Prognostic evaluation of HCC patients undergoing surgical resection: an analysis of 8 different staging systems

Jan Bednarsch 1, Zoltan Czigany 1, Daniel Heise 1, Katharina Joechle 1, Tom Luedde 2, Lara Heij 1,3, Philipp Bruners 4, Tom Florian Ulmer 1, Ulf Peter Neumann 1,5, Sven Arke Lang 1

Received: 2 September 2020 / Accepted: 2 December 2020 / Published online: 9 December 2020
© The Author(s) 2020

Abstract

Purpose No consensus exists regarding the most appropriate staging system to predict overall survival (OS) for hepatocellular carcinoma (HCC) in surgical candidates. Thus, we aimed to determine the prognostic ability of eight different staging systems in a European cohort of patients undergoing liver resection for HCC.

Methods Patients resected for HCC between 2010 and 2019 at our institution were analyzed with Kaplan-Meier and Cox regression analyses. Likelihood ratio (LR) $\chi^2$ (homogeneity), linear trend (LT) $\chi^2$ (discriminatory ability), and Akaike Information Criterion (AIC, explanatory ability) were used to determine the staging system with the best overall prognostic performance.

Results Liver resection for HCC was performed in 160 patients. Median OS was 39 months (95% confidence interval (CI): 32–46 months) and median RFS was 26 months (95% CI: 16–34 months). All staging systems (BCLC, HKLC, Okuda, CLIP, ITA.LI.CA staging and score, MESH, and GRETSCH) showed significant discriminatory ability regarding OS, with ITA.LI.CA score (LR $\chi^2$ 30.08, LT $\chi^2$ 13.90, AIC 455.27) and CLIP (LR $\chi^2$ 28.65, LT $\chi^2$ 18.95, AIC 460.07) being the best performing staging systems.

Conclusions ITA.LI.CA and CLIP are the most suitable staging system to predict OS in European HCC patients scheduled for curative-intent surgery.

Keywords HCC · Staging system · ITA.LI.CA · CLIP · Liver resection

Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| AIC | Akaike Information Criterion |
| ALD | Alcoholic liver disease |
| ALT | Alanine aminotransferase |
| AP | Alkaline phosphatase |
| ASA | American Society of Anesthesiologists |
| AST | Aspartate aminotransferase |
| BCLC | Barcelona Clinic Liver Cancer |
| BMI | Body mass index |
| CLIP | Cancer of Liver Italian Program |
| CPS | Child-Pugh score |
| CI | Confidence interval |
| CRP | C-reactive protein |
| CT | Computed tomography |
| CUPI | Chinese University Integrated System |
| CUSA | Cavitation Ultrasonic Surgical Aspirator |
| CVP | Central venous pressure |
| DFS | Disease-free survival |
| ECOG | Eastern Cooperative Oncology Group |
| ESMO | European Society for Medical Oncology |
| FFP | Fresh frozen plasma |
| FLR | Future liver remnant |
| GCP | Good clinical practice |
| GGT | Gamma glutamyltransferase |

* Sven Arke Lang
  svang@ukaachen.de

1 Department of Surgery and Transplantation, University Hospital RWTH Aachen, Pauwelsstrasse 30, 52074 Aachen, Germany
2 Department of Medicine III, University Hospital RWTH Aachen, Aachen, Germany
3 Institute of Pathology, University Hospital RWTH Aachen, Aachen, Germany
4 Department of Radiology, University Hospital RWTH Aachen, Aachen, Germany
5 Department of Surgery, Maastricht University Medical Centre (MUMC), Maastricht, Netherlands
Introduction

Hepatocellular carcinoma (HCC) is a major global health burden being the third most common cause of cancer-associated mortality worldwide [1, 2]. In the majority of patients, HCC arises on a background of chronic liver disease. Hence, liver transplantation is often considered treatment of choice since it addresses both the underlying liver and the oncological disease [3]. However, a large proportion of HCC patients is too old for transplantation, has major comorbidities, or presents with other contraindications, e.g., active alcoholism as well as advanced tumor stages precluding this approach. Moreover, the limited availability of liver grafts which result in strict allocation regulation and the excellent oncological outcome of surgery in small, solitary HCC underline the importance of liver resections as a major therapeutic option in patients suffering from HCC [4, 5]. Despite recent advances in HCC surgery including the use of dynamic liver function tests, e.g., LiMAX (maximum liver function capacity) or indocyanine green (ICG) and the increasing implementation of minimally invasive liver surgery, a significant proportion of patients is usually regarded as not ideal candidates for surgery based on a high risk of post-hepatectomy liver failure or poor overall oncological prognosis [6–11]. Therefore, identifying preoperative characteristics associated with a higher perioperative risk and prognostic value for oncological outcome has been in the center of interest in HCC. The latter has led to the development of various staging systems aiming to support clinical decision-making in HCC patients.

The most widespread preoperative staging systems are the Milan criteria and the Barcelona Clinic Liver Cancer (BCLC) staging system. While the Milan criteria is used to predict the outcome in HCC undergoing transplantation based on radiologic features, the BCLC system stratifies patients based on radiologic features, physical performance, and liver function. Hence, BCLC is commonly adopted in therapy guidelines, giving distinct recommendations regarding the treatment of choice for each subgroup of patients [12, 13]. In particular, BCLC allocates patients with early stage tumors to curative-intent surgery, while more oncological progressed individuals or patients with impaired liver function are scheduled to interventional or systemic therapy [12]. This traditional paradigm has been challenged by recent reports indicating a survival benefit of liver resection over other treatment modalities regardless of the pre-hepatectomy BCLC stage [14, 15]. Subsequently, various other staging systems have been proposed to overcome limitations of the BCLC staging systems, e.g., Cancer of Liver Italian Program (CLIP) score, Groupe d’Etude et de Traitement du Carcinome Hepatocellulaire (GRETCH) score, Italian Liver Cancer (ITA.LI.CA) tumor staging and score, Hong Kong Liver Cancer (HKLC) staging and score, model to estimate survival in hepatocellular carcinoma (MESH) score, or Okuda staging (Table 1) [12, 13, 16–21].

Given the variety of staging systems, we aimed to evaluate the prognostic ability of each staging systems to determine the “best” performing model in a European cohort of patients undergoing curative-intent surgery for HCC.

Material and methods

Patients

The study comprised one hundred sixty (n = 160) consecutive HCC patients who underwent curative-intend surgery at the University Hospital RWTH Aachen (UH-RWTH) between 2010 and 2019. Clinical staging was performed according to international guidelines, and all individuals had localized tumors without signs of systemic disease. The study was conducted at the UH-RWTH in accordance with the requirements of the Institutional Review Board of the RWTH-Aachen University (EK 115/20), the current version of the Declaration of Helsinki, and the good clinical practice guidelines (ICH-GCP).

Staging and surgical technique

All patients who were referred for surgical treatment to our institution underwent a detailed clinical work-up as previously
described [2, 4]. Therefore, the number, size, and location of tumor nodules as well as the presence of distant metastases were evaluated by magnetic resonance imaging (MRI) or computed tomography (CT). The preoperative risk assessment was carried out based on the American Society of Anesthesiologists (ASA) and the Eastern Cooperative Oncology Group (ECOG) performance status, calculation of the future liver remnant (FLR), as well as parenchymal liver function as assessed by standard laboratory parameters and the LiMax test (Humedics® GmbH, Berlin, Germany) [22]. Non-invasive liver function tests were routinely carried out, but no preoperative liver biopsies were obtained to assess the quality of the liver parenchyma. Patients staged BCLC A to BCLC C without any evidence of extrahepatic spread as well as compensated liver function were considered candidates for surgery as primary treatment. The definitive decision for hepatectomy was made by a staff hepatobiliary surgeon and approved by the institutional interdisciplinary tumor board in every patient. Liver resection was carried out in accordance with common clinical standards [2, 4]. In brief, an intraoperative ultrasound was performed to visualize the local tumor spread and other suspicious lesions. The decision for either anatomic resections or non-anatomic atypical wedge resections with an adequate resection margin was based on the surgeon’s preference. Parenchymal transection was carried out using the Cavitron Ultrasonic Surgical Aspirator (CUSA®, Integra LifeSciences®, Plainsboro NJ, USA) with low CVP and intermittent Pringle maneuvers if necessary in open hepatectomy. In laparoscopic hepatectomy, parenchymal transection was commonly performed by Thunderbeat® (Olympus K.K., Tokyo, Japan), Harmonic Ace® (Ethicon Inc. Somerville, NJ, USA), or laparoscopic CUSA (Integra LifeSciences, NJ, USA) in combination with vascular staplers (Echelon, Ethicon, Somerville, NJ, USA) or polymer clips (Teleflex Inc., PA, USA). The anesthesiologic management was based on a restrictive fluid intervention strategy ensuring a low central venous pressure (CVP) during parenchymal dissection.

### Statistical analysis

The primary endpoint of this study was to identify the staging system with the best prognostic ability for OS. Overall performance was defined by homogeneity (small differences in OS among patients within the same stage), discriminatory ability (great differences in OS among patients within different stages), and monotonicity of gradients (longer OS in patients in earlier stages than in more advanced stages within the same system) as previously described [23]. Therefore, Cox regression models of each staging systems were established and subsequently used to calculate the likelihood ratio (LR) $\chi^2$ to determine homogeneity, linear trend (LT) $\chi^2$.

| Synopsis of staging systems for hepatocellular carcinoma | Radiologic characteristics | Laboratory characteristics | Patient characteristics |
|----------------------------------------------------------|----------------------------|----------------------------|------------------------|
| Stage system | Nodule size and/or count or tumor burden | Vascular invasion | Encephalopathy | Performance status | Metastasis | Ascites | Laboratory values | Enzymes and biomarkers | AFP | AP | Bilirubin | INR | PIVKA II | Albumin | PT |
| Milan | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| BCLC | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| HKLC | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| Okuda | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| CLIP | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| ITA.LI.CA staging | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| ITA.LI.CA score | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| MESH | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| GRETCH | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |

The included characteristics of each staging system are displayed [12, 13, 16-21]. Milan criteria were included for reference reasons. AFP, alpha-fetoprotein; AP, alkaline phosphatase; BCLC, Barcelona Clinic Liver Cancer; CLIP, Cancer of Italian Liver Program; GRETCH, Groupe d’Etude et de Traitement du Carcinome Hepatocellulaire; ITA.LI.CA, Italian Liver Cancer; HKLC, Hong Kong Liver Cancer; MESH, model to estimate survival in hepatocellular carcinoma.
to assess discriminatory ability, and both LR $\chi^2$ and LT $\chi^2$ to measure monotonicity of gradients as well as Akaike Information Criterion (AIC) to describe the explanatory ability of the particular staging system [23]. The degrees of freedom were set to 1 in all calculations to allow the comparison of prognostic systems with a different total number of stages. Higher LR $\chi^2$ and LT $\chi^2$ as well as lower AIC indicate a better fitting model to predict OS in this statistical approach. Milan criteria were included in the analysis for reference reasons. Data derived from continuous variables are presented as median and interquartile range. Survival curves were generated by the Kaplan-Meier method and compared with the log-rank test. Median follow-up was accessed with the reverse Kaplan-Meier method. Complications are reported as in-hospital morbidity and in-hospital mortality. Perioperatively deceased patients were included in all survival analyses. The level of significance was set to $p < 0.05$, and $p$ values are given for two-sided testing. Analyses were performed using SPSS Statistics 24 (IBM Corp., Armonk, NY, USA).

**Results**

**Preoperative, operative, and postoperative data**

A total of 160 patients with a median age of 68 years and median body mass index (BMI) of 26 kg/m$^2$ who underwent curative-intent surgery for HCC at our institution from 2011 to 2019 were included in this study with more than half of the study cohort (60.0%, 96/160) belonging to the performance status category ASA III or higher. The vast majority of the patients were categorized as Child-Pugh A (93.1%, 149/160) with a median Child-Pugh score (CPS) of 5 and a median model for end-stage liver disease (MELD) score of 6. The median nodule count was 1 (range 1–7), and the median largest tumor diameter is 55 mm (range 6–228 mm). A tumor burden > 50% was detected in 5.0% (8/160) of the patients, while an overall invasion to major vessels was observed in 26.3% (42/160). Of all patients, 3.1% (5/160) underwent transarterial chemoembolization (TACE) and 1.3% (2/160) transarterial radioembolization (TARE) prior to surgery. A minority of the individuals treated for HCC (27.5%, 44/160) underwent laparoscopic liver resection, and the median operative time was 207 min. R0 resection was achieved in 95.6% (153/160) of the patients. Median hospital stay was 9 days. No complications were detected in 47% (75/160) of the patients. In contrast, 46 patients (28.8%) experienced major postoperative complications (Clavien-Dindo $\geq$ 3) and 10 patients (6.3%) deceased perioperatively. More clinicopathological and perioperative characteristics are outlined in Table 2, and a detailed overview of the applied staging systems (Milan criteria, BCLC, HKLC, Okuda, CLIP, ITA.LI.CA staging, Table 2  Clinical and perioperative characteristics

| Demographics |
|-------------|
| Gender, m/f (%) | 115 (71.9) / 45 (28.1) |
| Age (years) | 68 (60–75) |
| BMI (kg/m$^2$) | 26 (23–29) |
| Preoperative PVE, n (%) | 8 (5.0) |
| Preoperative TACE, n (%) | 5 (3.1) |
| Preoperative TARE, n (%) | 2 (1.3) |
| Preoperative TACE and TARE, n (%) | 1 (0.6) |
| ASA, n (%) | |
| I | 3 (1.9) |
| II | 61 (38.1) |
| III | 91 (56.9) |
| IV | 5 (3.1) |
| V | 0 |
| Liver disease, n (%) | |
| ALD | 37 (23.1) |
| NAFLD | 66 (41.3) |
| Viral | 43 (26.9) |
| Cryptogenic/others | 14 (8.8) |

**Preoperative liver function**

| | |
|---|---|
| MELD score | 6 (6–6) |
| AFP (ng/ml) | 9 (3–93) |
| Albumin (g/dl) | 4.0 (3.6–4.4) |
| AST (U/l) | 42 (30–62) |
| ALT (U/l) | 35 (23–56) |
| GGT (U/l) | 101 (52–205) |
| Total bilirubin (mg/dl) | 0.52 (0.40–0.80) |
| Platelet count (/nl) | 221 (170–282) |
| Alkaline Phosphatase (U/l) | 99 (73–139) |
| Prothrombin time (%) | 93 (85–100) |
| INR | 1.04 (0.98–1.10) |
| Creatinine (mg/dl) | 0.85 (0.70–1.04) |
| Hemoglobin (g/dl) | 13.3 (11.7–14.6) |
| Child-Pugh, n (%) | |
| A | 149 (93.1) |
| B | 11 (6.9) |
| C | 0 |
| Child-Pugh score | 5 (5–5) |

**Preoperative imaging features**

| Number of nodules | 1 (1–2) |
| Largest nodule diameter (mm) | 55 (38–81) |
| Tumor burden > 50%, n (%) | 8 (5.0) |
| Overall macrovascular invasion, n (%) | 42 (26.3) |
| Portal vein invasion, n (%) | 27 (16.9) |
| Extrahepatic vascular invasion, n (%) | 11 (6.9) |
| Portal vein thrombosis, n (%) | 8 (5.0) |
| Ascites, n (%) | 6 (3.8) |

**Operative data**

| Laparoscopic resection, n (%) | 44 (27.5) |
| Conversation rate, n (%) | 4 (9.1) |
Follow-up data is presented as median and interquartile range if not noted otherwise.

Postoperative data
Intensive care stay, days 1 (1–1)
Hospitalization, days 9 (6–15)
Postoperative complications, n (%) No complications 75 (46.9)
Clavien-Dindo I 20 (12.5)
Clavien-Dindo II 19 (11.9)
Clavien-Dindo IIIa 18 (11.3)
Clavien-Dindo IIIb 9 (5.6)
Clavien-Dindo IVa 7 (4.4)
Clavien-Dindo IVb 2 (1.3)
Clavien-Dindo V 10 (6.3)
PHELF 50-50 criteria*, n (%) 2 (1.3)
PHELF ISGLS*, n (%) 36 (22.5)
ISGLS grade, n (%) A 24 (66.7) B 5 (13.9) C 7 (19.4)
Postoperative blood transfusion 29 (18.7)
Postoperative FFP 13 (8.4)
Postoperative platelet transfusion 4 (2.6)
Follow-up data
Recurrence-free survival (months) 26 (16–34)
Overall survival (months) 39 (32–46)
Liver transplantation, n (%) 2 (1.3%)

---

### Table 3: Staging systems for HCC guiding clinical management and predicting survival

| Milan criteria | n (%) | Median OS (95% CI) | p value |
|----------------|-------|--------------------|---------|
| Yes            | 48 (30.0) | 58 (24–92) | .012 |
| No             | 112 (70.0) | 31 (22–40) | |
| BCLC           | n (%) | Median OS (95% CI) | p value |
| 0*             | 6 (3.8) | 62 (40–85) | .001 |
| A              | 89 (55.6) | 55 (32–78) | |
| B              | 38 (23.8) | 23 (9–37) | |
| C              | 27 (16.9) | 15 (4–22) | |
| D              | 0         | n.a. | |
| HKLC           | n (%) | Median OS (95% CI) | p value |
| I              | 56 (35.0) | 48 (27–69) | .001 |
| IIa            | 2 (1.3) | n.a. | |
| IIb            | 51 (31.9) | 66 (44–88) | |
| IIIa           | 7 (4.4) | 3 (0–8) | |
| IIIb           | 33 (20.6) | 20 (11–29) | |
| IVa            | 11 (6.9) | 38 (13–61) | |
| IVb-Vb         | 0         | n.a. | |
| Okuda staging  | n (%) | Median OS (95% CI) | p value |
| I              | 138 (86.3) | 42 (28–56) | .001 |
| II             | 22 (13.8) | 12 (6–18) | |
| III            | 0         | n.a. | |
| CLIP           | n (%) | Median OS (95% CI) | p value |
| 0*             | 57 (50.4) | 91 (68–114) | .001 |
| 1              | 42 (37.2) | 38 (12–64) | |
| 2              | 7 (6.2) | 15 (4–26) | |
| 3              | 6 (5.3) | 3 (0–9) | |
| 4              | 1 (0.9) | n.a. | |
| 5–6            | 0         | n.a. | |
| ITALLICA staging | n (%) | Median OS (95% CI) | p value |
| 0*             | 6 (3.8) | 63 (40–85) | .016 |
| A              | 41 (25.6) | 58 (33–83) | |
| B1             | 52 (32.5) | 41 (23–59) | |
| B2             | 15 (9.4) | 33 (0–72) | |
| B3             | 35 (21.9) | 20 (13–27) | |
| C              | 11 (6.9) | 38 (13–63) | |
| ITALLICA score | n (%) | Median OS (95% CI) | p value |
| 0              | 1 (0.9) | n.a. | .001 |
| 1              | 10 (8.8) | 42 (36–48) | |
| 2              | 27 (23.9) | 130 (22–238) | |
| 3              | 21 (18.6) | 41 (0–84) | |
| 4              | 22 (19.5) | 38 (0–77) | |
| 5              | 18 (15.9) | 17 (9–25) | |
| 6              | 3 (2.7) | n.a. | |
| 7              | 7 (6.2) | 8 (5–11) | |
| 8              | 3 (2.7) | 6 (0–12) | |
| 9              | 1 (0.9) | 2 (0–2) | |
| 10–13          | 0         | n.a. | |
| MESH           | n (%) | Median OS (95% CI) | p value |
| 0*             | 19 (16.8) | 88 (63–114) | .001 |

---

Data presented as median and interquartile range if not noted otherwise.

Follow-up data is presented as median and 95% CI. *Postoperative liver failure was assessed by the 50-50-criteria and the ISGLS definition [11, 24]. ALD alcoholic liver disease, ALT alanine aminotransferase, ASA American Society of Anesthesiologists classification, AST aspartate aminotransferase, BCLC Barcelona Clinical Liver Cancer staging system, BMI body mass index, CI confidence interval, FFP fresh frozen plasma, GGT gamma glutamyltransferase, INR international normalized ratio, ISGLS International Study Group of Liver Surgery, MELD model of end-stage liver disease, MWA microwave ablation, NAFLD non-alcoholic fatty liver disease, PHLF post-hepatectomy liver failure, PVE portal vein embolization, TACE transarterial chemoembolization, TARE transarterial radioembolization.
The BCLC staging system is traditionally considered to guide treatment decision in European patients and provides the basis of the current guideline of the European Society for Medical Oncology (ESMO) [25]. BCLC assigns early stage HCC in patients with compensated liver function and good performance status to surgical therapy, while more advanced disease stages or more compromised individuals are referred to ablative and locoregional therapies or palliative treatment [12, 13]. However, such conservative interpretation of the BCLC staging has to be viewed critically in the era of modern HCC surgery using minimal invasive techniques and novel liver function tests which resulted in significantly improved patient selection and perioperative outcomes [6–11]. Correspondingly, two multicenter studies have shown that liver resection for HCC patients results in survival benefit over and subset of patients (n = 113) for all staging systems separately (Table 4).

In the overall cohort, BCLC performed best among the analyzed staging system with a LR $\chi^2$ of 20.48, LT $\chi^2$ of 13.27, and AIC of 764.21 outreaching all other staging systems in each criterion. In contrast, Milan criteria showed the lowest LR $\chi^2$ (6.21) and highest AIC (775.91), while HKLC was the staging system with the lowest LT $\chi^2$ (4.35). In the sub cohort, ITA.LI.CA score presented with the highest LR $\chi^2$ (30.08) and lowest AIC (455.27) of all staging systems. The highest LT $\chi^2$ (18.95) was determined for CLIP. Similar to the overall cohort, HKLC (LR $\chi^2$ of 9.23, LT $\chi^2$ of 2.25, and AIC of 475.78) and Milan criteria (LR $\chi^2$ of 1.42, LT $\chi^2$ of 1.27, and AIC of 480.30) were the least fitting models to predict OS. More details regarding the different staging systems are outlined in Table 4.

**Discussion**

HCC represents one of the major global health issues with liver resection being the treatment of choice in patients with compensated liver function [1–3]. Given this importance of the disease, a variety of staging systems reflecting oncological prognosis and guiding treatment decisions have been proposed, but no international consensus has been achieved which staging system should be preferred [12, 13, 16–21]. In a European cohort of patients, we were able to demonstrate a superiority of ITA.LI.CA score and CLIP over various other staging systems in their prognostic ability for OS after surgical resection. Our data does further suggest that staging systems incorporating biochemical markers of tumor biology (AFP) provide more solid estimates for OS in surgical patients than staging systems focusing on radiological characteristics only. This suggests ITA.LI.CA score and CLIP as the preferable staging systems for preoperative risk assessment balancing oncological outcome with perioperative risks in patients with HCC scheduled for liver surgery.

The BCLC staging system is traditionally considered to guide treatment decision in European patients and provides the basis of the current guideline of the European Society for Medical Oncology (ESMO) [25]. BCLC assigns early stage HCC in patients with compensated liver function and good performance status to surgical therapy, while more advanced disease stages or more compromised individuals are referred to ablative and locoregional therapies or palliative treatment [12, 13]. However, such conservative interpretation of the BCLC staging has to be viewed critically in the era of modern HCC surgery using minimal invasive techniques and novel liver function tests which resulted in significantly improved patient selection and perioperative outcomes [6–11]. Correspondingly, two multicenter studies have shown that liver resection for HCC patients results in survival benefit over
medical or interventional therapy regardless of their BCLC stage [14, 15]. These observations are further reassured by a randomized trial indicating better survival in BCLC B patients undergoing liver resection compared to TACE [26]. Therefore, more patients with higher BCLC stages are nowadays considered candidates for surgery provided that severe liver dysfunction and a significantly impaired performance status are absent [14, 15]. Despite this accepted expansion of the BCLC criteria, there is lacking evidence whether higher BCLC stages—which were originally not developed to predict survival in patients undergoing surgery—can be used to provide a basis for the selection of surgical candidates. Interestingly, our data does suggest significant differences in survival in between patients categorized BCLC 0/A compared to BCLC B/C but less discriminative value of the staging system in BCLC B and C categories (Fig. 2b).
One might argue that for patients undergoing surgery, the pathology-based Union for International Cancer Control (UICC) staging systems provides an excellent stratification for estimated postoperative OS. The TNM staging usually does predict OS well in HCC patients but does not take the underlying liver disease into account which limits its overall prognostic ability [27]. Therefore, some staging systems do incorporate the pathological TNM staging and add individual patient characteristics to overcome this major limitation (e.g.,

Japanese Integrated System (JIS) or Chinese University Prognostic Index (CUP)) [28, 29]. However, pathological staging only allows a post hoc assessment and is not available for the preoperative decision-making and patient selection. We therefore decided not to include staging systems requiring data based on postoperative pathological examinations into our present analysis.

Similar to the BCLC system, all of the reported staging systems are originally designed to cover the whole disease
spectrum of HCC but not exclusively to predict OS in surgical candidates [12, 13, 16–21]. Furthermore, the initial publications regarding development and validation of the different prognostic staging systems are based on large heterogeneous cohorts using various treatment approaches including palliation. This explains the observation that advanced stages of some staging systems are not represented in our analysis (Table 3, Fig. 2). In addition, previous literature predominantly compares various prognostic scores using patient cohorts with a broad disease spectrum and different treatment modalities and, therefore, might have limited use for the selection and stratification of surgical candidates [30]. Nonetheless, to the best of our knowledge, this analysis is the first report in the literature comparing a plethora of staging systems in a European cohort of HCC patients undergoing curative-intent surgery. Based on our comparative analysis, we identified the ITA.LI.CA score and CLIP as preferable staging system for patients scheduled for liver resection. CLIP was introduced in 1998 and can be considered an advancement over the older Okuda staging which was published in 1985 [19, 20]. CLIP basically uses the same set of variables as the Okuda staging but added AFP and the presence of portal vein thrombosis to the assessed patient characteristics [20]. Despite being relatively old and simple compared to some more novel staging systems, CLIP performed well against other staging systems in comparative analyses. In particular, CLIP outranked BCLC, HCLC, JIS, GRETCH, CUPI, Okuda staging, and TNM staging in its overall prognostic performance in a large Taiwanese study with 3000 patients [31]. Another study from China also demonstrated a higher prognostic value of CLIP for 3- and 6-month OS compared to other staging systems [32]. Based on the small variable set included into calculating the score (tumor volume compared to liver volume, Child-Pugh category, AFP, and the presence of portal vein thrombosis), CLIP is feasible to be used in surgical candidates and does display decent prognostic ability in our current study. Although showing the best discriminatory ability among the investigated staging systems, CLIP was inferior to the novel ITA.LI.CA score in terms of homogeneity and overall explanatory ability. ITA.LI.CA is a complex system-based ITA.LI.CA staging which stratifies patients with respect to size and number of tumor nodules, vascular invasion, and metastasis into four main and some sub-stages [18]. Interestingly, ITA.LI.CA staging performed inferior to the standard BCLC staging in our analysis. The ITA.LI.CA score utilizes the ITA.LI.CA staging and adds functional status, Child-Pugh score, and AFP to calculate a score ranging from 0 to 13 corresponding to overall prognosis in HCC patients [18]. Of note, ITA.LI.CA score has already been validated with an external cohort of patients in a study analyzing 1500 patients undergoing various treatments and showed prognostic superiority over CLIP, HKLC, JIS, ITA.LI.CA staging, and BCLC [33].

Despite showing the best mathematical abilities to predict survival, ITA.LI.CA and CLIP are certainly not perfect from a
theoretical point of view as illustrated by the fact that patients with moderate HCC stages could have inferior OS compared to patients with higher cancer stages especially in the ITA.LI.CA staging (Fig. 2). This underlines that staging systems are helpful regarding patient selection but are just one of multiple characteristics guiding decision-making in this complex disease.

Interestingly, the two staging systems which showed the best prognostic ability (CLIP and ITA.LI.CA score) in our setting were originally developed using European patient cohorts. In contrast, the HKLC staging which is based on Asian HCC patients performed only slightly better than the Milan horts. In contrast, the HKLC staging which is based on Asian setting were originally developed using European patient co-

systems are helpful regarding patient selection but are just one of multiple characteristics guiding decision-making in this complex disease.

Interestingly, the two staging systems which showed the best prognostic ability (CLIP and ITA.LI.CA score) in our setting were originally developed using European patient cohorts. In contrast, the HKLC staging which is based on Asian HCC patients performed only slightly better than the Milan criteria which we have included in our analysis for reference reasons [13, 21]. These observations may suggest a potential difference between Asian and European patient cohorts and its impact on the prognostic ability of the various staging systems. General disease etiology and even genomic characteristics vary between Asian and European patients [34]. Also, the general approach to HCC seems to be more aggressive in Asian cohorts. This might partially be explained by the larger proportion of viral etiology in Eastern patients which results in a generally younger HCC population with often less severe underlying cirrhosis and fewer comorbidities [35]. Therefore, staging systems developed for European cohorts might be more suitable for European patients. The same is true for Eastern patients as recently demonstrated within a large Singaporean cohort. In a comparative analysis of Selby et al. comprising 716 patients, HKLC showed a better performance in guiding treatment compared to BCLC [36]. These considerations do also imply limitations when the results of comparative analyses of staging systems in Eastern patients are directly transferred to Western HCC patients.

Among the staging systems that do not include AFP to correlate radiological and clinical patient characteristics with tumor biology, BCLC showed good results in our cohort. However, in our subgroup of patients with available information on AFP levels, CLIP and ITA.LI.CA score provided a better overall staging performance. Based on this, AFP seems to be a major contributor for accurate staging of HCC patients undergoing surgery. AFP is a known predictor of OS in various clinical situations of HCC patients and characteristics of the tumor [37]. Thus, it is not surprising that staging systems incorporating AFP might be superior in their overall prognostic performance. This observation does further underline the importance of tumor biology and the individual genetic pathogenesis of HCC. Nault et al. have recently proposed a gene score including 5 genes to predict OS and demonstrated significant prognostic accuracy in a surgical cohort of patients [38]. It is therefore important that future staging systems integrate novel biomarkers to further increase the prognostic value of pretreatment staging in HCC patients.

Like any other clinical study, our analysis has certain inherent limitations. All HCC patients analyzed in this study underwent treatment in a monocentric setting reflecting our individual clinical approach to this particular disease, and the study is based on a retrospective data collection which was not obtained in a controlled prospective clinical trial. This also results in large proportion of ASA III patients and individuals with higher BCLC stages due to our liberal department policy. Further, our data set appears small compared to some other studies especially from Asian cohorts. Most importantly, however, the majority of studies focusing on staging systems for HCC comprise heterogeneous cohorts in which curative treatments are carried out in the minority of patients.

Notwithstanding the aforementioned limitations, we here provide a detailed analysis of a plethora of HCC staging systems in a European cohort of patients who underwent curative-intent liver resection, demonstrating ITA.LI.CA score and CLIP to be the most suitable staging systems for surgical candidates.

Conclusion

All staging systems assessed showed certain discriminatory ability regarding OS of patients undergoing liver resection for HCC. However, the ITA.LI.CA score and CLIP demonstrated a superior prognostic ability compared to other staging systems in our European cohort.

Authors’ contributions All authors contributed significantly to this manuscript and are in agreement with the content. The authors contributed as follows: study conception and design, JB, UPN, and SAL; acquisition of data, JB, ZC, ZC, DH, KJ, and PB; analysis and interpretation of data, JB, TL, LH, TFU, SAL, and UPN; drafting of manuscript, JB, UPN, and SAL; and critical revision of manuscript, ZC, DH, KJ, TL, LH, PB, and TFU.

Funding Open Access funding enabled and organized by Projekt DEAL.

Compliance with ethical standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was conducted at the UH-RWTH in accordance with the requirements of the Institutional Review Board of the RWTH-Aachen University (EK 115/20). Informed consent was obtained from all individual participants included in the study. All authors declare no conflict of interest.

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a
Banting S, Usatoff V, Nagino M, Maddern G, Hugh TJ, Vauthey JN, Greig P, Rees M, Yokoyama Y, Fan ST, Nimura Y, Figueiras J, Capussotti L, Buchler MW, Weitz J (2011) Posthepatectomy liver failure: a definition and grading by the International Study Group of Liver Surgery (ISGLS). Surgery 149(5):713–724. https://doi.org/10.1016/j.surg.2010.10.001

25. Vogel A, Cervantes A, Chau I, Daniele B, Llovet JM, Meyer T, Nault JC, Neumann U, Ricke J, Sangro B, Schirmacher P, Verslype C, Zech CJ, Arnold D, Martinelli E, clinicalguidelines@esmo.org EGCea (2018) Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol: Off J Eur Soc Med Oncol / ESMO 29(Suppl 4):iv238–iv255. https://doi.org/10.1093/annonc/mdy308

26. 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. https://doi.org/10.1016/j.jhep.2014.03.012

27. Marsh JW, Dvorchik I, Bonham CA, Iwatsuki S (2000) Is the pathologic TNM staging system for patients with hepatoma predictive of outcome? Cancer 88(3):538–543. https://doi.org/10.1002/(sici)1097-0142(20000201)88:3<538::aid-cncr7>3.0.co;2-h

28. 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. https://doi.org/10.1002/cncr.10384

29. 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. https://doi.org/10.1007/s005350300038

30. Tellapuri S, Sutphin PD, Beg MS, Singal AG, Kalva SP (2018) Staging systems of hepatocellular carcinoma: a review. Indian J Gastroenterol: Off J Indian Soc Gastroenterol 37(6):481–491. https://doi.org/10.4103/ijg.ijg.2018.21664-018-0915-0

31. Liu PH, Hsu CY, Hsia CY, Lee YH, Su CW, Huang YH, Lee FY, Lin HC, Huo T (2016) Prognosis of hepatocellular carcinoma: assessment of eleven staging systems. J Hepatol 64(3):601–608. https://doi.org/10.1016/j.jhep.2015.10.029

32. Chen ZH, Hong YF, Lin J, Li X, Wu DH, Wen JY, Chen J, Ruan DY, Lin Q, Dong M, Wei L, Wang TT, Lin ZX, Ma XK, Wu XY, Xu R (2017) Validation and ranking of seven staging systems of hepatocellular carcinoma. Oncol Lett 14(1):705–714. https://doi.org/10.3892/ol.2017.6222

33. Borzio M, Dionigi E, Rossini A, Marignani M, Sacco R, De Sio I, Bertolini E, Francica G, Giacomini A, Parisi G, Vicari S, Tolda A, Salmi A, Boccia S, Mitra F, Fornari F (2018) External validation of the ITA.LI.CA prognostic system for patients with hepatocellular carcinoma: a multicenter cohort study. Hepatology 67(6):2215–2225. https://doi.org/10.1002/hep.29662

34. Choo SP, Tan WL, Goh BKP, Tai WM, Zhu AX (2016) Comparison of hepatocellular carcinoma in Eastern versus Western populations. Cancer 122(22):3430–3446. https://doi.org/10.1002/cncr.30237

35. Vibert E, Ishizawa T (2012) Hepatocellular carcinoma: Western and Eastern surgeons’ points of view. J Visceral Surg 149(5):e302–e306. https://doi.org/10.1016/j.jviscsurg.2012.05.001

36. Selby LK, Tay RX, Woon WW, Low JK, Wei B, Shetal VG, Pang TC, Jannarkar SP (2017) Validity of the Barcelona Clinic Liver Cancer and Hong Kong Liver Cancer staging systems for hepatocellular carcinoma in Singapore. J Hepato-Biliary-Pancreat Sci 24(3):143–152. https://doi.org/10.1002/jhbp.423

37. Bai DS, Zhang C, Chen P, Jin SJ, Jiang GQ (2017) The prognostic correlation of AFP level at diagnosis with pathological grade, progression, and survival of patients with hepatocellular carcinoma. Sci Rep 7(1):12870. https://doi.org/10.1038/s41598-017-12834-1

38. Nault JC, De Reynies A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, Decaens T, Franco D, Imbeaud S, Rousseau F, Azoulay D, Sarie J, Blanc JF, Balabaud C, Bioulac-Sage P, Laurent A, Laurent-Puig P, Llovet JM, Zucman-Rossi J (2013) A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. Gastroenterology 145(1):176–187. https://doi.org/10.1053/j.gastro.2013.03.051

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.