Research article

Evaluation of staging systems to predict prognosis in hepatocellular carcinoma patients treated with radioembolization

Ece Meram [a,*], Colin Longhurst [b], Baran Umut Vardar [a], Kerim Karaoglu [a], Paul F. Laeseke [a], Orhan S. Ozkan [a]

[a] Department of Radiology, University of Wisconsin-Madison, Madison, WI, USA
[b] Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

HIGHLIGHTS

- There is no consensus on utility of staging systems to predict prognosis of HCC patients.
- Data for HCC patients undergoing RE is only available from a single center.
- Basic discriminatory ability (log-rank) was significant only for Child-Pugh Numeric and CLIP.
- Further analyses were performed to rank the staging systems.
- The results suggest that Child-Pugh Numeric has an overall greater prognostic value.

ABSTRACT

Purpose: To compare the prognostic accuracy of nine staging systems, some of which are well-known and some of which have only been more recently described, for patients with unresectable HCC treated with radioembolization (RE).
Materials and methods: Individual scores or classes for the following staging systems were recorded or calculated for patients (n = 89) with unresectable HCC who underwent RE at a single tertiary care center from January 2008 to October 2016: Eastern Cooperative Oncology Group, Barcelona Clinic Liver Cancer, Hong Kong Liver Cancer, Okuda, Cancer of the Liver Italian Program (CLIP), Model for End Stage Liver Disease, Child-Pugh (CP) Categorical and Numeric, and Albumin-Bilirubin. For each staging system, a cox proportional hazards regression model was fit to the data and log-rank test statistics, concordance indices, Akaike Information Criteria (AIC) and other diagnostic statistics were calculated.
Results: Of the nine staging systems analyzed, the basic discriminatory ability assessed with the log-rank test (rejected at the $\alpha = .05$-level) was significant for two of the systems: CP Numeric (p < .001) and CLIP (p < .05). Out of these two systems, CP Numeric system had a higher prognostic accuracy than CLIP with the lowest AIC (464.90), the highest optimism-corrected pseudo R^2 (0.16), and the highest estimated concordance index (0.64).
Conclusion: As applied to our patient population, the CP Numeric system contained the most predictive prognostic information for patients with HCC undergoing radioembolization. However, all evaluated staging systems performed suboptimally, and the relative superiority of any of the systems remains unclear when ranking them according to common practice. Further evaluation of current ranking methodologies is recommended.

1. Introduction

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide with an increasing incidence [1]. It is responsible for approximately 700,000 deaths annually [2]. Multiple factors influence patient survival; presence of a large tumor mass, vascular invasion, extrahepatic metastasis and a poor functional status are associated with a worse outcome in patients with HCC [3]. In addition, the underlying liver dysfunction, extent of cirrhosis and presence of hepatitis

* Corresponding author.
  E-mail address: ecemeram@gmail.com (E. Meram).

B or C virus are important prognostic parameters influencing survival [3].

The management of HCC depends on the tumor stage at the time of diagnosis. In an effort to predict treatment outcomes and prognosis of patients with HCC, multiple staging systems have been developed. Some of these systems are tumor specific, while others rely on liver dysfunction due to the fact that liver disease can be a major determinant of prognosis. However, there is no consensus on which staging system to use and when to use it. Despite well-known limitations and newer staging systems with reported higher prognostic value such as Hong Kong Liver Cancer classification [4], the Barcelona Clinic Liver Cancer (BCLC) Classification is a commonly used staging system, especially in Western countries, because of its link to available treatment strategies [5]. The American Association for the Study of Liver Disease (AASLD) suggests the use of BCLC system when classifying and managing patients with HCC [3]. In addition to treatment guidelines, local expertise in treating HCC is another major factor influencing patient management. The treatment algorithms of different institutions vary considerably, likely because a number of treatment options for patients with HCC have been associated with a survival benefit [6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

When considering the effect of treatment on prognosis, it is necessary to use a staging system tailored to a given treatment modality for predicting the prognosis of patients with HCC. The predictive ability of staging systems for prognosis has been investigated for patients with unresectable HCC treated with transarterial chemoembolization (TACE) [16]. Attempts have also been made to find the most appropriate staging system to predict survival for patients with advanced HCC receiving systemic therapy [17]. More recently, staging systems were compared for their prognostic ability in patients with unresectable HCC who were treated with radioembolization (RE) [12]. A recent review on staging systems also highlights that incorporating another staging system such as NIACE improves prognostic accuracy for patients with HCC within the same BCLC category [18].

Radioembolization is a locoregional treatment that uses intra-arterial injection of microspheres loaded with the radionuclide yttrium-90 to induce tumor necrosis [19, 20]. It is a relatively recent addition as a treatment option for HCC and its place in treatment algorithms is evolving. As such, there is a need for additional data from a variety of practice settings on patient and treatment factors correlating with improved outcomes. Although a number of staging systems for patients with HCC undergoing RE were investigated previously [12], newly described systems, such as the Hong Kong Liver Cancer classification and Albumin-Bilirubin scoring, were not included. In addition, more studies may yield information on the applicability of previously reported results to other practices. The purpose of this study was to evaluate the prognostic accuracy of commonly used and newly developed staging systems for patients with unresectable HCC who underwent RE at our institution.

2. Materials and methods

2.1. Patient selection

This retrospective single-center study was approved by the Institutional Review Board (2016-0418), compliant with the HIPAA guidelines, and conducted with no requirement of a consent from the participants. All patients with HCC who underwent yttrium-90 radioembolization between January 2008 and November 2016 were identified retrospectively from a RedCAP database (Vanderbilt University, TN, USA). The diagnosis of HCC was made according to the diagnostic criteria of the AASLD [21], and the decision to perform radioembolization for unresectable HCC was made by a multidisciplinary liver tumor board. Twenty-five patients, who underwent previous treatments at the time of RE, were included in the study as all patients were re-evaluated and re-classified before undergoing radioembolization. In total, 103 patients underwent 140 ^{90}Y RE procedures with median number of one treatment session. If patients underwent RE after disease progression, the RE was

recorded as an additional treatment along with any other type of HCC treatment such as ablation, chemoembolization, or sorafenib. If the patients underwent multiple RE procedures as a staged treatment in a treatment cycle (eg. in bilobar disease), the baseline parameters from the first procedure were used. Any patient who lacked the data required for the staging or classification systems was excluded. See Table 1 for patient demographics.

2.2. Yttrium-90 radioembolization

All patients underwent mesenteric and hepatic angiography and ^{99}mTc-macroaggregated albumin scanning to determine the hepatopulmonary shunt fraction. All ^{90}Y radioembolization treatments were performed using glass microspheres (TheraSphere, BTG, UK). Prescribed radiation doses (median 120 Gy, range 75–135 Gy) varied based on patient and tumor characteristics as well as the shunt fraction. The tumor response to treatment was evaluated with mRECIST criteria.

2.3. Patient assessment and follow-up

Using the last pre-procedural cross sectional imaging and biochemical tests to the date of first RE, individual scores or classes for the following staging systems were recorded or calculated as described: Eastern Cooperative Oncology Group (ECOG) [22], Barcelona Clinic Liver Cancer (BCLC) [4], Hong Kong Liver Cancer (HKLC) [23], Okuda [24], Cancer of the Liver Italian Program (CLIP) [25, 26], Model for End Stage Liver

Table 1. Basic characteristics of 89 patients included in the analysis.

Variable	n	%
Total	**89**	**100**
Sex		
Male	80	92
Female	9	11
Age> 65	42	47
Etiology		
HCV	36	40
Alcohol	18	20
Cryptogenic	8	9
NASH	7	8
HBV	6	7
Hemochromatosis	2	2
No Cirrhosis	12	13
Tumor Type		
Solitary	11	12
Multifocal	52	58
Infiltrative	26	29
AFP >400 ng/mL	28	31
PVT	27	30
Prior Treatment	26	29
Prior Resection	12	13
Prior Ablation	15	17
Prior Chemotherapy	5	6
Multiple RE	28	31
Tumor Response		
CR	14	16
PR	7	8
SD	18	20
PD	50	56
Additional Treatment		
Ablation	3	3
Chemotherapy	11	12
TACE	18	20

Disease (MELD) [27], Child-Pugh (CP) categorical and CP numeric [28, 29], and Albumin-Bilirubin (ALBI) [30]. In total, 68 patients were deceased at the time of analysis. Overall survival (OS) was defined from the first RE procedure to the date of death or last follow-up. A total of 39 patients underwent additional treatments after radioembolization. Three patients who underwent liver transplantation after their RE were censored at the date of transplant. Considering the high proportion of patients with additional treatments, censoring at the time of any other additional treatment was not performed. A similar approach was taken in previous studies due to the difficulty of finding a patient population that underwent only radioembolization without pursuing further treatments [12].

2.4. Statistical analysis

The reverse Kaplan-Meier method was used to calculate median follow-up time. For all classification systems, any distinct class consisting of less than six observations was removed from the data due to difficulties in proper parameter estimation via the cox proportional hazards model. For the HKLC staging system, subclasses IIa-IIb and IIIa-IIIb were merged into two distinct classes due to sample size restrictions and based on the findings of Yau et al. [23], which suggested a similar overall survival for patients in the HKLC subclasses. Patients within removed classes were excluded from all statistical analyses. These data were excluded instead of merged in order to maintain the homogeneity of each class within a staging system, resulting in 89 patients that were included in the statistical analyses.

For each of the nine systems, a cox proportional hazards model was fit to the data (with the staging system as the only predictor) using the "survival" package (V 2.41) in R (V 3.3.2, R Core Team, Vienna, Austria). To test for homogeneity between survival curves for each system stage, log-rank test p-values were calculated for each model. To test for a linear trend in hazard rates between stages of the same system, linear trend X^2 p-values were calculated for each model. Rejecting the linear trend hypothesis test indicates that there is at least one strict inequality in the hazard rates of the stages (in an ordinal fashion). The Akaike Information Criterion (AIC) was also calculated to measure the relative quality of the fitted model. A lower AIC indicates a better fit to the data, relative to other models fit to the same data.

To assess the predictive ability of the different staging systems, Nagelkerke's pseudo R^2 was computed. To adjust for the potential over-optimism in our R^2 estimates as described by Harrell, a bias-corrected estimate of R^2 was found via bootstrap validation (B = 10,000) using the "rms" (V 5.1) package [31, 32]. Estimated concordance indices (C-indices) were calculated for each system along with 95% studentized bootstrap (B = 10,000) confidence intervals to assess the discriminatory abilities of the different staging systems.

The staging systems which rejected the log-rank test at the $\alpha = .05$-level were analyzed further with two more direct statistical comparisons. First, the differences in C-indices between the two models were tested via a bootstrap simulation (B = 10,000) modeled after the algorithm described by Kang et al [33]. Second, to compare the relative predictive information contained in the two systems, pairwise adequacy indices were calculated for each system using the method described in Harrell [31]. An adequacy index near one indicates that that single system contains almost as much predictive information as a model which uses both systems in conjunction.

3. Results

Median follow up time of the cohort was 40.8 months (IQR; 71.83-21.30). Median overall survival time of 89 patients was 13.6 months (95% CI; 11.0–19.1 months) with a 74 % mortality rate at the time of analysis. Among the nine staging systems tested, the log rank null

Table 2. Patient survival within each staging system based on log-rank test of KM plots.

Staging System	n	%	Median Survival (months)
Total	**89**	**100**	**13.6**
Child-Pugh Categorical			p = 0.3539
A	71	80	14.3
B	18	20	7.13
Child-Pugh Numeric			**p < .0001**
5	32	36	34.5
6	39	44	11.6
7	12	13	5.2
8	6	7	24.2
MELD			p = 0.0760
≤9	62	70	14.8
>9	27	30	10.8
BCLC			p = 0.3764
B	31	35	18.8
C	58	65	11.6
HKLC			p = 0.3356
II	24	27	19.1
III	65	73	13.6
ECOG			p = 0.5444
0	39	44	18.8
1	50	56	11.5
ALBI			p = 0.1456
A1	12	13	11.33
A2	70	79	14.83
A3	7	8	5.23
Okuda			p = 0.0726
1	62	70	17.47
2	27	30	7.83
CLIP			**p = 0.0320**
0	9	10	44.0
1	41	46	22.3
2	33	37	10.6
3	6	7	15.7

hypothesis was rejected for only two of the systems, CP Numeric and CLIP (p < .05, Table 2). Kaplan Meier plots for these two systems can be seen in Figures 1 and 2.

In regard to the statistical analyses commonly seen in system-ranking studies, a system which rejects the log-rank and linear trend X^2, has a lower AIC, and a higher C-index is typically preferred. Regarding the further analyses performed in this study, the system with a higher pseudo-R^2 and optimism-adjusted R^2 is favored. In Table 3, it can be seen that the CP Numeric system had the lowest log rank test p-value (p < .001), the lowest AIC (464.90), the highest estimated concordance index (0.64) and the highest R^2 (adjusted: 0.16; unadjusted: 0.22). However, the CP Numeric system did have a higher linear trend p-value than other systems (p > .05), indicating a lack of evidence for ordinally decreasing hazard rates between the subclasses. When the C-indices of CP Numeric and CLIP were directly compared via bootstrap simulation, no statistically significant differences were found between the two systems (p > .05). Nevertheless, the relative superiority of CP Numeric was further demonstrated by the comparative adequacy indices seen in Table 4, where the calculated adequacy index for CP Numeric and CLIP are 0.77 and 0.30, respectively. This means that in a hypothetical system incorporating both CP Numeric and CLIP, most of the predictive information being contributed comes from CP Numeric.
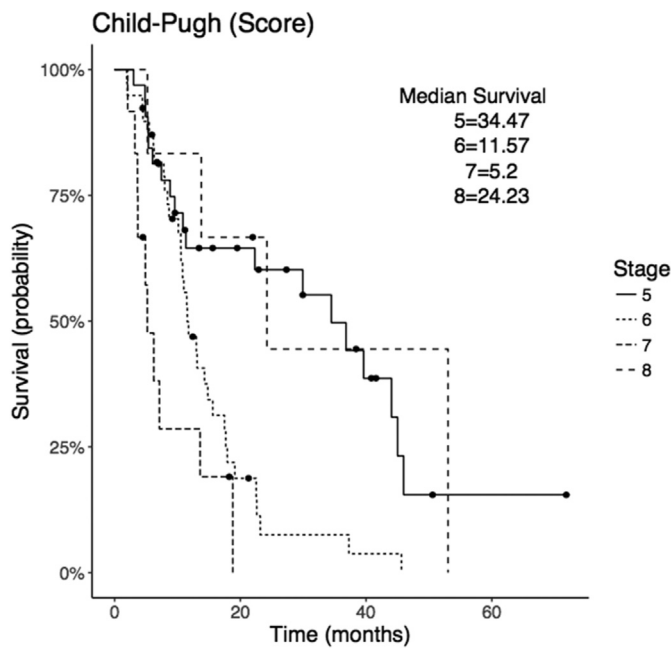
Figure 1. The Kaplan-Meier plot of Child-Pugh Numeric system with median overall survival of each group.
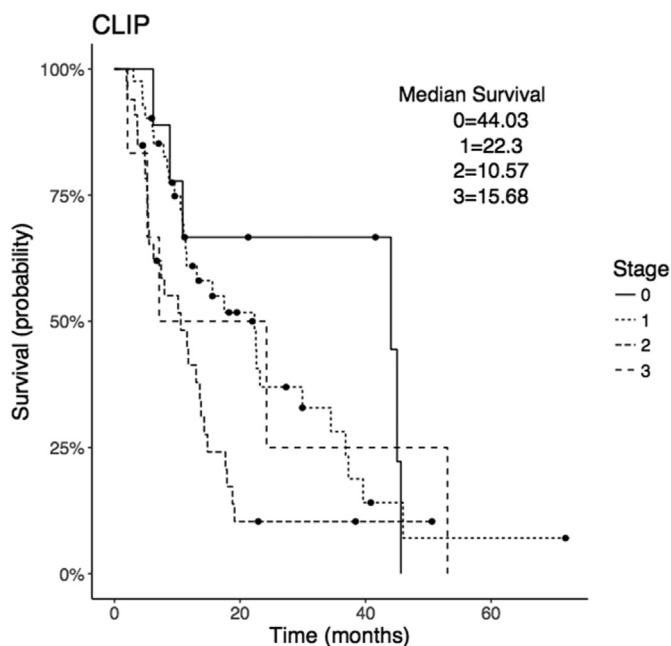

Figure 2. The Kaplan-Meier plot of CLIP system with median overall survival of each group.

4. Discussion

Staging systems aid to determine how to manage a patient's disease and to predict how the disease will progress. The strength of a staging system comes from how well it incorporates the influential parameters of a disease or a condition in order to discriminate the survival of different groups within it [21, 34]. In the case of hepatocellular carcinoma, despite the use of multiple staging systems, it can be challenging to predict the prognosis of a patient with HCC. Although the BCLC system is widely used when determining the management of a patient, it has been criticized for not being discriminative enough, especially for patients with

intermediate disease [35]. The use of the BCLC system has been less useful for predicting the prognosis of patients with unresectable HCC [12, 16, 17, 18].

In addition to the tumor characteristics and the underlying liver function, the choice of treatment has a major influence on prognosis due to a treatment-related survival benefit [6, 7, 8, 9, 10, 11, 13]. Using the same staging system across all treatment modalities may not have the same applicability in different treatment strategies. Child-Pugh nominal, CUPI and Tokyo scores were reported to provide the best prognostic accuracy for HCC patients treated with TACE. Although Child-Pugh scoring is not a tumor classification system, it was shown to perform better than other investigated systems including the staging systems that incorporates tumor characteristics [16]. For advanced HCC patients receiving systemic treatments, CLIP, CUPI and GETCH were the most informative staging systems in predicting survival [17]. More recently, M-SIRT was suggested as a new prognostic tool specifically for HCC patients undergoing radioembolization [36].

Radioembolization is a relatively new treatment modality for patients with unresectable HCC and its role in the management of patients with HCC is still being defined. To date, data on the prognostic accuracy of staging systems for patients with HCC undergoing RE is only available from a single center and for a subset of available systems. CLIP was found to be the most accurate system among them [12]. However, it should be noted that Child-Pugh was not investigated as a numeric system, and only Child-Pugh Class was used in the aforementioned study. Moreover, the performance of newer staging systems compared to the previously analyzed systems has yet to be evaluated. The prognostic ability of nine staging systems investigated for this study were found to be suboptimal as applied to our patient population. Although all of these systems are well-validated, none of them were developed for or incorporate radio-embolization in the treatment of HCC. Radioembolization may cause a change in the survival trajectory of patients for better or worse, therefore, the pre-procedural survival estimation with these staging systems may not be a good, post-treatment estimate of prognosis.

Using previously described methodologies for comparing staging systems, it was not immediately clear which system performed best at predicting survival in our HCC patient population undergoing radio-embolization. Although the CP Numeric system was found to have the lowest AIC, the smallest log-rank p-value, and the highest estimated C-index, it also had a large linear trend p-value (p = 0.065) relative to other systems analyzed. It seems that CP Numeric system performed better regarding prognostic accuracy than the other staging systems analyzed for this patient population, but still had some incongruities. The incongruities in model metrics present a unique difficulty in declaring superiority among the systems. Should candidate systems be pared down in a hierarchical manner (first discard systems which don't reject the log-rank null hypothesis, then discard models which don't reject the linear trend hypothesis then compare the C-index) or should the systems be evaluated in a more nuanced fashion?

Using the common metrics of the log-rank test, linear trend test, AIC and C-index, many system-ranking discontinuities can be seen in the results. For example, as seen in Table 3, CP Numeric has a lower log-rank yet a higher linear trend p-value than CLIP; ALBI has a lower linear trend p-value than ECOG yet also has a smaller C-index; CP Class has a lower AIC than ECOG, yet also a smaller C-index. Given these results, it seems that the ranking of the staging systems may not be as objective as previously thought. Further research into different statistical methodologies is recommended.

In Table 3, with the exception of CP Numeric, all systems had low R^2 and even lower optimism-corrected R^2 suggesting that those systems contained little predictive information as applied to our patient population. Unlike similar studies, pseudo R^2, optimism-corrected R^2 and bootstrapped confidence intervals about the C-indices were also calculated to provide a more comprehensive overview of model performance. From Table 3 it is seen that all C-index confidence intervals overlap and that our bootstrapped p-value from the direct comparison of the C-index

Table 3. The prognostic value of staging systems based on described statistical methods.

System	Log-rank p	Linear trend p	AIC	R^2	R^2_c	C-index	95% CI of C-index
CP (Score)	0.0000	0.0645	464.8736	0.2160	0.1637	0.6420	(0.573,0.717)
CLIP	0.0320	0.0239	478.1723	0.0890	0.0272	0.6199	(0.560,0.694)
Okuda	0.0726	0.0363	479.4190	0.0334	0.0192	0.5870	(0.498,0.654)
MELD	0.0760	0.0380	479.5293	0.0322	0.0205	0.5590	(0.508,0.623)
BCLC	0.3764	0.1885	481.6301	0.0089	-0.0023	0.5588	(0.491,0.617)
HKLC	0.3356	0.1680	481.4618	0.0108	0.0013	0.5575	(0.494,0.607)
ECOG	0.5444	0.2726	482.0565	0.0041	-0.0078	0.5530	(0.486,0.619)
CP (Class)	0.3539	0.1773	481.6158	0.0091	-0.0066	0.5465	(0.488,0.614)
ALBI	0.1456	0.1396	481.4666	0.0328	0.0012	0.5425	(0.509,0.607)

Table 4. Adequacy index and bootstrapped p-value for the direct comparison of best performing two staging systems.

Staging System	p-value$_{boot}$	Adequacy Index
CP (Score) vs. CLIP	0.3903	0.7722 vs. 0.2957

from CP Numeric and CLIP is quite large. These results offer support to Harrell's previous assertions that the C-index lacks the sensitivity to be used for the purpose of directly comparing multiple models [31].

This study does not attempt to define what exactly constitutes a "superior" staging system in the context of a cox regression model. However, based on the statistics calculated, the CP Numeric system does seem to have the most predictive and discriminatory ability, followed by CLIP. Although CLIP does have a lower linear trend p-value than CP Numeric, their direct comparison via the adequacy index suggests that CP Numeric, nevertheless, contains more predictive information than CLIP.

The findings of the study, however, need to be considered in the context of the study limitations. The primary limitation is a somewhat small samples size (n = 89), which created difficulties in parameter estimation for some of the cox regression models. There is a lack of variability of stages in this study; however, it might be a more accurate representation of patient populations, for whom radioembolization is typically prescribed in most practices. In addition, almost half (n = 39) of our patient population underwent additional treatments for HCC which could potentially confound the relationship between assessed stage and overall survival time. Some of the staging systems are nested in other staging systems. While not investigated here or in the aforementioned papers, nesting could potentially complicate the statistical analyses. All of the statistical methods used (C-index, AIC, R^2 etc.) are sensitive to censoring and the observed rate of censoring (~26%) in this study is non-negligible. Finally, although newer staging systems such as ALBI and HKLC were included, not all of the available staging systems for HCC were evaluated in this study. Some staging systems were found to lack an adequate sample size in majority of the stages. This could be a result of the limited overall sample size or a selection bias on the part of clinicians who choose radioembolization as the treatment of HCC. Among the systems that were analyzed, the small number of observations in some subclasses may have influenced the resulting model metrics.

In conclusion, multiple staging systems have been proposed for hepatocellular carcinoma in order to better classify the tumors and the patients. The superiority of some systems has been shown previously for multiple treatment modalities, including radioembolization. Although prevailing ranking methods do not provide an objectively clear answer in regard to staging system superiority, this study suggests that the CP Numeric system has an overall greater prognostic value than the other investigated systems for patients with unresectable HCC undergoing radioembolization. A complementary use of CP Numeric system to BCLC classification, instead of CP Classes, could be informative in the clinical setting to help predict the prognosis of patients with unresectable HCC when undergoing radioembolization. However, further validation is still necessary, and further evaluation of current ranking methodologies is recommended.

Declarations

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare the following conflict of interests: Paul F Laeseke is a Consultant at NeuWave Medical, a Consultant and Shareholder at Elucent Medical, and a Shareholder of HistoSonics and McGinley Orthopeadic Innovations.

Additional information

No additional information is available for this paper.

References

[1] H.B. El-Serag, Hepatocellular carcinoma, N. Engl. J. Med. 365 (12) (2011) 1118–1127.

[2] J. Ferlay, I. Soerjomataram, R. Dikshit, et al., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int. J. Cancer 136 (5) (2015) E359–386.

[3] J. Bruix, M. Sherman, Management of hepatocellular carcinoma: an update, Hepatology 53 (3) (2011) 1020–1022.

[4] N. Parikh, S. Scaglione, Y. Li, et al., A comparison of staging systems for hepatocellular carcinoma in a multicenter US cohort, Clin. Gastroenterol. Hepatol. 16 (5) (2018) 781–782.

[5] J.M. Llovet, C. Bru, J. Bruix, Prognosis of hepatocellular carcinoma: the BCLC staging classification, Semin. Liver Dis. 19 (3) (1999) 329–338.

[6] J. Bruix, J.L. Raoul, M. Sherman, et al., Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial, J. Hepatol. 57 (4) (2012) 821–829.

[7] J.M. Llovet, Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival, Hepatology 37 (2) (2003) 429–442.

[8] J.M. Llovet, M.I. Real, X. Montana, et al., Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial, Lancet 359 (9319) (2002) 1734–1739.

[9] J.M. Llovet, Z. Ricci, V. Mazzaferro, et al., Sorafenib in advanced hepatocellular carcinoma, N. Engl. J. Med. 359 (4) (2008) 378–390.

[10] C.M. Lo, Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma, Hepatology 35 (5) (2002) 1164–1171.

[11] V. Mazzaferro, C. Spasito, S. Bhoori, et al., Yttrium-90 radioembolization for intermediate-advanced hepatocellular carcinoma: a phase 2 study, Hepatology 57 (5) (2013) 1826–1837.

[12] K. Memon, L. Kulik, R. Lewandowski, et al., Comparative study of staging systems for hepatocellular carcinoma in 428 patients treated with radioembolization, J. Vasc. Intervent. Radiol. 25 (7) (2014) 1056–1066.

[13] R. Salem, A. Gordon, S. Mouli, et al., Y90 radioembolization significantly prolongs time to progression compared with chemoembolization in patients with hepatocellular carcinoma, Gastroenterology 151 (6) (2016) 1155–1163.

[14] R. Salem, R. Lewandowski, M.F. Mulcahy, et al., Radioembolization for hepatocellular carcinoma using Yttrium-90 microspheres: a comprehensive report of long-term outcomes, Gastroenterology 138 (1) (2010) 52–64.

[15] B. Sangro, L. Carpanese, R. Cianni, et al., Survival after yttrium-90 resin microsphere radioembolization of hepatocellular carcinoma across Barcelona clinic liver cancer stages: a European evaluation, Hepatology 54 (3) (2011) 868–878.

[16] C.S. Georgiades, E. Liapi, C. Frangakis, et al., Prognostic accuracy of 12 liver staging systems in patients with unresectable hepatocellular carcinoma treated with transarterial chemoembolization, J. Vasc. Intervent. Radiol. 17 (10) (2006) 1619–1624.

[17] F.D. Huitzil-Melendez, M. Capanu, E.M. O'Reilly, et al., Advanced hepatocellular carcinoma: which staging systems best predict prognosis? J. Clin. Oncol. 28 (17) (2010) 2889–2895.

[18] X. Adhoute, G. Penaranda, J.L. Raoul, et al., Usefulness of staging systems and prognostic scores for hepatocellular carcinoma treatments, World J. Hepatol. 8 (17) (2016) 703–715.

[19] J.L. Raoul, E. Boucher, Y. Rolland, E. Garin, Treatment of hepatocellular carcinoma with intra-arterial injection of radionuclides, Nat. Rev. Gastroenterol. Hepatol. 7 (1) (2010) 41–49.

[20] B. Sangro, M. Inarrairaegui, J.I. Bilbao, Radioembolization for hepatocellular carcinoma, J. Hepatol. 56 (2) (2012) 464–473.

[21] A. Forner, J.M. Llovet, J. Bruix, Hepatocellular carcinoma, Lancet 379 (9822) (2012) 1245–1255.

[22] M.M. Oken, R.H. Creech, D.C. Tormey, et al., Toxicity and response criteria of the eastern cooperative Oncology group, Am. J. Clin. Oncol. 5 (6) (1982) 649–655.

[23] T. Yau, V.Y.F. Tang, T.J. Yao, S.T. Fan, J.M. Lo, R.T.P. Poon, Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma, Gastroenterology 146 (7) (2014) 1691–1700.

[24] K. Okuda, T. Ohtsuki, H. Obata, et al., Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients, Cancer 56 (4) (1985) 918–928.

[25] J.M. Llovet, J. Bruix, Prospective validation of the Cancer of the Liver Italian Program (CLIP) score: a new prognostic system for patients with cirrhosis and hepatocellular carcinoma, Hepatology 32 (3) (2000) 679–680.

[26] G. Capuano, B. Daniele, G.B. Gaeta, C. Gallo, F. Perrone, A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators, Hepatology 28 (1998) 751–755.

[27] P.S. Kamath, W.R. Kim, The model for end-stage liver disease (MELD), Hepatology 45 (3) (2007) 797–805.

[28] C.G. Child, J.G. Turcotte, Surgery and portal hypertension, Major Probl. Clin. Surg. 1 (1964) 1–85.

[29] R.N. Pugh, I.M. Murray-Lyon, J.L. Dawson, M.C. Pietroni, R. Williams, Transection of the oesophagus for bleeding oesophageal varices, Br. J. Surg. 60 (8) (1973) 646–649.

[30] P.J. Johnson, S. Berhane, C. Kagebayashi, et al., Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade, J. Clin. Oncol. 33 (6) (2015) 550–558.

[31] F.E. Harrell, Regression modeling strategies : with applications to linear models, logistic regression, and survival analysis, in: Springer Series in Statistics, 2nd. ed, Springer, New York, 2015, p. 568, xxii.

[32] F.E. Harrell, Rms: Regression Modeling Strategies, 2017.

[33] L. Kang, W. Chen, N.A. Petrick, B.D. Gallas, Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach, Stat. Med. 34 (4) (2015) 685–703.

[34] A. Forner, M.E. Reig, C. Rodriguez de Lopa, J. Bruix, Current strategy for staging and treatment: the BCLC update and future prospects, Semin. Liver Dis. 30 (1) (2010) 61–74.

[35] A. Burroughs, J.F. Dufour, P. Galle, et al., Heterogeneity of patients with intermediate (BCLC B) Hepatocellular Carcinoma: proposal for a subclassification to facilitate treatment decisions, Semin. Liver Dis. 32 (4) (2012) 348–359.

[36] M. Op den Winkel, D. Nagel, M.A. de la Torre Aláez, et al., A simple prognostic scoring system for hepatocellular carcinoma treated with selective internal radiation therapy, Dig. Dis. (2021 Jun 10). Epub ahead of print. PMID: 34111866.