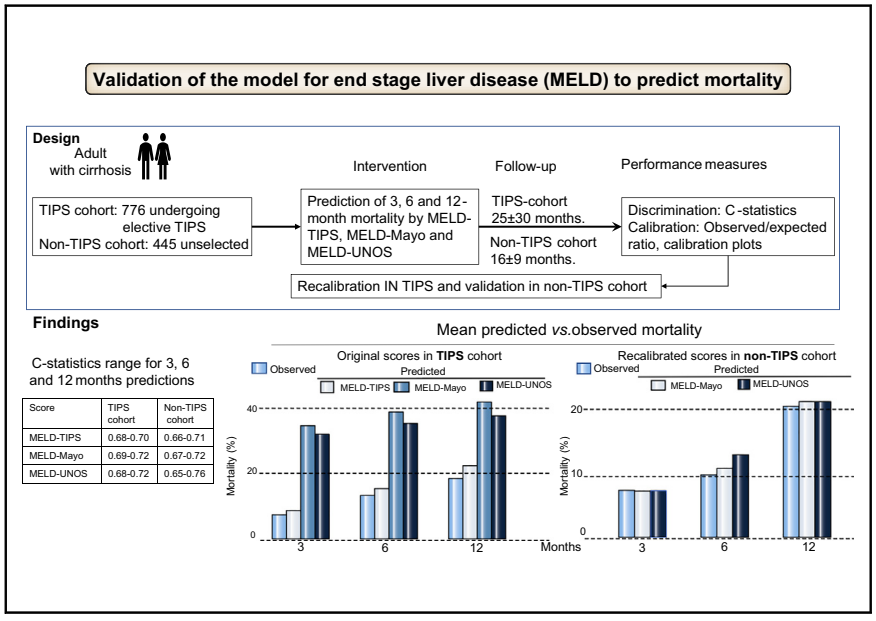
Research Article
Cirrhosis and Liver Failure

JOURNAL
OF HEPATOLOGY

Performance of the model for end-stage liver disease score for mortality prediction and the potential role of etiology

Graphical abstract



Highlights

- Discrimination of MELD is widely reported as fair to good, although its calibration is still unclear.
- In 2 cirrhosis cohorts we found barely acceptable c-statistics, which were significantly worse in patients with non-viral etiology.
- Calibration was largely unsatisfactory with the Mayo and UNOS MELD versions.
- Validated recalibrations of MELD-Mayo and UNOS versions are presented which allow reliable predictions for clinical practice.
- Age, albumin and ascites as the indication for TIPS are candidate variables for an update to the MELD-TIPS score.

Authors

Gennaro D'Amico, Luigi Maruzzelli, Aldo Airoldi, …, Pietro Pozzoni, Agostino Colli, Luca Saverio Belli

Correspondence

gedamico@libero.it (G. D'Amico).

Lay summary

While the discriminative performance of the model for end-stage liver disease (MELD) score is credited to be fair to good, its calibration, the correspondence of observed to predicted mortality, is still unsettled. We found that application of 3 different versions of the MELD in 2 independent cirrhosis cohorts yielded largely imprecise mortality predictions particularly in non-viral cirrhosis. Thus, we propose a recalibration and suggest candidate variables for an update to the model.

https://doi.org/10.1016/j.jhep.2021.07.018

Research Article
Cirrhosis and Liver Failure

JOURNAL OF HEPATOLOGY

Performance of the model for end-stage liver disease score for mortality prediction and the potential role of etiology

Gennaro D'Amico[1,2,*], Luigi Maruzzelli[3], Aldo Airoldi[4], Ioannis Petridis[5], Giulia Tosetti[6], Antonio Rampoldi[7], Mario D'Amico[8,9], Roberto Miraglia[3], Stella De Nicola[4], Vincenzo La Mura[10], Marco Solcia[7], Riccardo Volpes[5], Giovanni Perricone[4], Cristiano Sgrazzutti[11], Angelo Vanzulli[11], Massimo Primignani[6], Angelo Luca[3], Giuseppe Malizia[1], Alessandro Federico[12], Marcello Dallio[12], Angelo Andriulli[13], Angelo Iacobellis[13], Luigi Addario[14], Matteo Garcovich[15], Antonio Gasbarrini[15], Luchino Chessa[16], Francesco Salerno[17], Giulia Gobbo[17], Manuela Merli[18], Lorenzo Ridola[19], Gianluca Svegliati Baroni[20], Giuseppe Tarantino[20], Nicola Caporaso[21], Filomena Morisco[21], Pietro Pozzoni[22], Agostino Colli[22], Luca Saverio Belli[4]

[1]Gatroenterology Unit, Azienda Ospedaliera Ospedali Riuniti Villa Sofia-Cervello, Palermo, Italy; [2]Gastroenterology Unit, Clinica La Maddalena, Palermo, Italy; [3]Radiology Service, Mediterranean Institute for Transplantation and Advanced Specialized Therapies (IRCCS-ISMETT), Palermo, Italy; [4]Hepatology and Gastroenterology Unit, ASST GOM Niguarda, Milan, Italy; [5]Hepatology Unit, Mediterranean Institute for Transplantation and Advanced Specialized Therapies (IRCCS-ISMETT), Palermo, Italy; [6]Gastroenterology and Hepatology Unit Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, CRC "A. M. and A. Migliavacca" Center for Liver Disease, Milan, Italy; [7]Interventional Radiology Unit, ASST GOM Niguarda, Milan, Italy; [8]Radiology Unit, Azienda Ospedaliera Ospedali Riuniti Villa Sofia-Cervello, Palermo, Italy; [9]Radiology Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; [10]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Fondazione Luigi Villa, Milan, Italy; [11]University of Milan, Department of Oncology and Hemato-oncology and Radiology, Department ASST Niguarda, Milan, Italy; [12]Hepato-Gastroenterology Unit, Department of Precision Medicine, University of Campania Luigi Vanvitelli, Naples, Italy; [13]Department of Gastroenterology and Endoscopy, Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy; [14]Hepatology Unit, Cardarelli Hospital, Naples, Italy; [15]Department of Internal Medicine and Gastroenterology, Policlinico Gemelli, Rome, Italy; [16]Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy; [17]Internal Medicine Unit, IRCCS Policlinico San Donato, Milano, Italy; [18]Gastroenterology and Hepatology Unit, Department of Translational and Precision Medicine, Università Sapienza, Roma, Italy; [19]Gastroenterology Unit, ASL Latina, Department of Translational and Precision Medicine, "Sapienza" University of Rome, Italy; [20]Liver Injury and Transplant Unit, Polytechnic University of Marche, Ancona, Italy; [21]Gastroenterology Unit, "Federico II" University, Naples, Italy; [22]General Medicine Unit,Presidio Ospedaliero, Azienda Socio Sanitaria Territoriale di Lecco, Lecco, Italy

Background & Aims: Although the discriminative ability of the model for end-stage liver disease (MELD) score is generally considered acceptable, its calibration is still unclear. In a validation study, we assessed the discriminative performance and calibration of 3 versions of the model: original MELD-TIPS, used to predict survival after transjugular intrahepatic portosystemic shunt (TIPS); classic MELD-Mayo; and MELD-UNOS, used by the United Network for Organ Sharing (UNOS). We also explored recalibrating and updating the model.

Methods: In total, 776 patients who underwent elective TIPS (TIPS cohort) and 445 unselected patients (non-TIPS cohort) were included. Three, 6 and 12-month mortality predictions were calculated by the 3 MELD versions: discrimination was assessed by c-statistics and calibration by comparing deciles of predicted and observed risks. Cox and Fine and Grey models were used for recalibration and prognostic analyses.

Results: In the TIPS/non-TIPS cohorts, the etiology of liver disease was viral in 402/188, alcoholic in 185/130, and non-alcoholic steatohepatitis in 65/33; mean follow-up±SD was 25±9/19±21 months; and the number of deaths at 3-6-12 months was 57-102-142/31-47-99, respectively. C-statistics ranged from 0.66 to 0.72 in TIPS and 0.66 to 0.76 in non-TIPS cohorts across prediction times and scores. A *post hoc* analysis revealed worse c-statistics in non-viral cirrhosis with more pronounced and significant worsening in the non-TIPS cohort. Calibration was acceptable with MELD-TIPS but largely unsatisfactory with MELD-Mayo and -UNOS whose performance improved much after recalibration. A prognostic analysis showed that age, albumin, and TIPS indication might be used to update the MELD.

Conclusions: In this validation study, the performance of the MELD score was largely unsatisfactory, particularly in non-viral cirrhosis. MELD recalibration and candidate variables for an update to the MELD score are proposed.

Lay summary: While the discriminative performance of the model for end-stage liver disease (MELD) score is credited to be fair to good, its calibration, the correspondence of observed to predicted mortality, is still unsettled. We found that application of 3 different versions of the MELD in 2 independent cirrhosis

Keywords: MELD; clinical prediction rule; cirrhosis; TIPS.
Received 17 February 2021; received in revised form 2 July 2021; accepted 16 July 2021; available online 30 July 2021
* Corresponding author. Address: Viale Cavarretta 34, 90151 Palermo, Italy.
E-mail address: gedamico@libero.it (G. D'Amico).
https://doi.org/10.1016/j.jhep.2021.07.018

cohorts yielded largely imprecise mortality predictions particularly in non-viral cirrhosis. Thus, we propose a recalibration and suggest candidate variables for an update to the model.

Introduction

The model for end-stage liver disease (MELD) score is used worldwide to predict the risk of mortality in patients with cirrhosis and to prioritize patients for orthotopic liver transplant (OLT). The original MELD was developed using Cox regression to predict survival after elective transjugular intrahepatic portosystemic shunt (TIPS) in patients with cirrhosis.[1] It included disease etiology, bilirubin, creatinine and international normalized ratio (INR), as predictors. We will refer to this score as the MELD-TIPS. Subsequently, MELD-TIPS was adapted by removing the predictor "etiology" and multiplying the predictors' coefficients by 10.[2] This is the classic MELD and is commonly adopted to predict mortality in a broader range of patients with advanced liver disease. We will refer to this as the MELD-Mayo score.

The MELD-Mayo score was later modified by the United Network for Organ Sharing (UNOS) in 2002 to restrict the range of possible predictions,[3] and in 2016 to account for hyponatremia.[4] This modified score, which we will refer to as the MELD-UNOS, is commonly used for organ allocation priority for OLT.

Therefore, 3 different versions of the MELD have entered clinical practice and online calculators are available for each of them,[5–7] while it is not always clear which one should be used.

Several studies have investigated the performance of the MELD (mostly MELD-Mayo) models, reporting promising discrimination with concordance statistic ranging from 0.66 to 0.83.[8,9] However, some studies have identified unsatisfactory performance in several patient subgroups, which prompted exceptions to the MELD and model revisions.[10,11] Moreover, studies of the correspondence between observed and expected mortality (calibration) at defined observation times are lacking.[9]

Therefore, while the MELD helps physicians in ranking patients according to risk, it is hardly applicable when mortality probability is a key for clinical decisions or simply to inform the patient on his expected survival.

In the present study we assessed the discrimination and calibration performance of mortality predictions by the 3 aforementioned MELD scores in 2 independent cohorts of patients with cirrhosis. We also explored recalibrating and updating the model.

Patients and methods

Study participants

Two independent patient cohorts were included.

TIPS cohort. A total of 776 patients with cirrhosis of any etiology consecutively undergoing elective TIPS for refractory variceal bleeding or refractory ascites from July 1,1999 to May 31, 2020 were included. Since the study was planned in January 2017, 234 patients were included prospectively and 542 retrospectively. Inclusion criteria were the same as in the MELD derivation study.[1] Therefore, patients with other indications (n = 199), including emergency, early or rescue TIPS (n = 83), were not included.

Patients still alive at inclusion gave oral informed consent to participate in the study. Those already deceased had previously given informed consent to use their collected data for clinical research.

Non-TIPS cohort. This cohort was enrolled in a prospective multicenter study of the clinical course of cirrhosis promoted by the Italian Association of the Study of the Liver and designed in 2009. Inclusion criteria were: newly diagnosed cirrhosis from any etiology or first decompensation of cirrhosis. Exclusion criteria were: hepatocellular carcinoma; previously known cirrhosis; previous decompensation; age <18 years. The study was approved by the local Ethics Committee at each participating center. A total of 445 consecutive participants were prospectively included after informed consent, between March 1, 2009 and June 30, 2015 at 11 centers.

For both cohorts, recorded patient information included demographic and clinical data, MELD and Child-Pugh[12] scores at the time of inclusion. Occurrence of any clinical event was recorded during follow-up. Patient records were converted to anonymous files before inclusion in the study dataset.

The study conduct complied with the ethical principles reported in the Declaration of Helsinki.[13] Patient flow across the study phases is shown in Fig. S1.

Follow-up and outcomes

Follow-up was retrospective in 432 patients and prospective in 344 in the TIPS cohort, until October 20, 2020, and was prospective for all the patients in the non-TIPS cohort with study end set at January 2018. In both cohorts, all the included patients underwent scheduled control visits at 6-month intervals or as clinically required up to the study end.

The outcome of interest was any-cause death at 3, 6 and 12 months. When missing, the date of death was ascertained by direct contact with patient relatives or family physician. As in the derivation study,[1] OLT was a censoring event to achieve a comparable outcome estimation. However, since censoring OLT may result in biased death risk estimation we also assessed the cumulative incidence function (CIF) of death with OLT as a competing risk.[14]

Prediction models

The 3 MELD versions were calculated using component variable values obtained at the time of inclusion in the study (Table S1). The 3 linear predictors were calculated according to the published formulas[1,2,4]:

- MELD-TIPS = $0.378 \times \log_e$(bilirubin mg/dl) + $1.120 \times \log_e$(INR) + $0.957 \times \log_e$ (creatinine mg/dl) + $0.643 \times$ (cause of cirrhosis)
- MELD-Mayo = $3.8^{*}\log_e$(serum bilirubin [mg/dl]) + $11.2^{*}\log_e$(INR) + $9.6^{*}\log_e$(serum creatinine [mg/dl]) + 6.4
- MELD-UNOS (i) = [$0.378 \times \log_e$(bilirubin mg/dl) + $1.120 \times \log_e$(INR) + $0.957 \times \log_e$(creatinine mg/dl) + 0.643]x10. Bilirubin, creatinine or INR values <1.0 are set to 1 in this formula and the maximum score is set at 40. Creatinine values >4 mg/dl are set to 4 as well as creatinine values of patients who underwent ≥2 dialysis treatments in the prior 7 days or who received 24 hours of continuous veno-venous hemodialysis in the prior 7 days. According to the UNOS/OPTN policy 9.1[7], in patients with MELD-UNOS(i) >11, the score was recalculated as follows:

MELD-UNOS= MELD-UNOS(i)+ 1.32 * (137-Na) - [0.033*MELD * (137-Na)]. For this calculation sodium values <125 mmol/L were set to 125, and values >137 mmol/L to 137.

Table 1. Major patient characteristics for the derivation and validation samples.

		Validation cohorts in the present study						
		TIPS cohort					Non-TIPS cohort	
Patient characteristics	**Derivation cohort (1)**	**Total**	**Ismett**	**Niguarda**	**Maggiore**	**m#**	**Multicenter**	**m#**
Patients, n	231	776	590	137	49		445	
Follow-up, months*	13.2 (1-45.6)	14.3 (1-170)	13.3 (1-170)	14.8 (1-52)	23.1 (2-95)	8	16.6 (1-67)	0
Age¶	56±12	59± 10	59±10	60± 9	59±9	7	60±11	0
Etiology, n (%)						1		22
Viral	24 (10.4)	402 (51.8)	337 (57.1)	46 (33.6)	19 (38.8)	–	188 (42.3)	–
Alcohol	142 (61.9)	185 (23.8)	108 (18.4)	57 (41.6)	20 (40.8)	–	130 (29.2)	–
NASH	NR	65 (8.4)	46 (7.8)	15 (10.9)	4 (8.2)	–	33 (7.4)	–
Cholestatic	23 (10.0)	18 (2.3)	11 (1.9)	4 (2.9)	3 (6.1)	–	12 (2.7)	–
Other or mixed	41 (17.8)	106 (13.7)	88 (14.9)	15 (10.9)	3 (6.1)	–	72 (16.2)	–
Ascites %	183 (79.4)	629 (84.6)	520 (88.2)	72 (63.1)	37 (75.5)	24	261 (58.6)	–
Hepatic encephalopathy %	139 (60.1)	224 (31.8)	168 (29.7)	49 (56.3)	6 (12.2)	72	60 (5.8)	
Albumin (g/dl)¶	2.7±0.6	3.1±0.6	2.9±0.5	3.4±0.5	3.8±0.5	49	3.1±0.7	3
Bilirubin (mg/dl)¶	3.9±4.6	1.7±1.2	1.8±1.3	1.3±0.8	1.3±1.0	0	2.8±4.3	2
INR¶	1.6±0.7	1.3±0.2	1.3±0.2	1.3±0.2	1.3±0.2	0	1.4±0.4	
Creatinine (mg/dl)¶	1.4±1.2	1.1±0.5	1.1±0.5	0.9±0.4	1.0±0.3	0	0.9±0.5	11
Sodium (mEq/L)¶	NR	136.2±4.8	135.6±4.7	138.4±4.2	137.2±4.2	7	137±7.9	26
Indication for TIPS								
Refractory bleeding	58 (25)	233 (30)	171 (29)	50 (36)	12 (25)	–	–	–
Refractory ascites	173 (75)	452 (58)	355 (60)	70 (51)	27 (55)	–	–	–
Bleeding and ascites	NR	90 (12)	63 (11)	17 (13)	10 (20)	–	–	–
PPG pre-TIPS, mmHg¶	23.5±8.2	18.0±5.4	17.6±5.2	18.0±5.2	23.0±5.8	7	–	–
PPG post-TIPS, mmHg¶	11.2±4.5	7.8±3.6	7.2±3.1	10.5±4.0	7.9±3.6	7	–	–
PPG % reduction post-TIPS, %¶		56	59	40	65	12	–	–
Pugh score (N observations)	9.8±2.1	8.2±1.6	8.4±1.5	7.9±1.6	6.9±1.1	35	7.5±2.1	2
Pugh A/B/C %	8/36/56	14/65/21	11/65/24	19/69/12	29/69/2	35	36/45/19	2
MELD score¶,‡								
MELD-TIPS	1.127±1.02‡	0.861±0.6‡	0.953±0.6	0.561±0.6	0.588±0.6	1	0.70±0.7	12
MELD-Mayo	–	10.269±5.1‡	10.810±5.1	8.395±4.4	8.999±4.2	0	9.96±6.4	11
MELD-UNOS	–	13.365±5.0‡	13.888 ± 5.2	11.683±3.8	11.667±3.6	4	13.121±6.6	12
Deaths, n (%)	110 (47)	274 (35.6)	230 (39.0)	31 (23.7)	13 (26.5)	6	288	0
3 months	70 (30)	57 (7.4)	49 (8.3)	8 (6.2)	0 (0)	–	31 (7.0)	–
6 months	89 (38)†	102 (13.3)	84 (14.2)	15 (11.6)	3 (6.1)	–	47 (10.6)	–
12 months	102 (44)†	142 (18.5)	115 (19.5)	22 (17,1)	5 (10.2)	–	99 (22.3)	–
OLT, n (%)	28 (12.1)	143 (18.7)	120 (20.4)	12 (9.5)	11 (22.5)	0	13 (2.9)	–
Kaplan-Meier survival⊣						8		0
3 months	0.707⊣	0.926⊣	0.917	0.928	0.979	–	0.930	–
6 months	0.621⊣	0.866⊣	0.856	0.876	0.936	–	0.894	–
12 months	0.551⊣	0.801⊣	0.788	0.813	0.883	–	0.773	–
CIF of death⊣⊢						8		0
3 months	–	0.071	0.079	0.063	0.021	–	0.070	–
6 months	–	0.128	0.136	0.112	0.06 3	–	0.106	–
12 months	–	0.188	0.199	0.168	0.109	–	0.214	–

(*continued on next page*)

JOURNAL OF HEPATOLOGY

Table 1. (*continued*)

Patient characteristics	Derivation cohort (1)	Validation cohorts in the present study							
		TIPS cohort						Non-TIPS cohort	
		Total	Ismett	Niguarda	Maggiore	m#	Multicenter	m#
CIF of OLT‡⊢						0		0
3 months	–	0.035	0.041	0.023	0.021	–	0.011	–
6 months	–	0.074	0.079	0.048	0.085	–	0.016	–
12 months	–	0.115	0.115	0.079	0.131	–	0.023	–

CIF, cumulative incidence function; INR, international normalized ratio; MELD, model for end-stage liver disease; NASH, non-alcoholic steatohepatitis; OLT, orthotopic liver transplant; PPG, portal pressure gradient; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.
PPG % reduction post-TIPS is calculated as: [(PPG pre – PPG post)/PPG pre]*100.
#Number of patients with missing information.
*median (range).
¶mean ± standard deviation. When not available for the derivation sample, standard deviations were derived from the reported mean values by normal distribution.
†Since number of deaths observed at 6 and 12 months in the derivation study (1) were not reported, this data is derived by the underlying risk reported in Table 5 of ref 1.
‡R_0 is the mean value of the score used for calculations of expected survival probability.
†Kaplan-Meier survival probability, which is the underlying survival probability (S_0) at the times of interest used for calculations of expected survival probability.
⊢Cumulative incidence by competing risks analysis with death and OLT as competing events.

Prediction time

Time zero was the date of TIPS placement or of inclusion for the non-TIPS cohort and times of death prediction were 3, 6 and 12 months.

Outcome prediction

Individual patient survival probability was calculated according to the Cox model,[15] as reported by Malinchoc[1]:

$S(t)=S_0(t)^{exp(R - R_0)}$ where:

$S(t)$ is the probability of survival at each of the times of interest; R is the score value in the individual patient in the present cohorts; R_0 is the score of the average patient in the derivation study,[1] 1.127; $S_0(t)$ is the underlying survival probability for an average patient undergoing elective TIPS in the derivation study (ref 1: Table 5): S_0 *(3 months)* = 0.707; S_0*(6 months)* = 0.621; S_0*(12 months)* = 0.551. Mortality prediction at the relevant times was calculated as 1-$S(t)$.

However, for comparison, we also calculated the individual patient outcome probability by using S_0 and R_0 from the present cohorts. For MELD-Mayo and MELD-UNOS we used S_0 and R_0 only from our study cohorts because no corresponding data from derivation studies are available. Moreover, to account for any potential bias derived from censoring OLT, we also calculated mortality predictions by the competing risks analysis, according to the formula:

$CIF(t) = 1-(1- CIF_0(t))^{exp(R - R 0)}$, where CIF(t) is the expected probability of death at time (t), and $CIF_0(t)$ is the baseline cumulative incidence of mortality.[16]

Statistical analysis

Case mix analysis was based on the MELD distribution and on the membership analysis[17] (supplementary material). An explorative analysis by the Cox model[15] was also performed to assess the prognostic value of the individual MELD components in the validation cohorts.

Overall MELD performance (discrimination and calibration) was assessed by the Nagelkerke's R^2 and by the rescaled Brier score.[18]

Discrimination has been assessed by the c-statistics[19] and by the Yates slope,[18] the difference between the mean death risk predicted in patients who remained alive and in those who died.

Calibration has been assessed by plots showing the relationship of the mean predicted death probability *vs*. the mean observed mortality in deciles of patients with increasing values of the predicted probability. Plots were drawn by a *loess smoother* algorithm,[18] to allow more insight in calibration analysis. Differences between predicted probability and observed mortality were assessed by the Hosmer-Lemeshow test.[20]

Calibration–in–the-large (Observed/Expected ratio, O/E) and calibration slope were assessed by logistic regression models and differences from their ideal values (0 and 1, respectively) were tested by the Wald test.[18]

In a *post hoc* analysis of potential factors influencing MELD performance, we assessed c-statistics and calibration-in-the-large according to the type of received stent, period of TIPS placement, type of indication for TIPS and etiology.

For the MELD versions with predictions beyond the 95% CI of observed rates in >5 deciles, a model recalibration was

Table 2. Performance of the 3 assessed scores.

Performance measure	MELD-TIPS ‡	MELD-TIPS ¶	MELD-Mayo	MELD-UNOS
TIPS cohort				
Patients, n		767	768	765
Score, median (range)		0.874 (-1.558 to 2.896)	10.2 (-9.2 to 29.0)	11.96 (6.43-30.01)
3-month prediction				
R^2 Nagelkerke, %		9.2	11.2	11.0
Brier scaled %		5.4	6.4	6.1
c-statistics		0.70 (0.62–0.78)	0.72 (0.65–0.80)	0.72 (0.64–0.79)
Discrimination slope (95% CI)	0.11 (0.08–0.15)	0.05 (0.03–0.06)	0.32 (0.21–0.44)	0.34 (0.23–0.45)
O/E ratio (OLT censored)#	0.21 (0.14–0.31)	1.316 (0.484–3.582)	0.062 (0.043–0.089)	0.072 (0.050–0.103)
O/E ratio (OLT competing)*		1.400 (0.503–3.899)	0.060 (0.042–0.087)	0.071 (0.049–0.101)
6–month prediction				
R^2 Nagelkerke, %		10.4	11.7	10.5
Brier scaled %		7.0	7.5	6.6
c-statistics		0.70 (0.64–0.75)	0.71 (0.66–0.77)	0.70 (0.64–0.75)
Discrimination slope (95% CI)	0.12 (0.09–0.15)	0.07 (0.05–0.09)	0.30 (0.21–0.39)	0.30 (0.21–0.39)
O/E ratio (OLT censored)#	0.26 (0.21–0.34)	0.986 (0.560–1.734)	0.119 (0.091–0.156)	0.153 (0.118–0.199)
O/E ratio (OLT competing)*		1.057 (0.589–1.890)	0.120 (0.092–0.157)	0.154 (0.118–0.200)
12–month prediction				
R^2 Nagelkerke, %		9.3	10.5	10.3
Brier scaled %		6.8	7.2	7.0
c-statistics		0.68 (0.63–0.73)	0.69 (0.64–0.74)	0.68 (0.63–0.73)
Discrimination slope (95% CI)	0.11 (0.08–0.14)	0.08 (0.06–0.10)	0.25 (0.17–0.33)	0.28 (0.20–0.36)
O/E ratio (OLT censored)#	0.30 (0.24–0.36)	0.714 (0.495–1.030)	0.180 (0.144–0.228)	0.211 (0.167–0.266)
O/E ratio (OLT competing)*	—	0.774 (0.525–1.139)	0.181 (0.144–0.228)	0.215 (0.170–0.272)
Non-TIPS cohort				
Patients, n		433	434	433
Score, median (range)		0.68 (-1.01 to 3.85)	9.16 (–3.76 to 38.51)	10.79 (3.56 to 39.02)
3–month prediction				
R^2 Nagelkerke, %		8.9	9.5	11.9
Brier scaled %		6.5	6.4	6.7
c-statistics		0.71 (0.61–0.81)	0.72 (0.63–0.82)	0.76 (0.67–0.85)
Discrimination slope (95% CI)	0.13 (0.07–0.18)	0.06 (0.03–0.09)	0.30 (0.14–0.45)	0.41 (0.26–0.56)
O/E ratio (OLT censored)#	0.14 (0.09–0.23)	0.51 (0.19–1.35)	0.07 (0.04–0.12)	0.08 (0.05–0.14)
O/E ratio (OLT competing)*	—	0.52 (0.20–1.38)	0.08 (0.05–0.13)	0.09 (0.05–0.15)
6–month prediction				
R^2 Nagelkerke, %		8.2	9.3	10.0
Brier scaled %		9.1	9.1	9.3
c-statistics		0.69 (0.61–0.78)	0.72 (0.64–0.80)	0.72 (0.64–0.80)
Discrimination slope (95% CI)	0.12 (0.07–0.18)	0.068 (0.035–0.102)	0.30 (0.17–0.43)	0.35 (0.22–0.48)
O/E ratio (OLT censored)#	0.16 (0.11–0.23)	0.44 (0.22–0.89)	0.10 (0.07–0.16)	0.11 (0.07–0.17)
O/E ratio (OLT competing)*	—	0.44 (0.22–0.88)	0.10 (0.06–0.15)	0.11 (0.07–0.017)
12–month prediction				
R^2 Nagelkerke, %		9.1	9.1	6.7
Brier scaled %		17	17	17
c-statistics		0.66 (0.59–0.72)	0.67 (0.61–0.73)	0.65 (0.59–0.72)
Discrimination slope (95% CI)	0.12 (0.08–0.16)	0.10 (0.07–0.14)	0.27 (0.18–0.37)	0.25 (0.16–0.35)
O/E ratio (OLT censored)#	0.38 (0.29–0.49)	0.55 (0.39–0.79)	0.22 (0.16–0.30)	0.31 (0.22–0.43)
O/E ratio (OLT competing)*	—	0.59 (0.40–0.85)	0.23 (0.17–0.31)	0.32 (0.23–0.44)

O/E, observed/expected; MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.
‡S_0 and R_0 from the derivation study.
¶S_0 and R_0 from the validation study.
#O/E ratio = calibration-in-the-large computed with underlying survival function obtained by censoring OLT.
*O/E ratio = calibration-in-the-large computed with underlying survival obtained considering OLT as a competing event with death.

performed by a proportional hazards model for competing risks[21] with MELD as the only covariate, mortality as the outcome of interest and OLT as a competing risk. The recalibration coefficient was derived in the TIPS cohort and validated in the non-TIPS cohort. Predicted mortality was estimated according to the proportional hazard model for competing risks as follows[16]:

$CIF_{(t)} = 1\text{-}(1\text{-}\ CIF_{0(t)})^{exp(MELD_r\ -\ MELD_{r0})}$, where $CIF_{(t)}$ is the expected probability of death at time (t), $CIF_{0(t)}$ is the baseline cumulative incidence of mortality, $MELD_r$ is the recalibrated MELD and $MELD_{r0}$ is the mean recalibrated MELD in the TIPS cohort or, respectively, in the non-TIPS validation cohort.

To explore the potential of updating the score, a multivariable analysis by the Fine and Grey model[21] was performed in the TIPS cohort by including the following variables together with MELD-TIPS: sex, age, serum-albumin, serum-sodium, ascites, hepatic encephalopathy, previous variceal bleeding, portal pressure gradient (PPG) before TIPS placement, PPG after

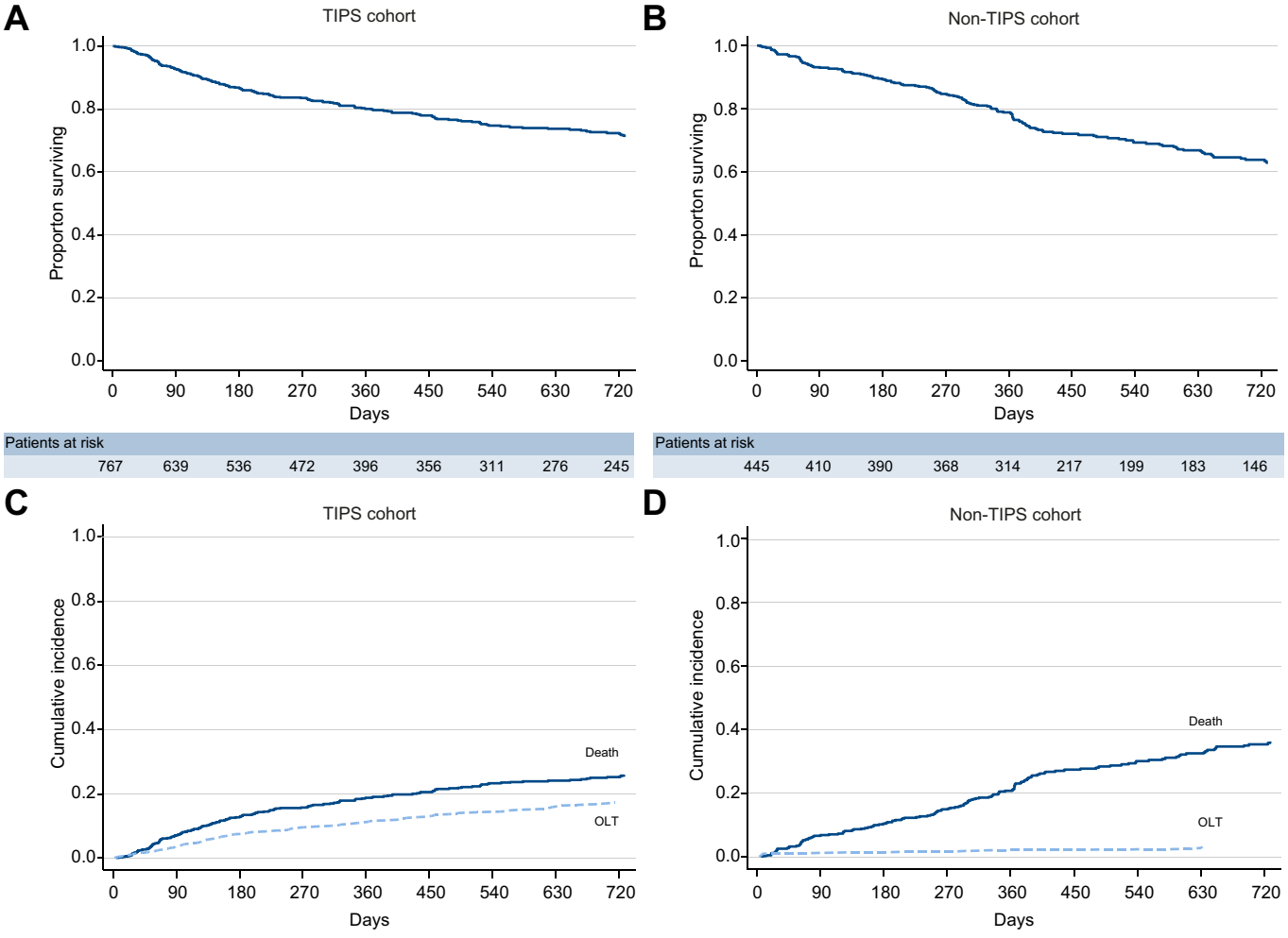


Fig. 1. Survival analysis. (A and B) Kaplan-Meier plots of survival analysis with OLT censored; (C and D) competing risks plots of cumulative incidences of death and OLT. Numbers between upper and lower panels are patients at risk. OLT, orthotopic liver transplant.

TIPS, post-TIPS %PPG reduction, bleeding as indication for TIPS, ascites as indication for TIPS, and type of TIPS (covered/uncovered). Quantitative variables were transformed to their natural logarithm to lessen the influence of extreme laboratory values, where appropriate. Single components of the MELD, as well as the Child-Pugh score, were not included to avoid redundancy.

All the performance and prognostic analyses were based on complete cases because missing data were very rare (Table 1).

Results

Case mix analysis showed significant differences between the derivation and the 2 validation cohorts (supplementary information), indicating suitability of this study for a score generalizability assessment.

Kaplan-Meier survival plots (censoring OLT) and the CIF of death and OLT (by competing risk analysis) are reported in Fig. 1, with 3-, 6- and 12-month values reported in Table 1. OLT was significantly more frequent in the TIPS cohort (enrolled at transplant centers) than in the non-TIPS cohort (143/776 *vs.* 13/432; p <0.0001).

Cox model analysis showed that among MELD variables, $\log_e$creatinine, $\log_e$ bilirubin and $\log_e$INR, but not etiology, were significantly associated with the risk of death in the TIPS cohort, while only $\log_e$bilirubin was significant in the non-TIPS cohort. Of note, MELD component coefficients were appreciably and variously different in the 2 validation cohorts compared with derivation cohort.[1] The Fine-Gray model showed only $\log_e$bilirubin being significant in both cohorts and $\log_e$creatinine in the TIPS cohort. Details are provided in the supplementary information.

Performance of the MELD scores

TIPS cohort

The median MELD-TIPS[1] in the 767 patients included in this analysis was 0.874 (range -1.558 to 2.896). C-statistics (95% CI) for 3-, 6- and 12-month mortality were 0.70 (0.62–0.78), 0.70 (0.64–0.75) and 0.68 (0.63–0.73), respectively. Other performance estimations (Table 2) including calibration-in-the-large were far from satisfactory. Calibration plots showed largely overestimated mortality probability when using S_0 and R_0 from the derivation study across all the assessed prediction times. However, when using S_0 and R_0 from the present cohorts, calibration improved considerably (Fig. 2A,B and Figs S3-4; Table S8).

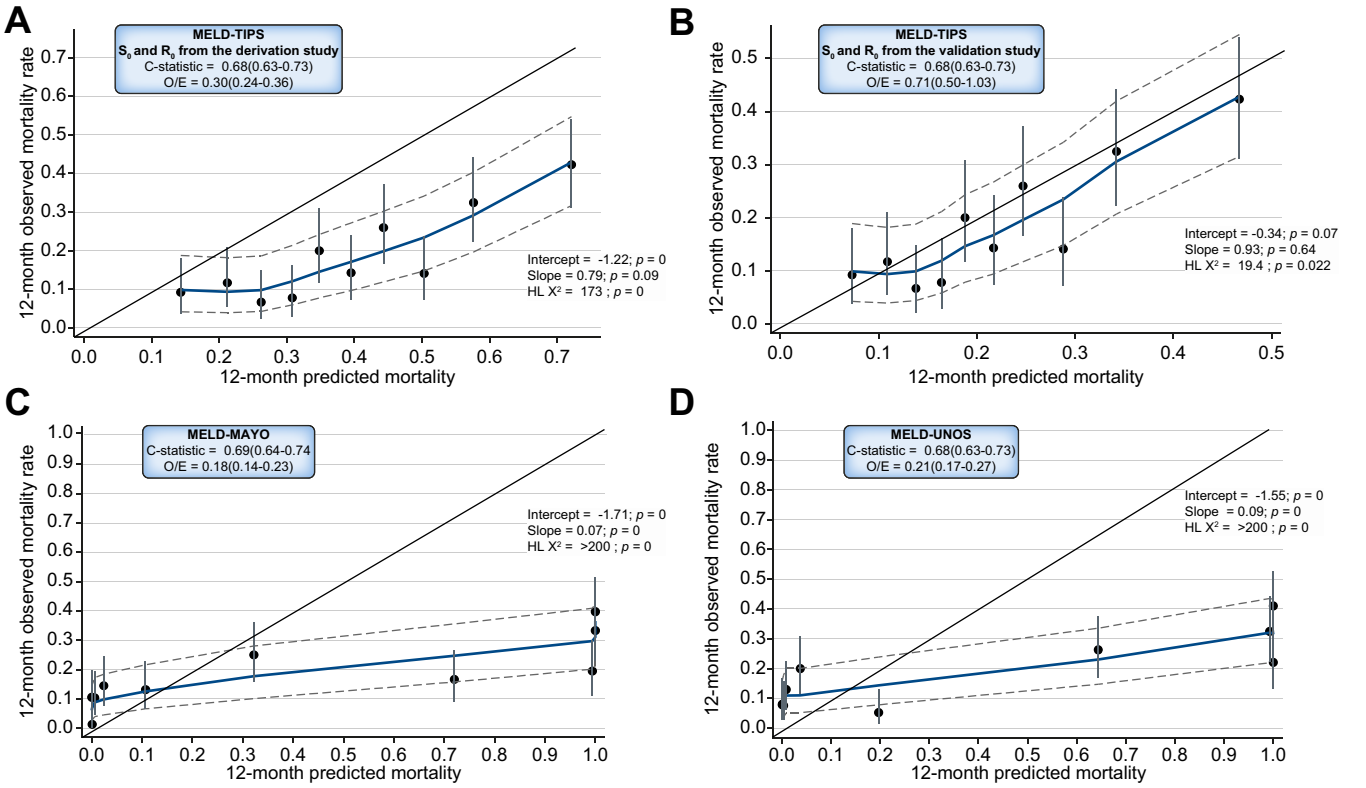


Fig. 2. Calibration plots for 12-month mortality prediction in TIPS cohort. Smoothed (loess) calibration plots with 95% confidence bounds (dashed lines) in deciles of patients ordered according to increasing predicted probability of death by the assessed scores. (A) MELD-TIPS with S_0 and R_0 from the derivation study; (B) MELD-TIPS, (C) MELD-Mayo, (D) MELD-UNOS; S_0 and R_0 from our TIPS cohort in (B-D). Vertical bars indicate the 95% CIs of observed mortality rates; the diagonal lines indicate the ideal line of perfect correspondence of predicted to observed mortality. *p* values for intercept and slope are from Wald test. HL, Hosmer-Lemeshow test; MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.

MELD-Mayo score[2] was assessed in 768 patients. The median score was 10.2 (range from -9.2 to 29.0). C-statistics (95% CI) for 3-, 6- and 12-month mortality were 0.72 (0.65–0.80), 0.71 (0.66–0.77), 0.69 (0.64–0.74), respectively, and other performance measures were mostly unsatisfactory (Table 2) except for calibration slope. Calibration plots showed extreme over- and underestimation of expected mortality (Fig. 2C and Figs S3-4; Table S9).

MELD-UNOS[4] performance was assessed in 765 patients. The median score was 11.96 (range 6.43–30.01). C-statistics (95% CI) for 3-, 6- and 12-month mortality were 0.72 (0.64–0.79), 0.70 (0.64–0.75) and 0.68 (0.63–0.73), respectively, slightly better than with the MELD-TIPS (Table 2). Other performance measures were generally unsatisfactory except for calibration slope. Mortality predictions were largely mis-calibrated (Fig. 2D and Figs S3-4; Table S10), with the best O/E ratio (95% CI) being 0.21 (0.17–0.27) across all the prediction times (Table 2).

Expected mortality computed with OLT as a competing event did not change appreciably (Tables S8-10). Therefore, calibration plots (not shown) were almost overlapping with those shown in Fig. 2 and Figs S3-4; calibration-in-the-large is shown in Table 2.

Explorative analyses of factors potentially influencing MELD performance are shown in Table 3. No significant influence was found for covered or uncovered stents, time period of TIPS placement, type of indication for TIPS and etiology. With MELD-Mayo and MELD-UNOS, calibration-in-the-large was better (Table S13) but still largely unsatisfactory for 6-month survival prediction when the indication for TIPS was refractory ascites rather than bleeding. Of note, an important worsening of c-statistics was consistently observed in the last 5 years and in patients with non-viral etiology (Table 3). This worsening seemingly parallels the reduction of hepatitis B or C from 65% to 35% and the increase of alcohol and non-alcoholic steatohepatitis (NASH) from 18% to 49% in patients enrolled in this period compared to before (p <0.0001). C-statistics were almost always lower in *non-viral* than in *viral* etiology, although not significantly, and were always ≤0.70 for non-viral etiology (Table 3); a similar trend was observed for O/E ratio (Table S13).

Non-TIPS cohort

Overall, there were 433 patients with complete data for performance assessment of MELD-TIPS and MELD-UNOS and 434 for MELD-Mayo (Table 2). Median (and range) score values were: MELD-TIPS 0.68 (-1.01 to 3.85), MELD-Mayo 9.16 (-3.76 to 38.51) and MELD-UNOS 10.79 (3.56 to 39.02). C-statistics ranged from 0.65 (0.59–0.72) to 0.76 (0.67–0.85) across the different prediction times and scores.

With OLT censored, calibration-in-the-large ranged from 0.07 (0.04–0.12) to 0.55 (0.39–0.79) and the other performance measures were almost all unsatisfactory (Table 2). Calibration plots are shown in Fig. 3 and Figs S5-6, with corresponding data in Tables S10-11. When OLT was considered a competing event,

Table 3. C-statistics for the 3 assessed scores in patient subgroups according to type of TIPS, date of placement, type of indication to TIPS and viral etiology.

Patient group*	Prediction time	MELD-TIPS (n = 767)	MELD-Mayo (n = 768)	MELD-UNOS (n = 765)
			C-statistics (95% CI)	
Type of stent				
Uncovered	3 months	0.66 (0.51–0.81)	0.69 (0.53–0.84)	0.80 (0.66–0.93)
Covered		0.70 (0.62–0.79)	0.73 (0.64–0.81)	0.70 (0.62–0.79)
Uncovered	6 months	0.72 (0.58–0.85)	0.73 (0.60–0.87)	0.78 (0.66–0.91)
Covered		0.70 (0.64–0.76)	0.71 (0.66–0.77)	0.69 (0.63–0.65)
Uncovered	12 months	0.65 (0.52–0.79)	0.64 (0.50–0.78)	0.66 (0.51–0.81)
Covered		0.68 (0.63–0.74)	0.70 (0.64–0.75)	0.69 (0.64–0.74)
TIPS date				
Before 2009	3 months	0.69 (0.57–0.81)	0.69 (0.56–0.82)	0.69 (0.57–0.82)
2009-2015		0.73 (0.62–0.85)	0.75 (0.63–0.86)	0.72 (0.61–0.84)
From 2016 on		0.57 (0.37–0.78)	0.68 (0.53–0.83)	0.69 (0.53–0.85)
Before 2009	6 months	0.69 (0.60–0.78)	0.71 (0.61–0.80)	0.70 (0.62–0.79)
2009-2015		0.76 (0.66–0.85)	0.75 (0.66–0.85)	0.73 (0.63–0.82)
From 2016 on		0.62 (0.51–0.74)	0.66 (0.57–0.76)	0.64 (0.54–0.74)
Before 2009	12 months	0.69 (0.62–0.77)	0.69 (0.71–0.77)	0.68 (0.60–0.76)
2009-2015		0.69 (0.60–0.78)	0.70 (0.62–0.79)	0.70 (0.62–0.78)
From 2016 on		0.61 (0.51–0.71)	0.65 (0.55–0.74)	0.63 (0.54–0.72)
Type of indication				
Bleeding	3 months	0.61 (0.27–0.94)	0.69 (0.38–1.0)	0.74 (0.45–1.0)
Ascites		0.67 (0.58–0.75)	0.68 (0.60–0.77)	0.65 (0.57–0.74)
Ascites+bleeding		0.75 (0.47–1.0)	0.73 (0.44–1.0)	0.79 (0.48–1.0)
Bleeding	6 months	0.63 (0.38–0.87)	0.66 (0.42–0.89)	0.72 (0.52–0.91)
Ascites		0.67 (0.60–0.73)	0.68 (0.61–0.74)	0.63 (0.57–0.70)
Ascites+bleeding		0.72 (0.52–0.92)	0.71 (0.52–0.91)	0.75 (0.55–0.93)
Bleeding	12 months	0.64 (0.49–0.80)	0.69 (0.55–0.83)	0.69 (0.54–0.83)
Ascites		0.64 (0.58–0.70)	0.64 (0.58–0–70)	0.63 (0.57–0.68)
Ascites+bleeding		0.73 (0.58–0.89)	0.71 (0.56–0.86)	0.73 (0.56–0.89)
Etiology				
Viral	3 months	0.74 (0.66–0.83)	0.74 (0.66–0.83)	0.72 (0.63–0.81)
Non-viral		0.63 (0.50–0.76)	0.67 (0.54–0.81)	0.70 (0.56–0.83)
Viral	6 months	0.75 (0.69–0.82)	0.75 (0.69–0.82)	0.72 (0.65–0.78)
Non-viral		0.64 (0.55–0.73)	0.66 (0.57–0.75)	0.66 (0.57–0.76)
Viral	12 months	0.70 (0.64–0.77)	0.70 (0.64–0.77)	0.67 (0.61–0.74)
Non-viral		0.64 (0.56–0.72)	0.66 (0.58–0.74)	0.69 (0.61–0.76)

MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.
*Number of patients in the shown analyses were as follows: uncovered stent n = 101, covered stent n = 675; TIPS before 2009, n = 252; from 2009 to 2015, n = 219; from 2016 on, n = 305; viral etiology, n = 402; non-viral etiology n = 374.

calibration did not change appreciably because only 13 patients were transplanted in this cohort (Table S10); calibration-in-the-large is reported in Table 2 (calibration plots, almost coincident with those with OLT censored, are not shown).

The reduction of discrimination performance in non-viral etiology was confirmed in the non-TIPS cohort and was significant at 12 months with all 3 scores (Fig. 4; Table S14); a similar trend was observed for O/E with MELD-Mayo and MELD-UNOS.

MELD recalibration and exploratory updating

We performed recalibration for MELD-Mayo and MELD-UNOS scores (details in the supplementary information) but considered that the calibration performance of the MELD-TIPS score was acceptable, with mortality predictions always within the 95% CI of observed mortality rates, with S_0 and R_0 from validation cohorts. The Fine and Grey model[21] in the TIPS cohort yielded the following formulas: recalibrated MELD-Mayo = 0.0745*(MELD-Mayo); recalibrated MELD-UNOS = 0.0716*(MELD-UNOS). Recalibration appreciably improved the performance of both scores in the TIPS cohort and even more in the non-TIPS cohort, which was used as an external independent validation cohort (Fig. 5 and Figs S7-8; Table S15).

The prognostic analysis aimed at updating the MELD-TIPS[1] score showed that age, albumin and ascites as the indication for TIPS were significant, together with MELD-TIPS (Table S16). The updated model was: *(0.383*MELD-TIPS) + (0.037*age) + (-0.451*albumin) + (0.744 if indication for TIPS was ascites).* The c-statistics (95% CI) were 0.72 (0.66–0.79) for 3-month, 0.73 (0.68–0.78) for 6-month and 0.70 (0.66–0.75) for 12-month survival prediction and were not significantly different from the original model. Corresponding values for calibration-in-the-large (95% CI) were 1.25 (0.46–3.37) for 3-month, 1.16 (0.66–2.01) for 6-month and 0.83 (0.57–1.19) for 12-month survival predictions, showing a consistent improvement of calibration compared to the original model. Calibration plots are shown in Fig. S9. Independent validation is needed to assess the performance and applicability of this explorative update to MELD-TIPS.

Discussion

A major result of this study is that in 2 independent cohorts of patients with cirrhosis, MELD performance was globally unsatisfactory either in terms of discrimination or in terms of calibration. Moreover, importantly, discrimination decreased over time, parallel to the relative reduction of viral and increase of alcoholic and NASH etiologies. The worst calibration

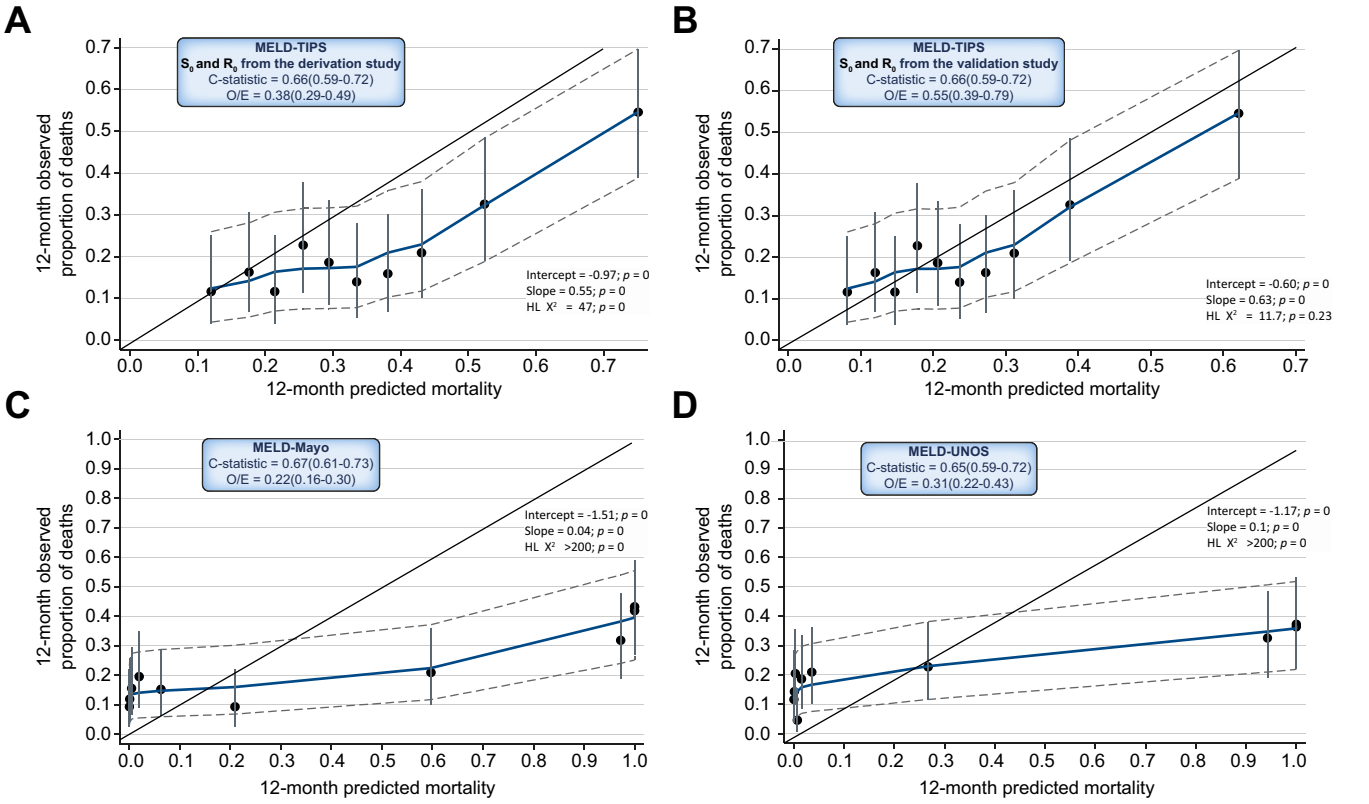


Fig. 3. Calibration plots for 12-month mortality prediction in non-TIPS cohort. Smoothed (loess) calibration plots with 95% confidence bounds (dashed lines) in deciles of patients ordered according to increasing predicted probability of death by the assessed scores. (A) MELD-TIPS with S_0 and R_0 from the derivation study; (B) MELD-TIPS, (C) MELD-Mayo, (D) MELD-UNOS; S_0 and R_0 from our non-TIPS cohort in (B-D). Vertical bars indicate the 95% CIs of observed mortality rates; the diagonal lines indicate the ideal line of perfect correspondence of predicted to observed mortality. *p* values for intercept and slope are from Wald test. HL, Hosmer-Lemeshow test; MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.

performance was found for the Mayo and UNOS versions of MELD. These results were almost overlapping in the 2 cohorts which were independently recruited and followed-up at different centers and by different physicians. Recalibration of these 2 scores enabled us to satisfactorily re-align predicted and observed mortality with both scores.

The interpretation of these results is that the MELD may not be used in populations with very different case mix distribution compared to the derivation study,[1] like the 2 included in the present study. Therefore, its use as a survival predictor seems not to be as generalizable as suggested by the study proposing the MELD-Mayo score.[2]

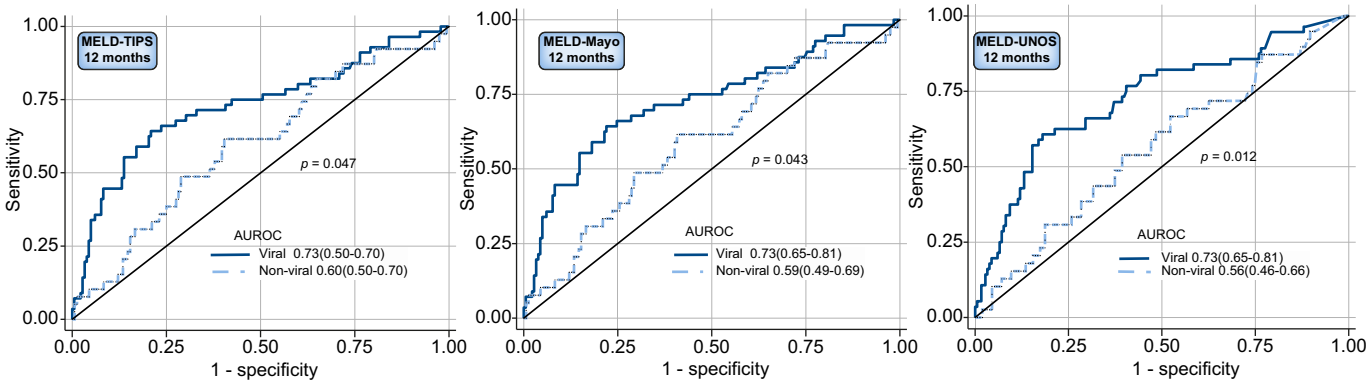


Fig. 4. MELD discrimination performance according to etiology of cirrhosis in non-TIPS cohort. Receiver-operating characteristics curves for 12-month survival prediction by the 3 assessed scores for patients with viral and non-viral etiology. Differences between curves were assessed by the DeLong test. MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt.

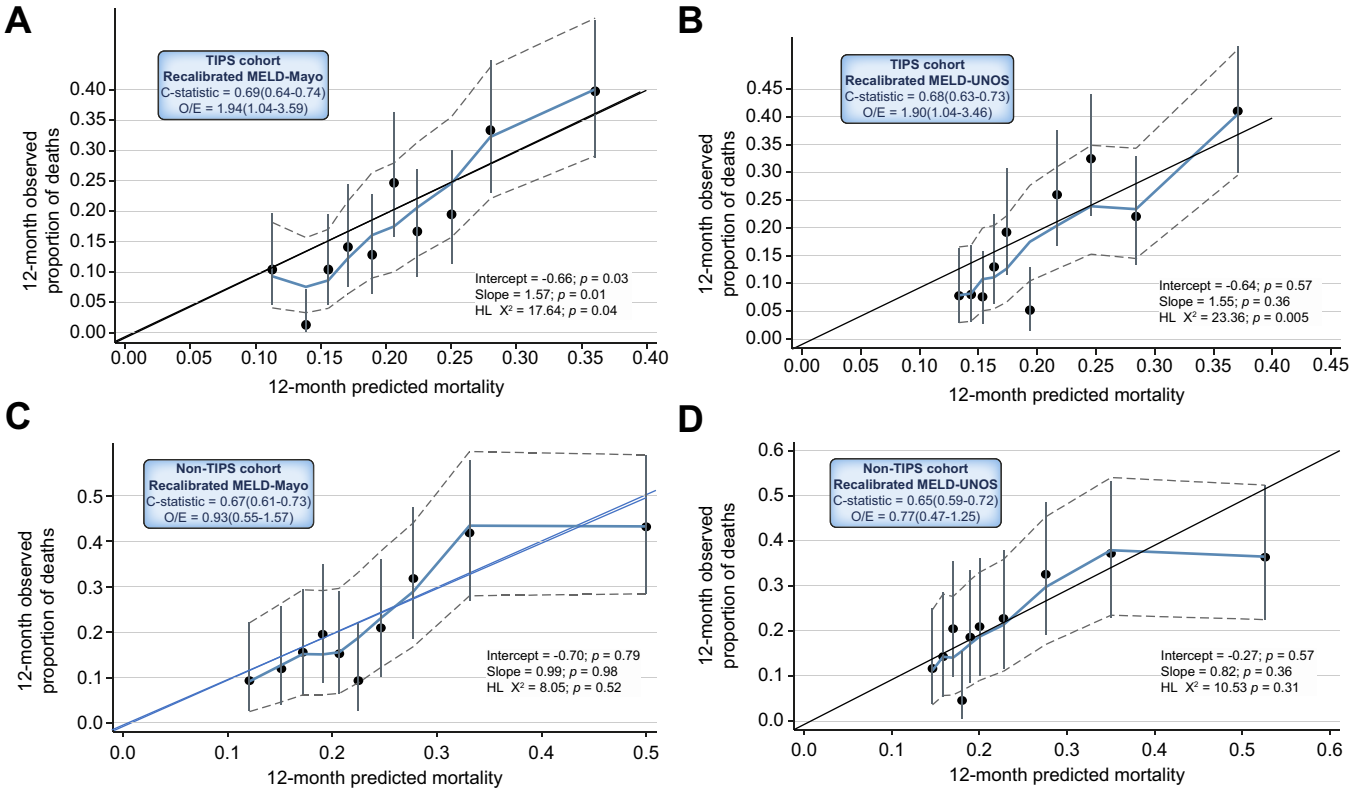


Fig. 5. Recalibration plots for 12-month mortality prediction. Smoothed (loess) calibration plots with 95% confidence bounds (dashed lines) in deciles of patients ordered according to increasing predicted probability of death by the recalibrated MELD-Mayo and MELD-UNOS scores. (A,B) TIPS cohort; (C,D) non-TIPS validation cohort. Vertical bars indicate the 95% CIs of observed mortality rates; the diagonal lines indicate the ideal line of perfect correspondence of predicted to observed mortality. MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.

On the other hand, the different patient case mix in our TIPS cohort compared to the derivation study is explained by the modification of patient selection criteria for TIPS over time, together with the use of covered stents, while the difference for the non-TIPS cohort is likely explained by the unselected admission in contrast to the derivation study where only patients with refractory ascites or bleeding were included.

In the present study, calibration was particularly poor with MELD-Mayo and MELD-UNOS, while it was still acceptable for MELD-TIPS if underlying survival (S_0) and mean score (R_0) from the present cohorts were used. Reasons for the large miscalibration with these 2 versions of the score are hard to detect and may lie in the score modifications without recalibration. It is notable, in this respect, that the MELD-Mayo validation was presented only in terms of discrimination in the proposing study[2] and subsequent calibration studies are scarce and show inconsistent results.[22–24]

In our study, the role of etiology on MELD performance is supported by temporal analysis showing that both discrimination and calibration of the score worsened in the last 5 years, parallel to a significant reduction of viral and increase of alcohol and NASH etiologies. Importantly this result was even more marked and statistically significant in the non-TIPS cohort whose recruitment started approximately 10 years later than in the TIPS cohort.

It is therefore likely that removing etiology in the MELD-Mayo score[2] without recalibrating the coefficients of the other component predictors may have contributed to worsening the model's calibration performance.

Moreover, censoring OLT in the MELD derivation study[1] instead of considering it as a competing event might have contributed to miscalibration by overestimating the risk of death. For this reason, we used a competing risks approach to recalibrate the Mayo and UNOS versions of the MELD in the TIPS cohort with substantial improvement of performance confirmed in the non-TIPS cohort.

A relevant issue raised from our study and also related to case mix, concerns the use of Cox model-based prognostic scores in clinical practice. This requires knowing the mean survival probability (S_0) and the mean score value (R_0) in the target population. However, physicians usually do not have these parameters to hand for their own patient population and use the parameters reported in the model derivation study. A typical example of this is the use the Mayo-Clinic calculator[5] or the corresponding nomogram[1] to predict mortality following TIPS. Both the web calculator and the nomogram are based on S_0 and R_0 from the derivation study.[1] However, our study shows how much predicted probabilities can deviate from observed outcomes when the underlying risk and score distribution of the target population are so different from the corresponding parameters in the derivation study. This finding calls for caution in using such prediction tools when the underlying risk and predictor distribution are not accounted for. In fact, the use of S_0 and R_0 from our cohorts resulted in acceptable calibration of the

MELD-TIPS with predictions always in the 95% CI boundaries of the observed rates.

Obviously, the proposed recalibrations for MELD-Mayo and UNOS, do not overcome the problem of S_0 and R_0, which should be derived from the target population whenever possible, but they allow for more reliable mortality predictions. For post-TIPS survival prediction, we have also proposed an updated MELD-TIPS score based on age, albumin and TIPS indication, together with a recalibrated MELD. Although it requires full external validation, the model is promising with good calibration and acceptable discrimination. Importantly, both albumin and age were also significant in a recently reported study proposing a new prognostic score for patient selection for TIPS, the FIPS score.[25]

Limitations of the present study are the partially retrospective patient enrollment and follow-up in the TIPS cohort. However, we are confident that the risk of bias is minimized by consecutive patient inclusion and the prospective nature of data collection even for patients observed before beginning the study. In fact, the very low number of missing values allowed for complete case analysis, avoiding data imputation. Moreover, the results in the TIPS cohort were fully replicated in the independent non-TIPS prospective cohort. A second limitation may be that the separate analyses of the influence of the indication for TIPS, placement date, type of stent and etiology of cirrhosis were planned after observing the unexpectedly low calibration performance of the MELD-Mayo and -UNOS versions of the score. The rationale for analyzing TIPS-related factors was based on the known technical improvements in TIPS over time, together with changes in indications and superiority of covered stents, which almost completely replaced uncovered stents in the early 2000s. The rationale for assessing the influence of etiology was the progressive reduction of viral and corresponding increase of non-viral etiologies of cirrhosis in the last years. Although we did not find any statistically significant difference between c-statistics for viral *vs*. non-viral etiology of cirrhosis in the TIPS cohort, we found that both discrimination and calibration-in-the-large of the 3 versions of the MELD were systematically lower for non-viral than viral etiologies. However, the effect of etiology on MELD performance was even more important and statistically significant in the non-TIPS cohort. This finding strengthens our conclusion that the Mayo and UNOS versions of the MELD are not so broadly generalizable as previously suggested, particularly in the face of the progressive change in cirrhosis etiology.

In conclusion, the present study provides evidence of largely unsatisfactory performance of MELD-Mayo and MELD-UNOS scores to predict mortality either in patients undergoing TIPS or in unselected patients, particularly in those with a non-viral etiology. Performance of MELD-TIPS is acceptable if underlying survival (S_0) and mean score value (R_0) from the target population are accounted for. A recalibration of both the MELD-Mayo and MELD-UNOS scores is proposed for when clinical decision making is based on the expected probability of death, or for patient prognostic information, until a valid MELD update or a new prognostic score become available.

Abbreviations

CIF, cumulative incidence function; MELD, model for end-stage liver disease; NASH, non-alcoholic steatohepatitis; O/E, observed/expected ratio; OLT, orthotopic liver transplant; OPTN, organ procurement and transplantation network; TIPS, transjugular intrahepatic portosystemic shunt; UNOS, United Network for Organ Sharing.

Financial support

The authors received no financial support to produce this manuscript.

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Gennaro D'Amico: study concept and design, analysis and interpretation of data, drafting of the manuscript; study supervision. Luigi Maruzzelli: protocol revision, TIPS placement, data collection. Aldo Airoldi: protocol revision, patient follow-up. Ioannis Petridis: patient follow-up. Giulia Tosetti: data collection, patient follow-up. Antonio Rampoldi: TIPS placement. Mario D'Amico: TIPS placement, data collection. Roberto Miraglia: protocol revision, TIPS placement, results interpretation, critical revision of the manuscript for important intellectual content. Stella De Nicola: data collection, patients follow-up. Vincenzo La Mura: data collection, patients follow-up, manuscript revision. Marco Solcia: TIPS placement. Riccardo Volpes: protocol revision, patients follow-up, results interpretation. Giovanni Perricone: data collection, patients follow-up. Angelo Vanzulli: protocol revision, results interpretation. Cristiano Sgrazzutti: TIPS placement. Massimo Primignani: patients follow-up, manuscript revision. Angelo Luca: protocol revision, results interpretation. Giuseppe Malizia: protocol revision, results interpretation, critical revision of the manuscript for important intellectual content. Alessandro Federico data collection, patients follow-up. Angelo Andriulli protocol revision. Angelo Iacobellis protocol revision, data collection, patients follow-up. Luigi Addario protocol revision, data collection, patients follow-up. Antonio Gasbarrini protocol revision, results interpretation. Matteo Garcovich data collection, patient follow-up. Luchino Chessa: protocol revision, data collection, patients follow-up. Francesco Salerno: protocol revision, results interpretation. Giulia Gobbo: protocol revision, data collection, patients follow-up. Manuela Merli: protocol revision, results interpretation. Lorenzo Ridola: protocol revision, data collection, patients follow-up, results interpretation. Gianluca Svegliati Baroni: protocol revision, results interpretation. Giuseppe Tarantino: data collection, patients follow-up. Nicola Caporaso: protocol revision. Filomena Morisco: data collection, patients follow-up. Pietro Pozzoni data collection, patients follow-up. Agostino Colli: protocol revision, results interpretation, critical revision of the manuscript for important intellectual content. Luca Saverio Belli: protocol revision, results interpretation, critical revision of the manuscript for important intellectual content.

Data availability statement

Data supporting the results of the present study are available upon request.

Acknowledgement

We are most grateful to Dr. Thomas Debray for his important conceptual contribute and thorough manuscript revision; to Dr. Davide Bernasconi for conceptual contribute to the application of the Fine and Gray proportional model for competing risks and for manuscript revision.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.07.018.

References

[1] 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;31:864–871.
[2] Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, et al. A model to predict survival in patients with end-stage liver disease. Hepatology 2001;33:464–470.
[3] Freeman RB, Wiesner RH, Harper A, McDiarmid SV, Lake J, Edwards E. The new liver allocation system: moving towards evidence based transplantation policy. Liver Transplant 2002;8:851–858.
[4] OPTN policy 9. https://optn.transplant.hrsa.gov/media/1200/optn_policies.pdf. Page 167. Accessed on October 22nd 2020.
[5] https://www.mayoclinic.org/medical-professionals/transplant-medicine/calculators/probability-of-mortality-following-transjugular-intrahepatic-portosystemic-shunts/itt-20434720 (accessed on june the 12th 2020).
[6] https://www.mayoclinic.org/medical-professionals/transplant-medicine/calculators/meld-model/itt-20434705 (accessed on june the 12th 2020).
[7] https://optn.transplant.hrsa.gov/resources/allocation-calculators/meld-calculator/(accessed on june the 12th 2020).
[8] Kamath PS, Kim WR. The model for end-stage liver disease (MELD). Hepatology 2007;45:797–805.
[9] Cholangitas E, Marelli L, Shusang V, Senzolo M, Rolles K, Patch D, et al. A systematic review of the performance of the Model for End-Stage Liver Disease (MELD) in the setting of liver transplantation. Liv Transpl 2006;12:1049–1061.
[10] Freeman Jr RB, Gish GR, Harper A, Davis GL, Vierling J, Lieblein L, et al. Model for end-stage liver disease (MELD) exception guidelines: results and recommendations from the MELD exception study group and conference (MESSAGE) for the approval of patients who need liver transplantation with diseases not considered by the standard MELD formula. Liver Transplant 2006;12:S128–S136.
[11] 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.
[12] Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the esophagus for bleeding oesophageal varices. Br J Surg 1973;60:646–649.
[13] World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310(20):2191–2194. https://doi.org/10.1001/jama.2013.281053.
[14] Pintile M. Competing risks. A practical perspective. Chichester: John Wiley & Sons; 2006.
[15] Cox DR. Regression models and life-Tables (with discussion). J R Statist Soc B 1972;34:187–220.
[16] Austin P, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med 2017;36. 4991-4400.
[17] Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015;68(3):279–289. https://doi.org/10.1016/j.jclinepi.2014.06.018.
[18] Steyerberger EW. Clinical prediction models. A practical approach to development, validation and updating. 2nd ed. © Springer Nature Switzerland AG; 2019.
[19] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
[20] Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 1997;16(9):965–980.
[21] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of competing risk. J Am Stat Ass 1999;94:496–509.
[22] Finkenstedt A, Dorn L, Edlinger M, Prokop W, Risch L, Griesmacher A, et al. Cystatin C is a strong predictor of survival in patients with cirrhosis: is a cystatin C-based MELD better? Liver Int 2012 Sep;32(8):1211–1216.
[23] Vilar Gomez E, Bertot LC, Oramas BG, Soler EA, Navarro RL, Elias JD, et al. Application of a biochemical and clinical model to predict individual survival in patients with end-stage liver disease. World J Gastroenterol 2009 Jun 14;15(22):2768–2777.
[24] Renfrew PD, Quan H, Doig CJ, Dixon E, Molinari M. The Model for End-stage Liver Disease accurately predicts 90-day liver transplant wait-list mortality in Atlantic Canada. Can J Gastroenterol 2011 Jul;25(7):359–364.
[25] Bettinger D, Sturm L, Plaff L, Hahn F, Kloekner R, Volkwein L, et al. Refining prediction of survival after TIPS with the novel Freiburg index of post-TTIPS survival. J Hepatol 2021;74:1362–1372. https://doi.org/10.1016/j.jhep.2021.05.042.