External validation and comparison of six prognostic models in a prospective cohort of HBV-ACLF in China


* Department of Epidemiology and Medical Statistics, School of Public Health, Nantong University, Jiangsu, China,
** Center for Liver Diseases, Nantong Third People's Hospital, Nantong University, Jiangsu, China.
*** Department of Pathology, Medical School, Nantong University, Jiangsu, China.
**** Department of Internal Medicine, Singapore General Hospital, Singapore.
♦ Yi Shen and Yan-Mei Liu contributed equally to this work

ABSTRACT

Background. Acute-on-chronic liver failure has high mortality. Currently, robust models for predicting the outcome of hepatitis B virus (HBV)-associated ACLF are lacking. Aim. To assess and compare the performance of six prevalent models for short- and long-term prognosis in patients with HBV-ACLF. Material and methods. The model for end-stage liver disease (MELD), MELD sodium (MELD-Na), MELD to sodium ratio (MESO), integrated MELD, Child-Turcotte-Pugh (CTP), and modified CTP (mCTP) were validated in a prospective cohort of 232 HBV-ACLF patients. The six models were evaluated by determining discrimination, calibration and overall performance at 3 months and 5 years. Results. According to the Hosmer-Lemeshow tests and calibration plots, all models could adequately describe the data except CTP at 3 months. Discrimination analysis showed that the iMELD score had the highest AUC of 0.76 with sensitivity of 62.6% and specificity of 80.2% for an optimal cut-off value of 52 at 3 months. It also had the highest AUC of 0.80 with sensitivity of 89.9% and specificity of 48.2% for an optimal cut-off value of 43 at 5 years. The overall performance of iMELD, assessed with Nagelkerke’s R² and the Brier score, was also the best among the six models. Conclusion. Integrated MELD may be the best model to predict short- and long-term prognosis in patients with HBV-ACLF.

Key words. Hepatitis B virus. Acute-on-chronic liver failure. Prognostic scores.

INTRODUCTION

Chronic hepatitis B virus (HBV) infection is a worldwide public health issue. Acute-on-chronic liver failure (ACLF) in patients with chronic hepatitis B (CHB) is usually caused by acute reactivation of CHB.1 If liver transplantation (LT) could not be arranged in time, the prognosis of patients with ACLF may be poor with high incidence of short-term (28-day) mortality of 30-40%.2

Before the Chinese Society of Hepatology (CSH) proposed the diagnostic and treatment guideline for ACLF formally in 2006,3 this serious medical ailment had been called severe chronic hepatitis in China.4 The definition of ACLF by CSH is “an acute decompensation (AD) in liver function in patients with previously diagnosed or undiagnosed chronic liver disease, manifesting as jaundice (serum total bilirubin, TBIL level greater than 10 mg/mL) and coagulopathy, complicated within 4 weeks by ascites and/or hepatic encephalopathy (HE).” This definition is close to that proposed by Asian Pacific Association for the Study of the Liver (APASL), except the cutoff value of jaundice (TBIL level greater than 5 mg/mL).5 However, as in many other aspects of life and medicine, there is a sharp East-West divide with respect to the definition of ACLF. The AASLD/EASL consensus defines it as “a syndrome that defines a subgroup of cirrhotic patients who develop organ failure following hospital admission with or without an identifiable precipitating event and have increased mortality rates.”6,7

Due to the extremely limited organ sources, liver transplantation could be regarded as a gift from God. Therefore, an objective numerical formula for predicting the prognosis of these patients may be essential for donor organ allocation. In 2000, Malinchoc first proposed the
model for end-stage liver disease (MELD) to predict 3-month survival rates in patients with chronic liver disease who also underwent transjugular intrahepatic portosystemic shunt. His insight has been substantiated by quite a few following investigations. MELD score system was adopted as the standard by which to determine the priority to allocate cadaveric livers to transplant candidates in February 2002 in USA and in many other countries thereafter. Meanwhile, several studies showed that certain subsets of patients with advanced liver disease may have high mortality in spite of low MELD scores. Thus, a number of attempts, including replacement of some MELD components, re-assignment of weights for existing components, incorporation of other variables into the model, were explored to improve the prognostic accuracy of this model, leading to several MELD-based models such as MELD sodium (MELD-Na), MELD to sodium ratio (MESO), integrated MELD (iMELD), updated MELD (uMELD), United Kingdom MELD (UKMELD) and donor MELD (D-MELD). Hypoalbuminaemia was associated with hepatic encephalopathy (HE), hepatorenal syndrome (HRS) and increased mortality in patients with cirrhosis. Older age was also an independent risk factor for mortality in these patients. Thus it is not surprising that MELD-Na and iMELD (incorporating both sodium and age) have been suggested to be the two best prognostic models to predict drop-out rates among patients awaiting LT. On the other hand, the Child-Turcotte-Pugh (CTP) system, which is used to stratify the severity of advanced cirrhosis, has its own advantage because it includes ascites and hepatic encephalopathy. In order to overcome its ceiling effect, additional points were advocated for exceptionally high serum total bilirubin (TBIL), long prothrombin time (PT) and low albumin levels in modified CTP model. More accurate, flexible, and easily accessible models are needed to simplify the practical task of prediction.

To date, only a few studies have evaluated the prognostic power of standard MELD compared to other MELD-based models in patients with advanced cirrhosis or ACLF. These studies have shown some limitations. First, most of the study subjects were transplantation candidates. The outcomes may be influenced by the surgery. Second, only short-term (6 months at most) survival of the patients was observed. In fact, the majority (63%) of patients with advanced liver disease waited almost one year for the donor liver. Besides, with the development of other treatment such as artificial liver support systems (ALSS), some subsets of patients with ACLF survived for 3 years or even longer. Another limitation lies on the etiology of ACLF. Alcoholic, autoimmune, hepatitis C virus (HCV) or HBV infection might contribute to the subtly different survival outcomes. It is plausible that the performance of these models may be influenced by the mixed spectrum of disease etiology.

Currently, robust models for predicting the outcome of hepatitis B virus (HBV)-associated acute-on-chronic liver failure (ACLF) are still lacking. In our previous study, we found that age, serum sodium and MELD scores were independent associated with 5-year mortality of HBV-ACLF patients who had no liver transplantation. These findings led us to consider, as an independent external validation, assessing and comparing of the short- and long-term prognostic performance of the standard MELD with five other scores, namely MELD-Na, MESO, iMELD, CTP and mCTP, in these patients.

MATERIAL AND METHODS

Study population

From January 2003 to December 2007, all patients admitted to our hospital with objective diagnosis of HBV-associated severe chronic hepatitis or HBV-associated ACLF, before or after the release of the Chinese ACLF guideline in 2006, were screened. We have adopted international classification of diseases the 10th edition (ICD-10) system since 2001 and the code for ACLF is K72.005. The following criteria were used to select eligible patients:

- Between 18 and 70 years of age.
- Presumptive diagnosis of hepatitis B surface antigen (HBsAg) carrier, CHB or HBV-related cirrhosis.
- Rapidly progressive hyperbilirubinemia with TBIL > 10 mg/dL.
- Coagulopathy with international normalized ratio (INR) > 1.5 or plasma prothrombin activity < 40%.
- Complicated within 4 weeks from symptom onset by ascites and/or hepatic encephalopathy (HE).

The exclusion criteria included: acute HBV infection, super-infection with other viruses, other etiology of chronic liver failure, coexistent hepatocellular carcinoma (HCC), severe gastrointestinal bleeding, pregnancy, or liver transplant recipients.

All study subjects were prospectively evaluated, with followed-up for at least five years or till death. Their medical profiles were retrospectively analyzed in this study. Our study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was duly approved by the institutional review board of Nantong Third People’s Hospital, Nantong University. Written informed consent for inclusion in the study was obtained from each patient (or his or her closest relative in some instances).
Calculation of scores

Based on laboratory results obtained at enrollment, the following six scores were calculated.

MELD score (on a scale of 6 to 40, with higher values indicating more severe disease) was calculated according to the standard formula: 12

\[
\text{MELD} = 11.2 \times \ln(\text{INR}) + 9.6 \times \ln(\text{creatinine (mg/dL)}) + 3.8 \times \ln(\text{bilirubin (mg/dL)}) + 6.4 \quad \text{(constant for liver disease etiology)}
\]

To avoid negative scores, laboratory values of bilirubin, INR or creatinine less than 1 were rounded off to 1. Creatinine > 4 mg/dL or with renal replacement therapy was capped at 4 mg/dL. Besides, the factor for etiology of liver disease was not used.

The MELD-Na equation was based on the formula: 12

\[
\text{MELD-Na} = \text{MELD} - \text{Na} - [0.025 \times \text{MELD} \times (140 - \text{Na})] + 140.
\]

The serum sodium concentration was bound between 125 and 140 mEq/L.

The MESO index was defined as: 13

\[
\text{MESO} = \left(\frac{\text{MELD}}{\text{Na}}\right) \times 10
\]

The iMELD equation on the basis of MELD, age (years), and Na (mEq/L) was calculated as follows: 14

\[
\text{iMELD score} = \text{MELD} + (\text{age} \times 0.3) - (\text{Na} \times 0.7) + 100
\]

The CTP score was calculated based on serum bilirubin and albumin levels, PT, and the presence and severity of ascites and HE. 21

The calculation of mCTP score was performed by combining traditional CTP score with additional points (3 at most) in patients whose serum bilirubin was > 8 mg/dL, PT prolongation > 11 sec, or albumin < 2.3 g/dL. 22

Statistical analysis

Continuous variables were expressed as mean ± standard deviation (SD) or median (range). The Cox proportional-hazards regression model was applied to calculate the hazard ratio (HR) and its 95% confidence interval (CI) for each score after adjusting for age and sex.

Calibration was assessed with Hosmer-Lemeshow (H-L) tests and calibration plots. For the H-L tests, the smaller the \( \chi^2 \) and the bigger the associated P value, the better the goodness of fit, that is, calibration. Acceptable calibration is represented by an H-L P value ≥ 0.05. Calibration plots could be described by an intercept a (“calibration-in-the-large”), which should be 0, and a calibration slope b, which should be 1.

Discrimination was evaluated by using concordance statistic (c-statistic) equivalent to the area under the receiver operating characteristic curve (AUC). AUC values range from 0.5 to 1, with values ≥ 0.7, ≥ 0.8, and ≥ 0.9 considered as satisfactory, good and excellent, respectively. 27 The AUCs were assessed with pairwise comparison adjusting for Bonferroni correction. Optimal cut-off values were derived from the Youden index \( J = \) (sensitivity + specificity - 1). The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated according to the optimal cutoff. The agreement between the observed and expected (O/E) mortality ratio was assessed with 95% CI calculated assuming a Poisson distribution.

Nagelkerke’s (Cragg-Uhler’s) R² and the Brier score were calculated to evaluate overall model performance. Nagelkerke’s R² can be readily applied to survival outcomes and has the advantage of being scaled from 1 to 100%. 28 The Brier score can range from 0 for a perfect model to 0.25 for a non-informative model. 29

All analyses were performed using the Stata version 12.0 (StataCorp, USA). Statistical significance was defined as P < 0.05.

RESULTS

Patient characteristics

A total of 283 patients presenting with HBV-ACLF were screened for inclusion in the study cohort. Of these, 51 patients were excluded from this study: 8 for super-infection with hepatitis E virus, 5 for age over 70 years, 2 for co-existence of HCC, 7 for liver transplantation, 29 for incomplete data. The final cohort compromised 232 patients with complete data. 26

The study cohort was 77% male with a median (range) age of 45 (21-69) years (Table 1). 112 (48.3%) patients had preexisting cirrhosis. The mean HBV viral load was 4.1 ± 2.5 lg copies/mL and 142 patients (61.2%) were positive for hepatitis B e antigen (HBeAg). In these patients, mean (SD) value of serum bilirubin was 22.2 (9.2) mg/dL, serum creatinine 0.93 (0.74) mg/dL, INR 4.2 (2.2), albumin 32.1 (5.0) g/L, and sodium 132.0 (8.9) mEq/L. The most common complications were ascites (194 patients; 83.6%), HE (64 patients; 27.67%), and HRS (37 patients; 16%).

Among these, 104 (44.8%) patients received ALSS treatment, and 93 (40.1%) patients had continuous antiviral treatment with nucleos(t)ide analogs (NAs) which initiated at any time and sustained for at least one month to the end of follow-up or till death.

During the first 3 months of follow-up, 111 patients died. At the end of follow-up in Dec 2012, 83 (35.8%) patients completed 5-year visits, including 15 (6.5%) patients...
with fulfilled 10-year visits. The Cox regression analysis showed that MELD and other five scores were significantly associated with mortality at 3 months and 5 years in patients with HBV-ACLF, adjusted for age and sex (Table 2).

**Calibration analysis**

The H-L goodness-of-fit tests revealed adequate calibration (P > 0.05) for all models, except for CTP at 3 months. Calibration for each score was further examined by plotting the observed against the expected mortality frequency (data not shown). For all the six models, the calibration intercepts were close to 0 and calibration slopes were close to 1.0, indicating good fit (Table 3).

**Discrimination analysis**

The AUCs computed from c-statistics are shown in table 4 and figure 1. AUCs at 3 months showed that all the models showed satisfactory discrimination (AUC ≥ 0.7) except MELD. With the highest AUC (0.76, 95%CI: 0.70-0.82), the iMELD showed the best diagnostic accuracy. In addition, the comparison between MELD and three other MELD-based models showed significant difference (P < 0.05), while the difference between MELD and CTP, mCTP was not significant (P = 0.154 and 0.089 respectively). According to AUCs for 5-year, all the models showed values in the range of clinical usefulness except

### Table 1. Demographic, clinical and laboratory features of the study cohort.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>232</td>
</tr>
<tr>
<td>Male/female</td>
<td>178 (76.7%)/54 (23.3%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>46.1 ± 10.5 (45; 21-69)</td>
</tr>
<tr>
<td>HBeAg positivity</td>
<td>142 (61.2%)</td>
</tr>
<tr>
<td>HBV DNA (lg copies/mL)</td>
<td>4.1 ± 2.5 (5.0; 0-9.9)</td>
</tr>
<tr>
<td>Bilirubin (mg/dL)</td>
<td>22.2 ± 9.2 (21; 7-52)</td>
</tr>
<tr>
<td>Creatinine (mg/dL)</td>
<td>0.93 ± 0.74 (0.8; 0.3-6.6)</td>
</tr>
<tr>
<td>INR</td>
<td>4.2 ± 2.2 (3.6; 1.7-17.7)</td>
</tr>
<tr>
<td>Albumin (g/L)</td>
<td>32.1 ± 5.0 (32; 15-48)</td>
</tr>
<tr>
<td>Serum sodium (mEq/L)</td>
<td>132.0 ± 8.9 (134; 96-148)</td>
</tr>
<tr>
<td>Preexisting cirrhosis</td>
<td>112 (48.3%)</td>
</tr>
<tr>
<td>Ascites</td>
<td>194 (83.6%)</td>
</tr>
<tr>
<td>HE</td>
<td>64 (27.6%)</td>
</tr>
<tr>
<td>HRS</td>
<td>37 (16.0%)</td>
</tr>
<tr>
<td>MELD</td>
<td>29.0 ± 5.4 (28; 17-40)</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>30.7 ± 5.0 (30; 20-40)</td>
</tr>
<tr>
<td>MESO</td>
<td>2.2 ± 0.5 (2.1; 1.3-4.2)</td>
</tr>
<tr>
<td>iMELD</td>
<td>50.4 ± 10.5 (49; 28-89)</td>
</tr>
<tr>
<td>CTP</td>
<td>11.6 ± 1.7 (12; 7-15)</td>
</tr>
<tr>
<td>mCTP</td>
<td>14.4 ± 2.0 (15; 8-18)</td>
</tr>
<tr>
<td>ALSS treatment</td>
<td>104 (44.8%)</td>
</tr>
<tr>
<td>Continuous NA treatment</td>
<td>108 (46.6%)</td>
</tr>
<tr>
<td>Transplant free survival</td>
<td>742 ± 840 (100; 3-1825)</td>
</tr>
</tbody>
</table>


### Table 2. Results of Cox proportional hazard regression analyses to assess the five score systems for 3-month and 5-year mortality.

<table>
<thead>
<tr>
<th>Model</th>
<th>3 months</th>
<th>5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95%CI)</td>
<td>P</td>
</tr>
<tr>
<td>MELD</td>
<td>1.12 (1.08-1.16)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>1.15 (1.10-1.19)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>MESO</td>
<td>4.65 (3.17-6.82)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>iMELD</td>
<td>1.09 (1.07-1.12)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>CTP</td>
<td>1.50 (1.32-1.71)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>mCTP</td>
<td>1.46 (1.30-1.65)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>


### Table 3. Calibration analysis for each score at 3 months and 5 years.

<table>
<thead>
<tr>
<th>Model</th>
<th>3 months</th>
<th>5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H-L test</td>
<td>Calibration plot</td>
</tr>
<tr>
<td></td>
<td>χ²  P</td>
<td>Intercept  Slope</td>
</tr>
<tr>
<td>MELD</td>
<td>10.78 0.21</td>
<td>0.037 0.93</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>3.27 0.92</td>
<td>0.021 0.97</td>
</tr>
<tr>
<td>MESO</td>
<td>4.58 0.80</td>
<td>0.022 0.96</td>
</tr>
<tr>
<td>iMELD</td>
<td>6.38 0.61</td>
<td>0.027 0.96</td>
</tr>
<tr>
<td>CTP</td>
<td>10.68 0.03</td>
<td>-0.008 1.01</td>
</tr>
<tr>
<td>mCTP</td>
<td>8.89 0.06</td>
<td>-0.006 0.99</td>
</tr>
</tbody>
</table>

MELD: Model for End-Stage Liver Disease. MELD-Na: sodium MELD. iMELD: integrated MELD. MESO: MELD to sodium ratio. CTP: Child-Turcotte-Pugh. mCTP: modified CTP.
MELD. The comparison among the AUCs suggested that all the other five models had better prognostic accuracy than MELD score. Similarly, the iMELD score had the highest AUC (0.80, 95%CI: 0.74-0.86).

From the AUCs at 3 months, the optimal cutoffs for the MELD, MELD-Na, MESO, iMELD, CTP and mCTP models were 30, 32, 2.6, 52, 11 and 14 respectively, according to Youden index. Meanwhile, the optimal cutoffs for the MELD, MELD-Na, MESO, iMELD, CTP and mCTP models at 5 years were 26, 26, 1.9, 43, 11 and 14 respectively. Table 5 reports their cutoff, sensitivity, specificity, positive (PPV), negative predictive values (NPV) and O/E ratios. With the cut-off value of 52, iMELD had sensitivity of 62.2%, specificity of 80.2%, positive predictive value (PPV) of 74.2%, negative predictive value (NPV) of 69.8% and an O/E ratio of 1.19 for 3-month mortality. With the cut-off value of 43, iMELD had sensitivity of 89.9%, specificity of 48.2%, PPV of 75.7%, NPV of 72.7% and an O/E ratio of 0.84 for 5-year mortality.

Actual survival curves showed that all these scores were able to identify the patients with the higher mortality risk (Figure 2). The 3-month mortality was 74.2% (69/93) for patients with iMELD scores \( \geq \) 52 compared with 30.2% (42/139) for patients with iMELD scores < 52. The 5-year mortality was 57.1% (101/177) for patients with iMELD scores \( \geq \) 43 compared with 18.2% (10/55) for patients with iMELD scores < 52 (data not shown).

### Overall performance analysis

Nagelkerke’s R\(^2\) and the Brier score were calculated for each model at 3 months and 5 years (Table 6). The Nagelkerke’s R\(^2\) values for iMELD were 28.6% and 34.5% at 3 months and 5 years respectively. The Brier scores for
Prognostic models for HBV-ACLF. 2016; 15 (2): 236-245

**Table 5.** Diagnostic accuracy of each model to predict 3-month and 5-year mortality rates with the optimal cutoffs.

<table>
<thead>
<tr>
<th>Model</th>
<th>Optimal cutoff</th>
<th>Sensitivity (95%CI)</th>
<th>Specificity (95%CI)</th>
<th>PPV (95%CI)</th>
<th>NPV (95%CI)</th>
<th>O/E (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MELD</td>
<td>30</td>
<td>55% (45.2-64.4%)</td>
<td>70.2% (61.3-78.2%)</td>
<td>62.9% (52.5-72.5%)</td>
<td>63% (54.2-71.1%)</td>
<td>1.14 (0.94-1.38)</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>32</td>
<td>55.9% (46.1-65.3%)</td>
<td>73.6% (64.8-81.2%)</td>
<td>66% (55.5-75.4%)</td>
<td>44.5% (55.9-72.4%)</td>
<td>1.19 (0.97-1.42)</td>
</tr>
<tr>
<td>IMELD</td>
<td>2.6</td>
<td>38.7% (29.6-48.5%)</td>
<td>69.6% (59.6-78.6%)</td>
<td>51.7% (42.3-61%)</td>
<td>71.3% (62.9-79.2%)</td>
<td>1.98 (1.63-2.39)</td>
</tr>
<tr>
<td>CTP</td>
<td>11</td>
<td>92.8% (86.3-96.8%)</td>
<td>55.1% (35.6-53.9%)</td>
<td>62.8% (52.8-68%)</td>
<td>47.2% (52.4-72%)</td>
<td>0.65 (0.54-0.79)</td>
</tr>
<tr>
<td>nCTP</td>
<td>14</td>
<td>91% (84.1-95.6%)</td>
<td>48.8% (39.6-58%)</td>
<td>62% (54-69.4%)</td>
<td>85.5% (75-92.8%)</td>
<td>0.68 (0.56-0.82)</td>
</tr>
<tr>
<td><strong>5 years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MELD</td>
<td>26</td>
<td>79.9% (72.5-86%)</td>
<td>47% (35.9-58.3%)</td>
<td>73% (65.5-79.7%)</td>
<td>56.5% (44-68.4%)</td>
<td>0.91 (0.77-1.07)</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>26</td>
<td>91.9% (86.4-95.8%)</td>
<td>32.5% (22.6-43.7%)</td>
<td>71% (64-77.3%)</td>
<td>69.2% (52.4-83%)</td>
<td>0.77 (0.65-0.91)</td>
</tr>
<tr>
<td>MESO</td>
<td>1.9</td>
<td>84.6% (77.7-90%)</td>
<td>42.2% (31.4-53.5%)</td>
<td>72.4% (65.1-78.9%)</td>
<td>60.3% (46.6-73%)</td>
<td>0.86 (0.72-1.01)</td>
</tr>
<tr>
<td>IMELD</td>
<td>43</td>
<td>89.9% (83.9-94.3%)</td>
<td>48.2% (37.1-59.4%)</td>
<td>75.7% (68.7-81.8%)</td>
<td>72.7% (59-83.9%)</td>
<td>0.84 (0.71-0.99)</td>
</tr>
<tr>
<td>CTP</td>
<td>11</td>
<td>88.6% (82.4-93.2%)</td>
<td>54.2% (42.9-65.2%)</td>
<td>77.6% (70.6-83.7%)</td>
<td>72.6% (59.8-83.1%)</td>
<td>0.88 (0.74-1.03)</td>
</tr>
<tr>
<td>nCTP</td>
<td>14</td>
<td>88.6% (80-91.6%)</td>
<td>59% (47.7-69.7%)</td>
<td>79.1% (72.1-85.1%)</td>
<td>71% (58.8-81.3%)</td>
<td>0.91 (0.77-1.07)</td>
</tr>
</tbody>
</table>

PPV: positive predictive value. NPV: negative predictive value. CI: confidence interval. MELD: model for end-stage liver disease. MELD-Na: sodium MELD. IMELD: integrated MELD. MESO: MELD to sodium ratio.

**DISCUSSION**

In the early years of the 21st century, MELD and MELD-based models have been developed to predict the short-term survival of patients with advanced liver disease and to define medical urgency for transplantation,8,11,13-16 of which the MELD is the most widely known. Serum creatinine, a key component in MELD, is not a good surrogate marker of renal function in patients with cirrhosis, while cystatin C may be a better predictor.30 However, substitution of creatinine by cystatin C or estimated glomerular filtration rate in the MELD formulas did not improve model performance compared with other models.31-33 In this study, we investigated five different models (MELD, MELD-Na, IMELD, MESO, CTP) to determine their predictive accuracy for mortality and compared their performance with other models such as mCTP and the predictive ability of each model was evaluated using receiver operating characteristic (ROC) curves. Our results showed that the predictive accuracy of the models was not significantly different. However, the AUCs of MELD-Na and MESO were 0.196 and 0.171 at 3 months and 5 years respectively. The overall performance was worse for the other five models.
242

Table 6. Overall performance of each model.

<table>
<thead>
<tr>
<th></th>
<th>MELD</th>
<th>MELD-Na</th>
<th>MESO</th>
<th>iMELD</th>
<th>CTP</th>
<th>mCTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nagelkerke’s $R^2$ (%) 3 month</td>
<td>11.2</td>
<td>16.5</td>
<td>18.3</td>
<td>28.6</td>
<td>23.2</td>
<td>24.9</td>
</tr>
<tr>
<td>Brier score</td>
<td>0.229</td>
<td>0.219</td>
<td>0.216</td>
<td>0.196</td>
<td>0.207</td>
<td>0.204</td>
</tr>
<tr>
<td>Nagelkerke’s $R^2$ (%) 5 year</td>
<td>9.1</td>
<td>16.7</td>
<td>17.0</td>
<td>34.5</td>
<td>30.1</td>
<td>30.0</td>
</tr>
<tr>
<td>Brier score</td>
<td>0.215</td>
<td>0.201</td>
<td>0.203</td>
<td>0.171</td>
<td>0.176</td>
<td>0.176</td>
</tr>
</tbody>
</table>

filtration rate (eGFR) does not improve the predictive power of MELD. Whether re-assigning the weights of MELD components or incorporating other variables such as serum sodium and age into this model could improve the prognostic accuracy remains controversial. All the MELD-based models were established on the data from American and European populations. For ACLF, only a few studies have applied these models to short-term prognosis, usually 3-month follow-up. Alcoholism, hepatitis C, and cholestasis are the leading causes of ACLF in Western countries, whereas HBV infection accounts for the majoritiy of ACLF cases in China. The enhanced efficacy of ALSS and NAs as important treatments for HBV-ACLF has led to improved prognosis of these patients. The proportion of patients who survived the liver failure and lived much longer has steadily increased. Therefore, these statistical models need to be evaluated, validated and/or improved in the specific HBV-ACLF population with long-term follow-up.

To the best of our knowledge, we for the first time conducted a long-term prospective cohort to compare the validation and accuracy of the six models as predictors for LT-free mortality of patients with HBV-ACLF. Although our study population was similar to that of some other Chinese centers, in terms of race, age and etiology, our observation period was extremely longer than theirs. The analysis of data from a single center should have ensured uniform enrollment requirements, diagnostic criteria, treatment protocols, and follow-up schedules throughout the study. Finally, the performances of the models for our patients with HBV infection as unique etiology would be more applicable to similar Asian patients.

The prognostic models should be as calibrated as possible to the study population. On the one hand, our calibration analysis showed that all the models had justification to be applied in this population, except CTP for 3-month data. On the other hand, our discrimination analysis indicated that iMELD predicted short- and long-term mortality rates better than the other five models. Among these models, only iMELD had AUC values exceeding 0.75 at both time points (3-month and 5-year), indicating best predictive accuracy. The overall performance analysis supported the results as well. This is not surprising because age and serum sodium were strongly associated with higher mortality in these patients. Actually, the iMELD model incorporating serum sodium and age has been proved better than original MELD in predicting 12-month mortality in patients with cirrhosis on the waiting list for LT. Our study further demonstrated that the iMELD model improves the predictive accuracy of time to death in patients with HBV-ACLF, compared with original MELD model.

The CTP score was originally introduced to predict prognosis for patients with liver disease undergoing surgery for portosystemic shunts. When used as a disease severity index to determine priority in organ allocation, it has two main limitations:

- Limited discriminatory ability.
- Subjective interpretation of some parameters.

Recently, a concern has been raised that the CTP system might be not inferior to the MELD system for outcome prediction in cirrhotic patients. To further improve, rather than to abandon, the CTP system which has been used for nearly half a century, the mCTP system was proposed to attenuate the ceiling effect by extending points up to 18 in cirrhotic patients awaiting LT in an Asian Center. Traditional and modified CTP systems may offer an advantage because MELD is not influenced by ascites and hepatic encephalopathy, given that the severity of the cirrhotic patients presenting these complications might be underestimated by MELD-based models. In our population under study, although the calibration performance of the two models was inferior to that of MELD-based models, they had good predictive accuracy for both short- and long-term prognosis.

Admittedly, our study has several limitations. First, although a new model might be proposed from our sample as a training cohort, a validation cohort for such a new model is lacking. Instead, we only conducted an external validation and comparison of prevalent models here. Second, the latest CLIF (Chronic Liver Failure)-SOFA (Sequential Organ Failure Assessment) model could not be tested for our study patients, because of different definition of ACLF we adopted. CLIF-SOFA has been proved...
to be strong predictors of short-term mortality in patients with ACLF satisfying criteria proposed by EASL-CLIF Consortium. However, our patients of ACLF diagnosed according to CSH guidelines did not take into account specific components of the CLIF-SOFA score such as circulatory and respiratory function. Third, our findings could not be readily applicable to the American or European patients where hepatitis C and alcoholism are predominant causes of their end-stage liver diseases. Finally, disease progression may be subtly variable depending on the therapeutic strategy as well as host factors. Whether the variable treatments interfere with the prognostic accuracy is unknown. Further studies are needed to clarify the potential impact of treatment factors on the outcome of the patients and to establish a more sensitive model specific to HBV-ACLF patients.

CONCLUSION

In conclusion, our findings indicate that integrated MELD score system is a markedly valued model in predicting the short- and long-term outcome for patients with HBV-ACLF. By incorporating two key factors (age and sodium) into traditional MELD system, the iMELD system significantly improves its predictive ability. We propose the iMELD score system might best serve as the prognostic model for HBV-ACLF patients without liver transplantation.

ABBREVIATIONS

- ACLF: acute-on-chronic liver failure.
- ALSS: artificial liver support system.
- AUC: area under the receiver operating characteristic curve.
- CHB: chronic hepatitis B.
- CI: confidence interval.
- CTP: Child-Turcotte-Pugh.
- D-MELD: donor MELD.
- eGFR: estimated glomerular filtration rate.
- HBeAg: hepatitis B e antigen.
- HBsAg: hepatitis B surface antigen.
- HBV: hepatitis B virus.
- HCV: hepatitis C virus.
- HE: hepatic encephalopathy.
- HR: hazard ratio.
- HRS: hepatorenal syndrome.
- iMELD: integrated MELD.
- INR: international normalized ratio.
- LT: liver transplantation.
- mCTP: modified CTP.
- MELD: model for end-stage liver disease.
- MELD-Na: MELD sodium.
- MESO: MELD to sodium ratio.
- NPV: negative predictive values.
- PPV: positive predictive value.
- PT: prothrombin time.
- SD: standard deviation.
- TBIL: total bilirubin.
- UKMELD: United Kingdom MELD.
- uMELD: updated MELD.

DISCLOSURE OF INTEREST

The authors declare that they have no conflicts of interest concerning this article.

ACKNOWLEDGEMENTS

This study was supported in part by grant number BK2012653 from the Natural Science Foundation of Jiangsu Province, China, by grant number 81370520 from National Natural Science Foundation of China (NSFC), by the Young Investigator Grant number Q201208 from the Department of Health, Jiangsu Province, China. We thank all patients and their families who participated in this study.

AUTHOR CONTRIBUTION

Yi Shen: study concept and design; analysis and interpretation of data; Yan-Mei Liu: field study, analysis and interpretation of data; Bin Wang: field study, analysis and interpretation of data; Yong-Gen Zhu: field study, valuable discussion; Yuan-Yuan Wang: field study, analysis and interpretation of data; Xu-Lin Wang: field study, analysis of data; Ju-Ling Ji: valuable discussion and support; Jian-Guo Shao: field study, valuable discussion and support; Yan Qin: overall study concept and design, critical revision of the manuscript for important intellectual content and final drafting of the manuscript; Gang Qin: overall study concept and design, analysis and interpretation of data, drafting of the manuscript.

REFERENCES


