LIVER TRANSPLANTATION 16:964–973, 2010

ORIGINAL ARTICLE

Six Score Systems to Evaluate Candidates with Advanced Cirrhosis for Orthotopic Liver Transplant: Which Is the Winner?

Maurizio Biselli,[1] Stefano Gitto,[1] Annagiulia Gramenzi,[1] Roberto Di Donato,[1] Lucia Brodosi,[1] Matteo Ravaioli,[2] Gian Luca Grazi,[2] Antonio Daniele Pinna,[2] Pietro Andreone,[1] and Mauro Bernardi[1]

[1]Department of Clinical Medicine and [2]Department of General Surgery and Organ Transplantation, Alma Mater Studiorum-Università di Bologna; Azienda Ospedaliero-Universitaria Policlinico S. Orsola-Malpighi, Bologna, Italy

Many prognostic systems have been devised to predict the outcome of liver transplantation (LT) candidates. Today, the Model for End-Stage Liver Disease (MELD) is widely used for organ allocation, but it has shown some limitations. The aim of this study was to investigate the performance of MELD compared to 5 different score models. We evaluated the prognostic ability of MELD, modified Child-Turcotte-Pugh, MELD-sodium, United Kingdom MELD, updated MELD, and integrated MELD in 487 candidates with cirrhosis for LT at the Bologna Transplant Centre, Bologna, Italy, between 2003 and 2008. Calibration analysis by Hosmer-Lemeshow test, calibration curves, and concordance *c*-statistics (area under the receiver operating characteristic curve [AUC]) were calculated at 3, 6, and 12 months. Actual cumulative survival curves, taking into account the event of interest in the presence of competing risk, were obtained using the best cutoffs identified by AUC. For each score, the Hosmer-Lemeshow test revealed a good calibration. Integrated MELD showed calibration curves closer to the line of perfect predicting ability, followed by MELD-sodium at 3 months and modified Child-Turcotte-Pugh at 6 months. MELD-sodium AUCs at 3 and 6 months (0.798 and 0.765, respectively) and integrated MELD AUC at 6 months (0.792) were better than standard MELD ($P < 0.05$). Actual survival curves showed that these 2 scores were able to identify the patients with the highest drop-out risk. In conclusion, MELD-sodium and integrated MELD were the best prognostic models to predict drop-out rates among patients awaiting LT. ***Liver Transpl 16:964-973, 2010.*** © **2010 AASLD.**

Received December 23, 2009; accepted April 23, 2010.

The Model for End-Stage Liver Disease (MELD) was first described in 2000 to predict 3-month survival rates in patients with chronic liver disease undergoing transjugular intrahepatic portosystemic shunt.[1] This model, which includes 3 objective measurements (serum creatinine, bilirubin, and prothrombin time international normalized ratio [INR]) was subsequently proved to predict waitlist mortality in the liver transplantation (LT) setting more precisely than the Child-Turcotte-Pugh (CTP) score.[2,3] At present, MELD is widely used for organ allocation,[4] but it has shown some limitations. First, MELD was devised to predict short-term survival,[3] whereas the time spent on the LT waitlist has increased over the past decade, reaching almost 1 year in about 63% of cases.[5] Moreover, MELD benefits patients with cholestasis or renal failure and is not directly influenced by other complications of cirrhosis associated with poor survival, such

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; HBV, hepatitis B virus; HCV, hepatitis C virus; HDV, hepatitis delta virus; HR, hazard ratio; iMELD, integrated MELD; INR, international normalized ratio; LT, liver transplantation; mCTP, modified Child-Turcotte-Pugh; MELD, Model for End-Stage Liver Disease; MELD-Na, MELD and serum sodium; MESO, MELD to serum sodium ratio; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic; SD, standard deviation; UKELD, United Kingdom MELD; uMELD, updated MELD.

Roberto Di Donato and Lucia Brodosi received research funding from Associazione per la Ricerca sulle Malattie Epatiche (ARME), Bologna, Italy.

Address reprint requests to Pietro Andreone, M.D., Dipartimento di Medicina Clinica, Alma Mater Studiorum, Università di Bologna, Azienda Ospedaliero-Universitaria Policlinico Sant'Orsola-Malpighi, Via Massarenti, 9, 40138 Bologna, Italy. Telephone: +390516363618; FAX: +39.051. 345806; E-mail: pietro.andreone@unibo.it

DOI 10.1002/lt.22093
View this article online at wileyonlinelibrary.com.
LIVER TRANSPLANTATION.DOI 10.1002/lt. *Published on behalf of the American Association for the Study of Liver Diseases*

as persistent ascites[6,7] and hyponatremia.[8] For this reason, many recent studies have evaluated the effect of incorporating other variables into the model, such as serum sodium and age.[9-13] On the other hand, the MELD formula has recently been modified by assigning a lower weight to creatinine and INR, and a higher weight to bilirubin, because LT candidates with higher bilirubin had higher waitlist mortality.[14] Moreover, to overrun the ceiling effect of standard CTP, which accounts for the inability to properly stratify the severity of advanced cirrhosis, additional points have also been advocated for very high bilirubin, very low albumin levels, and markedly prolonged prothrombin time.[15]

To date, only 2 studies have evaluated the prognostic power of standard MELD compared to other MELD-based score systems.[16,17] Namely, 3 scores were evaluated: MELD incorporating serum sodium (MELD-Na), integrated MELD (iMELD) incorporating sodium and age, and MELD to serum sodium ratio (MESO). These studies[16,17] found that MELD-Na and iMELD predicted the outcome of patients with decompensated cirrhosis better than standard MELD, but none specifically addressed the issue of test performances in the entire spectrum of disease severity within the studied patient population.

The purpose of the present study was to compare the short-term and intermediate-term prognostic ability of the standard MELD with respect to 5 alternative scores[10-15] in patients with advanced cirrhosis awaiting LT. The performance of these score systems in relation to the varying severity of cirrhosis was specifically assessed by calibration analysis.

PATIENTS AND METHODS

Patients

This study included 487 consecutive patients with cirrhosis, ≥18 years of age, and listed for LT at our center from January 2003 to December 2008. The minimal criteria for LT listing were those reported by the American Society of Transplant Physicians and the American Association for the Study of Liver Diseases.[18] In this time period, all patients were prioritized by use of MELD score. In the same period, 466 patients with liver diseases worthy of MELD exceptions and added points,[19,20] such as fulminant hepatic failure, retransplantation, hepatocellular carcinoma, primary sclerosing cholangitis, need for combined liver-kidney transplantation, presence of pulmonary hypertension, and genetic liver disease (polycystic liver disease and amyloidosis) were listed at our center but were not included in the present study. Patients infected by human immunodeficiency virus were also excluded.

Patients with a history of alcohol consumption or drug abuse were only listed after an abstinence period for at least 6 months. At the time of listing, demographic, clinical and laboratory data, etiology, and history of complications were recorded and available for analysis. The presence and severity of ascites (or hydrothorax) were assessed by physical examination, ultrasonography, or computed tomography and classified as either controlled or uncontrolled by diuretic treatment, assigning 2 or 3 points, respectively, according to CTP.[21,22] Hepatic encephalopathy was stratified into 3 grades: absent, low-grade, and high-grade.[23]

Survival was calculated from the time of listing to drop-out, LT, or end of the observation period (December 31, 2008). Drop-out included patients removed from the list because of either death or worsening of their disease up to the point that they were too sick to undergo LT. We considered these 2 events as a combined endpoint because they both represent a failure of the allocation system. Patients removed from the active waiting list because of disease improvement continued to be considered in active follow-up and were monitored at our transplant center.

All data were aggregated in a database and were rendered anonymous before the study. This procedure was notified to the Ethics Committee of the Institution. The retrospective study conformed to the ethical guidelines of the Declaration of Helsinki.

Calculation of Scores

The standard MELD score[1] was computed by a formula that adds multiples of the natural logarithm (ln) of the values for INR, creatinine, and bilirubin as follows:

$$11.2 \times \ln(\text{INR}) + 9.57 \times \ln[\text{creatinine(mg/dL)}] + 3.78 \times \ln[\text{bilirubin(mg/dL)}] + 6.43(\text{an intercept})$$

with a lower limit of 1 for all variables and with creatinine capped at 4 (creatinine was set at 4 if the patient was receiving renal-replacement therapy). The resulting score was rounded to the nearest integer, with higher values indicating more severe disease. The United Kingdom MELD (UKELD) score[10,11] was calculated as follows:

$$\begin{aligned}\text{UKELD} = [&(5.395 \times \ln(\text{INR})) \\ &+ (1.485 \times \ln(\text{creatinine}, \mu\text{mol/L})) \\ &+ (3.13 \times \ln(\text{bilirubin}, \mu\text{mol/L})) \\ &- (81.565 \times \ln(\text{Na}, \text{mmol/L}))] + 435\end{aligned}$$

The iMELD equation[12] was based on MELD, age (years), and Na (mEq/L) calculated as follows:

$$\text{original MELD score} + (\text{age} \times 0.3) - (0.7 \times \text{Na}) + 100.$$

The MELD-Na equation[13] was based on the formula:

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

where the Na concentration is bound between 125 and 140 mmol/L. The updated MELD (uMELD)[14] assigns a lower weight to creatinine and INR and a higher weight to bilirubin, and does not incorporate the lower bounds for bilirubin and INR, and both the lower and upper bound of creatinine. uMELD was calculated as follows:

$$1.266 \times \ln(1 + \text{creatinine}, \text{mg/dL}) + 0.939 \times \ln(1 + \text{bilirubin}, \text{mg/dL}) + 1.658 \times \ln(1 + \text{INR})$$

The modified CTP (mCTP)[15] was obtained by assigning an additional point in patients whose serum

bilirubin was >8 mg/dL, prothrombin time prolongation >11 seconds, or albumin <2.3 g/dL; accordingly, a mCTP score of 16-18 was defined as mCTP class D, which identifies severely decompensated cirrhosis.

Statistical Methods

Continuous variables were expressed as mean ± standard deviation (SD). Comparisons between groups of patients were made by the χ^2 test or Fisher's exact test (2-tailed) for qualitative variables, whereas the Mann-Whitney U test was performed for quantitative variables. The accuracy of outcome prediction was assessed in terms of calibration and discrimination. Calibration was evaluated with Hosmer-Lemeshow goodness-of-fit χ^2 estimates[24] and with calibration curves at 3 and 6 months. The Hosmer-Lemeshow test is based on grouping cases into 10 equal sets (deciles) of risk and comparing the observed probability with the expected probability of an event within each decile. Thus, this test is a measure of the discrepancy between the observed and predicted events. A better-calibrated model would have a small discrepancy between the observed and predicted events; thus, a high *P* value suggests good calibration whereas a small *P* indicates poor calibration. Calibration curves are based on linear regression analysis. The observed drop-out rates are plotted against predicted drop-out rates. The R-squared value represents the proportion of variation of the dependent variable (observed drop-out rate) that is predicted from the independent variable (predicted drop-out rate). An R-squared value of 1.0 indicates that all plotted points lie on a straight line and therefore the independent variable predicts the dependent variable with certainty. If the plotted points (predicted and observed drop-out rates) lie on a 45-degree line (slope = 1), with intercept = 0, the predictive score fits the study data well. An upward shift of the line implies that the score underestimates the actual drop-out rate whereas a downward shift represents an overestimation of the actual drop-out rate. Discrimination of each score in predicting the risk of dropout at 3 and 6 months was assessed by *c*-statistics equivalent to the area under the receiver operating characteristic (ROC) curve (AUC).[25] AUCs were evaluated with pairwise comparison using the algorithm described by Hanley and McNeil.[26] When a significant difference between AUCs was found, the sample size was evaluated with type 1 error alpha set to 0.05 and type 2 error beta set to 0.1. Finally, the cutoffs ensuring the lowest false negative and false positive results of the scores showing a better AUC than standard MELD at 3 and 6 months were used to calculate the actual cumulative patient survival probability. The same analysis was performed in patients with standard MELD score cutoff at 15 and 18 points on the basis of evidence emerging with the transplant benefit.[27] All the actual cumulative survival curves took into account the presence of a second event, namely transplantation, which was defined as a competing risk event.[28] The comparisons between the actual cumulative survival probability curves in the presence of competing risk were based on the class of tests proposed by Gray.[29] All statistical analyses were performed using SPSS for Windows version 13.0 (SPSS Inc., Chicago, IL), MedCalc for Windows version 9.2.1.0 (MedCalc Software, Mariakerke, Belgium), and an add-on package of R.[30] Statistical significance was defined as $P < 0.05$.

TABLE 1. Baseline Demographics of the Study Patients

Characteristic	Value
Number of patients	487
Male/female	339 (69.6%)/ 148 (30.4%)
Age (years)	51.5 ± 9.7
Blood group	
A	218 (44.7%)
B	52 (10.7%)
AB	10 (2.1%)
0	207 (42.5%)
Etiology of cirrhosis	
Virus	281 (57.7%)
HCV	212 (43.5%)
HBV	30 (6.2%)
HBV/HDV	26 (5.3%)
HBV/HCV or HBV/HDV/HCV	13 (2.7%)
Alcohol	112 (23%)
Virus and alcohol	44 (9%)
Other	50 (10.3%)
Serum Biochemistry	
Bilirubin (mg/dL)	5.96 ± 7.61
Creatinine (mg/dL)	1.06 ± 0.43
INR	1.68 ± 0.52
Albumin (g/dL)	3.11 ± 0.53
Serum sodium (mEq/L)	137 ± 4.7
MELD at listing	18.1 ± 6 (median: 17; range: 8-45)
Scores at listing	
mCTP	9.7 ± 2
UKELD	56.1 ± 5
iMELD	37.7 ± 7.7
MELD-Na	20 ± 6.1
uMELD	4 ± 0.9
MELD at transplant (n = 88)*	24.2 ± 6.5 (median: 24; range: 8-43)

NOTE: Values are expressed as number (%) or mean ± SD.
*Patients who underwent LT within 6 months after listing.

RESULTS

Table 1 shows the baseline characteristics of the study population. Patients were predominantly male and almost all had cirrhosis due to hepatitis virus infection and/or alcohol abuse. The 487 patients included in this study were followed up for a median period of 12.1 months (range = 0.1-71.5 months). Among them, 159 (32.6%) underwent LT, 157 (32.3%) dropped out of the waiting list because of either 127

TABLE 2. Calibration Performance (Hosmer-Lemeshow Goodness-of-Fit Test) for Each Score at 3 and 6 Months

	3 Months		6 Months	
Score	HL χ^2	HL *P*	HL χ^2	HL *P*
MELD	8.122	0.422	10.647	0.222
mCTP	2.881	0.718	6.101	0.296
MELD-Na	5.993	0.648	9.381	0.311
iMELD	5.856	0.663	11.213	0.19
UKELD	10.866	0.209	8.553	0.381
uMELD	3.59	0.892	4.048	0.853

NOTE: "HL *P*" indicates where Hosmer-Lemeshow (HL) probability > χ^2 value.

TABLE 3. Calibration Performance (Linear Regression Analysis) for Each Score at 3 and 6 Months

	R-Squared Value	
Score	3 Months	6 Months
MELD	0.726	0.74
mCTP	0.756	0.799
MELD-Na	0.78	0.787
iMELD	0.78	0.811
UKELD	0.752	0.761
uMELD	0.728	0.737

NOTE: The R-squared value represents the proportion of variation between the observed and the expected drop-out rate. An R-squared value of 1.0 indicates that the predicted drop-out matched the observed with certainty.

deaths (26.1%) or 30 patients with extremely poor clinical condition (6.2%). The overall actual survival rates were 89.7%, 85%, 79.3%, and 72.6% at 3, 6, 12, and 24 months, respectively. The actual survival curves obtained by patient stratification according to the etiology of liver disease did not differ significantly (data not shown).

Calibration Analysis

The Hosmer-Lemeshow goodness-of-fit test revealed a good calibration ($P > 0.05$) for all the scores at 3 and 6 months (Table 2). Calibration for all the scores was further explored by plotting the observed against the expected drop-out frequency for every score. Our data showed that all scores predicted the proportion of variation of observed drop-out rate with good estimate at 3 and 6 months (Table 3). The scores with the most accurate calibration curve at 3 months were MELD-Na and iMELD (Fig. 1). However, the stratified predicted risk of MELD-Na and iMELD at 3 months did not completely overlap with the diagonal identity line (predicted = observed), thus overestimating the probability of drop-out for low-scoring patients and underestimating the drop-out risk in high-scoring patients. At 6 months, the best calibrated score was iMELD.

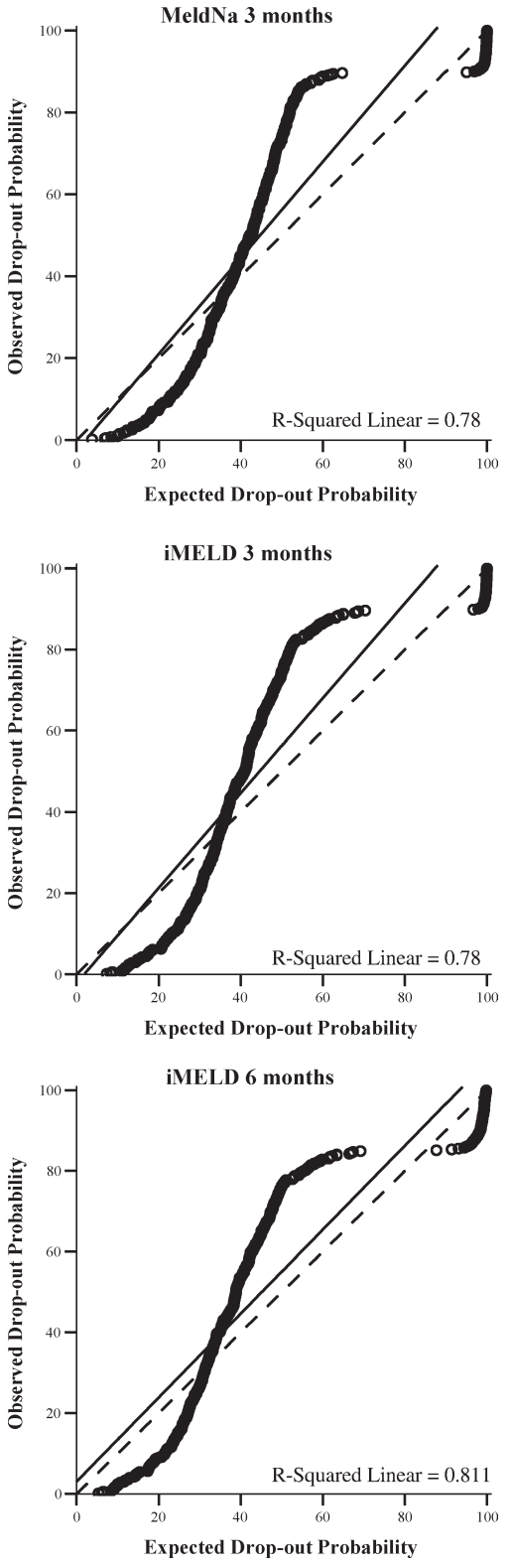

Figure 1. Calibration curves for MELD-Na at 3 months and iMELD at 3 and 6 months constructed by plotting observed drop-out rates against predicted drop-out rates. The dotted lines show perfect agreement between observed and expected drop-out estimates. The continuous lines are fit lines that represent data trend.

TABLE 4. Pairwise Comparison of the AUC to Predict 3-Month and 6-Month Drop-Out Rates Between MELD and the Other 5 Scores Analyzed

				Scores		
Time Point	MELD	mCTP	MELD-Na	iMELD	UKELD	uMELD
						3-months
AUC	0.759	0.769	0.798	0.806	0.781	0.757
95% CI	0.719-0.797	0.729-0.806	0.76-0.833	0.768-0.841	0.741-0.817	0.716-0.794
P values MELD	–	0.738	**0.03**	0.097	0.489	0.842
						3-months
AUC	0.724	0.757	0.765	0.792	0.748	0.727
95% CI	0.682-0.764	0.716-0.794	0.725-0.802	0.753-0.827	0.707-0.786	0.685-0.766
P values MELD	–	0.203	**0.011**	**0.006**	0.388	0.845

The bold p values are statistically significative.

Even in this case, the fit line (Fig. 1) shows that iMELD overestimates drop-out probability for the entire range of patient risk.

Discrimination Analysis

Comparison Between Standard MELD and the Other 5 Scores (c-Statistics)

The AUCs for standard MELD and the other 5 scores were evaluated at 3 and 6 months.

(1) Three-month analysis: 49 of 487 patients died (10.1%), 1 was judged too sick for LT (0.2%), 61 underwent transplantation (12.5%), and 376 survived (77.2%). AUCs computed from *c*-statistics are listed in Table 4. All the scores showed a good diagnostic accuracy.

The iMELD had the highest AUC, showing an excellent diagnostic accuracy, followed by MELD-Na, but the comparison between AUCs showed that only MELD-Na had a better prognostic power than the standard MELD because of a very small standard error (0.018) in the difference between the areas (0.039, 95% confidence interval [CI] = 0.004-0.074) (Table 4, Fig. 2).

(2) Six-month analysis: 68 of 487 patients died (14%), 5 were judged too sick for LT (1%), 88 underwent transplantation (18.1%), and 325 survived (66.7%). AUCs computed from *c*-statistics are listed in Table 4. All the scores had values in the range of clinical usefulness. The iMELD had the highest AUC, followed by MELD-Na. The comparison between AUCs showed that only iMELD and MELD-Na had a better prognostic power than the standard MELD (Table 4, Fig. 2).

Comparison Between Actual Survival Curves Obtained by MELD and 2 MELD-Based Scores (iMELD and MELD-Na)

Based on the analysis of ROCs at 3 and 6 months, we compared the actual survival curves of the scores with the predictive accuracy of patient drop-out higher than standard MELD at the 2 time points. The cutoff deriving from ROCs with the best ability to predict the 3-month and 6-month drop-out rate was 22 for MELD-Na, whereas the cutoff with the best ability to predict the 6-month drop-out rate was 39 for iMELD. Table 5 reports the sensitivity, specificity, positive (PPV) and negative predictive values (NPV) for these 2 scores. A MELD-Na cutoff of 22 and iMELD cutoff of 39 discriminated patients who would survive from those who would drop out (Fig. 3A) ($P < 0.001$). Interestingly, these 2 groups of patients were comparable in terms of age, sex, and blood group (data not shown), regardless of the score system employed and the cutoff analyzed. As expected, the only exception was age, which was higher in patients with elevated iMELD (53.8 ± 8.6 years versus 50.2 ± 10.1 years, $P < 0.001$).

Using the same cutoffs, we also compared the actual survival rates of a MELD-Na cutoff of 22 and an iMELD cutoff of 39 with patients with standard MELD score set to 15 (Fig. 3B) and to 18 points (Fig. 3C). Both the MELD-Na cutoff of 22 and iMELD cutoff of 39 performed similarly with standard MELD score set to 15 or 18 points, identifying patients with a good prognosis at 6 months ($P > 0.05$ in all cases; Fig. 3B,C). On the other hand, a MELD-Na cutoff of 22 classified listed patients with a significantly worse prognosis at 6 months with respect to standard MELD set both to 15 ($P < 0.001$; Fig. 3B) and 18 ($P = 0.034$; Fig. 3C). Furthermore, an iMELD cutoff of 39 identified listed patients with a worse prognosis more reliably than standard MELD set to 15 ($P = 0.005$; Fig. 3B), whereas no significant difference was found with respect to standard MELD set to 18 ($P = 0.234$; Fig. 3C).

DISCUSSION

The enhanced efficacy of LT as a treatment for end-stage liver disease has led to a progressive discrepancy between supply and demand for donor livers. As a result, the proportion of patients dying while on the waitlist has steadily increased.[31] In an attempt to reduce waitlist mortality, a new allocation policy replacing the CTP with MELD has been adopted since

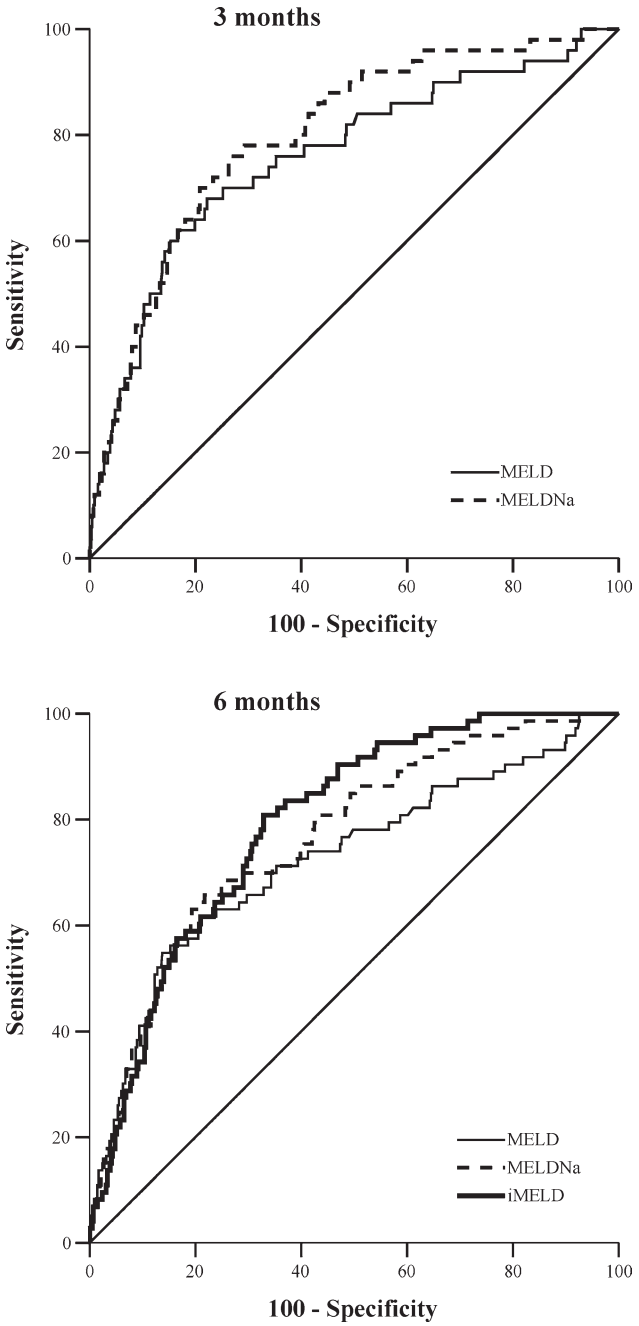

Figure 2. Comparison of ROC curves for scores with higher predictive accuracy of drop-out rates with respect to MELD at 3 and 6 months.

2002. Indeed, by allowing available grafts to be allocated to the sickest patients, the MELD system has led to a decrease in waitlist mortality[32] without impairing the transplant outcome.[33] Nevertheless, the MELD system does not take into account important prognostic factors. In particular, the role of hyponatremia as an independent predictor of mortality has been convincingly demonstrated,[8] and some studies assessed the prognostic value of a new score deriving from the integration of sodium in the MELD score.[34,35] The applicability of sodium-based MELD scoring systems in organ allocation has some limitations due to interlaboratory variability and the potential variability of serum sodium concentration after simple therapeutic maneuvers such as the administration of diuretics or intravenous hypotonic fluids or plasma volume expansion. Despite these caveats, Na-based MELD scoring systems represent a major advance in the prognostic assessment of patients with cirrhosis.[36]

To date, only two studies[11,13] with an adequate sample size have evaluated the impact of modified MELD score on waitlist mortality, and both reported that the incorporation of Na into the MELD score may enhance prognostic accuracy. One study elaborated the MELD-Na formula, based on data from the huge register of the U.S. Organ Procurement and Transplantation Network[14] and the other proposed the UKELD score, which is currently used to prioritize patients on the LT waiting list in the United Kingdom.[11] Recently, based on the observation that Na inversely correlated with the severity of cirrhosis, a further score derived from the ratio between MELD and Na concentration (MESO) has been proposed, but it was tested and validated in patients not listed for LT.[9,37] Other MELD-based models have also been devised that incorporate Na concentration and add either age[12] or presence of ascites.[38] Indeed, ascites associated with hyponatremia reflects a severe hemodynamic derangement,[39] a condition entailing a high risk of developing hepatic encephalopathy and hepatorenal syndrome, with both implying a poor prognosis.[7] However, the addition of ascites in a MELD-based score enhanced its prognostic ability only in patients with low standard MELD,[38] and its applicability to the entire spectrum of listed patients needs further assessment. On the contrary, iMELD that incorporates sodium and age was validated in a whole cohort of patients on the waiting list who were evaluated consecutively in an Italian transplant center.[12]

TABLE 5. Sensitivity, Specificity, NPV, and PPV for the Scores Better than MELD to Predict Drop-Out Rates with the Best Predictive Cutoffs at 3 and 6 Months

Score (Cutoff)	Sensitivity (95% CI)	Specificity (95% CI)	NPV (95% CI)	PPV (95% CI)
				3 months
MELD-Na (22)	76% (64.2-87.8%)	73.7% (69.6-77.8%)	96.4% (94.4-98.4%)	24.9% (18.1-31.8%)
				6 months
MELD-Na (22)	68.5% (57.9-79.2%)	75.1% (70.9-79.3%)	93.1% (90.4-95.8%)	32.7% (25.3-40.1%)
iMELD (39)	76.7% (67-86.4%)	67.6% (63.1-72.1%)	94.3% (91.7-96.9%)	29.4% (22.6-35.9%)

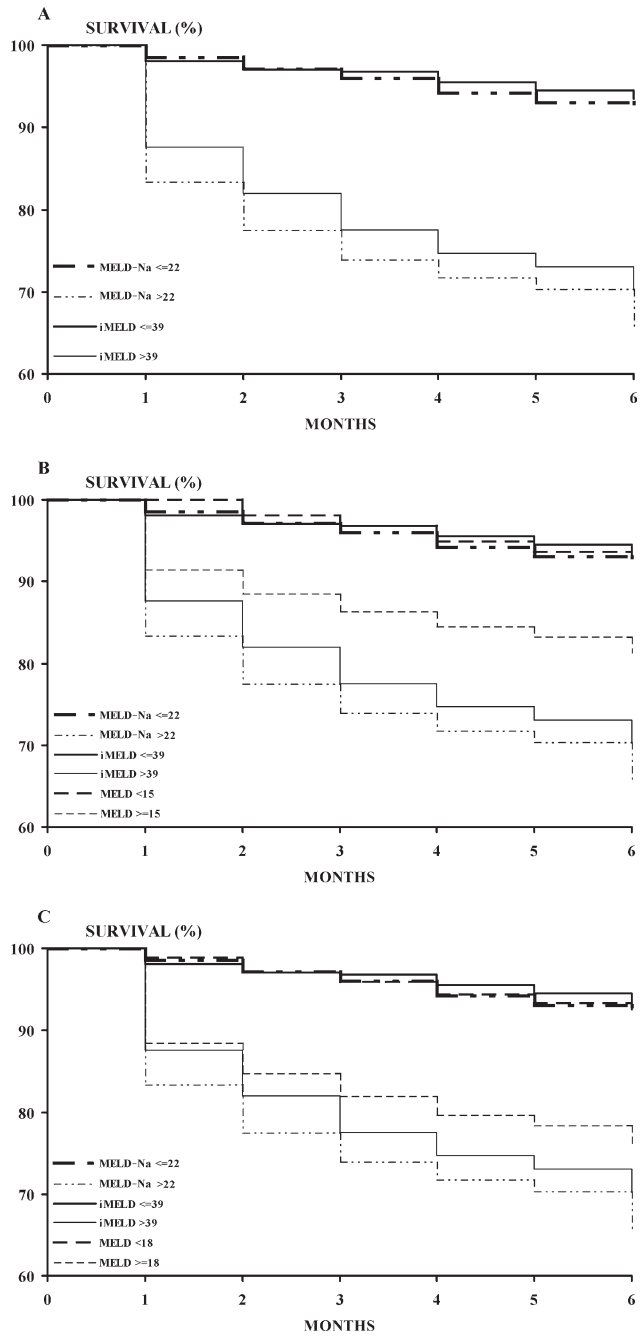

Figure 3. (A) Actual cumulative waiting list survival curves for scores with the predictive accuracy of drop-out higher than standard MELD. (B) Actual cumulative survival curves of MELD-Na and iMELD resulting from cutoffs of AUCs evaluated at 6 months, compared with standard MELD score set to 15 points. (C) Actual cumulative survival curves of MELD-Na and iMELD resulting from cutoffs of AUCs evaluated at 6 months, compared with standard MELD score set to 18 points. The thick lines represent patients with the best scores, whereas the thin lines represent patients with the worst scores.

The present study only assessed those prognostic models that had been evaluated in entire patient populations in the transplant setting, ie, MELD-Na, UKELD, iMELD, uMELD, and mCTP, and compared their performance to that of standard MELD.

To the best of our knowledge, only 2 studies have compared the performance of different scores. However, one study suffered from a small sample size,[17] whereas the other enrolled patients who were rather old for LT and mostly had HBV-related cirrhosis,[16] which contrasts with the usual prevalent etiologies seen at transplant centers in the United States and Europe. Both studies calculated MELD-Na using the formula proposed by Biggins et al.[35] instead of the more recent one elaborated by Kim et al.[13] Finally, neither study assessed the test performance at different degrees of cirrhosis.

Our study is the first to compare all the latest prognostic scores tested on candidates for LT, predicting short-term and medium-term drop-out rates in cirrhosis in the same population, that includes both MELD-based scores and a CTP-based score. To avoid a confrontation bias between MELD and the other scores, all patients affected by liver diseases with MELD exceptions and added points were excluded.[19,20] In fact, patients with MELD exceptions undergo LT after a time spent on the waiting list that is relatively unrelated to their standard MELD score. This could make it impossible to correlate the event of interest, namely the drop-out from the waiting list, with the score. Our study population was similar to that of other European and American centers[4,5,12,40-42] in terms of age, etiology, time spent on the waiting list, and mortality rate throughout the observation period.

In addition, the mean MELD at listing was 18, the minimal value from which the survival benefit at 1 year has been clearly demonstrated,[27] and the mean MELD at transplant represented an unquestionable indication for LT. Moreover, the etiology of cirrhosis did not modify the actual survival rate of listed patients, which then allowed an assessment of the prognostic ability of the scores not influenced by the etiology of liver disease. The analysis of data from a single center should have ensured uniform allocation criteria, follow-up schedule, diagnosis and treatment of pre-LT complications, and prioritization on the waiting list on the basis of MELD throughout the study. Finally, the issue of assessing the test performances in the entire spectrum of disease severity within our patients was specifically addressed.

Our discrimination analysis showed that 2 of 3 MELD-based scores incorporating sodium, namely MELD-Na and iMELD, predicts drop-out rates better than MELD, whereas the performances of UKELD is similar. Among the MELD-based scores, MELD-Na and iMELD were comparable, having an AUC value exceeding 0.75 at any time point (3 and 6 months), indicating good diagnostic accuracy. These 2 models could be considered similarly accurate in prediction outcome. However, MELD-Na better predicts drop-outs at each time point with respect to standard MELD, confirming that serum sodium concentration is a strong predictor of waitlist mortality.[8,34] In the medium-term evaluation, age emerged as the other predictor of patient drop-out. This is not surprising because age, which is incorporated into iMELD along

with sodium, is strongly associated with higher mortality in cirrhosis.[43] However, age inclusion in this setting could raise ethical issues because LT in elder recipients has been associated with lower patient and graft survival,[44] namely in patients with end-stage liver disease.[45] Nonetheless, the improved results of LT and the growing age of U.S. and European populations has led to a steadily increasing demand for LT in older recipients.[5,40] Thus, age could be included as a variable in a MELD-based evaluation in the coming years.

Our study also analyzed the performance of 2 scores that do not include sodium, ie, uMELD and mCTP. Because of the hypothesis that, given the same MELD, the mortality risk of patients with renal failure differs from that of patients with normal renal function, Sharma et al.[14] recently tried to improve MELD performance by modifying the three coefficients of the formula (uMELD) using data from the Scientific Registry of Transplant Recipients for all listed adult candidates in the United States. In our hands, however, uMELD and standard MELD had comparable predictive values at 3, 6, and 12 months.[14] Such variant results could likely be explained by differences among enrolled patients. Although our patients had mean creatinine within the normal range with a narrow SD, those enrolled by Sharma et al.[14] showed a slightly elevated mean creatinine with a wide SD. Therefore, creatinine may have influenced the score in their study.

The mCTP score[15] was devised to attenuate the ceiling effect of the traditional CTP by extending points up to 18 with the addition of a further class (class D) in patients with cirrhosis listed for LT in an Asian Center. This new score may offer an advantage because MELD is not influenced by hepatic encephalopathy and ascites, so that candidates for LT presenting these complications might not receive timely transplants.[46] However, in our study population, the prognostic power of mCTP did not differ from that of MELD, confirming a previous report,[47] even if the calibration of the score, evaluated as the match between observed and predicted drop-out events, seems to improve over time (Table 3). For this reason, this score could be useful for application in individual cases when the LT waiting time is prolonged.

Having found that the 3-month and 6-month AUCs of MELD-Na and 6-month AUC of iMELD were significantly better than MELD, we identified their respective cutoffs with best NPV and PPV. Interestingly, the two scores had a relatively low PPV (25%-33%), but a high NPV (>90%) in all cases. Such a feature may serve to select and prioritize LT candidates, because it would help to avoid futile transplants. Indeed, the 3-month and 6-month actual survival curves of patients with a score lower than the best predictive cutoffs of the 2 mentioned scores showed a high probability of survival, comparable to that reported in patients who underwent transplantation.[5,40] Although this dichotomous approach is not the standard procedure in the organ allocation process, it could be helpful in selecting LT candidates. Interestingly, both the MELD-Na and iMELD cutoffs identified patients with a good prognosis in a similar way with respect to standard MELD score set to either 15 or to 18 points. However, both a MELD-Na cutoff of 22 and iMELD cutoff of 39 were more able to identify patients with a worse prognosis than standard MELD set to 15 points, and a MELD-Na cutoff of 22 also did so with respect to standard MELD set to 18 points.

Despite these good results, discrimination is not the only factor determining the applicability of a prognostic model. A model could exhibit a good discrimination but still be useless for application in individual cases. In fact, these scores were implemented to stratify the drop-out risk in large cohorts of LT candidates who show widely different degrees of disease severity. For this reason, these scores must be as calibrated as possible to the population under study.

All the analyzed scores showed a good calibration. This emphasizes their usefulness in clinical practice. In particular, iMELD and MELD-Na at 3 months and iMELD at 6 months showed the best calibration curves, meaning that they possess the greatest ability to predict drop-out in single patients.

In conclusion, according to both calibration and discrimination analysis, among the scores proposed for selecting and prioritizing LT candidates, some of those incorporating sodium, namely MELD-Na and iMELD, are the most accurate in predicting the drop-out rate of patients with cirrhosis from the waiting list. MELD-Na was the best drop-out predictor at 3 months, whereas both MELD-Na and iMELD emerged as highly performing scores in the medium term.

ACKNOWLEDGMENTS

The authors thank their colleagues at the Hepatology Units of Azienda Ospedaliero-Universitaria Policlinico S. Orsola-Malpighi of Bologna who gave their support in patient management: Giorgio Ballardini, Sonia Berardi, Francesco Bianco Bianchi, Gian Paolo Bianchi, Luigi Bolondi, Paolo Caraceni, Matteo Cescon, Antonio Colecchia, Paolo Costigliola, Alessandro Cucchetti, Alessandro Dazzi, Massimo Del Gaudio, Giorgio Ercolani, Giuliano La Barba, Augusto Lauro, Marco Lenzi, Francesca Lodato, Stefania Lorenzini, Giuseppe Mazzella, Maria Cristina Morelli, Fabio Piscaglia, Enrico Roda, Claudia Sama, Maria Rosa Tamè, Fabio Tumietto, Franco Trevisani, Giovanni Varotti, Gaetano Vetrone, Giovanni Vitale, Marco Vivarelli, Matteo Zanello, and Marco Zoli.

We also thank Caterina Liberati, Ph.D. (Dipartimento di Scienze Statistiche, Università di Bologna, Bologna, Italy) for providing assistance in the statistical evaluation, and Professor Calogero Cammà (Cattedra ed Unità Operativa di Gastroenterologia, Università di Palermo, Palermo, Italy) for advice about the statistical methods.

REFERENCES

1. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology 2000;31:864-871.

2. Wiesner RH, McDiarmid SW, Kamath PS, Edwards EB, Malinchoc M, Kremers WK, et al. MELD and PELD: application of survival models to liver allocation. Liver Transpl 2001;7:567-580.
3. Forman LM, Lucey MR. Predicting the prognosis of chronic liver disease: an evolution from Child to MELD. Mayo End-stage Liver Disease. Hepatology 2001;33: 473-475.
4. Wiesner RH, Edwards E, Freeman R, Harper A, Kim R, Kamath P, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. Gastroenterology 2003;124:91-96.
5. Berg CL, Steffick DE, Edwards EB, Heimbach JK, Magee JC, Washburn WK, et al. Liver and intestine transplantation in the United States 1998-2007. Am J Transplant 2009;9(Part 2):907-931.
6. D'Amico G, Garcia-Tsao G, Magliaro L. Natural history and prognostic indicators of survival in cirrhosis. A systematic review of 118 studies. J Hepatol 2006;44: 217-231.
7. Sanyal AJ, Genning C, Reddy KR, Wong F, Kowdley KW, Benner K, et al. The North American study for the treatment of refractory ascites. Gastroenterology 2003;124: 634-641.
8. Biggins SW, Rodriguez HJ, Bacchetti P, Bass NM, Roberts JP, Terrault NA. Serum sodium predicts mortality in patients listed for liver transplantation. Hepatology 2005;41:32-39.
9. Huo TI, Wang YW, Yang YY, Lin HC, Lee PC, Hou MC, 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.
10. Barber KM, Pioli SE, Blackwell JE, Collett D, Neuberger JM, Gimson AE. Development of a UK score for patients with end-stage liver disease [Abstract]. Hepatology 2007; 46(Suppl. 1):510A.
11. Neuberger J, Gimson A, Davies M, Akyol M, O'Grady J, Burroughs A, et al. Selection of patients for liver transplantation and allocation of donated livers in the UK. Gut 2008;57:252-257.
12. Luca A, Angermayr B, Bertolini G, Koenig F, Vizzini G, Ploner 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; 13:1174-1180.
13. Kim WR, Biggins SW, Kremers W, Wiesner R, Kamath PS, Benson JT, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. N Engl J Med 2008;359:1018-1026.
14. Sharma P, Shaubel DE, Sima CS, Merino RM, Lok AS. Re-weighting the model for end-stage liver disease score components. Gastroenterology 2008;135:1575-1581.
15. Huo TI, Lin HC, Wu JC, Lee FY, Hou MC, Lee PC, 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.
16. Huo TI, Lin HC, Huo SC, Lee PC, Wu JC, Lee FY, et al. Comparison of four models for end-stage liver disease-based prognostic system for cirrhosis. Liver Transpl 2008;14:837-844.
17. Jiang M, Liu F, Xiong WJ, Zhong L, Chen XM. Comparison of four models for end-stage liver disease in evaluating the prognosis of cirrhosis. World J Gastroenterol 2008;14:6546-6550.
18. Lucey MR, Brown KA, Everson GT, Fung JJ, Gish R, Keefe EB, et al. Minimal criteria for placement of adults on the liver transplant waiting list: a report of a national conference organized by the American Society of Transplant Physicians and the American Association for the Study of Liver Diseases. Transplantation 1998;66:956-962.
19. Ravaioli M, Masetti M, Dazzi A, Romano A, Spaggiari M, Grazi GL, et al. Model for end-stage liver disease (MELD) system to allocate and share livers: experience of two Italian centers. Transplant Proc 2008;40:1814-1815.
20. Dawwas MF, Gimson AE. Candidate selection and organ allocation in liver transplantation. Semin Liver Dis 2009; 29:40-52.
21. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg 1973;60:646-649.
22. Kuiper JJ, De Man RA, Van Buuren HR. Review article: management of ascites and associated complications in patients with cirrhosis. Aliment Pharmacol Ther 2007; 26(Suppl. 2):183-193.
23. Häussinger D, Schliess F. Pathogenetic mechanisms of hepatic encephalopathy. Gut 2008;57:1156-1165.
24. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982;115:92-106.
25. Hanley JA, McNeil BJ. The meaning and use of area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-843.
27. Merion RM, Schaubel DE, Dykstra DM, Freeman RB, Port FK, Wolfe RA. The survival benefit of liver transplantation. Am J Transplant 2005;5:307-313.
28. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007; 26:2389-2430.
29. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988; 16:1140-1154.
30. R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org. Accessed May 2010.
31. Annual Report of the U.S. Scientific Registry for Organ Transplantation and the Organ Procurement and Transplantation Network. Transplant Data 1990-1999. UNOS, Richmond, VA, and the Division of Transplantation, Bureau of Health Resources and Services Administration, US Department of Health and Human Services, Rockville, MD; 2000.
32. Austin MT, Poulose BK, Ray WA, Arbogast PG, Feurer ID, Pinson CW. Model for end-stage liver disease: did the new liver allocation policy affect waiting list mortality? Arch Surg 2007;142:1079-1085.
33. Kanwal F, Dulai GS, Spiegel BM, Yee HF, Gralnek IM. A comparison of liver transplant outcomes in the pre- vs post-MELD era. Aliment Pharmacol Ther 2005;21: 169-177.
34. Ruf AE, Kremers WK, Chavez LL, Descalzi VI, Podesta LG, Villamil FG. Addition of serum sodium into the MELD score predicts waiting list mortality better than MELD alone. Liver Transpl 2005;11:336-343.
35. Biggins SW, Kim WR, Terrault NA, Saab S, Balan V, Schiano T, et al. Evidence-based incorporation of serum sodium concentration into MELD. Gastroenterology 2006; 130:1652-1660.
36. Cárdenas A, Ginés P. Predicting mortality in cirrhosis–serum sodium helps. N Engl J Med 2008;358:1060-1062.
37. Lv XH, Liu HB, Wang Y, Wang BY, Song M, Sun MJ. Validation of model for end-stage liver disease score to serum sodium ratio index as a prognostic predictor in patients with cirrhosis. J Gastroenterol Hepatol 2009;24: 1547-53.

38. Heuman DM, Abou-assi SG, Habib A, Williams LM, Stravitz RT, Sanyal 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;40:802-810.
39. Ginés P, Berl T, Bernardi M, Bichet DG, Hamon G, Jiménez W, et al. Hyponatremia in cirrhosis: from pathogenesis to treatment. Hepatology 1998;28:851-864.
40. Adam R, McMaster P, O'Grady J, Castaing D, Klempnauer JL, Jamieson N, et al. Evolution of liver transplantation in Europe: report of the European Liver Transplant Registry. Liver Transpl 2003;9:1231-1243.
41. Londoño MC, Cárdenas A, Guevara M, Quintó L, de las Heras D, Navasa M, et al. MELD score and serum sodium in the prediction of survival of patients with cirrhosis awaiting liver transplantation. Gut 2007;56:1283-1290.
42. Silberhumer GR, Hetz H, Rasoul-Rockenschaub S, Peck-Radosavljevic M, Soliman T, Steininger F, et al. Is MELD score sufficient to predict not only death on waiting list, but also post-transplant survival? Transpl Int 2006;19: 275-281.
43. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44: 217-231.
44. Kemmer N, Safdar K, Kaiser TE, Zacharias V, Neff GW. Liver transplantation trends for older recipients: regional and ethnic variations. Transplantation 2008;86:104-107.
45. Levy MF, Somasundar PS, Jennings LW, Jung GJ, Molmenti EP, Fasola CG, et al. The elderly liver transplant recipient: a call for caution. Ann Surg 2001;233:107-113.
46. Yoo HY, Edwin D, Thuluvath PJ. Relationship of the model for end-stage liver disease (MELD) scale to hepatic encephalopathy, as defined by electroencephalography and neuropsychometric testing, and ascites. Am J Gastroenterol 2003;98:1395-1399.
47. Angermayr B, Cejna M, Karnel F, Gschwantler M, Koenig F, Pidlich J, et al. Child-Pugh versus MELD score in predicting survival in patients undergoing transjugular intrahepatic portosystemic shunt. Gut 2003;52:879-885.