



Edeline, J. et al. (2016) A multicenter comparison between Child Pugh and ALBI scores in patients treated with sorafenib for hepatocellular carcinoma. *Liver International*, 36(2), pp. 1821-1828. (doi:[10.1111/liv.13170](https://doi.org/10.1111/liv.13170))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/119466/>

Deposited on: 23 May 2016

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

A multicenter comparison between Child Pugh and ALBI scores in patients treated with sorafenib for Hepatocellular Carcinoma

Julien Edeline¹, Jean-Frédéric Blanc², Philip Johnson^{3,4}, Boris Campillo-Gimenez¹, Paul Ross⁵, Yuk Ting Ma^{6,7}, Judy King⁸, Richard A. Hubner⁹, Kate Sumpter¹⁰, Suzanne Darby¹¹, Jeff Evans¹², Chinenye Iwuji¹³, Daniel Swinson¹⁴, Peter Collins¹⁵, Kinnari Patel¹⁶, Iqtedar Muazzam¹⁷, Daniel H Palmer^{3,4} and Tim Meyer^{7,18}

¹ Centre Eugène Marquis, Rennes, France; ²Centre Hospitalier Universitaire de Bordeaux, France; ³University of Liverpool, UK; ⁴Clatterbridge Cancer Centre, Liverpool, UK; ⁵King's College Hospital NHS Foundation Trust, London, UK; ⁶University of Birmingham, UK; ⁷University Hospital Birmingham NHS Foundation Trust, UK; ⁸Department of Oncology, Royal Free London NHS Foundation Trust, UK; ⁹The Christie NHS Foundation Trust, Manchester, UK; ¹⁰The Newcastle upon Tyne NHS Foundation Trust, UK; ¹¹Weston Park Hospital, Sheffield Teaching Hospitals NHS Foundation Trust, UK; ¹²University of Glasgow; Beatson West of Scotland Cancer Centre, Glasgow, UK; ¹³Leicester Royal Infirmary, UK; ¹⁴Leeds Teaching Hospitals NHS Trust, UK; ¹⁵University Hospitals Bristol NHS Foundation Trust, UK; ¹⁶Oxford University Hospitals NHS Foundation Trust, UK; ¹⁷Castle Hill Hospital, Hull and East Yorkshire Hospitals NHS Trust, UK; ¹⁸UCL Cancer Institute, London.

Address for correspondence:

Prof Tim Meyer, UCL Cancer Institute, University College London

72 Huntley Street, London WC1E 6BT

email; t.meyer@ucl.ac.uk Tel; +44 0207 679 6731, Fax; +44 0203 108 2025

Keywords: liver function; prognosis; albumin; bilirubin; cirrhosis

List of abbreviations:

HCC: Hepatocellular Carcinoma

CP: Child-Pugh

ALBI: Albumin-Bilirubin

INR: International Normalized Ratio

BCLC: Barcelona Clinic for Liver Cancer

OS: Overall Survival

AIC: Akaike Information Criterion

HR: Hazard Ratio

AASLD: American Association for Study of Liver Disease

NICE: National Institute for Health and Care Excellence

SNFGE: Société Nationale Française de Gastro-Entérologie

We report no relevant conflict of interest

Julien Edeline was funded by Fondation de France during the conduct of this study

Word Count: 3 728 in text, references and legends

Figures and tables: 5 Tables and 3 Figures, +1 supplementary Table

Authors' contributions: All authors contributed to the conception and design of the study. All authors contributed to the acquisition of data. Julien Edeline, Boris

Campillo-Gimenez and Tim Meyer analyzed the data, drafted a first version of this article, which was subsequently reviewed and corrected by all authors. The final

version was approved by all authors, and all authors agreed to be accountable for all

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Abstract

Background & aims: The ALBI grade was proposed as an objective means to evaluate liver function in patients with Hepatocellular Carcinoma (HCC). ALBI grade 1 vs 2 were proposed as stratification factors within the Child Pugh (CP) A class. However, the original publication did not provide comparison with the sub-classification by points (5 to 15) within the CP classification.

Methods: We retrospectively analyzed data from patients treated with sorafenib for HCC from 17 centers in United Kingdom and France. Overall survival (OS) was analyzed with the Kaplan-Meier method and a Cox regression model. Discriminatory abilities of the classifications were assessed with the log likelihood ratio, Harrell's C statistics and Akaike information criterion.

Results: Data from 1,019 patients were collected, of which 905 could be assessed for both scores. 92% of ALBI grade 1 were CP A5 while ALBI 2 included a broad range of CP scores of which 44% were CP A6. Median OS was 10.2, 7.0 and 3.6 months for CP scores A5, A6 and >A6, respectively ($P<0.001$), Hazard Ratio (HR)=1.60 (95%CI: 1.35-1.89, $P<0.001$) for A6 vs A5. Median OS was 10.9, 6.6 and 3.0 months for ALBI grade 1, 2 and 3, respectively ($P<0.001$), HR=1.68 (1.43-1.97, $P<0.001$) for grade 2 vs 1. Discriminatory abilities of CP and ALBI were similar in the CP A population, but better for CP in the overall population.

Conclusions: Our findings support the use CP class A as an inclusion criterion, and ALBI as a stratification factor in trials of systemic therapy.

Key points:

- This series offers a comparison of CP by points and ALBI in a large population of patients treated with sorafenib
- CP A class was a better inclusion criteria than ALBI grade 1 or ALBI grade ≤ 2
- ALBI grade and CP sub classification by points offered similar discriminative abilities when analyzed within the CP A class
- We proposed that CP class A should be used as an inclusion criteria, and ALBI grade as a stratification factor in clinical trials of systemic therapy

Introduction

Hepatocellular Carcinoma (HCC) is the second most common cause of cancer death worldwide [1]. As most cases arise in patients with chronic liver disease, the prognosis and treatment algorithms need to take into account both tumoral characteristics (such as size, number of lesions, portal vein thrombosis, extra-hepatic spread) and liver function [2]. A number of different composite scoring systems are currently used to define the extent of underlying liver dysfunction but the most widely is the Child-Pugh (CP) [2–5].

However, CP score has many limitations, including the empiric development of the score, and the subjectivity of some parameters including the clinical assessment of ascites and encephalopathy [3]. The Albumin-Bilirubin (ALBI) grade was recently proposed as an objective alternative to the CP classification in patients with HCC [6]. The ALBI grade was developed with the inclusion of variables representing liver functions (excluding tumoral characteristics) that were independently associated with survival, and was validated in large databases of patients treated with different treatment modalities in different countries. The score is calculated from serum albumin and bilirubin concentrations analyzed as continuous variables, and is then categorized in three grades. The calculation of the score requires a complex formula, but an approximation can be estimated using a nomogram, and the grade can also be assigned using a heat map, both provided in the original publication. As only albumin and bilirubin values are required, ALBI grade is entirely objective.

ALBI grade was proposed as an alternative to the CP score, with an emphasis put on the ability of ALBI grade 1 vs 2 to discriminate prognosis in patients with CP score A, thus potentially allowing stratification in clinical trials [6]. However, a limitation of the original publication was the absence of comparison with the CP sub-classification by

points (5 to 15). Hence it was not possible to determine the extent to which ALBI grade 1 and 2 differed from CP 5 and 6 with respect to sub-classifying CP A grade patients. Understanding this comparison has important implications for the implementation of ALBI. Our aim here was to compare the prognostic value of ALBI classification with CP sub-classification by points, in patients treated with sorafenib for HCC. The study was compliant with the STROBE guidelines [7].

Methods

Patients:

We retrospectively collected data from patients treated with sorafenib for advanced HCC from 17 centers in the United Kingdom and France. Data was acquired under an ethically approved protocol (REC reference 12/LO/1088). As a retrospective audit, we were advised by the ethical committee that written informed consent was not required. The data bases were anonymized. All consecutive patients treated with sorafenib were entered in the databases. HCC was confirmed either by biopsy or by radiologic criteria [2]. Relevant authorizations were obtained from institutional and ethical review boards for use of the data. Data collected included age, gender, cause of underlying liver disease, previous treatment for HCC, presence of extrahepatic spread, presence of portal vein thrombosis, performance status, alpha-fetoprotein, prothrombin time or International Normalized Ratio (INR), albumin, bilirubin, ascites and encephalopathy (as coded by centers for the CP classification), Barcelona Clinic for Liver Cancer (BCLC) classification and CP score as provided by centers. All data were obtained at the initiation of sorafenib.

Calculation of the liver functions scores:

In an effort to control for bias coming from miscoding, the CP score was then calculated from raw data, according to the original publication and adaptation for normalization of units used (Table 1) [3,5,8,9]. When both INR and prothrombin time were provided, INR was preferred. If the patient was recorded as receiving anticoagulation treatment, the CP coagulation score was assumed to be 1. The ALBI score was calculated using the formula: $ALBI\ score = (\log_{10}\ bilirubin \times 0.66) +$

(albumin \times - 0.085), and grades were attributed as follows: grade 1 if score \leq -2.60; grade 2 if score $>$ -2.60 but \leq -1.39; grade 3 if score $>$ -1.39 [6].

Statistical analysis:

Statistical analyses and graphs were performed on R statistical software version 3.1.1 (2014-07-10). Analyses were conducted on all patients, then on patients with a calculated CP class A only. Patients' characteristics were described by medians and range, and frequencies. Overall survival (OS) was calculated from the start of treatment with sorafenib to death, survival curves were estimated with the Kaplan-Meier approach, and compared with Log-Rank tests. A p value $<$ 0.05 was considered as statistically significant for all analyses.

Reliability of CP score and ALBI grade was evaluated using contingency tables, raw concordance rates and quadratic weighted kappa coefficients with adjusted bootstrap percentile confidence interval.

The performance of a prognostic system has been shown to be related to homogeneity (small differences in survival among patients in the same class within each system) and discriminatory ability (greater differences in survival among patients in different stages within each system) [10,11].

Harrell's C statistics was used to measure the discriminatory ability of CP scores and the ALBI score. A higher Harrell's C statistics indicates higher discriminative ability.

The log likelihood ratio was calculated with Cox regression to determine homogeneity. A higher log likelihood ratio indicates higher homogeneity of survival between patients classified in a same class. In addition, the results of Cox regression were expressed using the Akaike information criterion (AIC), which shows how the

model, and so how the explanatory variable (the staging system) is informative, meaning that the staging system explains by itself most of the difference in prognosis between patients. The lower the AIC, the lower the model loses information and the better the goodness of fit. We calculated the relative likelihood of ALBI vs CP using the formula: $\exp((AIC_{CP}/AIC_{ALBI})/2)$. The relative likelihood represents the probability that ALBI minimizes information as effectively as CP, and could thus be interpreted as a p value for the comparison of both AIC. For these analyses, the classifications were compared as ALBI grade 1 vs 2 vs 3 and CP A5 vs A6 vs >A6 for the whole population, and ALBI grade 1 vs 2 and CP A5 vs A6 for the CP A population. We grouped together all patients with >A6 score due to low numbers in each class, usual grouping of CP B patients in clinical studies, and similar survival of patients with CP B7 and >B7.

Results

Patient characteristics and Child-Pugh scoring

From February 2003 to August 2014, 1019 patients were treated with sorafenib for HCC. The characteristics of the cohort are presented in Table 2. Median follow-up was 14.3 months, and 843 patients (82.7%) died.

All data for calculation of CP scores were available in 905 patients (88.8%). As shown in the flow-chart diagram (Figure 1), most missing values were coagulation, followed by encephalopathy and albumin. As presented in the Supplementary Table, there were discrepancies between CP score calculated from raw data using our CP classification and CP score provided by centers. Overall, results were discrepant in 109 out of 904 patients (12.1%). However, the overall concordance stayed excellent, with a kappa of 0.90 (95% Confidence Interval (CI): 0.87-0.92). The cause of misclassification could be determined when the CP score provided was 5 but our calculation based on raw data was >5: in the 42 patients with a calculated CP of 6, the abnormal value was albumin in 39 and bilirubin in 3 cases; in the 4 patients with CP calculated at 7, abnormal values were albumin and bilirubin both misclassified in 3, and bilirubin alone misclassified in 1. Using the CP given by our calculations, abnormal values leading to CP A6 score (n=264) were albumin in 195 (73.9%), ascites in 36 (13.6%), bilirubin in 28 (10.6%), encephalopathy in 4 (1.5%) and coagulation in 1 (0.4%). Values leading to CP B7 (n=101) involved albumin in 81 (80.2%), bilirubin in 49 (48.5%), ascites in 31 (30.7%), coagulation in 5 (5.0%) and encephalopathy in 3 (3.0%). Thus, in this population, albumin was the main driver of CP A5 vs A6 vs B7 classification, followed by bilirubin and ascites.

Child-Pugh and ALBI grade comparison

ALBI grade could be calculated for 962 patients (94.4%), including all patients with allocated CP score. Median ALBI score was -2.36 (range: -4.39/-0.35). Overall, 327 patients (34.0%) were ALBI grade 1, 574 (59.7%) were grade 2 and 61 (6.3%) were grade 3. Comparison between ALBI grade and CP scores are reported in Table 3 for the 905 patients with both scores. While 91.7% of ALBI grade 1 patients were classified as CP A5, ALBI grade 2 comprised patients with CP A6 in 44.3%, CP A5 in 33.3%, and CP >A6 in 22.4%. ALBI grade 3 patients were all classified as CP >A6. Two patients with ALBI grade 1 were classified as CP B7, and 6 patients with ALBI grade 2 were classified as CP B9; such discrepancies were always explained by 3 points given to ascites, except for one ALBI grade 2 case with 2 points given to ascites and 2 points to coagulation.

CP scores were significantly associated with OS (log-rank $p < 0.001$), with median OS of 10.2 months (95%CI: 8.6-11.7) in CP A5 ($n=458$), 7.0 (95%CI: 6.1-7.9) in CP A6 ($n=264$, $p < 0.001$ vs CP A5), 3.6 (95%CI: 3.1-4.2) in CP >A6 ($n=183$, $p < 0.001$ vs CP A6) (Figure 2A). If separated from patients with CP >B7, CP B7 patients ($n=101$) still had a dismal prognosis, with median OS of 3.9 months (95%CI: 2.7-5.1, $p=0.009$ vs CP A6, $p=0.57$ vs CP B8). The Hazard Ratio (HR) for OS between A6 vs A5 was 1.60 (95%CI: 1.35-1.89, $p < 0.001$), and between >A6 vs A5 was 2.49 (95%CI: 2.07-3.00, $p < 0.001$).

ALBI grades were also significantly associated with OS (log-rank $p < 0.001$), with median OS of 10.9 months (95%CI: 9.2-12.6) for ALBI grade 1 ($n=302$), 6.6 (95%CI: 5.9-7.3) for ALBI grade 2 ($n=544$, $p < 0.001$ vs ALBI grade 1), and 3.0 (95%CI: 2.1-3.8) for ALBI grade 3 ($n=59$, $p < 0.001$ vs ALBI grade 3) (Figure 2B). The HR between ALBI grade 2 vs grade 1 was 1.68 (95%CI: 1.43-1.97, $p < 0.001$), and between ALBI

grade 3 vs grade 1 was 3.36 (95%CI: 2.48-4.56, $p < 0.001$). Focusing on the CP class A population ($n=722$), median OS was 8.8 months (95%CI: 8.0-9.6), and when the group was split according to ALBI grade was 10.9 months (95%CI: 9.2-12.7) for ALBI grade 1 and 7.5 (95%CI: 6.7-8.3, $p < 0.001$) for ALBI grade 2, with a HR=1.54 (95%CI: 1.30-1.82, $p < 0.001$).

In patients classified as ALBI grade 1 ($n=302$), CP A5 patients ($n=277$) had significantly better survival than CP A6 patients ($n=23$), with respective median OS of 11.6 (95%CI: 9.9-13.3) and 7.9 months (95%CI: 4.4-11.3) ($p=0.008$, Figure 3A). In patients classified as ALBI grade 2 ($n=544$), median OS were 8.6 months (95%CI: 6.9-10.3) in CP A5 ($n=181$), 6.9 (95%CI: 6.0-7.9, $p=0.016$ vs CP A5) in CP A6 ($n=241$) and 3.9 (95%CI: 3.2-4.6, $p=0.002$ vs CP A6) in patients with CP >A6 ($n=122$) (Figure 3B).

In patients classified as CP A5 ($n=458$), ALBI grade 1 patients ($n=277$) had significantly better survival than ALBI grade 2 patients ($n=181$), with respective median OS of 11.6 (95%CI: 9.9-13.3) and 8.6 months (95%CI: 6.9-10.3) ($p=0.003$, Figure 3C). However, in patients classified as CP A6 ($n=264$), there was no significant difference between patients classified as ALBI grade 1 ($n=23$) and grade 2 ($n=241$), with respective median OS of 7.9 (95%CI: 4.4-11.3) and 6.9 months (95%CI: 6.0-7.9) ($p=0.96$, Figure 3D). In patients with CP >A6 ($n=183$), there was no difference between ALBI grade 2 ($n=122$) and grade 3 ($n=59$), with respective median OS of 3.9 (95%CI: 3.2-4.6) and 3.0 months (95%CI: 2.1-3.8) ($p=0.17$, Figure 3E).

Combining both classifications might offer a broader range of prognosis (Figure 3F and Table 4), classifying patients with progressively worse prognosis as CP A5-ALBI 1, CP A5-ALBI 2, CP A6-ALBI 1, CP A6-ALBI 2, CP >A6-ALBI 2 and CP >A6-ALBI 3.

However, numbers in certain subgroups (especially CP A6-ALBI 1) are too low to draw definitive conclusions.

Results regarding the discriminative abilities of both scoring systems are reported on Table 5. CP was associated with less information loss than ALBI in the overall population as evidenced by a lower AIC with a relative likelihood <0.001 . However, this was not the case when only the CP class A population was considered.

Moreover, discriminative abilities of CP and ALBI were very similar as evidenced by similar Harrell's C statistics, either in the overall population or the CP class A population, showing that both staging systems are equally able to accurately differentiate the prognosis of patients. Likewise, homogeneity of classes appeared different in the overall population, as illustrated by higher likelihood ratio score for CP representing better homogeneity of the CP system in this population, meaning that a same ALBI grade might group together patients with different prognosis. However, homogeneity was similar between both scoring systems when focusing in the CP class A population. Similar results were found if CP scores provided by centers were used, rather than CP scores calculated. Hence, the CP score might be more informative than ALBI in the overall population, but the CP score and the ALBI grade showed very similar prognostic abilities in the CP class A population.

Discussion

In this large retrospective series of European patients treated with sorafenib for advanced HCC, we found that ALBI grade and CP sub-classification by points provide similar prognostic information when focusing on CP A population. The majority of patients classified as ALBI grade 1 were also classified as CP A5, while the patients classified as ALBI grade 2 showed a broader range of classification within CP scores, including many patients with CP B. To our knowledge, this is the first multicenter study comparing ALBI and CP sub-classification by points in a large number of European patients, and our findings are supported by those of two recently published studies from Asia [12,13]. Moreover, two recent studies originating from the same consortium showed that the incorporation of ALBI into the BCLC or Japan Integrated Staging (JIS) classification could add information over CP [14,15].

The ALBI grade was developed as an evidence-based scoring system specifically for assessing liver function in patient with HCC and uses only albumin and bilirubin which are independently associated with OS [6]. The formula derived from the model avoids subjectively chosen thresholds so that albumin and bilirubin are analyzed as continuous variable using international units rather than categorical variables. However, thresholds were introduced to allow the allocation of grade based on risk. Using a Japanese training set, the grades were defined to classify the 25% with the lowest risk of death as grade 1, those 10% with highest risk of death as grade 3 and those in between as grade 2. The model was then validated in independent international cohorts using different treatment modalities. The use of a formula might be judged cumbersome, but a nomogram and a heat map are provided, thus allowing easier use in a clinical setting.

Conversely, and as already discussed in previous publications [3,6], the CP scoring system has several limitations for use in patients with HCC. First, it was developed using an empiric approach rather than evidence-based and was intended to be used to define the operative risk in cirrhotic patients with esophageal varices. Therefore it is not strictly applicable to non-cirrhotic patients with HCC. Second, two of the five factors that contribute to the CP score -clinical ascites and encephalopathy- are subjective and prone to error. Finally, the score gives similar weighting to parameters which may not be of similar prognostic importance. Despite these limitations, CP score consistently emerges as a significant prognostic factor in clinical series evaluating outcomes from therapeutic interventions. Given its widespread use, most publications do not define the CP [16] yet we found surprisingly significant discrepancies in scoring, illustrated by a 12% discordance between our score based on the raw data and the score provided by the participating centers. At least two explanations could be found for such high discrepancies. First, there is surprisingly no international consensus about CP scoring system, as illustrated by different thresholds proposed for bilirubin (34, 35, 50 or 51 $\mu\text{mol/L}$) [5,9,17–21]. Second, miscoding errors at the thresholds values are frequent, as presented by 39% of the patients with errors in our cohorts. Using only two variables, ALBI might be less prone to such coding-related errors. However, despite these miscoding of CP scores, the concordance between provided and calculated CP scores remained excellent with a kappa at 0.90, which is superior than kappa obtained for inter-observer variability judged acceptable in radiological evaluation, in the range of 0.50 to 0.60 [22,23].

The respective roles of ALBI and CP in clinical practice and research need to be considered in view of our results. Broadly, these classifications have a potential role

in patient selection and stratification, and to date, the majority of clinical trials in advanced HCC have included CP A disease as a key inclusion criteria.

Our results suggest that ALBI grade should not be used instead of CP as an inclusion criteria for clinical trials. First, the ALBI grade 2 has an intermediate prognosis (median OS of 6.6 months), but seems to encompass patients with quite different prognosis, as shown by higher loss of information and lower homogeneity for ALBI as compared with CP in the overall population, and illustrated by a median of 8.9 months for ALBI grade 2 – CP A5 vs 3.9 months for ALBI grade 2 – CP >A6. This is probably related to the chosen design of ALBI grade 2, comprising patients between 25% and 90%-risk of death. Moreover, one might advocate to select only ALBI grade 1 CP A5 patients, as this population appears to have the best prognosis. However, that could exclude a significant number of patients who might otherwise benefit, as was demonstrated in the SHARP trial in which a survival advantage was demonstrated in both CP A5 and A6 patients. Second, excluding only patients with ALBI grade 3 would probably be less efficient than excluding patients with CP class B, as ALBI grade 2 also included 22% of patients classified as CP class B, associated with a worse prognosis with a median OS of only 3.9 months. Another area of debate is the appropriate evaluation of CP class B patients. This criterion is frequently used for exclusion from clinical trials. Some authors advocate similar treatment strategies for CP B7 patients as for CP A patients, when ascites is absent [24]. However, in our cohort, we did not find any statistically or clinically significant difference in OS between patients classified as CP B7 vs CP >B7. The role of sorafenib in HCC patients with CP B cirrhosis is under evaluation in 2 ongoing randomized controlled trials (PRODIGE 21 and BOOST trials) [25,26].

Conversely, the comparison of both scores when restricted to CP class A patients showed that taken as a whole, ALBI grade 1 vs grade 2 was able to provide similar information as CP A5 vs A6, with the use of only 2 objective parameters rather than 5 parameters, some of them subjective, which suggests that ALBI might be preferred as a stratification factor.

Limitations of our study comprise its retrospective nature, and the discrepancies between calculated and provided CP scores. Moreover, this study is limited to European centers, and patients treated with sorafenib, and results should be confirmed in other contexts.

In conclusion, our findings support the continued use of CP class A as an inclusion criteria, but ALBI grade 1 vs 2 as a stratification factor for clinical trials of systemic therapy. Prospective studies will help to define further the relative benefits of ALBI and CP in different therapeutic contexts and inform clinical guidelines.

Acknowledgments

Tim Meyer is part funded by the NIHR Biomedical Research Centre at University College London Hospitals NHS Foundation Trust and University College London.

Julien Edeline received a research grant from Fondation de France.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012: Global Cancer Statistics, 2012. *CA. Cancer J. Clin.* 2015; 65:87–108.
2. EASL, EORTC, others. EASL–EORTC Clinical Practice Guidelines: management of hepatocellular carcinoma. *J. Hepatol.* 2012; 56:908–943.
3. Durand F, Valla D. Assessment of Prognosis of Cirrhosis. *Semin. Liver Dis.* 2008; 28:110–122.
4. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatol.* 2000; 31:864–871.
5. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br. J. Surg.* 1973; 60:646–649.
6. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of Liver Function in Patients With Hepatocellular Carcinoma: A New Evidence-Based Approach--The ALBI Grade. *J. Clin. Oncol.* 2015; 33:550–558.
7. STROBE Statement: Home [Internet]. [cited 2015 Dec 9]; at <<http://strobe-statement.org/index.php?id=strobe-home>>
8. Trey C, Burns DG, Saunders SJ. Treatment of hepatic coma by exchange blood transfusion. *N. Engl. J. Med.* 1966; 274:473–481.
9. Garcia-Tsao G, Sanyal AJ, Grace ND, Carey W, Practice Guidelines Committee of the American Association for the Study of Liver Diseases, Practice Parameters Committee of the American College of Gastroenterology. Prevention and

management of gastroesophageal varices and variceal hemorrhage in cirrhosis. *Hepatology*. 2007; 46:922–938.

10. Ueno S, Tanabe G, Sako K, Hiwaki T, Hokotate H, Fukukura Y, et al. Discrimination value of the new western prognostic system (CLIP score) for hepatocellular carcinoma in 662 Japanese patients. *Cancer of the Liver Italian Program*. *Hepatology*. 2001; 34:529–534.

11. Marrero JA, Fontana RJ, Barrat A, Askari F, Conjeevaram HS, Su GL, et al. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. *Hepatology*. 2005; 41:707–716.

12. Ogasawara S, Chiba T, Ooka Y, Suzuki E, Kanogawa N, Saito T, et al. Liver function assessment according to the Albumin-Bilirubin (ALBI) grade in sorafenib-treated patients with advanced hepatocellular carcinoma. *Invest. New Drugs*. 2015; 33:1257–1262.

13. Hiraoka A, Kumada T, Michitaka K, Toyoda H, Tada T, Ueki H, et al. Usefulness of albumin-bilirubin (ALBI) grade for evaluation of prognosis of 2584 Japanese patients with hepatocellular carcinoma. *J. Gastroenterol. Hepatol*. 2015; doi:10.1111/jgh.13250

14. Chan AWH, Chong CCN, Mo FKF, Wong J, Yeo W, Johnson PJ, et al. Applicability of Albumin-Bilirubin-based Japan Integrated Staging (ALBI-T) score in hepatitis B-associated hepatocellular carcinoma. *J. Gastroenterol. Hepatol*. 2016; doi:10.1111/jgh.13339

15. Chan AWH, Kumada T, Toyoda H, Tada T, Chong CCN, Mo FKF, et al. Integration of albumin-bilirubin (ALBI) score into Barcelona clinic liver cancer (BCLC)

system for hepatocellular carcinoma. *J. Gastroenterol. Hepatol.* 2016;
doi:10.1111/jgh.13291

16. Okajima C, Aii S, Tanaka S, Matsumura S, Ban D, Ochiai T, et al. Prognostic role of Child-Pugh score 5 and 6 in hepatocellular carcinoma patients who underwent curative hepatic resection. *Am. J. Surg.* 2015; 209:199–205.

17. Wikipedia. Wikipedia Child-Pugh entry: available at:
http://en.wikipedia.org/wiki/Child-Pugh_score. 2015;

18. MedCalc. Medical Calculator, available at: <http://www.mdcalc.com/child-pugh-score-for-cirrhosis-mortality/>. 2015;

19. NICE. NICE: What is the Child-Pugh score: available at:
<http://www.evidence.nhs.uk/search?q=%22What+is+the+Child+Pugh+score%22>.
2014;

20. SNFGE. Score de Child-Pugh; available at: www.snfge.org/download/file/fid/420
[Internet]. 2008; at <www.snfge.org/download/file/fid/420>

21. Murray KF, Carithers RL, AASLD. AASLD practice guidelines: Evaluation of the patient for liver transplantation. *Hepatol.* 2005; 41:1407–1432.

22. Suzuki C, Torkzad MR, Jacobsson H, Aström G, Sundin A, Hatschek T, et al. Interobserver and intraobserver variability in the response evaluation of cancer therapy according to RECIST and WHO-criteria. *Acta Oncol.* 2010; 49:509–514.

23. Seyal AR, Gonzalez-Guindalini FD, Arslanoglu A, Harmath CB, Lewandowski RJ, Salem R, et al. Reproducibility of mRECIST in assessing response to transarterial

radioembolization therapy in hepatocellular carcinoma. *Hepatol.* 2015;

doi:10.1002/hep.27915

24. Piscaglia F, Terzi E, Cucchetti A, Trimarchi C, Granito A, Leoni S, et al.

Treatment of hepatocellular carcinoma in Child-Pugh B patients. *Dig. Liver Dis.* 2013;
45:852–858.

25. Clinicaltrials.gov. PRODIGE 21 trial: available

at:<https://clinicaltrials.gov/ct2/show/NCT01357486?term=child+b+sorafenib&rank=1>.

2015;

26. Clinicaltrials.gov. BOOST trial: available at:

<https://clinicaltrials.gov/ct2/show/NCT01405573?term=child+b+sorafenib&rank=3>.

2015;

Table 1: Scoring used for the Child-Pugh classification (3, 5, 8)

	1 point	2 points	3 points
Albumin	> 35 g/L	35-28 g/L	< 28 g/L
Bilirubin	< 34 mcmol/L	34-51 mcmol/L	> 51 mcmol/L
Coagulation: -INR -Prothrombin Time, as a percentage relative to control	< 1.7 > 50%	1.7-2.3 40-50%	> 2.3 < 40%
Ascites	None	Medically controlled	Refractory
Encephalopathy [8]	None	Grade 1 or 2 (or medically controlled)	Grade 3 or 4 (or refractory)

Table 2: Characteristics of the cohort

Age (n=1019) median (range)	67 (17-89)
Gender (n=945), male / female	810 (85.7%) / 135 (14.3%)
Excessive alcohol consumption (n=921)	375 (40.7%)
Hepatitis B Virus (n=921)	81 (8.8%)
Hepatitis C Virus (n=921)	164 (17.8%)
Non-Alcoholic Steato-Hepatitis (n=921)	228 (24.8%)
No identified cause (n=921)	258 (28.0%)
Albumin, g/L (n=967), median (range)	37 (17-61)
Bilirubin, mcmol/L (n=989), median (range)	15 (3-436)
Previous treatment for HCC (n=857)	495 (57.8%)
Extra-Hepatic Spread (n=910)	338 (37.1%)
Portal Vein Thrombosis (n=997)	375 (37.6%)
Alpha-Feto Protein, ng/mL (n=978), median (range)	128 (0-849,553)
Performance status (n=902): 0/1/2/3	389 (43.1%) / 399 (44.2%) / 109 (12.1%) / 5 (0.6%)
BCLC stage (n=910): A/B/C/D	11 (1.2%) / 178 (19.6%) / 711 (78.1%) / 10 (1.1%)

BCLC: Barcelona Clinic for Liver Cancer; CP: Child-Pugh

Table 3: Correspondences between calculated Child-Pugh scores and ALBI grades

(n=905):

	CP A5	CP A6	CP B7	CP B8	CP B9	CP C10	CP C11
ALBI grade 1	277	23	2	0	0	0	0
ALBI grade 2	181	241	87	29	6	0	0
ALBI grade 3	0	0	12	22	12	11	2

Table 4: Median overall survival according to both Child-Pugh and ALBI classifications

Class	n	median OS (95%CI)
CP5-ALBI1	277	11.6 (9.8-13.9)
CP5-ALBI2	181	8.6 (7.4-10.6)
CP6-ALBI1	23	7.9 (5.8-NA)
CP6-ALBI2	241	6.9 (6.1-7.7)
CP>6-ALBI2	122	3.9 (3.3-5.4)
CP>6-ALBI3	59	3.0 (2.1-4.0)

CP: Child-Pugh; OS: Overall Survival; NA: Not assessable

Table 5: Discriminative abilities of Child-Pugh and ALBI scoring systems. A higher Harrell's C statistics indicates higher discriminative ability (patients of different risk groups have higher difference in survival). A higher log likelihood ratio indicates higher homogeneity (similar survival between patients classified in a same class). A lower Akaike information criterion indicates lower loss of information (the classification explaining most of the difference in prognosis between patients). The relative likelihood calculated represents the probability that ALBI minimizes information loss as effectively as CP.

	All patients (n=905)		CP A patients (n=722)	
	CP	ALBI	CP	ALBI
Harrell's C statistic	0.61	0.60	0.56	0.57
Log likelihood ratio	90.7	68.3	29.0	26.3
AIC	8876	8898	6649	6652
Relative likelihood of ALBI equivalence of AIC comparatively to CP	<0.001		0.22	

AIC: Akaike Information Criterion; CP: Child-Pugh

Figures legends:

Fig. 1: Flow-chart diagram of patients included in the study.

Fig. 2: Overall survival according to (A) Child-Pugh scores and (B) ALBI grade.

Fig. 3: Overall Survival according to Child-Pugh score (A and B) in patients classified as (A) ALBI grade 1 and (B) ALBI grade 2, and according to ALBI grade (C, D and E) in patients classified as (C) Child-Pugh A5, (D) Child-Pugh A6 and (E) Child-Pugh B7 or more.