The Challenge of Prognosis and Staging for Hepatocellular Carcinoma

JORGE A. MARRERO,[a] MASATOSHI KUDO,[b] JEAN-PIERRE BRONOWICKI[c]

[a]Multidisciplinary Liver Tumor Program, University of Michigan, Ann Arbor, Michigan, USA; [b]Department of Gastroenterology and Hepatology, Kinki University School of Medicine, Osaka, Japan; [c]Department of Gastroenterology and Hepatology, INSERM U954, University Hospital of Nancy, University Henri Poincaré, Vandœuvre-lès-Nancy, France

ABSTRACT

Hepatocellular carcinoma (HCC) is a heterogeneous condition, with multiple confounding factors making patient assessment extremely complex. Tumor burden, the presence of symptoms, liver function, and comorbidities must all be considered to ensure accurate patient assessment, thereby providing physicians with a common language on which to base treatment decisions and guide research. Although many staging classifications have been developed, there is no consensus on the best classification to use. The Barcelona Clinic Liver Cancer system is a promising candidate for a standard western classification, because it has been externally validated and is endorsed by the European Association for the Study of the Liver and the American Association for the Study of Liver Diseases. Similarly, the biomarker-combined Japanese Integrated Staging (JIS) score is the most promising candidate for a standard Asia-Pacific classification, because it has been externally validated and shown to be superior to conventional JIS. Because risk factors vary significantly by region, so too does the predictive power of current staging classifications; any standard global staging classification would need to be validated in both western and Asia-Pacific patients. To date, no such globally validated classification exists. Findings from scientific research have improved our understanding of HCC and enabled us to refine current classifications. The role of tumor markers to predict survival was recently reported, and α-fetoprotein, *lens culinaris* agglutinin-reactive α-fetoprotein, and des-γ-carboxyprothrombin have now been incorporated into some classifications. Molecular markers have also been linked with poor outcomes and will likely play a role in future classifications. Although more work is required, it is hoped that these and other ongoing research efforts will eventually enable the development of a global staging classification. *The Oncologist* 2010;15(suppl 4):23–33

INTRODUCTION

Hepatocellular carcinoma (HCC) is a heterogeneous condition with multiple variables that vary from region to region, complicating diagnosis, prognosis, and treatment recommendations. The presence of comorbidities is a common confounding factor that can compromise liver function and affect outcomes. For example, 80% of patients present with liver cirrhosis [1] and 85.5% of patients are carriers of either hepatitis B virus (HBV), which is particularly prevalent in Africa and Asia, or hepatitis C virus (HCV), prevalent in western countries and Japan [2, 3]. The characteristics of HCC also vary with geographic location. In rural South Africa, HCC is commonly diagnosed at a more advanced stage than in North America [4]. Because HBV is often acquired at an early age in Africa and Asia, HCC may also develop in younger patients and in the absence of liver cirrhosis [5]. Conversely, in North America, many patients have long-term liver cirrhosis and subsequently develop HCC. Clinical presentation in these patients is therefore dominated by complications of cirrhosis. These confounding factors mean that multiple variables must be considered when assessing patients with HCC.

The aims of HCC staging classifications are to: stratify patients to determine their overall survival (OS) probability prior to treatment, facilitate treatment, and enable objective comparison among the outcomes of research studies. What separates HCC from other solid tumors is that the presence of chronic liver disease and cirrhosis affects OS and the ability to treat this tumor. Therefore, liver disease is a very important variable, together with the overall health of the patient [6, 7]. In considering all these variables, it is hoped that accurate and consistent assessment of all patients can be achieved, thereby providing a common language for physicians as well as the broader multidisciplinary team. This, in turn, should facilitate appropriate treatment selection and ensure optimum patient management. However, with >15 HCC staging classifications available, each measuring a range of different factors and developed in different patient populations, physicians are faced with the complicated task of choosing which classification to use.

In this article, we review the major HCC staging classifications used globally and examine the factors assessed, as well as how each of the staging classifications was developed and validated. We also provide an overview of comparisons among various staging classifications reported in the literature. The paper does not aim to assess the relative values of individual classifications, nor to provide any endorsement of one system over another. However, we suggest possible areas for improvements that are necessary if we are to achieve a globally applicable HCC staging classification.

STAGING CLASSIFICATIONS IN HCC

The factors influencing the development of HCC and its disease course vary considerably from region to region. As a result, various staging classifications have been developed that take into account a range of factors (Table 1), and although some classifications appear to be effective across broad regions, such as western or Asia-Pacific patient populations, others have been evaluated only in a single country. However, there is no globally applicable staging classification, and thus no common language on which to base treatment decisions and guide research.

Tumor–Node–Metastasis Staging System

The first staging classification for solid tumors was developed >50 years ago by the French surgeon Pierre Denoix [8]. In 1968, his recommendations for various tumors were compiled and published by the International Union Against Cancer and the American Joint Committee on Cancer in the first edition of the tumor–node–metastasis (TNM) staging system. Since then, this staging classification has undergone several amendments, and the most recent, sixth edition, was published in 2003 [9, 10].

The TNM staging classification provides an assessment of solid tumors based only on size and extent of invasion. This is measured according to the size of the primary tumor (T), presence of tumor in the regional lymph nodes (N), and presence of metastatic spread beyond the lymph nodes (M). Assessment of TNM staging can be prior to treatment (clinical staging) or after surgery (pathologic staging) [8]. Clinical staging is performed using imaging procedures, but in patients with HCC, the presence of cirrhosis and/or swelling of the lymph nodes as a result of chronic liver disease may prevent accurate assessment. Pathologic staging is therefore needed, but this may not be possible in the majority of patients because very few undergo surgical therapies that allow appropriate sampling.

The prognostic value of the sixth edition of the TNM staging system was compared with three other staging classifications (the Okuda, Cancer of the Liver Italian Program [CLIP], and Chinese University Prognostic Index [CUPI] classifications) in 234 patients with HCC who underwent curative resection at the Southwestern Hospital in China. Both the Okuda and the TNM systems were better at stratifying patients according to survival than the CLIP or CUPI system. However, the TNM classification was also better for predicting prognosis than the three other classifications, and was significantly better than the CLIP score ($p < .05$) [11]. The sixth edition of the TNM staging system also proved to be more effective than six other classifications (the Okuda, Barcelona Clinic Liver Cancer [BCLC], Japanese Integrated Staging [JIS], CLIP, and Groupe d'Etude et

Table 1. Key characteristics of various staging classifications available to assess the prognosis of patients with hepatocellular carcinoma

Staging classification	Variables measured				Year published	Study
	Tumor staging	Liver function	Performance status	Serum tumor markers		
CLIP	Tumor morphology (uninodular and extension ≤50%, multinodular and extension ≤50%, massive or extension >50%), portal vein thrombosis	Child-Pugh	No	AFP	1998	CLIP Investigators [16]
BCLC	Tumor size, number of nodules, portal vein thrombosis	Child-Pugh, bilirubin, portal hypertension	PST	No	1999	Llovet et al. [24]
GRETCH	Portal vein thrombosis	Bilirubin, alkaline phosphatase	Karnofsky	AFP	1999	Chevret et al. [34]
U.S. nomogram	Resection margin status, tumor size >5 cm, satellite lesions, vascular invasion	No	Age, operative blood loss	AFP	2008	Cho et al. [45]
Okuda	Tumor size (</>50% of liver)	Ascites, albumin, bilirubin	No	No	1985	Okuda et al. [13]
CUPI	TNM fifth edition	Ascites, bilirubin, alkaline phosphatase	Presence of symptoms	AFP	2002	Leung et al. [37]
JIS	Japanese TNM fourth edition	Child-Pugh	No	No	2003	Kudo et al. [23]
bm-JIS	Japanese TNM fourth edition	Child-Pugh	No	AFP, AFP-L3, DCP	2008	Kitai et al. [39]
SLiDe	Stage and liver damage categories from the Japanese TNM fourth edition	No	No	DCP	2004	Omagari et al. [41]
Tokyo	Size and number of tumors	Albumin, bilirubin	No	No	2005	Tateishi et al. [42]
BALAD	No	Albumin, bilirubin	No	AFP, AFP-L3, DCP	2006	Toyoda et al. [44]
ALCPS	Tumor size, portal vein thrombosis, lung metastases	Ascites, Child-Pugh, alkaline phosphatase, bilirubin, urea	Abdominal pain, weight loss	AFP	2008	Yau et al. [46]

Abbreviations: AFP, α-fetoprotein; AFP-L3, *lens culinaris* agglutinin-reactive AFP; ALCPS, Advanced Liver Cancer Prognostic System; BALAD, bilirubin, albumin, AFP-L3, AFP, DCP; BCLC, Barcelona Clinic Liver Cancer; bm-JIS, biomarker-combined JIS; CLIP, Cancer of the Liver Italian Program; CUPI, Chinese University Prognostic Index; DCP, des-γ-carboxyprothrombin; GRETCH, Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire; JIS, Japanese Integrated Staging; PST, performance status test; SLiDe, Stage, Liver damage, DCP; TNM, tumor–node–metastasis. From Meier V, Ramadori G. Clinical staging of hepatocellular carcinoma. Dig Dis 2009;27:131–141. Reproduced with permission from S. Karger AG, Basel, Switzerland.

de Traitement du Carcinome Hépatocellulaire [GRETCH] classifications) at assessing prognosis in 163 patients with HCC following resection in a retrospective study at a single institution in Korea [12]. Those studies were limited to the postsurgery setting, and evaluation in a larger sample size and broader patient population is still required.

Okuda Classification

The Okuda classification was published in 1985 and was the first staging system to include parameters related to tumor size (>50% versus <50% of the liver involved) and liver function (albumin, ascites, bilirubin) [13]. Its ability to predict prognosis according to treatment was evaluated as part of a retrospective analysis among 850 patients treated in three different institutes in Japan, with patients stratified into three stages (I, not advanced; II, moderately advanced; III, very advanced). These findings showed that surgically treated patients had a longer survival time than medically treated patients, and that medical treatment prolonged survival in stage II and stage III patients but not in stage I pa-

tients. However, because stage at diagnosis as well as the available medical interventions have moved on since the time this staging classification was developed, stratifying patients to receive radical or palliative therapies using this system alone would not be appropriate. Moreover, although its simplicity makes it clinically attractive, its ability to predict prognosis is relatively modest [5]. Indeed, in a retrospective study in Canada, the Okuda classification failed to identify two thirds of the 37 patients with a poor prognosis who were identified by the CLIP criteria [14]. Furthermore, in an evaluation of staging systems for HCC patients undergoing surgery, the Okuda system was not superior to TNM staging [15].

CLIP Scoring

The CLIP scoring system was derived from a retrospective analysis of 435 patients with HCC from 16 Italian institutions and was published in 1998 [16]. Here, four independent predictive factors of survival were identified (Child-Pugh score, tumor morphology, α-fetoprotein [AFP], and

portal vein thrombosis), and a simple linear scoring system (0, 1, or 2) was assigned to the covariates in order to give patients a total score of 0–6. This scoring system was subsequently validated by the same group in a prospective trial of 196 patients with HCC and cirrhosis [17] and was also shown to be effective in predicting survival among a group of 145 patients in the Middle East [18] and in 662 Japanese patients [19]. However, whereas the median survival time associated with each CLIP score (0–6) appears to be similar between patients included in the prospective validation conducted by the founding group and those included in the study conducted in the Middle East [18, 20], the median survival times reported for Japanese patients were higher for all CLIP scores [19], and it has been suggested that these findings could compromise the external validation of the CLIP scoring system [21].

In a comparison of the CLIP, BCLC, and Okuda staging systems using a pooled database from two randomized trials of French patients with mainly alcoholic HCC, the performances of all three systems were disappointing; different systems performed differently according to patient populations and for individual prognostic factors. None clearly emerged as an unquestionable reference [22]. However, for all statistics, the CLIP system had better prognostic ability. The authors concluded that the CLIP staging seems to be most adapted to the palliative setting and that it could be improved by associating World Health Organization performance status.

A number of limitations of the CLIP scoring system have been reported [23]. First, the tumor morphology categories used may be too general to be globally applicable, particularly in countries such as Japan, where more patients are diagnosed with very small solitary tumors, largely because of the established screening programs in place. Secondly, although patient populations with different CLIP scores appear to be well discriminated from each other, there is no clear difference among patient populations with CLIP scores of 4–6 [17]. Indeed, in the prospective validation of this scoring system performed by the founding group [20], they grouped patients with a CLIP score of 4–6 into one group. Finally, all studies evaluating the CLIP score reported to date show that a high proportion of patients are categorized as CLIP score 0–2, suggesting poor stratification ability with this system.

Taken together, these findings suggest that, although the CLIP scoring system is associated with a good prognostic ability, this staging system may not be sensitive enough to be applicable to all patient populations and cannot easily be applied to a patient's management.

BCLC Staging

The BCLC staging classification was proposed by Llovet and colleagues in 1999 [24]. One of the most important observations for the development of the BCLC staging system came from the follow-up of patients with nonresectable and nontransplantable HCC who were randomized to placebo in two different clinical trials [25]. In that study, the multivariate analysis identified performance status, constitutional syndrome, vascular invasion, and extrahepatic spread as independent predictors for mortality. The authors showed that the 1-, 2-, and 3-year survival rates for the 48 patients without predictors of mortality (i.e., intermediate stage) were 80%, 65%, and 50%, respectively, and these were 29%, 16%, and 8% in the 54 patients with at least one adverse factor (i.e., advanced stage). This has been externally validated [26]. This allowed patients to be divided into different categories based on tumor stage (tumor size, number of nodules, and presence of portal vein thrombosis), liver function (Child-Pugh score, portal hypertension, bilirubin level), physical status (performance status test), and cancer-related symptoms. Furthermore, four categories were created (A, early; B, intermediate; C, advanced; D, end-stage disease). It is also unique in that it is the only system that provides treatment recommendations for each of the assigned stages based on the best treatment options currently available. The BCLC staging classification has been externally validated in the U.S. [6], Europe [27, 28], and Taiwan [29] and has demonstrated superior survival stratification and prognosis prediction over a range of other classifications, including the Okuda, TNM, CLIP, GRETCH, CUPI, and JIS classifications [28, 30]. Moreover, BCLC staging is endorsed by both the European Association for the Study of the Liver (EASL) [5] and the American Association for the Study of Liver Diseases (AASLD) [31], and it is emerging as a standard staging classification in western populations [32]. The most important aspect of this staging classification is that it is linked to an evidence-based treatment algorithm and can easily be used in a clinical setting. However, it should be noted that, in a study investigating which of the available staging systems was the most informative for the medical oncologist [33], the BCLC system was found to be less informative than the GRETCH and CLIP classifications when ranked using a concordance index, a likelihood ratio, and the Akaike information criterion. However, that study mostly evaluated patients with advanced tumors and may not be generally applicable.

GRETCH Scoring

The GRETCH scoring system was based on findings from a prospective study among 761 patients from 24 western medical centers and was published in 1999 [34]. The aim of

the study was to compile a classification system for predicting survival among these patients using a multivariate Cox model. Five prognostic factors were selected (Karnofsky index <80%, bilirubin >50 μmol/l, alkaline phosphatase $\geq$2$\times$ the upper limit of normal, AFP >25 μg/l, and ultrasonographic portal vein obstruction) in order to divide patients in the study training sample ($n = 506$) into three prognostic classification groups (A, B, C). The 1-year survival rates associated with these three groups were derived (72%, 34%, and 7% for groups A, B, and C, respectively) and independently validated in the study test sample (79%, 31%, and 4% for groups A, B, and C, respectively; $n = 255$). This system has not been validated in nonwestern patient populations. Furthermore, because this system originated from a multivariate analysis, it may not be reproducible or easily used in clinical practice.

Liver Cancer Study Group of Japan TNM Staging

In 1965, the Liver Cancer Study Group of Japan (LCSGJ) started a nationwide registration of clinicopathologic and prognostic data from patients with primary liver cancer, and using data collected in this database they introduced the Japanese version of the TNM staging system in 1983. This has subsequently undergone a number of revisions, and in 2007 the LCSGJ evaluated data from their database of 63,736 patients with primary liver cancer, 13,772 of whom underwent curative resection, in order to present evidence to develop and validate this staging classification [35]. Based on univariate and multivariate survival analyses, they selected three factors (vascular or bile duct invasion, tumor diameter $\leq$2 cm versus >2 cm, and number of tumors—single versus multiple), and classified patients as T1–T4 based on the number of adverse factors present (patients with none were considered T1, those with one were T2, those with two were T3, and those with three were T4). Significant survival differences were demonstrated among patients in each of the four assigned stages, with 5-year survival rates of 70% (T1), 58% (T2), 41% (T3), and 24% (T4) ($p < .0001$). A potential weakness of the LCSGJ staging system is that it assumes equal weight for growth pattern, size, and vascular or bile duct invasion. No external validation has been reported to date.

The Vauthey Simplified Staging System

In 2002, Vauthey and colleagues evaluated the efficacy of using the TNM's T categories to stratify patients according to survival and assessed a range of independent prognostic factors among 557 patients undergoing resection [36]. Independent predictors of death in that study were major vascular invasion, microvascular invasion, severe fibrosis/

cirrhosis of the liver, multiple tumors, and tumors >5 cm. Based on these findings, Vauthey and colleagues proposed a simplified model of patient stratification using vascular invasion, tumor number and size, and the effect of fibrosis on survival. Patients were divided into three stages (I, II, III) and these were associated with a significant survival difference, with 5-year survival rates of 55% (I), 37% (II), and 16% (III) ($p < .001$) [36]. This is limited to postsurgery patients and has not been externally validated.

CUPI Score

The CUPI score was developed at the Chinese University in Hong Kong and was published in 2002 [37]. In that study, 19 potential prognostic factors were evaluated in a multivariate analysis using a Cox regression model among 926 Chinese patients, mostly with HBV-associated HCC. From this, five additional prognostic factors (asymptomatic disease at presentation, AFP, total bilirubin, alkaline phosphatase, and ascites) were added to the fifth edition of the TNM staging classification. Patients were divided into three risk groups (high, medium, and low risk for dying within 3 months), and highly significant differences in survival were observed among these groups ($p < .00001$). Findings from that study also showed that the CUPI system was better at classifying patients into different risk groups than the TNM staging system alone, or the Okuda or CLIP scoring systems, although the authors advise that validation across broader patient populations is needed. In a more recent study, the CUPI staging system was compared with the Okuda, CLIP, and sixth edition of the TNM staging systems among 234 Chinese patients who underwent resection [11]. The authors concluded that the TNM sixth edition was superior in discriminating survival among patients stratified into different stages, and suggested that a possible limitation of the CUPI score is that it is based on the fifth edition of the TNM. The CUPI system has not been externally validated.

JIS Score

In 2003, an integrated prognostic classification system was published by Kudo and colleagues [23]. This scoring system combines the Japanese TNM staging (stages I, II, III, and IV are converted to scores 0, 1, 2, and 3, respectively) and the conventional Child-Pugh (stages A, B, and C are converted to scores 0, 1, and 2, respectively) to produce a JIS score of 0–5. This scoring system was evaluated in 722 Japanese patients with HCC, and statistically significant differences were observed in the survival curves among JIS scores of 0–3, but not among scores of 4–6 [23]. It has been noted that the JIS system may be limited in its ability to stratify patients with advanced scores because it uniformly

assigns tumor stage and liver function [35]. However, this system has been externally validated [38] and it appears to be one of the most promising candidates for a standard classification system across the Asia-Pacific region. However, it has not been validated in a western patient population.

The JIS staging classification was further modified by Kitai and colleagues to include evaluation of three tumor markers for HCC, namely AFP, *lens culinaris* agglutinin-reactive AFP (AFP-L3), and des-γ-carboxyprothrombin (DCP). This biomarker-combined JIS (bm-JIS) scoring system was evaluated in 1,924 patients with HCC, and findings published in 2008 showed that the bm-JIS scoring system had superior stratification ability and was a better predictor of prognosis than the conventional JIS scoring system [39]. This system has now been externally validated but still requires validation in a western patient population [40].

STAGE, LIVER DAMAGE, DCP STAGING SYSTEM

The stage, liver damage, DCP (SLiDe) staging system was established in 2004 when Omagari and colleagues evaluated a range of prognostic markers in univariate and multivariate analyses using the medical records of 177 patients with HCC from the Nagasaki University School of Medicine in Japan [41]. In that analysis, only the "stage" and "liver damage" categories from the fourth edition of the Japanese TNM staging classification, as well as serum DCP, remained significant prognostic factors of survival. Thus, in the SLiDe staging system, patients were assigned a score based on these covariates (0, 1, 2, or 3), and findings from this retrospective analysis showed that there was clear discrimination among the survival curves plotted for patients with different SLiDe scores [41]. Although the authors concluded that this is a useful system to assess the prognosis of patients, they also advised that, because the Japanese TNM staging classification must be used, which includes some parameters that are not routinely assessed in other parts of the world, external validation in a large patient population would be needed before this system could be adopted.

Tokyo Classification

In a study published in 2005, 403 patients with HCC treated with percutaneous ablation at the University of Tokyo were used as a training sample to identify prognostic factors and to develop the Tokyo score based on four factors (albumin, bilirubin, and size and number of tumors) [42]. Prognostic factors were then analyzed in a testing sample of 203 patients with HCC who had undergone resection. Clear survival differences were demonstrated among Tokyo scores, with 5-year survival rates of 78.7% (0), 62.1% (1), 40.0%

(2), 27.7% (3), and 14.3% (4–6). This system was validated by the same group, whereby it showed similar predictive ability to the CLIP scoring system and superior predictive ability to the BCLC staging classification. However, in a comparison of the JIS, BCLC, and Tokyo classifications in a Japanese cohort of HCC patients mainly with early-stage disease treated with radical therapy, the JIS score provided the best prognostic stratification [43]. Further external validation of the Tokyo classification in different patient populations is needed.

Bilirubin, Albumin, AFP-L3, AFP, DCP Score

The bilirubin, albumin, AFP-L3, AFP, DCP (BALAD) score, published by Toyoda and colleagues in 2006 [44], is a staging classification devised using only serum markers (bilirubin, albumin, AFP-L3, AFP, DCP). This scoring system, calculated as the sum of the remnant liver function score (i.e., albumin and bilirubin scoring, as devised by Tateishi and colleagues [42]) plus the tumor progression score (measured as the number of elevated tumor markers), was evaluated among 2,600 patients with HCC from five institutions. Patients were divided into six groups on the basis of the five laboratory values, with clear survival differences observed among the groups. Toyoda and colleagues also compared the BALAD scoring system with two staging classifications that consider both tumor progression and liver function factors (the JIS and CLIP classifications). They demonstrated that all three systems showed comparable prediction and discrimination of patient survival [44]. However, in a study comparing the BALAD scoring system with the JIS and bm-JIS systems conducted by Kitai and colleagues [40], there were significant differences between the BALAD and bm-JIS scores and the BALAD and JIS scores, even though all three systems effectively predicted patient survival. The authors concluded that the bm-JIS classification was superior to both the JIS and BALAD scoring systems, especially among patients with a good prognosis [40].

A U.S.-Based Prognostic Nomogram

In a recent study published in 2008, 184 patients with HCC undergoing resection at a single institution in the U.S. were classified according to eight staging classifications [45]. The ability of these classifications to predict postoperative survival was evaluated in randomly selected pairs using Harrell's concordance index. A novel nomogram was then developed using age, AFP level, operative blood loss, surgical resection margin status, tumor size, satellite lesions, and vascular invasion. Using this nomogram, survival could be predicted with a higher concordance level between randomly tested pairs than with any of the eight conven-

tional classification systems tested (concordance index of 0.74 for the nomogram versus 0.54–0.59 for the eight staging classifications tested) [45]. That analysis relied on a single institutional data set of HCC patients, which may introduce selection bias. These findings have not yet been externally validated and this nomogram is not currently used clinically.

Advanced Liver Cancer Prognostic System Score

Because patients with advanced HCC who are not amenable to locoregional therapy are candidates for inclusion in clinical trials providing they have a good 3-month survival probability, the advanced liver cancer prognostic system (ALCPS) scoring system was devised to objectively predict the 3-month survival probability among these patients [46]. In a study by Yau and colleagues published in 2008, the prognostic significance of a range of factors was evaluated by univariate and multivariate Cox regression analyses in a training set of 1,109 patients. From this, 11 significant prognostic factors were identified (ascites, abdominal pain, weight loss, Child-Pugh score, alkaline phosphatase, total bilirubin, AFP, urea level, tumor size, portal thrombosis, and lung metastases) and assessed to provide patients with a score of 0–39 (with a higher score being associated with a lower survival probability). These scores were then divided into three groups in order to categorize patients as having a good (ALCPS score, 0–8), intermediate (ALCPS score, 9–15), or poor (ALCPS score, 16–39) probability of surviving at least 3 months. Patients assessed in the training set were stratified according to their ALCPS score, and Kaplan–Meier estimates for each group showed clear survival differences, with median OS times of 7.9 months, 3.2 months, and 1.4 months for the good, intermediate, and poor groups, respectively. In the same study, ALCPS scores were subsequently assessed in a validation sample of 320 patients, and outcomes very similar to the testing sample were reported (median OS time, 7.5 months, 3.2 months, and 1.2 months for the good, intermediate, and poor groups, respectively) [46]. Moreover, patients in the validation set were also assessed by the Okuda and CLIP scoring systems, and the discriminatory ability of each prognostic scoring system, assessed by constructing receiver-operating characteristic curves, showed that the ALCPS scoring system had significantly better predictive power than either the Okuda (area under the curve [AUC], 0.77 versus 0.66 for the ALCPS and Okuda classifications, respectively; $p <$ .001) or CLIP (AUC, 0.77 versus 0.71 for ALCPS and CLIP classifications, respectively; $p =$ .002) scoring systems. It must be noted that the data set used to construct ALCPS system was from a single institute, consisting predominantly of an HBV-prevalent Chinese population. It is not

known whether ALCPS system can be applied to other populations.

SUMMARY OF STAGING CLASSIFICATIONS: WHAT IS THE BEST SYSTEM AVAILABLE?

The number of staging classifications for HCC has increased in recent years, and more recent classifications have demonstrated better prognostic ability than earlier systems (Table 2). However, improvements are still ongoing and there is no agreement on a standard classification that could be used globally.

Earlier classifications, such as the TNM staging system, only considered tumor staging factors, and as such their prognostic ability was regarded as limited. Given the impact that HCC and common comorbidities such as cirrhosis, HBV, and HCV have on liver function, most classifications now consider both tumor staging factors and liver function to predict patient outcomes. In recent years, there has been increasing interest in the role of biomarkers to predict survival. However, although adding further parameters to staging classifications may help improve the accuracy of these systems, it is important to ensure we do not create systems that are overly complex, because this may limit their clinical utility.

One of the goals of staging systems today is to provide an evidence-based treatment guide [6, 7, 21]. Although all staging classifications have been designed to predict prognosis, the BCLC staging classification is currently the only system that also provides a recommended treatment algorithm linked to each stage of disease [24]. However, the main strength of the BCLC staging system is that the four categories of patients have distinct natural histories and it is easy to apply clinically. Whether the treatment that is linked to each BCLC stage is used will depend on factors such as institutional strength and patient selection.

Because most patients with HCC present with advanced disease, many of the staging classifications, including the CUPI, CLIP, GRETCH, and ALCPS classifications, were constructed among this patient group [16, 34, 37, 46]. This could represent a limitation of these systems in terms of the accuracy of predicting prognosis in patients with earlier-stage HCC. Thus, systems such as the Japanese TNM staging system, which was constructed based on a large database of clinicopathologic data from patients at all stages of disease, including 13,772 who were eligible for curative resection, may be more appropriate for assessing patients with earlier-stage disease [35].

Because there are significant regional differences in HCC in terms of tumor morphology and the presence of comorbidities, which affect the disease course and ultimately patient prognosis, a staging classification needs to be validated in both western and Asia-Pacific patient populations

Table 2. Comparison of externally validated staging classifications available for hepatocellular carcinoma

Staging classification	Region developed	*n* of patients	Validation studies	Comparator staging classifications used	*n* of patients	Main outcomes
CLIP [16]	Italy	435	Italy [17, 20]	CLIP, Okuda	196	CLIP demonstrated greater survival predictive power than Okuda
			Middle East [18]	CLIP, Okuda	145	CLIP was more reliable than Okuda in predicting survival
			Japan [19]	CLIP, TNM, Okuda	662	CLIP had the highest stratification ability. Median survival times greater in this study than two previous studies
BCLC [24]	Spain	239	USA [6]	Okuda, TNM, BCLC, CLIP, GRETCH, CUPI, JIS	239	BCLC demonstrated the best independent predictive power for survival
			Italy [28]	Okuda, CLIP, Child-Pugh, BCLC, CUPI	187	BCLC was the best prognostic system among patients suitable for resection or ablation
			Italy [30]	Okuda, TNM, BCLC, CLIP, GRETCH, CUPI, JIS	112	BCLC showed superior discriminatory power among a group of patients who underwent radiofrequency ablation therapy
JIS [23]	Japan	722	Japan [38]	JIS, CLIP	4,525	The prognostic predictive power of JIS was superior to that of CLIP JIS score was simple to obtain and remember
bm-JIS [39]	Japan	1,924	Japan [40]	JIS, bm-JIS, BALAD	1,173	bm-JIS score showed good stratification ability and was superior in predicting prognosis, especially among patients with a good prognosis

Abbreviations: BALAD, bilirubin, albumin, *lens culinaris* agglutinin-reactive α-fetoprotein, α-fetoprotein, des-γ-carboxyprothrombin; BCLC, Barcelona Clinic Liver Cancer; bm-JIS, biomarker-combined JIS; CLIP, Cancer of the Liver Italian Program; CUPI, Chinese University Prognostic Index; GRETCH, Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire; JIS, Japanese Integrated Staging; TNM, tumor–node–metastasis.

before it can be considered globally applicable. Unfortunately, none of the staging classifications currently available has been validated in all these patient populations, and as such none can be recommended for worldwide use. However, the BCLC system has been validated in the U.S., Europe, and Taiwan, and it is the only system that has so far been validated in three continents.

A number of studies have been conducted to compare various staging classifications in the same patient population (Table 3), and findings suggest that the staging classification to show superior predictive power depends on the region. In western patient populations, the BCLC staging system appears to be superior based on findings in separate studies (two conducted in Italy, one in Taiwan, and one in North America) [6, 28, 29, 30]. In Japan, Kudo and colleagues demonstrated that the JIS scoring system was superior to the CLIP classification among 4,525 patients with HCC [38]. However, it has not been validated outside Japan.

Taken together, these findings show that, as our knowledge of this complex disease improves, staging classifica-

tions continue to be refined. As more is known about the pathogenesis of HCC and molecular markers, better staging systems will be developed.

CONCLUSIONS

HCC is a heterogeneous condition, with multiple confounding factors making assessment of these patients extremely complex. Many elements, including tumor burden, the presence of symptoms, liver function, comorbidities, and the likely effect of treatment, need to be considered in order to ensure accurate and consistent assessment of all patients, thereby providing physicians with a common language on which to base treatment decisions and guide research. This review examines each classification but does not assess their relative value. Although many different staging classifications have been developed and there is currently no consensus on the best classification to use, the BCLC staging classification is emerging as a promising candidate for a standard classification in western regions, because it has been externally validated [6, 28, 30] and it is also endorsed

Table 3. Comparison of studies evaluating different staging classifications in the same hepatocellular carcinoma patient population

Study	*n* of patients	Region	Staging classifications compared	Superior classification identified
Choi et al. [12]	163	Korea	TNM (fifth and sixth editions), Okuda, BCLC, CLIP, GRETCH, JIS	TNM sixth edition
Lu et al. [11]	234	China	Okuda, CLIP, TNM, CUPI	TNM sixth edition
Cillo et al. [28]	187	Italy	Okuda, BCLC, CLIP, GRETCH, CUPI	BCLC
Guglielmi et al. [30]	112	Italy	Okuda, TNM, BCLC, CLIP, GRETCH, CUPI, JIS	BCLC
Marrero et al. [6]	239	USA	Okuda, TNM, BCLC, CLIP, GRETCH, CUPI, JIS	BCLC
Kitai et al. [40]	1,173	Japan	JIS, bm-JIS, BALAD	bm-JIS
Kudo et al. [38]	4,525	Japan	CLIP, JIS	JIS

Abbreviations: BALAD, bilirubin, albumin, *lens culinaris* agglutinin-reactive α-fetoprotein, α-fetoprotein, des-γ-carboxyprothrombin; BCLC, Barcelona Clinic Liver Cancer; bm-JIS, biomarker-combined JIS; CLIP, Cancer of the Liver Italian Program; CUPI, Chinese University Prognostic Index; GRETCH, Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire; JIS, Japanese Integrated Staging; TNM, tumor–node–metastasis.

by both the EASL [5] and the AASLD [31]. However, because risk factors vary significantly from region to region, any standard global staging classification needs to be validated in both western and Asia-Pacific patient populations; to date, no such staging classification exists.

Continued research efforts have improved our understanding of this complex disease, which has allowed us to refine staging classifications and improve our therapeutic approach. In recent years, a significant amount of research has reported on the role of tumor markers to predict survival in HCC, and the markers AFP, AFP-L3, and DCP have now been incorporated into some staging classifications. In addition, molecular markers such as hepatocyte growth factor, vascular endothelial growth factor, and transforming growth factor β 1 have been linked with poor outcomes in HCC patients [47], and so may play a role in helping us to further improve staging classifications. In addition to the added information that tumor and molecular markers bring, data from ongoing studies may contribute. The Global Investigation of therapeutic DEcisions in hepatocellular carcinoma and Of its treatment with sorafeNib (GIDEON) study is a large global, noninterventional study of patients with unresectable HCC receiving sorafenib (Nexavar®; Onyx Pharmaceuticals, Inc., Emeryville, CA; Bayer HealthCare Pharmaceuticals, Inc., Wayne, NJ; Bayer Schering Pharma AG, Berlin, Germany) therapy. That study will collect details of local, regional, and global methods of patient evaluation, diagnosis, and follow-up, and assess comorbidities and their influence on treatment and outcome. Information collected in this database may be of value in further refining current staging classifications. However, further research efforts are needed for us to gain a full understanding of the factors that affect the prognosis of patients with HCC.

AUTHOR CONTRIBUTIONS

Conception/Design: Jean-Pierre Bronowicki, Masatoshi Kudo, Jorge A. Marrero
Data analysis and interpretation: Jean-Pierre Bronowicki, Masatoshi Kudo, Jorge A. Marrero
Manuscript writing: Jean-Pierre Bronowicki, Masatoshi Kudo, Jorge A. Marrero
Final approval of manuscript: Jean-Pierre Bronowicki, Masatoshi Kudo, Jorge A. Marrero

REFERENCES

1 Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. Lancet 2003; 362:1907–1917.

2 Parkin DM, Bray F, Ferlay J et al. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.

3 Poon D, Anderson BO, Chen LT et al. Management of hepatocellular carcinoma in Asia: Consensus statement from the Asian Oncology Summit 2009. Lancet Oncol 2009;10:1111–1118.

4 Rilling WS, Drooz A. Multidisciplinary management of hepatocellular carcinoma. J Vasc Interv Radiol 2002;13:S259–S263.

5 Bruix J, Sherman M, Llovet JM et al. Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. J Hepatol 2001;35:421–430.

6   Marrero JA, Fontana RJ, Barrat A et al. Prognosis of hepatocellular carci-
    noma: Comparison of 7 staging systems in an American cohort. Hepatology
    2005;41:707–716.

7   Llovet JM, Fuster J, Bruix J et al. The Barcelona approach: Diagnosis, stag-
    ing, and treatment of hepatocellular carcinoma. Liver Transpl 2004;10(2
    suppl):S115–S120.

8   Greene FL, Sobin LH. The staging of cancer: A retrospective and prospec-
    tive appraisal. CA Cancer J Clin 2008;58:180–190.

9   Green F. Liver (including intrahepatic bile ducts). In: Green F, Page D,
    Fleming I, eds. AJCC Cancer Staging Handbook, Sixth Edition. New York:
    Springer, 2002:131–144.

10 Sobin LH. TNM, sixth edition: New developments in general concepts and
    rules. Semin Surg Oncol 2003;21:19–22.

11 Lu W, Dong J, Huang Z et al. Comparison of four current staging systems
    for Chinese patients with hepatocellular carcinoma undergoing curative re-
    section: Okuda, CLIP, TNM and CUPI. J Gastroenterol Hepatol 2008;23:
    1874–1878.

12 Choi SB, Lee JG, Kim KS et al. The prognosis and survival analysis ac-
    cording to seven staging systems of hepatocellular carcinoma following
    curative resection. Hepatogastroenterology 2008;55:2140–2145.

13 Okuda K, Ohtsuki T, Obata H et al. Natural history of hepatocellular car-
    cinoma and prognosis in relation to treatment. Study of 850 patients. Can-
    cer 1985;56:918–928.

14 Levy I, Sherman M; Liver Cancer Study Group of the University of To-
    ronto. Staging of hepatocellular carcinoma: Assessment of the CLIP,
    Okuda, and Child-Pugh staging systems in a cohort of 257 patients in To-
    ronto. Gut 2002;50:881–885.

15 Huang YH, Chen CH, Chang TT et al. Evaluation of predictive value of
    CLIP, Okuda, TNM and JIS staging systems for hepatocellular carcinoma
    patients undergoing surgery. J Gastroenterol Hepatol 2005;20:765–771.

16 The Cancer of the Liver Italian Program (CLIP) Investigators. A new prog-
    nostic system for hepatocellular carcinoma: A retrospective study of 435
    patients. Hepatology 1998;28:751–755.

17 Llovet JM, Bruix J. Prospective validation of the Cancer of the Liver Italian
    Program (CLIP) score: A new prognostic system for patients with cirrhosis
    and hepatocellular carcinoma. Hepatology 2000;32:679–680.

18 Siddique I, El-Naga HA, Memon A et al. CLIP score as a prognostic indi-
    cator for hepatocellular carcinoma: Experience with patients in the Middle
    East. Eur J Gastroenterol Hepatol 2004;16:675–680.

19 Ueno S, Tanabe G, Sako K et al. Discrimination value of the new western
    prognostic system (CLIP score) for hepatocellular carcinoma in 662 Japa-
    nese patients. Cancer of the Liver Italian Program. Hepatology 2001;34:
    529–534.

20 The Cancer of the Liver Italian Program (CLIP) Investigators. Prospective
    validation of the CLIP score: A new prognostic system for patients with
    cirrhosis and hepatocellular carcinoma. Hepatology 2000;31:840–845.

21 Pons F, Varela M, Llovet JM. Staging systems in hepatocellular carcinoma.
    HPB (Oxford) 2005;7:35–41.

22 Collette S, Bonnetain F, Paoletti X et al. Prognosis of advanced hepatocel-
    lular carcinoma: Comparison of three staging systems in two French clini-
    cal trials. Ann Oncol 2008;19:1117–1126.

23 Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular
    carcinoma (CLIP score): Its value and limitations, and a proposal for a new
    staging system, the Japan Integrated Staging score (JIS score). J Gastroen-
    terol 2003;38:207–215.

24 Llovet JM, Brú C, Bruix J. Prognosis of hepatocellular carcinoma: The
    BCLC staging classification. Semin Liver Dis 1999;19:329–338.

25 Llovet JM, Bustamante J, Castells A et al. Natural history of untreated non-
    surgical hepatocellular carcinoma: Rationale for the design and evaluation
    of therapeutic trials. Hepatology 1999;29:62–67.

26 Beaugrand M, Sala M, Degos F et al. Treatment of advanced hepatocellular
    carcinoma (HCC) by seocalcitol (a Vit D analogue): An international ran-
    domized double-blind placebo-controlled study in 747 patients [abstract
    37]. J Hepatol 2005;42(suppl 2):17.

27 Cillo U, Vitale A, Grigoletto F et al. Prospective validation of the Barcelona
    Clinic Liver Cancer staging system. J Hepatol 2006;44:723–731.

28 Cillo U, Bassanello M, Vitale A et al. The critical issue of hepatocellular
    carcinoma prognostic classification: Which is the best tool available?
    J Hepatol 2004;40:124–131.

29 Wang JH, Changchien CS, Hu TH et al. The efficacy of treatment schedules
    according to Barcelona Clinic Liver Cancer staging for hepatocellular car-
    cinoma: Survival analysis of 3892 patients. Eur J Cancer 2008;44:1000–
    1006.

30 Guglielmi A, Ruzzenente A, Pachera S et al. Comparison of seven staging
    systems in cirrhotic patients with hepatocellular carcinoma in a cohort of
    patients who underwent radiofrequency ablation with complete response.
    Am J Gastroenterol 2008;103:597–604.

31 Bruix J, Sherman M; Practice Guidelines Committee, American Associa-
    tion for the Study of Liver Diseases. Management of hepatocellular carci-
    noma. Hepatology 2005;42:1208–1236.

32 Meier V, Ramadori G. Clinical staging of hepatocellular carcinoma. Dig
    Dis 2009;27:131–141.

33 Huitzil FD, Capanu M, O'Reilly E et al. Ranking and improvement of stag-
    ing systems (SS) in advanced hepatocellular carcinoma (AHCC) [abstract
    210]. Presented at the American Society of Clinical Oncology Gastrointes-
    tinal Cancers Symposium, Orlando, FL, January 25–27, 2008.

34 Chevret S, Trinchet JC, Mathieu D et al. A new prognostic classification for
    predicting survival in patients with hepatocellular carcinoma. J Hepatol
    1999;31:133–141.

35 Minagawa M, Ikai I, Matsuyama Y et al. Staging of hepatocellular carci-
    noma: Assessment of the Japanese TNM and AJCC/UICC TNM systems in
    a cohort of 13,772 patients in Japan. Ann Surg 2007;245:909–922.

36 Vauthey J-N, Lauwers GY, Esnaola NF et al. Simplified staging for hepa-
    tocellular carcinoma. J Clin Oncol 2002;20:1527–1536.

37 Leung TWT, Tang AMY, Zee B et al. Construction of the Chinese Univer-
    sity 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. Can-
    cer 2002;94:1760–1769.

38 Kudo M, Chung H, Haji S et al. Validation of a new prognostic staging sys-
    tem for hepatocellular carcinoma: The JIS score compared with the CLIP
    score. Hepatology 2004;40:1396–1405.

39 Kitai S, Kudo M, Minami Y et al. A new prognostic staging system for hep-
    atocellular carcinoma: Value of the biomarker combined Japan Integrated
    Staging score. Intervirology 2008;51(suppl 1):86–94.

40 Kitai S, Kudo M, Minami Y et al. Validation of a new prognostic staging
    system for hepatocellular carcinoma: A comparison of the biomarker-
    combined Japan Integrated Staging score, the conventional Japan Inte-
    grated Staging score and the BALAD score. Oncology 2008;75(suppl 1):
    83–90.

41 Omagari K, Honda S, Kadokawa Y et al. Preliminary analysis of a newly
    proposed prognostic scoring system (SLiDe score) for hepatocellular car-
    cinoma. J Gastroenterol Hepatol 2004;19:805–811.

42 Tateishi R, Yoshida H, Shiina S et al. Proposal of a new prognostic model

for hepatocellular carcinoma: An analysis of 403 patients. Gut 2005;54:
419–425.

43 Chung H, Kudo M, Takahashi S et al. Comparison of three current staging
systems for hepatocellular carcinoma: Japan Integrated Staging score, new
Barcelona Clinic Liver Cancer staging classification, and Tokyo score. J
Gastroenterol Hepatol 2008;23:445–452.

44 Toyoda H, Kumada T, Osaki Y et al. Staging hepatocellular carcinoma by
a novel scoring system (BALAD score) based on serum markers. Clin Gas-
troenterol Hepatol 2006;4:1528–1536.

45 Cho CS, Gonen M, Shia J et al. A novel prognostic nomogram is more ac-
curate than conventional staging systems for predicting survival after re-
section of hepatocellular carcinoma. J Am Coll Surg 2008;206:281–291.

46 Yau T, Yao TJ, Chan P et al. A new prognostic score system in patients with
advanced hepatocellular carcinoma not amendable to locoregional therapy:
Implication for patient selection in systemic therapy trials. Cancer 2008;
113:2742–2751.

47 Gomaa AI, Khan SA, Leen ELS et al. Diagnosis of hepatocellular carci-
noma. World J Gastroenterol 2009;15:1301–1314.