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

Yi Shen,* Yan-Mei Liu,* Bin Wang,** Yong-Gen Zhu,** Yuan-Yuan Wang,*
Xu-Lin Wang,* Ju-Ling Ji,*** Jian-Guo Shao,** Yan Qin,**** Gang Qin**

* 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 R^2 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.[8] 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.[9] Meanwhile, several studies showed that certain subsets of patients with advanced liver disease may have high mortality in spite of low MELD scores.[10] 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),[11,12] MELD to sodium ratio (MESO),[13] integrated MELD (iMELD),[14] updated MELD (uMELD),[15] United Kingdom MELD (UKMELD)[16] and donor MELD (D-MELD).[17] Hyponatraemia was associated with hepatic encephalopathy (HE), hepatorenal syndrome (HRS) and increased mortality in patients with cirrhosis.[18] Older age was also an independent risk factor for mortality in these patients.[15] 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.[19,20] 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.[21] 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.[22] 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.[19,20,23] 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.[24] 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.[25] 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.[26] 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,[3] 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] MELD = 11.2 x ln(INR) + 9.6 x ln[creatinine (mg/dL)] + 3.8 x ln[bilirubin (mg/dL)] + 6.4 (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]

MELD-Na = MELD - Na - [0.025 x MELD x (140 - Na)] + 140.

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

The MESO index was defined as:[13]

MESO = (MELD/Na) x 10

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

iMELD score = MELD + (age x 0.3) - (Na x 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 χ^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) R^2 and the Brier score were calculated to evaluate overall model performance. Nagelkerke's R^2 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

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

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

INR: international normalized ratio. HE: hepatic encephalopathy. MELD: Model for Endstage Liver Disease. MELD-Na: sodium MELD. iMELD: integrated MELD. MESO: MELD to sodium ratio. CTP: Child-Turcotte-Pugh. mCTP: modified CTP. ALSS: artificial liver support system. NA: nucleos(t)ide analogues.

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

Model	3 months		5 years	
	HR[a] (95%CI)	P	HR[a] (95%CI)	P
MELD	1.12 (1.08-1.16)	< 0.001	1.10 (1.06-1.13)	< 0.001
MELD-Na	1.15 (1.10-1.19)	< 0.001	1.12 (1.09-1.16)	< 0.001
MESO	4.65 (3.17-6.82)	< 0.001	3.89 (2.75-5.50)	< 0.001
iMELD	1.09 (1.07-1.12)	< 0.001	1.09 (1.07-1.11)	< 0.001
CTP	1.50 (1.32-1.71)	< 0.001	1.50 (1.34-1.67)	< 0.001
mCTP	1.46 (1.30-1.65)	< 0.001	1.44 (1.30-1.59)	< 0.001

HR: hazard ratio. MELD: Model for End-Stage Liver Disease. iMELD: integrated MELD. MELD-Na: sodium MELD. MESO: MELD to sodium ratio. CTP: Child-Turcotte-Pugh. mCTP: modified CTP. [a] Adjusted for age and sex.

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

	3 months				5 years			
	H-L test		Calibration plot		H-L test		Calibration plot	
	χ^2	P	Intercept	Slope	χ^2	P	Intercept	Slope
MELD	10.78	0.21	0.037	0.93	5.79	0.67	0.055	0.98
MELD-Na	3.27	0.92	0.021	0.97	4.25	0.81	0.001	1.01
MESO	4.58	0.80	0.022	0.96	3.55	0.90	0.015	0.98
iMELD	6.38	0.61	0.027	0.96	3.72	0.88	0.029	0.96
CTP	10.68	0.03	-0.008	1.01	5.45	0.24	-0.025	1.04
mCTP	8.89	0.06	-0.006	0.99	5.79	0.22	-0.028	1.04

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.

Table 4. Pairwise comparison of the AUC to predict 3-month and 5-year mortality between MELD and the other five models.

	MELD	MELD-Na	MESO	iMELD	CTP	mCTP
						3 months
AUC	0.65	0.70	0.70	0.76	0.73	0.74
95%CI	0.58-0.72	0.63-0.77	0.63-0.76	0.70-0.82	0.67-0.79	0.67-0.80
P values MELD*	-	0.003	0.002	<0.001	0.154	0.089
						5 years
AUC	0.65	0.71	0.70	0.80	0.78	0.78
95%CI	0.58-0.72	0.64-0.78	0.64-0.77	0.74-0.86	0.72-0.84	0.71-0.84
P values MELD*	-	<0.001	<0.001	<0.001	0.003	0.002

Adjusted for Bonferroni correction. AUC: area under ROC curve. CI: confidence interval. 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.
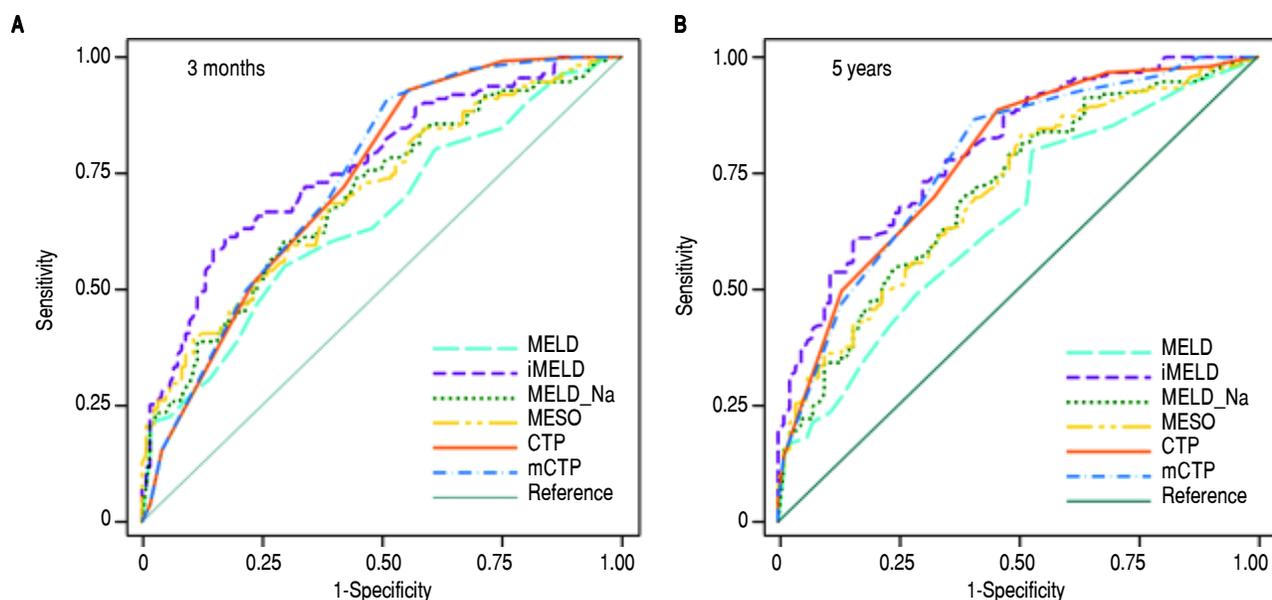
Figure 1. *Comparison of ROC curves for six score systems, with 3-month (**A**) and 5-year (**B**) mortality as primary outcome measures.*

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 ≥ 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 ≥ 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

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

Model	Optimal cutoff	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	O/E (95% CI)
						3 months
MELD	30	55% (45.2-64.4%)	70.2% (61.3-78.2%)	62.9% (52.5-72.5%)	63% (54.2-71.1%)	1.14 (0.94-1.38)
MELD-Na	32	55.9% (46.1-65.3%)	73.6% (64.8-81.2%)	66% (55.5-75.4%)	44.5% (55.9-72.4%)	1.18 (0.97-1.42)
MESO	2.6	38.7% (29.6-48.5%)	89.3% (82.3-94.2%)	76.8% (63.6-87%)	61.4% (53.7-68.6%)	1.98 (1.63-2.39)
iMELD	52	62.2% (52.5-71.2%)	80.2% (71.9-86.9%)	74.2% (64.1-82.7%)	69.8% (61.4-77.3%)	1.19 (0.98-1.44)
CTP	11	92.8% (86.3-96.8%)	44.6% (35.6-53.9%)	60.6% (52.8-68%)	87.1% (76.1-94.3%)	0.65 (0.54-0.79)
mCTP	14	91% (84.1-95.6%)	48.8% (39.6-58%)	62% (54-69.4%)	85.5% (75-92.8%)	0.68 (0.56-0.82)
						5 years
MELD	26	79.9% (72.5-86%)	47% (35.9-58.3%)	73% (65.5-79.7%)	56.5% (44-68.4%)	0.91 (0.77-1.07)
MELD-Na	26	91.9% (86.4-95.8%)	32.5% (22.6-43.7%)	71% (64-77.3%)	69.3% (52.4-83%)	0.77 (0.65-0.91)
MESO	1.9	84.6% (77.7-90%)	42.2% (31.4-53.5%)	72.4% (65.1-78.9%)	60.3% (46.6-73%)	0.86 (0.72-1.01)
iMELD	43	89.9% (83.9-94.3%)	48.2% (37.1-59.4%)	75.7% (68.7-81.8%)	72.7% (59.8-83.9%)	0.84 (0.71-0.99)
CTP	11	88.6% (82.4-93.2%)	54.2% (42.9-65.2%)	77.6% (70.6-83.7%)	72.6% (59.8-83.1%)	0.88 (0.74-1.03)
mCTP	14	88.6% (80-91.6%)	59% (47.7-69.7%)	79.1% (72.1-85.1%)	71% (58.8-81.3%)	0.91 (0.77-1.07)

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.

iMELD were 0.196 and 0.171 at 3 months and 5 years respectively. The overall performance was worse for the other five models.

DISCUSSION

In the early years of 21st century, MELD and MELD-based models have been developed to predict the short-term survival of patient 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
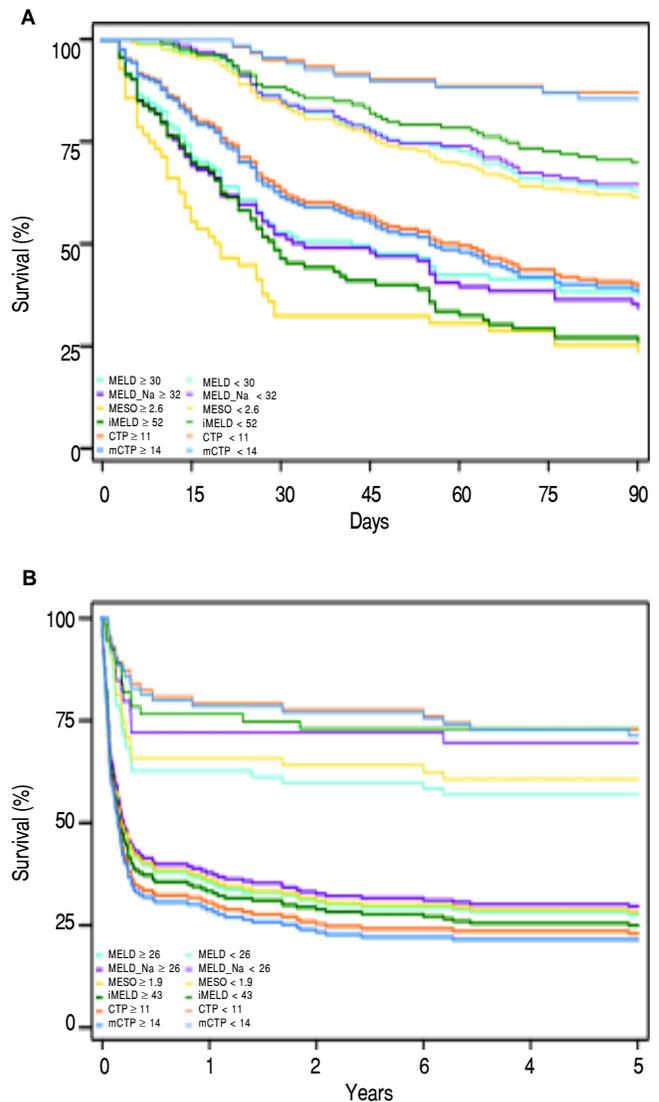
Figure 2. *Actual cumulative survival curves of the six models with optimal cutoffs of AUCs evaluated at 3 months (**A**) and 5 years (**B**).*

Table 6. Overall performance of each model.

	MELD	MELD-Na	MESO	iMELD	CTP	mCTP
						3 month
Nagelkerke's R^2 (%)	11.2	16.5	18.3	28.6	23.2	24.9
Brier score	0.229	0.219	0.216	0.196	0.207	0.204
						5 year
Nagelkerke's R^2 (%)	9.1	16.7	17.0	34.5	30.1	30.0
Brier score	0.215	0.201	0.203	0.171	0.176	0.176

filtration rate (eGFR) does not improve the predictive power of MELD.[31,32] 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.[14,15,33-35] 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.[23] Alcoholism, hepatitis C, and cholestasis are the leading causes of ACLF in Western countries, whereas HBV infection accounts for the majority 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.[36-38] The proportion of patients who survived the liver failure and lived much longer has steadily increased.[25,39] 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,[23,33,40] 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 surpris-

ing because age and serum sodium were strongly associated with higher mortality in these patients.[26] 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.[14] 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.[21] 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.[41] 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.[42] 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

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

1. Seto WK, Lai CL, Yuen MF. Acute-on-chronic liver failure in chronic hepatitis B. *J Gastroenterol Hepatol* 2012; 27: 662-9.
2. Arroyo V, Moreau R, Jalan R, Gines P, Study E-CCC. Acute-on-chronic liver failure: A new syndrome that will re-classify cirrhosis. *J Hepatol* 2015; 62: S131-S143.
3. Li LJ. Diagnostic and treatment guidelines for liver failure. *Zhonghua Gan Zang Bing Za Zhi* 2006; 14: 643-6.
4. Si CW, Zhuang H. Prevention and cure project of viral hepatitis. *Zhonghua Gan Zang Bing Za Zhi* 2000; 8: 324-9.
5. Sarin SK, Kumar A, Almeida JA, Chawla YK, Fan ST, Garg H, de Silva HJ, et al. Acute-on-chronic liver failure: con-

244

Shen Yi, et al. Annals of Hepatology, 2016; 15 (2): 236-245

sensus recommendations of the Asian Pacific Association for the study of the liver (APASL). *Hepatol Int* 2009; 3: 269-82.

6. Olson JC, Kamath PS. Acute-on-chronic liver failure: concept, natural history, and prognosis. *Curr Opin Crit Care* 2011; 17: 165-9.

7. Jalan R, Gines P, Olson JC, Mookerjee RP, Moreau R, Garcia-Tsao G, Arroyo V, et al. Acute-on chronic liver failure. *J Hepatol* 2012; 57: 1336-48.

8. 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: 864-71.

9. Zhang Q, Yang F, Li X, Wang LW, Chu XG, Zhang H, Gong ZJ. Trichostatin A Protects Against Experimental Acute-on-Chronic Liver Failure in Rats Through Regulating the Acetylation of Nuclear Factor-kappaB. *Inflammation* 2015.

10. Cholongitas E, Senzolo M, Triantos C, Samonakis D, Patch D, Burroughs AK. MELD is not enough-enough of MELD? *J Hepatol* 2005; 42: 475-7; author reply 478-9.

11. Biggins SW, Kim WR, Terrault NA, Saab S, Balan V, Schiano T, Benson J, et al. Evidence-based incorporation of serum sodium concentration into MELD. *Gastroenterology* 2006; 130: 1652-60.

12. Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, Edwards E, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med* 2008; 359: 1018-26.

13. Huo TI, Wang YW, Yang YY, Lin HC, Lee PC, Hou MC, Lee FY, et al. Model for end-stage liver disease score to serum sodium ratio index as a prognostic predictor and its correlation with portal pressure in patients with liver cirrhosis. *Liver Int* 2007; 27: 498-506.

14. Luca A, Angermayr B, Bertolini G, Koenig F, Vizzini G, Ploner M, Peck-Radosavljevic M, et al. An integrated MELD model including serum sodium and age improves the prediction of early mortality in patients with cirrhosis. *Liver Transpl* 2007; 13: 1174-80.

15. Sharma P, Schaubel DE, Sima CS, Merion RM, Lok AS. Reweighting the model for end-stage liver disease score components. *Gastroenterology* 2008; 135: 1575-81.

16. Neuberger J, Gimson A, Davies M, Akyol M, O'Grady J, Burroughs A, Hudson M, et al. Selection of patients for liver transplantation and allocation of donated livers in the UK. *Gut* 2008; 57: 252-7.

17. Halldorson JB, Bakthavatsalam R, Fix O, Reyes JD, Perkins JD. D-MELD, a simple predictor of post liver transplant mortality for optimization of donor/recipient matching. *Am J Transplant* 2009; 9: 318-26.

18. Koller H, Rosenkranz A. Hyponatremia to be an excellent predictor of outcome in patients with advanced cirrhosis. *Liver Transpl* 2005; 11: 1001; author reply 1001-2.

19. Biselli M, Gitto S, Gramenzi A, Di Donato R, Brodosi L, Ravaioli M, Grazi GL, et al. Six score systems to evaluate candidates with advanced cirrhosis for orthotopic liver transplant: Which is the winner? *Liver Transpl* 2010; 16: 964-73.

20. Marroni CP, de Mello Brandao AB, Hennigen AW, Marroni C, Zanotelli ML, Cantisani G, Fuchs SC, et al. MELD scores with incorporation of serum sodium and death prediction in cirrhotic patients on the waiting list for liver transplantation: a single center experience in southern Brazil. *Clin Transplant* 2012; 26: E395-E401.

21. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973; 60: 646-9.

22. Huo TI, Lin HC, Wu JC, Lee FY, Hou MC, Lee PC, Chang FY, et al. Proposal of a modified Child-Turcotte-Pugh scoring system and comparison with the model for end-stage liver disease for outcome prediction in patients with cirrhosis. *Liver Transpl* 2006; 12: 65-71.

23. Xun YH, Shi JP, Li CQ, Li D, Shi WZ, Pan QC, Guo JC, et al. Prognostic performance of a series of model for end-stage liver disease and respective Delta scores in patients with hepatitis B acute-on-chronic liver failure. *Mol Med Rep* 2014; 9: 1559-68.

24. Berg CL, Steffick DE, Edwards EB, Heimbach JK, Magee JC, Washburn WK, Mazariegos GV. Liver and intestine transplantation in the United States 1998-2007. *Am J Transplant* 2009; 9: 907-31.

25. Hessel FP, Bramlage P, Wasem J, Mitzner SR. Cost-effectiveness of the artificial liver support system MARS in patients with acute-on-chronic liver failure. *Eur J Gastroenterol Hepatol* 2010; 22: 213-20.

26. Qin G, Shao JG, Wang B, Shen Y, Zheng J, Liu XJ, Zhang YY, et al. Artificial Liver Support System Improves Short-and Long-Term Outcomes of Patients With HBV-Associated Acute-on-Chronic Liver Failure: A Single-Center Experience. *Medicine (Baltimore)* 2014; 93: e338.

27. Hosmer DWJ, Lemeshow S, Sturdivant R. Applied Logistic Regression. New York: John Wiley & Sons, 2013.

28. Nagelkerke NJ. A note on a general definition of the coefficient of determination. *Biometrika* 1991; 78: 691-2.

29. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; 21: 128-38.

30. Wan Z, Wu Y, Yi J, You S, Liu H, Sun Z, Zhu B, et al. Combining serum cystatin C with total bilirubin improves short-term mortality prediction in patients with HBVRelated acute on-chronic liver failure. *PLoS One* 2015; 10: e0116968.

31. Finkenstedt A, Dorn L, Edlinger M, Prokop W, Risch L, Griesmacher A, Graziadei I, et al. Cystatin C is a strong predictor of survival in patients with cirrhosis: is a cystatin C-based MELD better? *Liver Int* 2012; 32: 1211-6.

32. Myers RP, Shaheen AA, Aspinall AI, Quinn RR, Burak KW. Gender, renal function, and outcomes on the liver transplant waiting list: assessment of revised MELD including estimated glomerular filtration rate. *J Hepatol* 2011; 54: 462-70.

33. Yan H, Wu W, Yang Y, Wu Y, Yang Q, Shi Y. A novel integrated Model for End-Stage Liver Disease model predicts short-term prognosis of hepatitis B virus-related acute-on-chronic liver failure patients. *Hepatol Res* 2014; 45(4): 405-14**.**

34. Londono MC, Cardenas A, Guevara M, Quinto L, de Las Heras D, Navasa M, Rimola A, et al. MELD score and serum sodium in the prediction of survival of patients with cirrhosis awaiting liver transplantation. *Gut* 2007; 56: 1283-90.

35. Cywinski JB, Mascha EJ, You J, Sessler DI, Kapural L, Argalious M, Parker BM. Pre-transplant MELD and sodium MELD scores are poor predictors of graft failure and mortality after liver transplantation. *Hepatol Int* 2011; 5: 841-9.

36. Tsubota A, Arase Y, Suzuki Y, Suzuki F, Sezaki H, Hosaka T, Akuta N, et al. Lamivudine monotherapy for spontaneous severe acute exacerbation of chronic hepatitis B. *J Gastroenterol Hepatol* 2005; 20: 426-32.

37. Xia Q, Dai X, Huang J, Xu X, Yang Q, Liu X, Chen Y, et al. A single-center experience of non-bioartificial liver support systems among Chinese patients with liver failure. *Int J Artif Organs* 2014; 37: 442-54.

38. Liu JP, Gluud LL, Als-Nielsen B, Gluud C. Artificial and bioartificial support systems for liver failure. Cochrane Database Syst Rev 2004:CD003628.

39. Cui YL, Yan F, Wang YB, Song XQ, Liu L, Lei XZ, Zheng MH, et al. Nucleoside analogue can improve the long-term prognosis of patients with hepatitis B virus infection-associated acute on chronic liver failure. *Dig Dis Sci* 2010; 55: 2373-80.
40. Wu SJ, Yan HD, Zheng ZX, Shi KQ, Wu FL, Xie YY, Fan YC, et al. Establishment and Validation of ALPH-Q Score to Predict Mortality Risk in Patients With Acute-on-Chronic Hepatitis B Liver Failure: A Prospective Cohort Study. *Medicine (Baltimore)* 2015; 94: e403.
41. Kim DH, Kim SH, Kim KS, Lee WJ, Kim NK, Noh SH, Kim CB. Predictors of mortality in cirrhotic patients undergoing extrahepatic surgery: comparison of Child-Turcotte-Pugh and model for end-stage liver disease-based indices. *ANZ J Surg* 2014; 84: 832-6.
42. Silva PE, Fayad L, Lazzarotto C, Ronsoni MF, Bazzo ML, Colombo BS, Dantas-Correa EB, et al. Single-centre validation of the EASL-CLIF consortium definition of acute-on-chronic liver failure and CLIF-SOFA for prediction of mortality in cirrhosis. *Liver Int* 2015; 35: 1516-23.

Correspondence and reprint request:
Gang Qin, M.D., Ph.D.
Center for Liver Diseases, Nantong Third People's Hospital,
Nantong University, 99 Mid-Youth Road, Nantong, 226006,
Jiangsu, China.
Tel.: +86-189-1228-8106
E-mail: tonygqin@ntu.edu.cn