Prognostic models for survival in patients with stable cirrhosis. A multicentre cohort study.

Maria Kalafateli, MD[1]*, Konstantinos Zisimopoulos, MD[1]*, Georgia Vourli, Biostatistician[2], Cristina Rigamonti, MD[3], John Goulis, MD[4], Emanuel Manesis, MD[5], Spilios Manolakopoulos, MD[5], Emmanuel Tsochatzis, MD[3], Aikaterini Georgiou, MD[6], Georgia Diamantopoulou, MD[1], Konstantinos Thomopoulos, MD[1], Charalambos Gogos, MD[7], Giota Touloumi, Biostatistician/Epidemiologist[2], Evangelos Akriviadis, MD[4], Chryssoula Lambropoulou-Karatza, MD[7], Christos Triantos, MD[1]

Kalafateli M and Zisimopoulos K contributed equally to this work

1. Department of Gastroenterology, University Hospital of Patras, Patras, Greece

2. Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece

3. The Royal Free Sheila Sherlock Liver Centre, Royal Free Hospital, London, United Kingdom

4. Fourth Department of Internal Medicine, Aristotle University Medical School, Hippokration General Hospital of Thessaloniki, Thessaloniki, Greece

5. 2nd Department of Internal Medicine, Athens University Medical School, Hippokration General Hospital of Athens, Athens, Greece

6. Department of Gastroenterology, Polyclinic General Hospital, Athens, Greece

7. Department of Internal Medicine, University Hospital of Patras, Patras, Greece

Kalafateli Maria: mariakalaf@hotmail.com

Zisimopoulos Konstantinos: condoczisimo@hotmail.com

Vourli Georgia: gvourli@med.uoa.gr

Rigamonti Cristina: crigamo@tin.it

Goulis John: igoulis@gmail.com

Manesis Emanuel: emanesis@med.uoa.gr

Manolakopoulos Spilios: spilios1@otenet.gr

Tsochatzis Emmanuel: e.tsochatzis@ucl.ac.uk

Georgiou Aikaterini: kthgeorgiou@yahoo.gr

Diamantopoulou Georgia: geodiamant@hotmail.com

Thomopoulos Konstantinos: kxthomo@hotmail.com

Gogos Charalambos: cgogos@med.upatras.gr

Touloumi Giota: gtouloum@med.uoa.gr

Akriviadis Evangelos: akriviev@med.auth.gr

Lambropoulou-Karatza Chryssoula: ckaratza@upatras.gr

Triantos Christos: chtriantos@hotmail.com

Correspondence to: Christos Triantos, MD, D. Stamatopoulou 4, Rio 26504, Patras, Greece.chtriantos@hotmail.com

Telephone: +306972894651

Fax: +302610625382

Supported by: Nothing to declare

Conflict-of-interest statement: The Authors declare that there is no conflict of interest.

ABSTRACT

BACKGROUND: Two models are mostly used to predict survival in cirrhosis, the Child-Pugh score (CP score) and the model for end-stage liver disease (MELD).

AIMS: The aim of this study is to evaluate the CP score and the MELD for short and long-term prognosis in cirrhosis, as well as CP-creatinine, MELD-Na and UKELD.

METHODS: 1047 patients from five referral centres were included. Men/women: 620/427, median age: 58 years (IQR: 48-66), median follow-up: 33 months (IQR: 12-74), CP (A/B/C): 493/357/147, CP score: 7 (IQR: 5-9), MELD: 12 (IQR: 9-16). The performance of each score was evaluated by the Cox hazard model in terms of their: discrimination ability (C-index and Sommer's D) and calibration (3, 12 months). Internal validation was done with bootstrapping (100 samples).

RESULTS: 352 patients (33.6%) died. All scores were significantly associated with overall mortality, when assessed by univariate Cox analysis. CP-creatinine score performed significantly better than all other scores (Bootstrap C-index 0.672, 95% CI 0.642-0.703, Bootstrap Sommer's D 0.344 (0.285-0.401)), apart from CP score, which showed similar performance. Inclusion in the multivariable Cox model of age together with CP-creatinine improved the discriminative ability of the model (Bootstrap C-index (95%CI): 0.700 (0.661-0.740)). In terms of calibration, CP-creatinine was the best for both 3-month and 12-month survival in the total population.

CONCLUSIONS: CP score and CP-creatinine have better prognostic value compared to MELD, MELD-Na and UKELD scores for predicting short and long term mortality in patients with stable cirrhosis.

KEY WORDS: cirrhosis, liver diseases, prognosis, survival, Cox Proportional Hazards Models.

INTRODUCTION

Liver cirrhosis has been sub-classified in clinical stages reflecting differences in mortality rates [1]. Thus, a score that can accurately predict mortality is of major importance. Two models are mostly used in clinical practice, Child-Pugh grade/score (CP score) and model for end-stage liver disease (MELD). However, both were developed for short-term prognosis following procedures for portal hypertension, namely portacaval shunt for CP score [2] and transjugular intrahepatic portosystemic shunt (TIPS) for MELD [3].

Several studies have compared the performance of both scoring systems in predicting survival with conflicting results. In a systematic review [4] comparing the prognostic accuracy of MELD versus CP score, only 4 of 11 studies including 12,532 patients showed a superiority of MELD compared to CP score in predicting 3-month survival. Moreover, the discriminative ability of MELD to predict mortality following liver transplantation was poor (concordance (c) statistic below 0.7 in all 6 studies).

Apart from the short-term prognostic utility, both models have limitations [5]. CP score has not been validated statistically and uses clinically assessed ascites and encephalopathy which can be influenced by subjective interpretation. On the other hand, MELD score needs computation and lacks the simplicity of the CP score [6]. These limitations have led to the development of variants of the original scoring systems, including MELD-Na (including serum sodium, a significant predictor of early mortality and indirect indicator of the presence of ascites) [7, 8], and CP-creatinine (which includes renal function in order to increase its predictive accuracy) [9].

We conducted this multicentre cohort study to evaluate MELD, CP score, MELD-Na, CP-creatinine and United Kingdom End-stage Liver disease (UKELD) scores as prognostic models for long-term prognosis as well as for 3-, 6-, and 12-

monthsurvival in patients with stable cirrhosis taking into consideration the conflicting results in the literature.

PATIENTS AND METHODS

Study population

This is a multicentre retrospective cohort study including consecutive outpatients with cirrhosis between1974 and 2012 in five referral centres: a) Department of Gastroenterology, University Hospital, Patras, Greece (n=496), b) The Royal Free Sheila Sherlock Liver Centre, Royal Free Hospital and University College School of Medicine, London, UK (n=341), c) 4th Department of Internal Medicine, Hippokration General Hospital of Thessaloniki, Medical School of Aristotle University, Greece (n=104). d) 2nd Academic Department of Internal Medicine, University of Athens, Hippokration General Hospital, Athens, Greece (n=102) and e) Department of Gastroenterology, Polyclinic General Hospital, Athens, Greece (n=92). The included patients were recruited from the hepatology clinics of the physicians that participated in the study. The study was compliant with the ethical standards of the Declaration of Helsinki. There were 1135 patients with cirrhosis. Patients with hepatocellular carcinoma (HCC) (n=78) at baseline and those with no information on any of the prognostic scores under investigation (n=10) were excluded from the analysis. None of the patients had concomitant human immunodeficiency virus (HIV) infection. The study cohort finally included 1047 patients.

Diagnosis of cirrhosis was based on clinical, laboratory and previous radiological and histological findings. The management according to etiology of cirrhosis was based on international practice guidelines [10-13]. None of the patients with HCV cirrhosis had sustained virological response before enrolment. In patients enrolled before 1990, diagnosis of HCV infection was documented later and these

patients were included in the 'HCV etiology' cohort or in 'other etiology' cohort, if HCV seropositivity was not confirmed. Medical history, physical examination and laboratory tests were available in all patients at their first admission. The definitions of diabetes mellitus, chronic kidney disease, hypertension, ischemic heart disease and chronic respiratory diseases (including chronic obstructive pulmonary disease, asthma, obstructive sleep apnea syndrome, bronchiectasis, pulmonary hypertension) were based on the workup of the attending physician according to the current internationally accepted guidelines. Patients were followed-up by experienced hepatologists until death, liver transplantation, study closure or last contact with the health system**.** The management of moderate/severe ascites, spontaneous bacterial peritonitis, hepatorenal syndrome, hepatic encephalopathy and acute variceal bleeding was according to the international practice guidelines as published at the time of enrolment. The final status (alive, transplanted, death) was recorded.

End-points and definitions

The primary end-point was the comparison of the prognostic accuracy of the different models regarding mortality at 3, 6 and 12 months and at the end of follow-up. We also performed a subgroup analysis for different aetiologies of liver disease and for patients with MELD≥ 15. CP score, MELD, MELD-Na and UKELD were assessed as previously described [2, 14-17] . CP-creatinine was calculated as CP score + 1, 2 or 3 (if creatinine<1.3, 1.3-1.5 and ≥1.5 mg/dl, respectively) [9]. Regarding PBC patients, the Mayo risk score [18] and the Royal Free score [19] were also calculated; CP score was assessed as suggested by Pugh *et al* [2]. Only PBC patients with histological staging of 3 and 4 were included [20].

Statistical Analysis

Categorical characteristics were summarized as counts and corresponding percentages, while medians and interquartile ranges (IQRs) were given for the continuous characteristics. Predictors of overall survival were evaluated using Cox proportional hazards models. Patients were censored at the last follow-up date or at the date of liver transplantation. The discrimination ability of each proposed prognostic score was assessed by both the C-index and the Sommer's D [21]. To adjust for overfitting, our results were internally validated with bootstrapping [22, 23]. Comparisons of the c-index across different models were performed using non-parametric tests [24]. As a second aspect of models' validity, the calibration plots at 3 and 12 months were presented [25].

For illustration, dynamic/incident Receiver-operating curve (ROC) suitable for survival data is presented at 3, 6 and 12 months of follow-up using the risk-set package in R, based on previously published work [26].

To account for known confounders, all scores were evaluated separately in multivariable models adjusted for age, sex, ischemic heart disease, total bilirubin, creatinine, sodium, INR, albumin, encephalopathy and ascites. The variables were adjusted depending on the components of the scores, excluding factors incorporated into the definition of the score; for instance, sodium, INR, bilirubin and creatinine were not included in the model when MELD-Na was evaluated.

Several sensitivity analyses were carried out: a) to account for potential differences across centers, analysis was repeated after including the different centers in both the univariable and multivariable models; b) subgroup analyses among those with and without decompensation as well as among those with and without anti-HBV treatment were performed. All analyses were performed with Stata11.2.

Institutional review board statement

The study was reviewed and approved by the Ethics committee of University Hospital of Patras.

RESULTS

A total of 1047 patients were included in the analysis. Patients' characteristics are presented in Table 1. Patients were followed-up for a median of 33 months (IQR 12-74). Fifty-nine deaths were observed within the first 3 months and 121 deaths within the first year of follow-up. During the entire follow-up, 352 patients (33.6%) died and 42 (4%) patients were transplanted.

Comparison of the predictive models (CP, CP-creatinine, MELD, MELD-Na and UKELD) for 3-, 6-, 12-month and overall survival

Total Population

Discrimination measures as estimated in the original sample did not differ substantially from the corresponding bootstrap estimates, indicating low overfitting. CP-creatinine score performed overall better than the other examined scores (bootstrap C-index: 0.672, 95% CI: 0.642-0.703, bootstrap Sommer's D: 0.344 (0.285-0.401)) (Table 2) but this difference was not significant when compared to CP score. Subgroup analysis showed that all scores performed better in patients with compensated compared to those with decompensated cirrhosis, although the CP-creatinine score performed overall better than the other examined scores, as in the main analysis (Suppl. Table 1). The C-indexes of CP score and CP-creatinine scores were significantly higher compared to MELD, MELD-Na or UKELD scores (Table 3). Sensitivity analysis after inclusion of the centers together with each predictive score in a model yielded similar results (data not shown).

The dynamic/incident Receiver Operating Characteristic (ROC) curves of CP score, CP-creatinine, MELD, MELD-Na and UKELD score for 3-month, 6-month and

12-month survival are shown in Figure 1 (a-c). CP-creatinine and CP scores had the best discriminative ability for all time points. Nevertheless, given that the C-index was relatively low in all cases (a C-index equal to 0.5 is expected by chance and the Somer's D was well below 1), the discrimination ability of all prognostic scores was moderate to low.

Patients with alcoholic, viral cirrhosis and MELD score ≥ 15

In the subgroup analysis of patients with alcoholic and viral cirrhosis, the results were the same as in the total population, with CP-creatinine and CP score performing better than the other prognostic scores (bootstrap C-index (95% CI): 0.630 (0.579-0.681) and 0.626 (0.569-0.684), respectively for alcoholic cirrhosis, and 0.691 (0.652-0.730) and 0.688 (0.643-0.732), respectively for viral cirrhosis) (Suppl. Table 2). In patients with MELD score ≥15, CP-creatinine score performed overall better than the other scores with a bootstrap C-index of 0.632 (95%CI: 0.576-0.688) (Suppl. Table 2). The comparison of the C-indexes between the scores for the subgroup analyses is shown in Suppl. Table 3. A separate analysis was performed for patients with HBV cirrhosis to explore if there were any differences in the prognostic accuracy of the scores between patients included in the study before and after the year 2000 in order to assess their performance before and after the era of nucleot(s)ide analogues. No differences were observed (data not shown). However, when data was divided according to anti-HBV therapy, all scores performed better in patient treated for HBV compared to untreated patients, but the order of the relative performance was similar to that of the main analysis (Suppl. Table 4).

Predictive factors for overall survival

Univariable analysis

Univariate analysis using Cox proportional hazards model showed that advanced age, male sex, ischemic heart disease at baseline, INR, total bilirubin, albumin, creatinine, sodium, encephalopathy and ascites at baseline were significantly associated with survival (Suppl. Table 5). Moreover, all the prognostic models (CP score, CP-creatinine, MELD, MELD-Na and UKELD) were significantly associated with overall mortality (Table 2).

Multivariable analysis

Total Population

Multivariate Cox regression analysis for overall survival including all significant baseline characteristics together with each prognostic score is shown in Suppl. Table 6. Inclusion in the Cox model of age together with CP-creatinine score improved the discrimination ability of the model substantially (bootstrap C-index (95%CI): 0.700 (0.661-0.740)). Similarly, inclusion of age, albumin and ascites together with MELD-Na (or UKELD) score increased the C-index and Somer's D of MELD-Na (or UKELD) score (bootstrap C-index (95%CI): 0.699 (0.663-0.735) and 0.698 (0.662-0.733), respectively). Sensitivity analysis after inclusion of the centers in all the aforementioned multivariable analyses yielded similar results (data not shown).

In all cases, differences between naïve and bootstrap estimates were small, indicating that optimism was negligible in all cases. There were no differences between the C-indexes derived by the multivariable models for all prognostic scores of interest (Suppl. Table 7).

Patients with viral, alcoholic cirrhosis and MELD score ≥ 15

When the same analysis was performed in patients with viral cirrhosis, the results were similar to the abovementioned with age being significant for CP-creatinine, and albumin, age and ascites for MELD-Na (or UKELD). In the subgroup analysis of patients with alcoholic cirrhosis and in those with MELD ≥15, age was the only factor found statistically significant for all prognostic scores (data not shown).

Primary Biliary Cirrhosis (PBC)

Comparison of the predictive models (CP, CP-creatinine, MELD, MELD-Na, UKELD, PBC-RFH and Mayo Risk score) for 3-, 6-, 12-month and overall survival

UKELD score had excellent predictive accuracy (C-index > 0.8) (bootstrap C-index 0.828, 95% CI 0.733-0.923, bootstrap Sommer's D 0.344 (0.285-0.401)), followed by Mayo Risk score (bootstrap C-index (95%CI): 0.794 (0.683-0.906)) and PBC RFH score (bootstrap C-index (95%CI): 0.784 (0.682-0.887)). CP, CP-creatinine and MELD had the lowest performance (Table 4). The comparison of the C-indexes between scores showed that UKELD, PBC RFH and Mayo risk scores had significantly higher diagnostic accuracy compared to CP score. The dynamic/incident ROC curves of CP, CP-creatinine, MELD, MELD-Na, UKELD, Mayo Risk and PBC RFH scores for 3-month, 6-month and 12-month survival in patients with PBC are shown in Figure 2 (a-c). UKELD had the highest discriminative ability for 3-month and 6-month survival, but in 12 months PBC RFH and Mayo Risk scores performed better than all other scores.

Predictive factors for overall survival

All prognostic scores were associated with overall mortality in the univariable analysis. Results from multivariable Cox regression analysis including all significant baseline characteristics together with each prognostic score are shown in Suppl. Table 8. Age was the only factor that was statistically significant for all prognostic scores, apart from Mayo Risk and PBC RFH scores which already incorporate age. The addition of age in the Cox multivariable model increased the C-index and the Somer's D of each predictive score (Suppl. Table 8). Regarding PBC RFH and Mayo Risk scores, creatinine was independently associated with overall survival. Male sex was also independently associated with worse survival when tested together with PBC RFH score. The comparison between the C-indexes of each multivariable model showed no statistical significance, apart from MELD-Na/Age model that was found significantly better than CP-creatinine/age model (bootstrap C-index (95%CI): 0.96 (0.896-1.02) and 0.897 (0.794-1.001), respectively) (Suppl. Table 8).

Calibration

Calibration between observed and predicted survival probabilities was good in the majority of the models, both at 3 and 12 months. The CP-creatinine score demonstrated the best calibration (Suppl. Figures 1a-l) when tested in the total population.

DISCUSSION

The present study clearly demonstrates the superiority of CP and CP-creatinine scores to predict both short and long-term survival in a large population of outpatients with cirrhosis compared to MELD, MELD-Na or UKELD scores. The main advantage of this study is that it takes into account other chronic diseases such as ischemic heart disease or diabetes mellitus that might independently affect mortality. The incorporation of age increased the predictive accuracy of the models. Advanced age is associated with a longer duration of chronic liver disease, reduced cellular regeneration and immune surveillance and there is data supporting its prognostic significance [1, 27].

The advantages of MELD score lie in the statistically weighted "markers" and the inclusion of creatinine [5], while CP score is a simple score without need for complex computations [5]. However, while CP score weighs all variables equally, MELD weighs them in a non-linear fashion. The performance of MELD and CP score has been compared with conflicting results [28-30]. The majority of the studies that included patients undergoing TIPS placement showed an equal performance for 3-month and 12-month mortality apart from a study by Salerno et al [28] in which MELD performed significantly better than CP score for 3-month mortality [31]. In the studies that included patients without TIPS placement, the predictive accuracies of the MELD and CP scores were equal, with higher performance at 3 and 6 months and progressively lower at longer time-periods [27, 31]. In the present study, CP and CP-creatinine had a significantly higher predictive accuracy than MELD or MELD-Na scores for overall mortality. The discrepancies between our study and the study by Said et al [27] can be related to the lower proportion of HCV and higher proportion of PBC patients, to the inclusion of outpatients, to the longer follow-up period and the

lower incidence of patients undergoing liver transplantation in our study. However, our findings are consistent to those reported by D'Amico et al [1].

Since no renal parameter was taken into account in the original CP score, a modified CP score including serum creatinine as a dichotomous variable was developed [32]. In the study by Angermayer et al [32], the c-statistics for 3-month survival for CP score, MELD and CP score-creatinine, in patients undergoing TIPS were similar. In contrary, Giannini et al [33] found that in the prediction of 3-month mortality, the c-statistic of MELD (0.947) was marginally better compared to CP score-creatinine, but significantly higher compared to CP score (0.757). Papatheodoridis et al [9] evaluated the predictive value of CP score, MELD and creatinine-modified CP score in 102 patients with decompensated cirrhosis. The AUROCs for 3, 6, 12 and 24-month mortality did not differ significantly among the four scores. In a recent publication, Kaplan et al adjusted CP score outpoints to improve predictive capacity and evaluated a modified CP-creatinine score [34]. They concluded that the modified CP score and creatinine-modified CP score models showed superiority over the original CP score and MELD in predicting 1-, 2-, 3-, 4- and 5-year transplant-free survival [34]. In our study, which included both compensated and decompensated cirrhotics, the addition of creatinine in CP score did not significantly improve its accuracy. When we divided our patients in two groups (with and without decompansation) and performed subgroup analysis, it had been shown that all scores performed better in patients without decompensation compared to those with compensated cirrhosis, although the CP-creatinine score performed better. These discrepancies might reflect the differences in the study populations, the cut-offs used for creatinine, the length of follow-up periods and the variations in creatinine measurements among different laboratories. It seems that the superiority of

MELD score in patients awaiting for liver transplantation [15] relies on the fact that this cohort has renal impairment [35]. Apparently, creatinine has little impact in the non-transplant setting and in stable cirrhosis.

It has been suggested that the inclusion of other factors to the original MELD may offer additional prognostic value [5, 15, 32, 36, 37]. Although mortality rates in patients with cirrhosis who develop complications are higher [38-40], MELD score fails to predict this impact on survival [1]. It has been suggested [8] that the addition of serum sodium to MELD can improve its accuracy [41, 42]. In the present study no differences were observed between the prognostic performances of MELD and MELD-Na or UKELD similarly to the study by Boursier et al [29]. Londono et al [43] found that MELD score and serum sodium concentration were the only independent predictors of 3-month and 12-month survival, but the addition of serum sodium did not significantly improve its accuracy. We did not observe independent associations between serum sodium and overall mortality.

In the present study, the performance of models predicting long-term survival was poor compared to other studies [29, 36]. However, the prognostic performance of both MELD and CP score is lessened over time [28].

Cirrhosis is a dynamic entity [44] and repeated assessment of scoring systems is needed [27]. In patients with MELD≥15, a superiority of CP score-creatinine was observed. This MELD cut-off was used considering that a major change in survival probability occurs at this threshold [45]. Most importantly, the incorporation of creatinine as a trichotomous variable in CP score seems to have a higher performance for long-term survival compared to its inclusion in MELD score as a continuous variable in high risk patients.

The data on the impact of etiology of liver disease is scarce. The original MELD [3] incorporated etiology, but it was removed later when Kamath et al [37] showed that etiology had a little impact on its accuracy [15]. In our study, in patients with viral and alcoholic etiology, CP-creatinine and CP scores had the highest predictive accuracy as in the total population. The performance of MELD, MELD-Na and UKELD was better in viral, compared to alcoholic cirrhosis. This suggests that in patients with alcoholic cirrhosis other factors such as active alcohol drinking contribute to long-term mortality.

Another important observation is that, in HBV cirrhosis, all scores performed equally irrespective of the year of enrollment, i.e. before or after the year 2000, when anti-HBV treatment with lamivudine was initiated. This implies that the models have equal diagnostic accuracy independent to the initiation of antiviral treatment. Moreover, when we divided the patients in two groups according to the receipt or not of anti-HBV therapy, all scores performed better in patient treated for HBV compared to untreated patients, however the order of the relative performance was similar to that of the main analysis.

In patients with PBC, UKELD, Mayo Risk and RFH-PBC scores had excellent and equal accuracy to predict long-term mortality and they performed significantly better compared to CP, MELD and their modified versions. In our knowledge, this is the first study to confirm the excellent performance of UKELD in PBC patients during a long follow-up period. In ROC curve analysis, the UKELD performance was better for the prediction of 3- and 6-month survival, while the Mayo Risk and RFH-PBC scored better for the 12-month mortality risk. This is not surprising considering that UKELD was developed specifically for patient awaiting liver transplantation in a

short-term follow-up, whereas the other two scores were built to predict 24-month mortality.

The present study is subject to certain limitations. Firstly, it has a retrospective/prospective design. However, all variables were collected prospectively by experienced physicians in well-organized medical databases. There were few missing values and our results are consistent to those observed by others [29]. Secondly, we included a large number of patients with PBC who have a survival benefit compared to alcoholic or viral-related cirrhotics [46]. However, PBC patients have been included in many other similar studies [9, 37, 47]. Thirdly, there was a small number of transplanted patients in our cohort compared to other studies [27]. However, due to the small number of liver transplant centers in Greece and the high demand for organs, a large number of patients will not be eventually transplanted**.** Thus, our results have also important implications in liver centers with no transplantation facilities. The confounding probability related to the various differences among the hospitals (standard of care, different distribution of disease etiology, differences in the laboratory tests used, time period of recruitment) was addressed in a sensitivity analysis. All conclusions regarding the effect of the prognostic scores in terms of hazard ratios, as well as their predictive ability in terms of C-indexes, remained intact indicating no confounding probability. Another limitation of this study is that we included only baseline characteristics as predictive factors of survival in order to reach conclusions useful in the everyday clinical practice and to avoid bias related to missing values.

In conclusion, CP score and CP-creatinine are superior to MELD, MELD-Na and UKELD for the prediction of short-term and long-term mortality. Although CP

score is likely to need statistical validation, it should not be abandoned in clinical practice.

REFERENCES

1 D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *Journal of hepatology* 2006; 44(1): 217-231 [PMID: 16298014 DOI: 10.1016/j.jhep.2005.10.013]

2 Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *The British journal of surgery* 1973; 60(8): 646-649 [PMID: 4541913]

3 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. *Hepatology* 2000; 31(4): 864-871 [PMID: 10733541 DOI: 10.1053/he.2000.5852]

4 Cholongitas E, Marelli L, Shusang V, et al. A systematic review of the performance of the model for end-stage liver disease (MELD) in the setting of liver transplantation. *Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society* 2006; 12(7): 1049-1061 [PMID: 16799946 DOI: 10.1002/lt.20824]

5 Durand F, Valla D. Assessment of prognosis of cirrhosis. *Seminars in liver disease* 2008; 28(1): 110-122 [PMID: 18293281 DOI: 10.1055/s-2008-1040325]

6 Cholongitas E, Senzolo M, Triantos C, Samonakis D, Patch D, Burroughs AK. MELD is not enough--enough of MELD? *Journal of hepatology* 2005; 42(4): 475-477; author reply 478-479 [PMID: 15763330 DOI: 10.1016/j.jhep.2005.02.002]

7 Heuman DM, Abou-Assi SG, Habib A, et al. Persistent ascites and low serum sodium identify patients with cirrhosis and low MELD scores who are at high risk for

early death. *Hepatology* 2004; 40(4): 802-810 [PMID: 15382176 DOI: 10.1002/hep.20405]

8 Biggins SW, Kim WR, Terrault NA, et al. Evidence-based incorporation of serum sodium concentration into MELD. *Gastroenterology* 2006; 130(6): 1652-1660 [PMID: 16697729 DOI: 10.1053/j.gastro.2006.02.010]

9 Papatheodoridis GV, Cholongitas E, Dimitriadou E, Touloumi G, Sevastianos V, Archimandritis AJ. MELD vs Child-Pugh and creatinine-modified Child-Pugh score for predicting survival in patients with decompensated cirrhosis. *World journal of gastroenterology : WJG* 2005; 11(20): 3099-3104 [PMID: 15918197 PMCID: 4305847]

10 European Association for the Study of the Liver. Electronic address eee. EASL Recommendations on Treatment of Hepatitis C 2015. *Journal of hepatology* 2015 [PMID: 25911336 DOI: 10.1016/j.jhep.2015.03.025]

11 European Association for the Study of L. EASL clinical practical guidelines: management of alcoholic liver disease. *Journal of hepatology* 2012; 57(2): 399-420 [PMID: 22633836 DOI: 10.1016/j.jhep.2012.04.004]

12 European Association For The Study Of The L. EASL clinical practice guidelines: Management of chronic hepatitis B virus infection. *Journal of hepatology* 2012; 57(1): 167-185 [PMID: 22436845 DOI: 10.1016/j.jhep.2012.02.010]

13 European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *Journal of hepatology* 2009; 51(2): 237-267 [PMID: 19501929 DOI: 10.1016/j.jhep.2009.04.009]

14 Kim WR, Biggins SW, Kremers WK, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. *The New England journal of*

medicine 2008; 359(10): 1018-1026 [PMID: 18768945 DOI: 10.1056/NEJMoa0801209]

15 Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003; 124(1): 91-96 [PMID: 12512033 DOI: 10.1053/gast.2003.50016]

16 Bonazzi PR, Bacchella T, Freitas AC, et al. Double-dose hepatitis B vaccination in cirrhotic patients on a liver transplant waiting list. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* 2008; 12(4): 306-309 [PMID: 19030730]

17 Allocation Calculators/MELD/PELD Calculators.

18 Murtaugh PA, Dickson ER, et al. Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits. *Hepatology* 1994; 20(1 Pt 1): 126-134 [PMID: 8020881]

19 Hughes MD, Raskino CL, Pocock SJ, Biagini MR, Burroughs AK. Prediction of short-term survival with an application in primary biliary cirrhosis. *Statistics in medicine* 1992; 11(13): 1731-1745 [PMID: 1485056]

20 Ludwig J, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Archiv A, Pathological anatomy and histology* 1978; 379(2): 103-112 [PMID: 150690]

21 RH S. A new asymmetric measure of association for ordinal variables. *American Sociological Review* 1962; 27: 799–811

22 Al-Qabandi W, Buhamrah E, Al-Abdulrazzaq D, Hamadi K, Al Refaee F. Celiac disease in children: is it a problem in Kuwait? *Clinical and experimental gastroenterology* 2015; 8: 43-48 [PMID: 25565879 PMCID: 4284061 DOI: 10.2147/CEG.S73067]

23 Moons KG, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012; 98(9): 683-690 [PMID: 22397945 DOI: 10.1136/heartjnl-2011-301246]

24 Newson R. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *The Stata Journal* 2002; 2(1): 45-64

25 Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? *Seminars in urologic oncology* 2002; 20(2): 96-107 [PMID: 12012295]

26 Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics* 2005; 61(1): 92-105 [PMID: 15737082 DOI: 10.1111/j.0006-341X.2005.030814.x]

27 Said A, Williams J, Holden J, et al. Model for end stage liver disease score predicts mortality across a broad spectrum of liver disease. *Journal of hepatology* 2004; 40(6): 897-903 [PMID: 15158328 DOI: 10.1016/j.jhep.2004.02.010]

28 Salerno F, Merli M, Cazzaniga M, et al. MELD score is better than Child-Pugh score in predicting 3-month survival of patients undergoing transjugular intrahepatic portosystemic shunt. *Journal of hepatology* 2002; 36(4): 494-500 [PMID: 11943420]

29 Boursier J, Cesbron E, Tropet AL, Pilette C. Comparison and improvement of MELD and Child-Pugh score accuracies for the prediction of 6-month mortality in cirrhotic patients. *Journal of clinical gastroenterology* 2009; 43(6): 580-585 [PMID: 19197195 DOI: 10.1097/MCG.0b013e3181889468]

30 Botta F, Giannini E, Romagnoli P, et al. MELD scoring system is useful for predicting prognosis in patients with liver cirrhosis and is correlated with residual

liver function: a European study. *Gut* 2003; 52(1): 134-139 [PMID: 12477775 PMCID: 1773509]

31 Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? *Alimentary pharmacology & therapeutics* 2005; 22(11-12): 1079-1089 [PMID: 16305721 DOI: 10.1111/j.1365-2036.2005.02691.x]

32 Angermayer B, Koenig F, Cejna M, et al. Creatinine-modified Child-Pugh score (CPSC) compared with MELD score to predict survival in patients undergoing TIPS. *Hepatology* 2002; 36: 378A

33 Giannini E, Botta F, Fumagalli A, et al. Can inclusion of serum creatinine values improve the Child-Turcotte-Pugh score and challenge the prognostic yield of the model for end-stage liver disease score in the short-term prognostic assessment of cirrhotic patients? *Liver international : official journal of the International Association for the Study of the Liver* 2004; 24(5): 465-470 [PMID: 15482344 DOI: 10.1111/j.1478-3231.2004.0949.x]

34 Kaplan DE, Dai F, Skanderson M, et al. Recalibrating the Child-Turcotte-Pugh Score to Improve Prediction of Transplant-Free Survival in Patients with Cirrhosis. *Dig Dis Sci* 2016; 61(11): 3309-3320 [PMID: 27405990 PMCID: PMC5067291 DOI: 10.1007/s10620-016-4239-6]

35 Fernandez-Esparrach G, Sanchez-Fueyo A, Gines P, et al. A prognostic model for predicting survival in cirrhosis with ascites. *Journal of hepatology* 2001; 34(1): 46-52 [PMID: 11211907]

36 Huo TI, Lin HC, Huo SC, Lee PC, Wu JC, Lee FY, Hou MC, Lee SD. Comparison of four model for end-stage liver disease-based prognostic systems for

cirrhosis. *Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society* 2008; 14(6): 837-844 [PMID: 18508377 DOI: 10.1002/lt.21439]

37 Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; 33(2): 464-470 [PMID: 11172350 DOI: 10.1053/jhep.2001.22172]

38 Arroyo V, Gines P, Gerbes AL, et al. Definition and diagnostic criteria of refractory ascites and hepatorenal syndrome in cirrhosis. International Ascites Club. *Hepatology* 1996; 23(1): 164-176 [PMID: 8550036 DOI: 10.1002/hep.510230122]

39 Graham DY, Smith JL. The course of patients after variceal hemorrhage. *Gastroenterology* 1981; 80(4): 800-809 [PMID: 6970703]

40 Altman C, Grange JD, Amiot X, et al. Survival after a first episode of spontaneous bacterial peritonitis. Prognosis of potential candidates for orthotopic liver transplantation. *Journal of gastroenterology and hepatology* 1995; 10(1): 47-50 [PMID: 7620107]

41 Borroni G, Maggi A, Sangiovanni A, Cazzaniga M, Salerno F. Clinical relevance of hyponatraemia for the hospital outcome of cirrhotic patients. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver* 2000; 32(7): 605-610 [PMID: 11142560]

42 Gines P, Berl T, Bernardi M, et al. Hyponatremia in cirrhosis: from pathogenesis to treatment. *Hepatology* 1998; 28(3): 851-864 [PMID: 9731583 DOI: 10.1002/hep.510280337]

43 Londono MC, Guevara M, Rimola A, et al. Hyponatremia impairs early posttransplantation outcome in patients with cirrhosis undergoing liver

transplantation. *Gastroenterology* 2006; 130(4): 1135-1143 [PMID: 16618408 DOI: 10.1053/j.gastro.2006.02.017]

44 Liaw YF, Sung JJ, Chow WC, et al. Lamivudine for patients with chronic hepatitis B and advanced liver disease. *The New England journal of medicine* 2004; 351(15): 1521-1531 [PMID: 15470215 DOI: 10.1056/NEJMoa033364]

45 Londono MC, Cardenas A, Guevara M, et al. MELD score and serum sodium in the prediction of survival of patients with cirrhosis awaiting liver transplantation. *Gut* 2007; 56(9): 1283-1290 [PMID: 17452425 PMCID: 1954951 DOI: 10.1136/gut.2006.102764]

46 Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology* 2009; 50(1): 291-308 [PMID: 19554543 DOI: 10.1002/hep.22906]

47 Giannini EG, Risso D, Caglieris S, Testa R. Longitudinal modifications of the MELD score have prognostic meaning in patients with liver cirrhosis. *Journal of clinical gastroenterology* 2005; 39(10): 912-914 [PMID: 16208118]

Table 1. Baseline characteristics of study population.

Sex (M/F), n (%)	620/427 (59.2/40.8)
Age (years), median (IQR)	58 (48, 66)
Cause of liver disease, n (%)	
HBV±HDV	141 (13.5)
HCV	128 (12.2)
Alcohol	347 (33.1)
PBC	151 (14.4)
PSC	2 (0.2)
AIH	56 (5.3)
HBV±HDV/HCV	10 (1)
HBV±HDV +Alcohol	44 (4.2)
HCV+Alcohol	54 (5.2)
HBV±HDV/HCV + Alcohol	4 (0.4)
Other	110 (10.5)

Diabetes mellitus, n (%)	176 (16.8)
Ischemic heart disease, n (%)	57 (5.4)
Chronic respiratory diseases, n (%)	46 (4.4)
Hypertension, n (%)	152 (14.5)
Smoking, n (%)	283 (27.0)
INR, median (IQR)	1.3 (1.1, 1.5)
Total bilirubin (mg/dl), median (IQR)	1.6 (0.9, 3.1)
Albumin (g/dl), median (IQR)	3.5 (3.0, 3.9)
Urea (mg/dl), median (IQR)	30.5 (20.0, 41.0)
Creatinine (mg/dl), median (IQR)	0.9 (0.7, 1.0)
Sodium (mmol/L), median (IQR)	139.0 (136.0, 141.0)
Encephalopathy, n (%)	281 (26.8)
Ascites, n (%)	318 (30.4)
CP score, median (IQR)	7.0 (5.0, 9.0)
CP A/B/C, n (%)	493/357/147 (47.1/34.1/14)
CP score -creatinine, median (IQR)	7.0 (5.0, 9.0)

MELD, median (IQR)	12.0 (9.0, 16.0)
MELD-Na, median (IQR)	14.0 (10.0, 20.0)
UKELD score, median (IQR)	52.0 (49.0, 57.0)
PBC Mayo risk score, median (IQR)	7.0 (5.8, 8.4)
PBC Royal Free Hospital Score	-1.0 (-1.9, -0.1)

M/F: male/female, PBC: primary billiary cirrhosis, PSC: primary sclerosing cholangitis, AIH: autoimmune hepatitis, INR: International normalized ratio, CP score: Child-Pugh score, CP: Child-Pugh class, MELD: Model for End-stage Liver Disease, UKELD: United Kingdom End-stage Liver Disease, IQR: interquartile range.

Table 2. Results based on the Cox proportional hazards models including the scores under investigation as the only predictor of survival in the total sample size. Hazard ratio estimates are based on the whole dataset. Discrimination measures, i.e C-index and Sommer's D were estimated a) using the whole sample (i.e naïve) and b) as the average of 100 bootstrap samples.

	Median (IQR)	HR (95% CI) p value	C-index		Somer's D	
			Naïve (95% CI)	Bootstrap (95% CI)	Naïve (95% CI)	Bootstrap (95% CI)
Child-Pugh	7.0 (5.0, 9.0)	1.26 (1.20-1.32) <0.001	0.667 (0.637-0.697)	0.664 (0.633,0.694)	0.333 (0.270-0.397)	0.329 (0.266-0.392)
Child-Pugh + Creatinine	7.0 (5.0, 9.0)	1.24 (1.19-1.29) <0.001	0.673 (0.642- 0.703)	0.672 (0.642-0.703)	0.345 (0.288-0.401)	0.344 (0.879-0.401)
MELD	12.0 (9.0, 16.0)	1.08 (1.06, 1.10) <0.001	0.643 (0.613-0.673)	0.643 (0.613,0.674)	0.286 (0.220-0.351)	0.285 (0.220-0.351)
MELD + Na	14.0 (10.0, 20.0)	1.07 (1.05-1.09) <0.001	0.623 (0.588-0.658)	0.625 (0.591-0.660)	0.246 (0.178-0.313)	0.246 (0.178-0.314)
UKELD	52.0 (49.0, 57.0)	1.06 (1.04-1.08) <0.001	0.605 (0.568-0.641)	0.605 (0.568-0.642)	0.209 (0.139-0.279)	0.208 (0.138-0.279)

Table 3. Comparison (p-values) between the c- indexes of the different prognostic scores regarding overall survival in the total population.

	Child-Pugh	**Child-Pugh+ Creatinine**	**MELD**	**MELD + Na**	**UKELD**
Child-Pugh	-				
Child-Pugh+ Creatinine	0.203	-			
MELD	0.033	0.001	-		
MELD + Na	0.026	0.001	0.764	-	
UKELD	0.001	<0.001	0.090	0.054	-

MELD: Model for end-stage liver disease, UKELD: United Kingdom End-stage Liver Disease.

Table 4. Results based on the Cox proportional hazards models including the scores under investigation as the only predictor of survival in patients with primary biliary cirrhosis (PBC). Hazard ratio estimates are based on the whole dataset. Discrimination measures, i.e C-index and Sommer's D were estimated a) using the whole sample (i.e naïve) and b) as the average of 100 bootstrap samples.

PBCN=151		C-index		Somer's D	
Score	**HR (95% CI) p**	**naive (95% CI)**	**bootstrap (95% CI)**	**naive (95% CI)**	**bootstrap (95% CI)**
Child-Pugh score	1.56 (1.18-2.06) 0.004	0.660 (0.532-0.798)	0.658 (0.525-0.791)	0.330 (0.078-0.581)	0.331 (0.079-0.582)
CP-creatinine score	1.58 (1.24-2.00) <0.001	0.670 (0.560-0.789)	0.687 (0.573-0.802)	0.349 (0.071-0.628)	0.349 (0.070-0.627)
MELD score	1.15 (1.06-1.24) 0.003	0.690 (0.536-0.835)	0.683 (0.533-0.833)	0.371 (0.081-0.661)	0.371 (0.080-0.661)
MELD-Na score	1.23 (1.11-1.38) <0.001	0.760 (0.571-0.959)	0.753 (0.559-0.947)	0.530 (0.158-0.901)	0.548 (0.176-0.920)
UKELD model	1.19 (1.09-1.29) <0.001)	0.830 (0.732-0.922)	0.828 (0.733-0.923)	0.654 (0.444-0.864)	0.655 (0.445-0.866)
RFH	2.45 (1.71-3.52 <0.001)	0.780 (0.681-0.887)	0.784 (0.682-0.887)	0.568 (0.341-0.795)	0.582 (0.355-0.809)
Mayo risk score	1.87 (1.43-2.43) <0.001	0.790 (0.678-0.902)	0.794 (0.683-0.906)	0.580 (0.359-0.801)	0.573 (0.353-0.794)

Table 5. Comparison (p-values) between the c- indexes of the different prognostic scores regarding overall survival in patients with Primary Biliary Cirrhosis.

	Child-Pugh	**Child-Pugh+ Creatinine**	**MELD**	**MELD+Na**	**UKELD**	**RFH score**	**Mayo risk score**
Child-Pugh	-						
Child-Pugh + Creatinine	0.361	-					
MELD	0.102	0.063	-				
MELD+Na	0.120	0.105	0.473	-			
UKELD	0.049	0.251	0.802	0.687	-		
RFH score	0.019	0.174	0.494	0.436	0.444	-	
Mayo risk score	0.030	0.173	0.527	0.463	0.477	0.612	-

Figure 1. Dynamic/incident ROC curves of Child-Pugh, Child-Pugh creatinine, MELD, MELD-Na and UKELD score for a) 3-month, b) 6-month and c) 12-month survival in the total population.

Figure 2. Dynamic/incident ROC curves of Child-Pugh, Child-Pugh creatinine, MELD, MELD-Na and UKELD score for a) 3-month, b) 6-month and c) 12-month survival in patients with Primary Biliary Cirrhosis.

Figure 1.

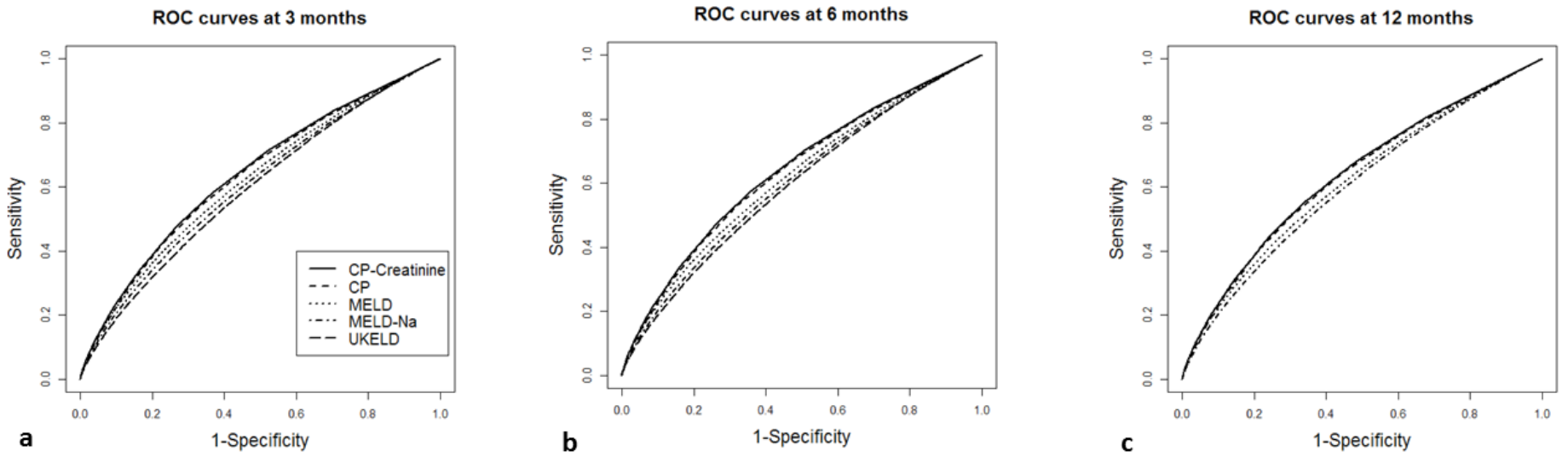
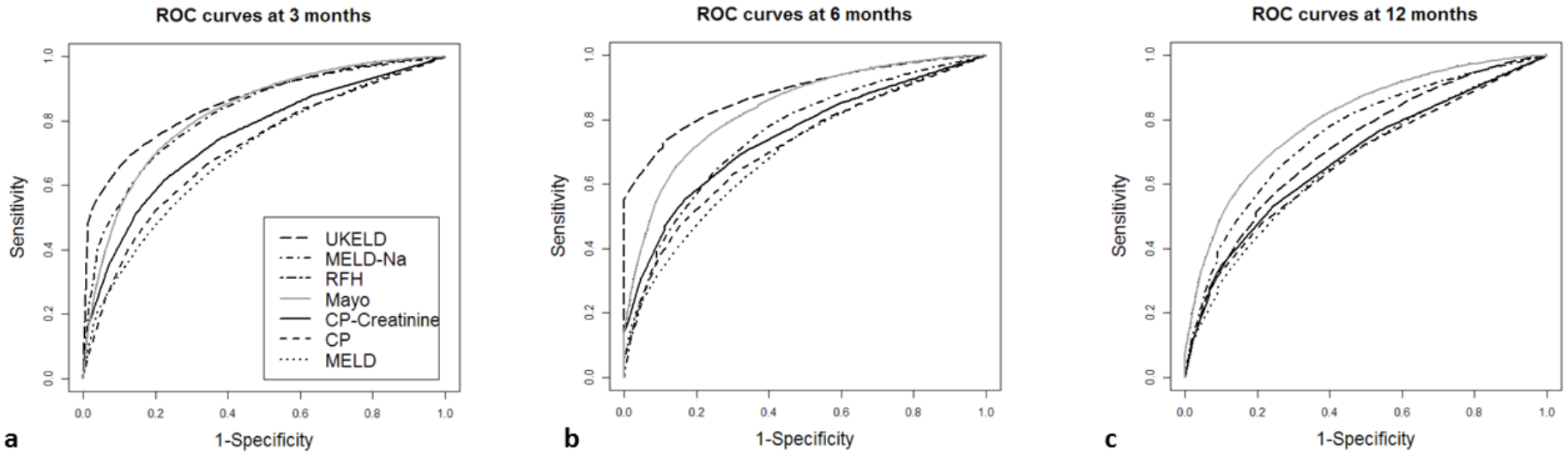

ROC curves at 3 months
Sensitivity
1-Specificity
CP-Creatinine
CP
MELD
MELD-Na
UKELD
a
ROC curves at 6 months
Sensitivity
1-Specificity
b
ROC curves at 12 months
Sensitivity
1-Specificity
c

Figure 2.


ROC curves at 3 months
Sensitivity
1-Specificity
UKELD
MELD-Na
RFH
Mayo
CP-Creatinine
CP
MELD
a
ROC curves at 6 months
Sensitivity
1-Specificity
b
ROC curves at 12 months
Sensitivity
1-Specificity
c