are reported. Continuous variables, such as prothrombin time and size (diameter) of the tumour, were transformed into

categorical variables. Each of these continuous variables was divided into two or three leveled categorical data by setting one or two break point(s), respectively, which were then represented by one or two binary variable(s); *P*-values were calculated for each set of break points with univariate or multivariate Cox-proportional hazard regression, and the set of break points with the lowest *P*-value was retained if the value reached significance. Univariate analysis of OS time was performed on the estimation set. The log-rank test was performed to detect significant parameters in the univariate analysis. Parameters with a *P*-value (log-rank) <0.05 in the univariate analyses were entered into the multivariate analysis.

Multivariate Cox-regression analysis with stepwise selection was performed to detect independent predictors of OS (entry criteria for selection into the final multivariate model was *P*<0.05). The regression coefficients (*B*) of the Cox-regression model were multiplied by 3 and rounded to the nearest unit (1.00 units) to obtain simple point numbers to facilitate the bedside calculation of the NSP score.

The abilities of the different systems to differentiate prognosis were compared by the area under the receiver operating characteristic (ROC) curve for each score (which is equivalent to the concordance statistic (c-statistic); Hanley and McNeil, 1982). To perform this test, patients censored before 1, 3, and 5 years were excluded from the analysis. To further validate the discriminative ability of the NSP score, the NSP score was analysed as a survival predictor in each subgroup of commonly used staging systems by log-rank test methodology.

To avoid overoptimistic results due to model development and evaluation using the same data set, the prognostic performance of the NSP score was assessed in an independent external validation cohort.

All reported *P*-values were the result of two-sided tests. A significance level of 0.05 was applied throughout. Statistical analyses were performed using IBM SPSS v.19.0 (SPSS, Armonk, NY, USA).

RESULTS

Clinical Characteristics of the Patients. In the training cohort, of 290 patients with intermediate-stage HCC who underwent HR at the Department of Hepatobiliary Oncology of Sun Yat-sen University Cancer Center between February 2005 and December 2012, 255 patients met the inclusion criteria (Figure 1A). For the external validation cohort, 169 patients with intermediate-stage HCC who received HR at the Nanfang Hospital of Southern Medical University between May 2005 and December 2012 were ultimately included for further analysis (Figure 1B). The baseline characteristics of the training cohort and the validation cohort are provided in Table 1.

At the time of censoring, 148 out of 255 (58.0%) patients in the training cohort and 110 out of 169 (65.1%) patients in the validation cohort had died. In the training cohort, 111 (75.0%) died of recurrence, 22 (14.9%) of liver failure, and 15 (10.1%) of other causes (haemorrhage = 8, infection = 3, gastric cancer = 1, unknown = 3). In the validation cohort, 86 (78.2%) died of recurrence, 15 (13.6%) of liver failure, and 9 (8.2%) of other causes (haemorrhage = 5, hepatitis = 2, renal failure = 1, unknown = 1).

OS and DFS in the Training and Validation Cohorts. The median follow-up periods for the training cohort and validation cohort were 26.6 (range, 1.3–96.7) months and 27.8 (range, 2.3–115.1) months, respectively. For the training cohort, the median OS was 31.5 (95% CI: 23.2–39.9) months, and the 1-, 3-, and 5-year OS were 77.5%, 46.5%, and 36.0%, respectively. For the validation cohort, the median OS for the validation cohort was 30.6

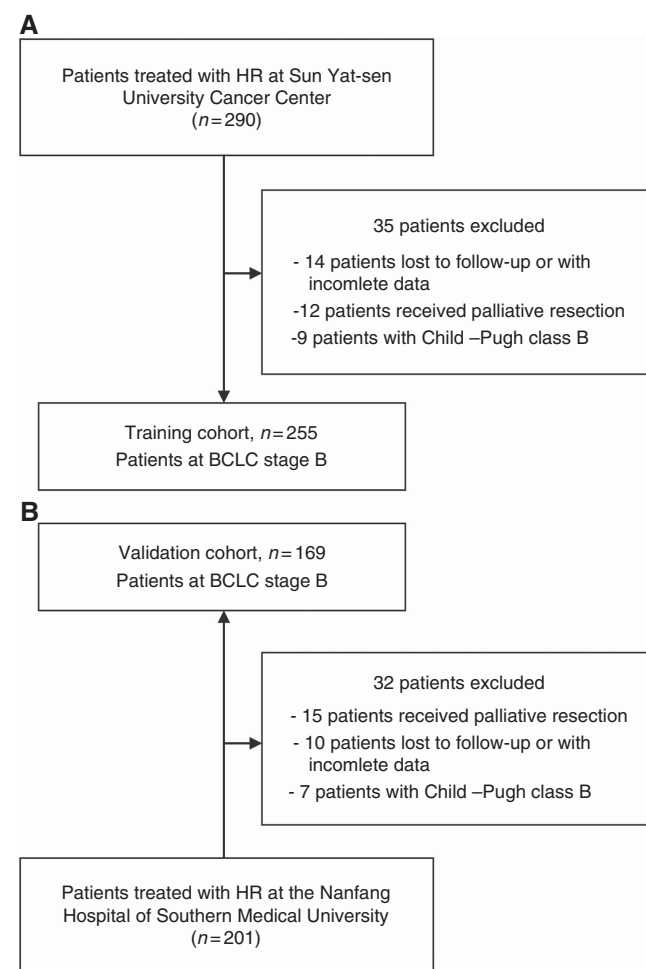


Figure 1. Flowchart in the training (A) and validation cohorts (B).

(95% CI: 23.9–37.3) months, and the 1-, 3-, and 5-year OS were 76.3%, 43.9%, and 32.1%, respectively.

During the follow-up period, HCC recurrence was identified in 190 out of 255 patients (74.5%) in the training cohort and 133 out of 169 patients (78.7%) in the validation cohort.

For the training cohort, the median DFS was 9.7 (95% CI: 7.1–12.3) months. For the validation cohort, the median DFS was 8.3 (95% CI: 4.9–11.8) months.

Univariate and Multivariable Cox-Regression Analyses in the Training Cohort.

The seven parameters of predictive value in the univariate analysis (Table 2) were entered into a Cox-regression analysis. After the stepwise removal of variables that were not significant (step 1, glutamyl transpeptidase; step 2, aspartate aminotransferase; step 3, alkaline phosphatase; step 4, alpha-fetoprotein), only prothrombin time (≤ 12 s/ > 12 s), tumour size ($\leq 6/6-9/ \geq 9$ cm), and tumour number ($\leq 3/ > 3$) remained as significant predictors of OS (Supplementary Figure 1A–C). The calculated regression coefficients (*B*-values) were multiplied by a factor of 3 and rounded to facilitate the NSP score calculation (Table 3). The NSP score for a patient can be calculated using the following equation, adding the sum of multiplying these three factors by their respective weights:

$$\text{NSP score} = \text{N}(\leq 3 = 0; > 3 = 2) + \text{S}(\leq 6 \text{ cm} = 0; 6-9 \text{ cm} = 1; \geq 9 \text{ cm} = 2) + \text{P}(\leq 1 \text{ 2 s} = 0; > 12 \text{ s} = 1).$$

NSP Score Predicts Survival in the Training and Validation Cohorts. In the training cohort, the NSP score identified six subgroups with different prognoses (Figure 2A). Using a threshold

Table 1. Baseline patient and disease characteristics in the training cohort

| Characteristic | Number (%) | |
|---|-----------------------------|-----------------------------|
| | Training cohort (n = 255) | Validation cohort (n = 169) |
| Year of resection | | |
| 2005–2008 | 121 (47) | 87 (51) |
| 2009–2012 | 134 (53) | 82 (49) |
| Age (years) | 49.0 ± 12.7 | 49.1 ± 12.7 |
| Male | 227 (89) | 154 (91) |
| Aetiology | | |
| HBV | 210 (82) | 135(80) |
| HCV | 21 (8) | 17 (10) |
| Alcohol | 9 (4) | 8 (5) |
| Others | 15 (6) | 9 (5) |
| Portal hypertension ^a | 56 (22) | 36 (21) |
| Liver cirrhosis | 181 (71) | 130 (77) |
| Platelet count, 10 ⁹ l ⁻¹ | 198.7 ± 79.5 | 191.5 ± 71.9 |
| Prothrombin time, s | 12.2 ± 1.2 | 12.3 ± 1.2 |
| Alanine transaminase, U l ⁻¹ | 45.7 ± 36.0 | 45.7 ± 30.1 |
| Aspartate aminotransferase, U l ⁻¹ | 46.6 ± 37.3 | 43.7 ± 24.3 |
| Total bilirubin, μmol l ⁻¹ | 14.0 ± 6.5 | 14.3 ± 6.5 |
| Serum albumin, g l ⁻¹ | 41.6 ± 3.8 | 41.9 ± 3.5 |
| Alkaline phosphatase, U l ⁻¹ | 90.0 ± 44.4 | 91.5 ± 34.0 |
| Glutamyl transpeptidase, U l ⁻¹ | 99.2 ± 98.2 | 97.0 ± 83.7 |
| MELD score | 4.2 ± 2.8 | 4.5 ± 2.5 |
| ALBI score | | |
| Grade 1 | 185 (73) | 127 (75) |
| Grade 2/3 | 70/0 (27/0) | 42/0 (25/0) |
| Alpha-fetoprotein, ng l ⁻¹ | | |
| ≤ 200 | 115 (45) | 76 (45) |
| > 200 | 140 (55) | 93 (55) |
| Tumour extent | | |
| Unilobar | 203 (80) | 128 (76) |
| Bilobar | 52 (20) | 41 (24) |
| Tumour number | | |
| 2 | 135 (53) | 100 (59) |
| 3 | 43 (17) | 29 (17) |
| ≥ 4 | 77 (30) | 40 (24) |
| Tumour size, cm ^b | | |
| ≤ 6 | 107 (42) | 76 (45) |
| 6–9 | 92 (36) | 64 (38) |
| ≥ 9 | 56 (22) | 29 (17) |
| Staging system | | |
| NDR score (≤2/>2) | 155/100 (61/39) | 106/63 (63/36) |
| BCLC-B sub-stage (B1/B2) | 51/204 (20/80) | 44/125 (26/74) |
| TNM stage (II/III) | 67/188 (26/74) | 51/118 (30/70) |
| CLIP score (0/1/2/3/4/5) | 128/121/6/0/0 (50/48/2/0/0) | 91/75/3/0/0 (54/44/2/0/0) |
| CUPI (I/II/III) | 233/22/0 (91/9/0) | 156/13/0 (92/8/0) |
| Okuda stage (I/II/III) | 224/31/0 (88/12/0) | 150/19/0 (89/11/0) |

Abbreviations: HBV = hepatitis B virus; HCV = hepatitis C virus; MELD = model for end-stage liver disease; ALBI score = albumin-bilirubin score; NDR score = tumour number (N) > 3, total tumour diameter (D) > 8 cm, and a ratio of largest/smallest diameter (R) > 6; BCLC = Barcelona Clinic Liver Cancer; TNM = tumour node metastasis; CLIP = Cancer of the Liver Italian Program; CUPI = Chinese University Prognostic Index. Values in parentheses are percentages unless indicated otherwise.

^aPortal hypertension was defined as the presence of oesophageal varices or a platelet count of < 100 × 10⁹ l⁻¹ in association with splenomegaly.

^bMaximum tumour size.

of 1, we defined two groups of patients with significantly different OS ($P < 0.001$), with a median OS of 61.3 months (95% CI, 44.6–78.1 months) for an NSP score ≤ 1 ($n = 125$), and a median OS of 19.3 months (95% CI, 14.5–24.2 months) for an NSP score > 1 ($n = 130$; Figure 2B). Importantly, the NSP score performed equally well in the validation cohort, with a median OS of 51.5 months (95% CI, 36.9–66.2 months) for an NSP score ≤ 1 ($n = 86$), and a median OS of 17.3 months (95% CI, 12.8–21.8 months) for an NSP score > 1 ($n = 83$; Figures 2C, $P < 0.001$).

In addition, the NSP score also had good performance in DFS prediction, with two significantly different prognostic subgroups in the training cohort (20.2 vs 4.7 months, $P < 0.001$, Figure 3A) and the validation cohort (14.2 vs 5.0 months, $P < 0.001$, Figure 3B).

Comparison of the Predictive Accuracy of the NSP System and the Current Commonly Used Staging Systems in the Training and Validation Cohorts. Kaplan–Meier curves were generated for the BCLC-B sub-classification, TNM 7th, Okuda, CLIP, CUPI,

Table 2. Univariate analysis of prognostic factors in the training cohort

| Variables | Overall survival (months) | | | |
|--|---------------------------|--------|-----------|--------------------|
| | N = 255 | Median | 95% CI | P-value (log-rank) |
| Year of resection | | | | |
| 2005–2008 | 121 | 31.0 | 19.5–42.4 | 0.610 |
| 2009–2012 | 134 | 32.6 | 21.9–43.3 | |
| Age (years) | | | | |
| ≤50 | 126 | 26.1 | 20.8–31.4 | 0.063 |
| >50 | 129 | 40.9 | 26.7–55.1 | |
| Sex | | | | |
| Male | 227 | 31.5 | 22.7–40.4 | 0.715 |
| Female | 28 | 28.2 | 7.8–48.6 | |
| Aetiology | | | | |
| Viral | 231 | 32.6 | 24.3–41.0 | 0.695 |
| Others | 24 | 17.0 | 9.9–24.1 | |
| Portal hypertension | | | | |
| Yes | 56 | 24.2 | 23.5–41.8 | 0.168 |
| No | 199 | 32.6 | 11.3–37.1 | |
| Liver cirrhosis | | | | |
| Yes | 181 | 30.6 | 21.7–39.5 | 0.190 |
| No | 74 | 34.5 | 7.1–61.8 | |
| Platelet count, 10⁹ l⁻¹ | | | | |
| ≤100 | 22 | 40.9 | 23.7–58.1 | 0.468 |
| >100 | 233 | 31.0 | 22.9–39.1 | |
| Prothrombin time, s | | | | |
| ≤12 | 122 | 46.6 | 25.9–67.3 | 0.008 |
| >12 | 133 | 24.2 | 14.8–33.6 | |
| Alanine transaminase, U l⁻¹ | | | | |
| ≤40 | 148 | 30.6 | 20.2–40.9 | 0.478 |
| >40 | 107 | 32.6 | 21.8–43.5 | |
| Aspartate aminotransferase, U l⁻¹ | | | | |
| ≤45 | 163 | 40.9 | 27.0–54.9 | 0.004 |
| >45 | 92 | 20.2 | 11.5–28.9 | |
| Total bilirubin, μmol l⁻¹ | | | | |
| ≤17 | 192 | 30.1 | 23.4–36.9 | 0.533 |
| >17 | 63 | 40.6 | 29.5–51.6 | |
| Serum albumin, g l⁻¹ | | | | |
| ≤36 | 22 | 52.7 | 0.0–106.5 | 0.826 |
| >36 | 233 | 31.5 | 23.9–39.1 | |
| Alkaline phosphatase, U l⁻¹ | | | | |
| ≤200 | 250 | 32.3 | 23.9–40.7 | 0.008 |
| >200 | 5 | 27.8 | 0.0–18.2 | |
| Glutamyl transpeptidase, U l⁻¹ | | | | |
| ≤100 | 172 | 42.7 | 28.5–56.9 | 0.001 |
| >100 | 83 | 19.3 | 12.9–25.8 | |
| MELD score | | | | |
| ≤4 | 134 | 30.1 | 22.6–37.7 | 0.962 |
| >4 | 121 | 35.0 | 25.4–44.6 | |
| ALBI score | | | | |
| Grade 1 | 185 | 34.5 | 24.6–44.3 | 0.147 |
| Grade 2/3 | 70 | 23.4 | 15.2–31.5 | |
| Alpha-fetoprotein, ng ml⁻¹ | | | | |
| ≤200 | 115 | 40.9 | 23.4–58.4 | 0.008 |
| >200 | 140 | 28.2 | 21.6–34.8 | |
| Tumour extent | | | | |
| Unilobar | 203 | 31.5 | 16.8–46.2 | 0.181 |
| Bilobar | 52 | 31.0 | 20.3–41.7 | |
| Tumour size (cm)^a | | | | |
| ≤6 | 107 | 51.0 | 33.4–68.6 | <0.001 |
| 6–9 | 88 | 28.2 | 21.1–35.2 | |
| ≥9 | 60 | 18.6 | 14.3–22.8 | |
| Tumour number | | | | |
| ≤3 | 178 | 49.4 | 34.7–64.2 | <0.001 |
| >3 | 77 | 17.6 | 15.1–20.2 | |

Abbreviations: CI = confidence interval; MELD = model for end-stage liver disease; ALBI score = albumin–bilirubin score.

^aMaximum tumour size.

Table 3. Multivariable Cox-regression analysis of prognostic factors for patients in the training cohort

| Overall survival | | | | | |
|-------------------------------------|------|-----------|------|---------------------------|---------|
| Variables | HR | 95% CI | B | Score points ^a | P-value |
| Tumour number (n) | | | | | |
| <0.001 | | | | | |
| ≤3 | 1 | | — | 0 | |
| >3 | 2.22 | 1.53–3.22 | 0.80 | 2 | |
| Tumour size (cm)^b | | | | | |
| 0.004 | | | | | |
| ≤6 | 1 | | — | 0 | |
| 6–9 | 1.30 | 1.16–1.93 | 0.26 | 1 | |
| ≥9 | 1.93 | 1.23–2.99 | 0.66 | 2 | |
| Prothrombin time, s | | | | | |
| 0.018 | | | | | |
| ≤12 | 1 | | — | 0 | |
| >12 | 1.48 | 1.06–2.05 | 0.39 | 1 | |

Abbreviations: HR = hazard ratio; CI = confidence interval.
^aThe regression coefficients (B) were multiplied by 3 and rounded to facilitate bedside calculation of the score.
^bMaximum tumour size.

and the NDR scoring system (Supplementary Figures 2A–F). The NDR scoring system, BCLC-B sub-classification, and TNM 7th staging systems showed clear different prognostic strata. The CLIP, CUPI, and Okuda systems presented overlapping survival curves or without statistical difference. When all the staging systems were introduced into the Cox-regression analysis, the NSP and NDR scoring system were selected as independent predictors in the training cohort (NSP score: HR = 1.99, 95% CI: 1.33–3.00, P = 0.001; NDR score: HR = 1.62, 95% CI: 1.09–2.41, P = 0.017). Only the NSP was selected as an independent predictor in the validation cohort (HR = 2.74, 95% CI: 1.86–4.04, P < 0.001).

We next determined which staging systems were the best at predicting survival evaluated by ROC curve area analysis. In the training cohort, the AUCs of the NSP system at 1, 3, and 5 years was 0.688, 0.739, and 0.732, respectively, and was greater than those of the six other commonly used staging systems for HCC (AUCs, 0.513 to 0.677; Figure 4A–C). In the validation cohort, the AUCs of the NSP system at 1, 3, and 5 years was 0.719, 0.750, and 0.718, respectively, and was greater than those of the six other

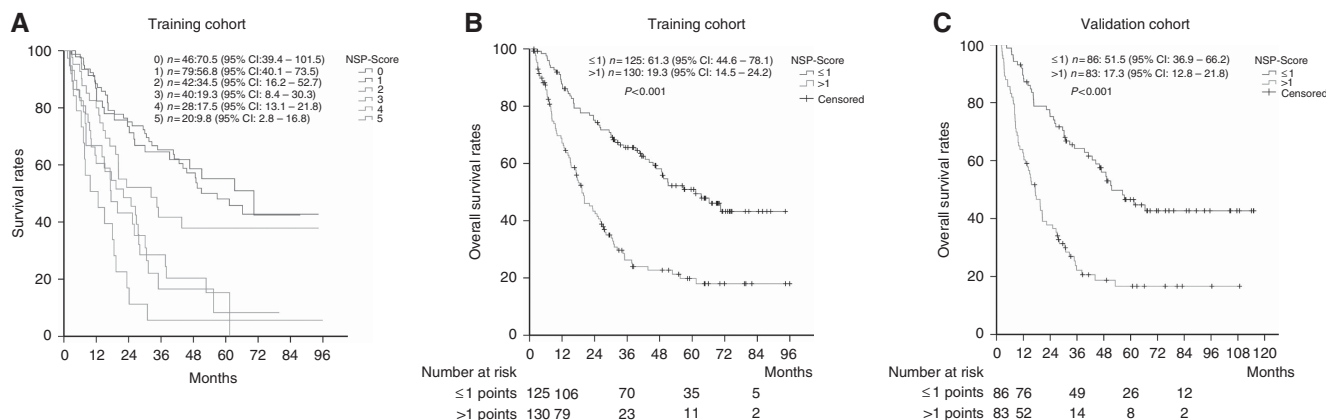


Figure 2. Kaplan-Meier estimated OS curves by NSP score. (A) Prognostic significance of the single-point scores. The dichotomous NSP score cut-off was established based on favourable median OS in the Kaplan–Meier curves. The prognostic significance of the two NSP-score groups (≤1 point, >1 point) for OS in the training (B) and validation cohorts (C). All analyses were performed using the Kaplan–Meier method (OS in months) and the log-rank test. A full colour version of this figure is available at the *British Journal of Cancer* journal online.

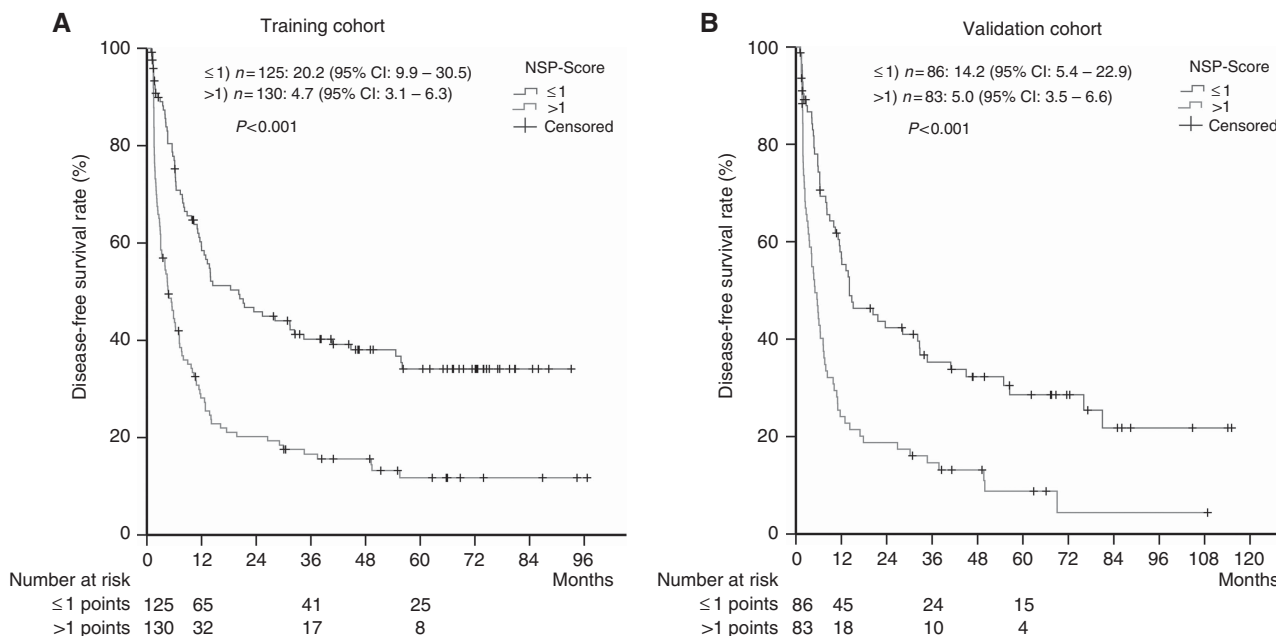


Figure 3. Kaplan-Meier estimated DFS curves by NSP score. The prognostic significance of the two NSP-score groups (≤1 point, >1 point) for DFS in the training (A) and validation cohorts (B). A full colour version of this figure is available at the *British Journal of Cancer* journal online.

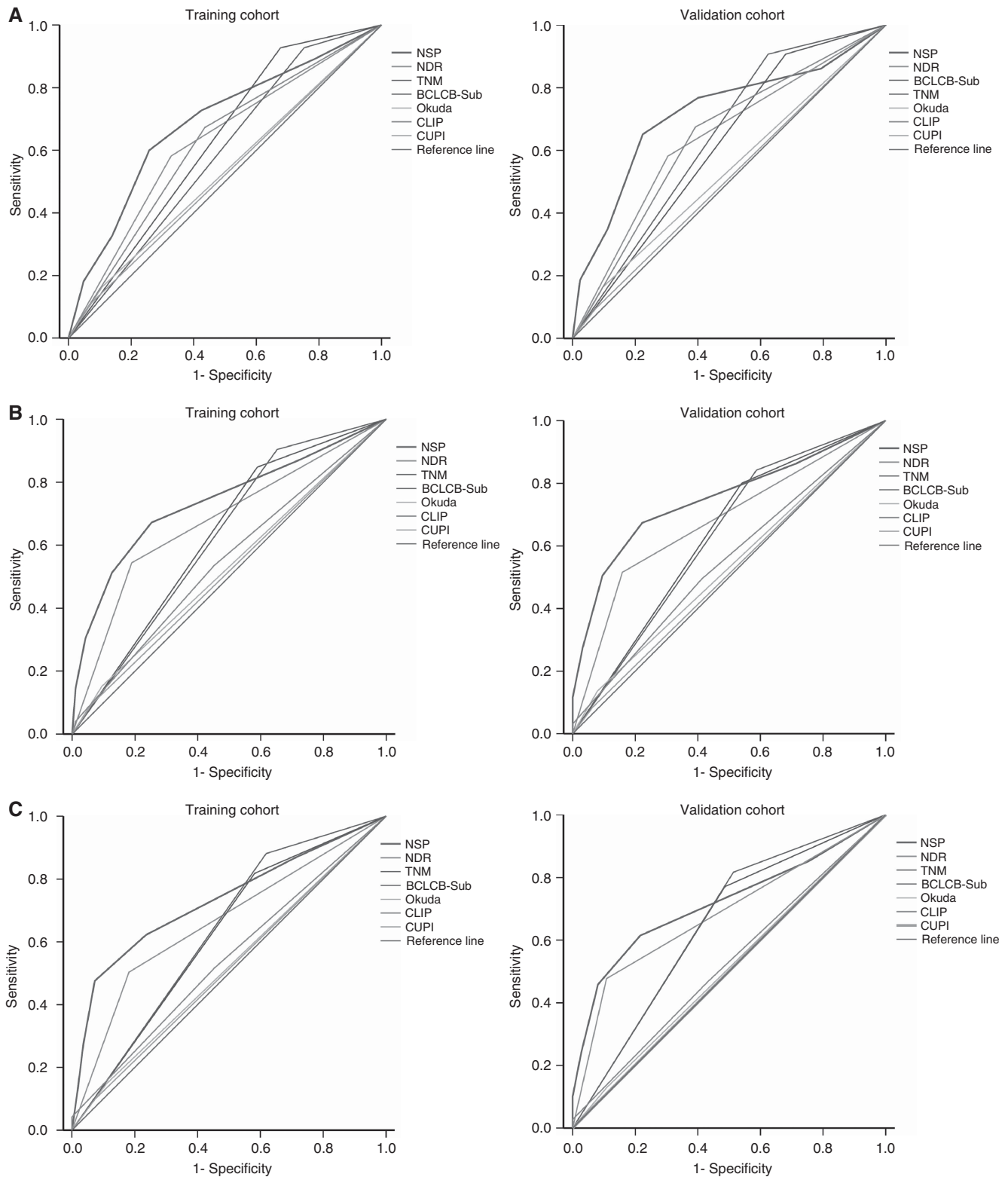


Figure 4. Receiver operating characteristic curves. The AUC of the NSP score system and the other six staging systems to predict OS in the training and validation cohorts at 1 (A), 3 (B), and 5 (C) years. A full colour version of this figure is available at the *British Journal of Cancer* journal online.

commonly used staging systems for HCC (AUCs, 0.510 to 0.684; Figure 4A–C).

NSP Score Predicts OS in Subgroups of the Current Commonly Used Staging Systems. When patients were further stratified by the other six staging systems, the NSP score remained a significant OS predictor in each subgroup in the training (Supplementary

Figure 3) and validation cohorts (Supplementary Figure 4). These results combined with the highest AUC value for the NSP score suggest NSP score is a better scoring system than the other six staging systems.

Furthermore, higher NSP score values were associated with an increased number of documented major adverse events within 90 days after HR in both cohorts (Table 4).

Table 4. Association of NSP score with adverse events after HR in the training and validation cohorts

| NSP score points | Training cohort (n = 255) ^a | | Validation cohort (n = 169) ^b | | P _T -value | P _V -value |
|---|--|-------------|--|-------------|-----------------------|-----------------------|
| | ≤1 N (%) | >1 N (%) | ≤1 N (%) | >1 N (%) | | |
| AE ≥ grade 3 within 90 days after HR | | | | | | |
| Present | 7 (5.6) | 18 (13.8) | 4 (4.7%) | 11 (13.3) | 0.027 | 0.049 |
| Absent | 118 (94.4) | 112 (86.2) | 82 (95.3) | 72 (86.7) | | |

Abbreviations: AE = adverse events; HR = hepatic resection.
^aNSP score ≤ 1: ascites/pleural effusion (n = 4), infection (n = 2), haemorrhage (n = 1); NSP score > 1: ascites/pleural effusion (n = 10), liver dysfunction = 4, haemorrhage (n = 3), bile leakage (n = 1).
^bNSP score ≤ 1: ascites/pleural effusion (n = 3), bile leakage (n = 1); NSP score > 1: ascites/pleural effusion (n = 7); liver dysfunction = 2, haemorrhage (n = 2).

DISCUSSION

Currently, many staging systems have been developed to classify patients with HCC. However, none of these staging systems were specifically developed to predict outcomes for intermediate HCC patients who received HR. In addition, the indication of liver resection for intermediate-stage HCC remains controversial (Forner *et al*, 2014, 2015; Zhong *et al*, 2015). Therefore, in clinical practice, a novel prognostic staging system specific for patients with intermediate-stage HCC is the key to selecting the best candidates for HR. Our study is the first to report and validate a scoring system (NSP) that predicts survival in patients with intermediate-stage HCC treated with HR. A low NSP score (≤ 1 point) was chosen as a cut-off because it reliably identified patients who were good candidates for HR based on their favourable median OS. The NSP score increases the pool of ideal HR candidates with intermediate-stage HCC by ~50% with favourable long-term survival.

Although our study confirmed that the other six staging systems can stratify patients treated with HR into distinct risk categories, the ability of these systems to predict survival was suboptimal. The reason for the inferiority of these staging systems might be that they were originally developed from mixed populations with a smaller number of patients receiving surgery (2.8% in CLIP, 8.5% in Okuda, and 10.4% in CUPI). Different treatments have a large influence on prognosis that should be considered in these staging systems. Furthermore, a major issue is that TNM staging and the NDR scoring system does not include measurements of liver function. The greatest criticism of the CLIP and Okuda systems is that their criteria of tumour morphology are too broad, reducing their value to intermediate-stage HCC patients. Although the BCLC-B sub-classification was specifically developed for intermediate-stage HCC, previous studies (Ha *et al*, 2014; Weinmann *et al*, 2015) indicated that there were no significant differences survival among several adjacent stages.

This study has several limitations. First, our effort is obviously limited by the retrospective nature of the analysis. However, the analysis was conducted based on prospectively collected data. Furthermore, we confirmed the results derived from the initial cohort with the validation cohort to increase reliability. Second, the purpose of this study was to identify a suitable subset of patients with intermediate-stage HCC for HR. The results do not suggest that patients with an NSP score ≤ 1 should or should not be treated with other therapies, such as liver transplantation, radio-frequency ablation, or TACE. Third, with the majority of the patients having evidence of HBV infection, our data require validation from other study groups in which HCV infection or

alcohol is the prevailing aetiology of chronic liver disease. Therefore, multicenter and prospective studies are needed to confirm our results.

In conclusion, we developed a preoperative, externally validated, simple objective prognostic score (NSP) to identify patients with intermediate-stage HCC who could benefit from aggressive surgical interventions. For intermediate-stage HCC, HR is recommended for patients with an NSP score ≤ 1. These findings should be validated in further prospective studies.

ACKNOWLEDGEMENTS

This study was supported by grants from the National Natural Science Foundation of China (No. 81172037/H1606, No. 81272639, and No. 81572385), Guangdong Province Science and Technology Project of China (No. 2013B021800159), and the Fundamental Research Funds for the Central Universities of China (No. 13ykzd24).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Bolondi L, Burroughs A, Dufour JF, Galle PR, Mazzaferro V, Piscaglia F, Raoul JL, Sangro B (2012) Heterogeneity of patients with intermediate (BCLC B) hepatocellular carcinoma: proposal for a subclassification to facilitate treatment decisions. *Semin Liver Dis* 32(4): 348–359.
- Bruix J, Sala M, Llovet JM (2004) Chemoembolization for hepatocellular carcinoma. *Gastroenterology* 127(5 Suppl 1): S179–S188.
- Bruix J, Sherman M (2005) Practice Guidelines Committee American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma. *Hepatology* 42(5): 1208–1236.
- Chow PK (2012) Resection for hepatocellular carcinoma: is it justifiable to restrict this to the American Association for the Study of the Liver/Barcelona Clinic for Liver Cancer criteria? *J Gastroenterol Hepatol* 27(3): 452–457.
- Colombo M, Sangiovanni A (2015) Treatment of hepatocellular carcinoma: beyond international guidelines. *Liv Int* 35(Suppl 1): 129–138.
- Cucchetti A, Djulbegovic B, Tsalatsanis A, Vitale A, Hozo I, Piscaglia F, Cescon M, Ercolani G, Tuci F, Cillo U, Pinna AD (2015) When to perform hepatic resection for intermediate-stage hepatocellular carcinoma. *Hepatology* 61(3): 905–914.
- Delis SG, Bakoyiannis A, Tassopoulos N, Athanassiou K, Kelekis D, Madariaga J, Dervenis C (2010) Hepatic resection for hepatocellular carcinoma exceeding Milan criteria. *Surg Oncol* 19(4): 200–207.
- Edge SB, Compton CC (2010) The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* 17(6): 1471–1474.
- European Association For The Study Of The Liver and European Organisation For Research And Treatment Of Cancer (2012) EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 56(4): 908–943.
- Forner A, Gilibert M, Bruix J, Raoul JL (2014) Treatment of intermediate-stage hepatocellular carcinoma. *Nat Rev Clin Oncol* 11(9): 525–535.
- Forner A, Gilibert M, Bruix J, Raoul JL (2015) Intermediate-stage HCC—upfront resection can be feasible. *Nat Rev Clin Oncol* 12: 5.
- Forner A, Reig ME, de Lope CR, Bruix J (2010) Current strategy for staging and treatment: the BCLC update and future prospects. *Semin Liver Dis* 30(1): 61–74.
- Ha Y, Shim JH, Kim SO, Kim KM, Lim YS, Lee HC (2014) Clinical appraisal of the recently proposed Barcelona Clinic Liver Cancer stage B subclassification by survival analysis. *J Gastroenterol Hepatol* 29(4): 787–793.
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1): 29–36.

- Jiayong L, Lunan Y, Wentao W, Yong Z, Bo L, Tianfu W, Minqing X, Jiaying Y (2014) Barcelona clinic liver cancer stage B hepatocellular carcinoma: transarterial chemoembolization or hepatic resection? *Medicine* **93**(26): e180.
- Kudo M, Chung H, Osaki Y (2003) 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* **38**(3): 207–215.
- Leung TW, Tang AM, Zee B, Lau WY, Lai PB, Leung KL, Lau JT, Yu SC, Johnson PJ (2002) 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* **94**(6): 1760–1769.
- Llovet JM, Bru C, Bruix J (1999) Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* **19**(3): 329–338.
- Luo J, Peng ZW, Guo RP, Zhang YQ, Li JQ, Chen MS, Shi M (2011) Hepatic resection versus transarterial lipiodol chemoembolization as the initial treatment for large, multiple, and resectable hepatocellular carcinomas: a prospective nonrandomized analysis. *Radiology* **259**(1): 286–295.
- Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, Ohnishi K (1985) Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* **56**(4): 918–928.
- Pomfret EA, Washburn K, Wald C, Nalesnik MA, Douglas D, Russo M, Roberts J, Reich DJ, Schwartz ME, Miele L, Lee FT, Florman S, Yao F, Harper A, Edwards E, Freeman R, Lake J (2010) Report of a national conference on liver allocation in patients with hepatocellular carcinoma in the United States. *Liver Transpl* **16**(3): 262–278.
- Roayaie S, Jibara G, Tabrizian P, Park JW, Yang J, Yan L, Schwartz M, Han G, Izzo F, Chen M, Blanc JF, Johnson P, Kudo M, Roberts LR, Sherman M (2015) The role of hepatic resection in the treatment of hepatocellular cancer. *Hepatology* **62**(2): 440–451.
- Shi M, Guo RP, Lin XJ, Zhang YQ, Chen MS, Zhang CQ, Lau WY, Li JQ (2007) Partial hepatectomy with wide versus narrow resection margin for solitary hepatocellular carcinoma: a prospective randomized trial. *Ann Surg* **245**(1): 36–43.
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA Cancer J Clin* **65**(2): 87–108.
- Torzilli G, Belghiti J, Kokudo N, Takayama T, Capussotti L, Nuzzo G, Vauthey JN, Choti MA, De Santibanes E, Donadon M, Morengi E, Makuuchi M (2013) A snapshot of the effective indications and results of surgery for hepatocellular carcinoma in tertiary referral centers: is it adherent to the EASL/AASLD recommendations? An observational study of the HCC East-West study group. *Ann Surg* **257**(5): 929–937.
- Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, Langer C, Murphy B, Cumberlin R, Coleman CN, Rubin P (2003) CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol* **13**(3): 176–181.
- Wang JH, Changchien CS, Hu TH, Lee CM, Kee KM, Lin CY, Chen CL, Chen TY, Huang YJ, Lu SN (2008) The efficacy of treatment schedules according to Barcelona Clinic Liver Cancer staging for hepatocellular carcinoma—Survival analysis of 3892 patients. *Eur J Cancer* **44**(7): 1000–1006.
- Weinmann A, Koch S, Sprinzel M, Kloeckner R, Schulze-Bergkamen H, Duber C, Lang H, Otto G, Worns MA, Galle PR (2015) Survival analysis of proposed BCLC-B subgroups in hepatocellular carcinoma patients. *Liver Int* **35**(2): 591–600.
- Wong JS, Wong GL, Tsoi KK, Wong VW, Cheung SY, Chong CN, Wong J, Lee KF, Lai PB, Chan HL (2011) Meta-analysis: the efficacy of anti-viral therapy in prevention of recurrence after curative treatment of chronic hepatitis B-related hepatocellular carcinoma. *Aliment Pharmacol Ther* **33**(10): 1104–1112.
- Yamakado K, Kudo M (2014) Treatment strategies of intermediate-stage hepatocellular carcinomas in Japan (Barcelona Clinic Liver Cancer stage B). *Oncology* **87**(Suppl 1): 78–81.
- Yang P, Wu D, Xia Y, Li J, Wang K, Yan Z, Shi L, Lau WY, Wu M, Shen F (2015) A prognostic scoring system for patients with multiple hepatocellular carcinomas treated by hepatectomy. *Ann Surg Oncol* **22**(3): 826–833.
- Yin L, Li H, Li AJ, Lau WY, Pan ZY, Lai EC, Wu MC, Zhou WP (2014) Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan Criteria: a RCT. *J Hepatol* **61**(1): 82–88.
- Zhong JH, Lu SD, Wang YY, Ma L, Li LQ (2015) Intermediate-stage HCC—upfront resection can be feasible. *Nat Rev Clin Oncol* **12**(5).
- Zhong JH, Xiang BD, Gong WF, Ke Y, Mo QG, Ma L, Liu X, Li LQ (2013) Comparison of long-term survival of patients with BCLC stage B hepatocellular carcinoma after liver resection or transarterial chemoembolization. *PLoS One* **8**(7): e68193.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)