TRANSPLANT INTERNATIONAL

Transplant International ISSN 0934-0874

ORIGINAL ARTICLE

Validation of a dropout assessment model of candidates with/without hepatocellular carcinoma on a common liver transplant waiting list

Christian Toso,[1,2] Pietro Majno,[1,2] Thierry Berney,[1,2] Philippe Morel,[1,2] Gilles Mentha[1,2] and Christophe Combescure[3]

1 Divisions of Transplant and Abdominal Surgery, Department of Surgery, University of Geneva Hospitals, Geneva, Switzerland
2 Hepato-Pancreato-Biliary Centre, University of Geneva Hospitals, Geneva, Switzerland
3 Division of Clinical Epidemiology, University of Geneva Hospitals, Geneva, Switzerland

Keywords
allocation, deMELD, MELD, score, validation.

Correspondence
Christian Toso, MD, PhD, Department of Surgery, Rue Gabrielle-Perret-Gentil, 1211 Geneva, Switzerland.
Tel.: 41-22-3723311;
Fax: 41-22-3727755;
e-mail: christian.toso@hcuge.ch

Conflicts of interest
The authors have declared no conflict of interests.

Received: 24 September 2013
Revision requested: 7 November 2013
Accepted: 16 March 2014
Published online: 25 April 2014

doi:10.1111/tri.12323

Summary

The model of end-stage liver disease (MELD) score is often used for liver graft allocation, and patients with hepatocellular carcinoma (HCC) receive exception points (22 in the US). A better model is desirable for patients with HCC as they tend to have a privileged access to transplantation, without taking HCC characteristics into account. A new simpler model designed from a training set of US patients (n = 49 026) was tested on two validation sets (US and UK patient cohorts with, respectively, n = 20 475 and n = 1781). The risk of dropout was between 3.2 and 7.8% at 3 months in patients with HCC, and was captured into a score, including HCC size, HCC number, AFP, and MELD (−37.8 + 1.9*MELD+5.9 if HCC Nb ≥ 2 + 5.9 if AFP > 400 + 21.2 if HCC size > 1 cm). This new model could be validated on external US and UK liver candidate cohorts. It provides a dynamic and more accurate assessment of dropout than the use of exception MELD (C-indices of 66.2–73.7% vs. 52.7–56.6%). In addition, the model shows a similar distribution as MELD for patients with non-HCC, and both scores could be used in parallel for the management of waiting-list patients with and without HCC.

Introduction

Liver transplantation is the best treatment for patients with end-stage liver failure and/or early nonresectable hepatocellular carcinoma (HCC). In most institutions, these two groups of patients are on a common liver transplant waiting list, competing for the same grafts.

The model of end-stage liver disease (MELD) score is based on the serum levels of creatinine, bilirubin, and the international normalized ratio (INR). It predicts the risk of death (or dropout from the list) at 3 months and is often used to prioritize liver candidates, transplanting the patient with the highest score first [1,2]. While most patients with HCC have a preserved liver function and a low MELD score, they receive artificial "exception" points to bring them higher on the list and offer them the chance of a transplant. To illustrate, US candidates with HCCs within Milan criteria (excluding those with a single HCC ≤2 cm) receive 22 points at the time of listing.

A number of challenges are linked to this type of allocation. Despite the good intention, the system as presently applied does not allow an equitable graft allocation between patients with and without cancer, as HCC candidates have been consistently shown to have easier access to transplantation (and lower rates of dropout) [3–5]. In addition, the system does not allow for an equitable graft distribution among patients with HCC themselves: all HCC patients receive the same number of points, but those with larger/more aggressive lesions have higher risks of HCC progression and dropout from the list, and response rates to waitlist local HCC treatment vary from one patient to the other [6].

In an effort to devise a more satisfactory distribution model, we developed a dropout equivalent MELD (deMELD)[5]. This model calculated an individualized estimate of the probability of dropout assessment for HCC candidates, similar to the MELD score for the patients with non-HCC. The key advantage of this score over previously proposed models was its scale, which was similar to the MELD one [3]. To illustrate, a deMELD of 15 and a MELD of 15 provided similar risks of dropout, respectively, in patients with HCC and non-HCC, thus allowing an equitable management of patients with both HCC and non-HCC on a common waiting list.

However, the design and the use of the proposed deMELD model were linked to a number of issues: (i) the inclusion of recipient age and liver disease in the deMELD was bringing forward a unique situation for patients with HCC, which was not matching the common (MELD) practice in patients without HCC (where MELD is used alone). This was a concern as both scores should be used in parallel for patients with HCC and non-HCC on a common waiting list; (ii) Candidate age and liver disease were not predicting the risk of dropout at 3 months [5]. Their inclusion into the model was therefore debatable; (iii) The use of variables such as candidate age and liver diagnosis also introduced ethical concerns, as older patients and patients with NASH had a higher hazard ratio for dropout and were favored over younger patients with HBV-induced, hemochromatosis or cholestatic liver disease. Such a model appeared debatable as favoring such patients was associated with an expected decreased transplant benefit. (iv) Finally, the proposed deMELD model has been designed only on the United States of America (US) population, without external validation.

The aim of the present study was to develop a more clinically relevant new model based only on HCC-related variables and MELD for comparing the transplant opportunities of patients with both HCC and non-HCC on a common waiting list, with external (more recent US and UK data) validations.

Patients and methods

Patients

The study was based on three patient populations. The training set included patients from the US Scientific Registry of Transplant Recipients (SRTR) listed for liver transplantation between January 1, 2004 and December 31, 2009, which was similar to the sample used previously [5]. The validation sets were from the US (SRTR, January 1, 2009–December 31, 2011) and the UK (NHSBT, July 1, 2008–September 30, 2011). To work on comparable samples of patients with HCC and non-HCC, only patients $\geq$45 years were included (patients with HCC were overall older than non-HCC ones, and <5% of them were removed based on this strategy). Patients with HCC were selected as having a diagnosis of "hepatocellular carcinoma" or "hepatoma."

The SRTR data system includes data on all donors, wait-listed candidates, and transplant recipients in the US, submitted by the members of the Organ Procurement and Transplantation Network (OPTN), and has been described elsewhere [7]. The Health Resources and Services Administration (HRSA), US Department of Health, and Human Services provide oversight to the activities of the OPTN and SRTR contractors.

Outcomes

The model was designed and validated according to the risk of dropout at 3 months. Dropout was linked to poor medical condition ("too sick for transplantation" in the US or "condition deteriorated/patient unfit/medical contra-indication" in the UK), HCC progression beyond transplant criteria or death. Of note, only the overall rate of dropout was taken into account to reflect the natural wait-list behavior of patients with and without HCC, and to allow a comparison between both groups.

Based on the OPTN/United Network of Organ Sharing (UNOS) rules, transplantation was only possible within Milan criteria, and a progression beyond Milan should exclude the patient from transplantation (a few regions are more liberal, accepting downstaged patients or selected patients beyond Milan, i.e., Region 5). In the UK, listing (and transplantation) was possible for patients with a single HCC $\leq$5 cm, with up to five HCCs each $\leq$3 cm or with a single HCCs >5 cm and $\leq$7 cm without evidence of progression (volume increase <20%) and without new nodule formation for more than 6 months. The detection of an extrahepatic HCC spread or of macroscopic vascular invasion was considered as a contra-indication in both countries, as was tumor rupture and AFP >10 000 ng/ml in the UK.

Potential predictors of dropout in HCC and non-HCC candidates

Studied listing variables included calculated MELD, number of tumors, maximum HCC size in cm, and alpha-fetoprotein level in ng/ml (AFP, analyzed as LnAFP). Of note, non-HCC-related variables, such as recipient age and type of liver disease, were not included as these variables are not taken into account for the wait-list management of non-HCC patients (which is only based on MELD). Continuous variables were used for all analyses.

Model predicting the risk of dropout in patients with HCC

A model predicting the risk of dropout in HCC candidates was developed on the US training set with the predictors of

dropout previously listed and following a methodology previously described [5]. Briefly, the association between HCC-related variables and the risk of dropout at 3 months was assessed with a multivariate proportional hazard competitive risk model [8]. This model was chosen because of the presence of these two competing outcomes (dropout and transplantation). Some variables (including MELD) had the potential to impact both on the risk of dropout and on the chance of being transplanted, and a classical multivariate Cox analysis would not have been appropriate.

As for classical Cox model, the model assumed that hazards were proportional and that, for continuous predictors, the increase in the risk of dropout was the same when the predictor was increased by one point whatever the reference level. The first assumption was assessed graphically by representing the complementary log-log of survival against time, and the second assumption was checked by assessing the hazard ratio for categories of the predictor and interpreting the increase of the HRs over categories. When the risk of dropout did not increase regularly with categories, the predictor was used in a categorized way selecting clinically relevant cutoffs. This assessment was specifically performed in the US training set in view of the design of the score.

The tested HCC-related variables and MELD (linear combination weighted by the regression parameters) were integrated in a model (MHCC) predicting the risk of dropout.

Derivation of a score predicting the dropout for patients with HCC

The purpose of the score was to determine a new value of MELD which integrates HCC-related factors and which predicts the same level of risk of dropout than a patient with non-HCC having the same value of MELD. This score has been called new dropout equivalent MELD (new deMELD). The new deMELD could be used for HCC candidates together with the calculated MELD for the patients with non-HCC, all patients being on the same waiting list. The methodology to establish the new deMELD has been described previously [5]. Briefly, the risk of dropout in HCC was modeled according to the MHCC: after applying a logit transformation on the probability (P) of dropout ($\log(P/(1-P))$), the relationship was linear and a linear regression model was performed. A similar regression model was performed in patients with non-HCC to model the risk of dropout according to MELD. A correspondence was further established between the model developed for patients with HCC and the one based on MELD for patients with non-HCC to design the new deMELD model. Analyses were performed using R for Windows (Version 2.12.0), with the packages "etm" version 0.5-2 (empirical transition matrix) [9] for the nonparametric analyses and "mstate" [8] for the competitive risk regression models.

Exception MELD

The currently used exception MELD points were calculated according to the OPTN/UNOS rules. Stage 2 patients with single HCC >2 cm and ≤5 cm, or with two or three HCCs each ≤3 cm, received 22 exception MELD points. Stage 1 patients (single HCC ≤2 cm) and patients with calculated MELDs higher than 22 were listed according to their natural calculated MELDs. Of note, exception MELD points are not in use nationally in the UK (each center manages its list according to a center allocation policy), but were used in the present study to further test the validity of the proposed new deMELD.

External validations of the model

The multivariate proportional hazard competitive risk model was tested. The risk of dropout was assessed in each of the three sets according to categories of MELD, new deMELD, exception MELD, and mixed new deMELD. The mixed new deMELD included the highest value between new deMELD and MELD for each individual patient, capturing the highest risk of dropout linked to liver failure or HCC. The ability of each score to predict dropout at 3 months was assessed by the C-index using the method proposed by Wolbers *et al.* [10]. C-index is commonly used for the assessment of the prognostic performance of models. The method used in the present study is best suited for the assessment of models with competitive events. A proportional subdistribution hazards regression model was fitted [11] with deMELD (or MELD or exception MELD) as predictor using the function crr of the R package cmprk, and the function C-index of the R package pec was applied to the fitted model. Overall, the C-index provides an estimate of the accuracy of the model in predicting dropout at 3 months, 100% reflecting an ideal model, and 50% a noninformative model. 95% confidence intervals were obtained by a bootstrap procedure with 5000 replications.

Other statistical assessments

Further statistical assessments included the Student t- and chi-square tests to study the demographic variables and the mean scores between groups. Results were shown as mean ±standard deviation. A standard alpha of 0.05 was set for statistical significance.

Results

Demographics

Overall, 49 026 patients with and without HCC were included in the US training set, 20 475 in the US and 1781 in the UK validation sets (Table 1). The proportions of candidates with HCC were 11.2% in the US training set, 20.4% in the US validation set, and 27.5% in the UK validation set, likely reflecting the increased proportion of transplantations for HCC in the most recent years [12]. Patient ages were similar between sets, with consistently younger candidates in the non-HCC groups (by 1.4 to 1.8 years in average). The proportions of female transplant candidates were also similar between population sets, with more females in the non-HCC groups. More patients had alcohol-related liver diseases in the UK, and patients with HCC were more often infected with HCV. As anticipated, HCC patients had lower calculated MELD scores than patients with non-HCC (by 5.2–6.7 points in average).

Hepatocellular carcinoma characteristics remained similar between sets, with limited number of HCCs (1.4–1.5 in average) and small lesions (2.7–2.9 in average). Listing AFP averaged between 130.8 and 312.2 ng/ml with wide variations between patients (large standard deviations) [13]. Among patients with documented data, 2801 of 4498 (62.3%) and 2822 of 3450 (81.8%) received at least one wait-list local HCC treatment in the US training and validation sets (such data were not available for the UK set).

Predictors of dropout in patients with HCC

Most transplant candidates either had dropped or were transplanted within the first 3 months after listing (Fig. 1). The dropout rates at 3 months were 8.7, 10.3, and 8.1% for non-HCC US training, US validation and UK validation candidates. They were, respectively, 3.2, 5.1, and 7.8% for HCC candidates. The chance of being transplanted was lower for patients with HCC in the US validation group.

In the HCC groups, the risk of dropout was independently predicted by MELD in the three population sets with similar HRs of 1.14–1.19 per MELD unit (multivariate competitive risk model, Table 2). Regarding HCC-related variables, HCC size was associated with a higher risk of dropout. The effect linked to the number of HCC was only observed in the HCC training group. Of note, the assumption of proportionality of hazards was respected.

In the non-HCC groups, MELD consistently predicted the risk of dropout at 3 months with HRs between 1.17 and 1.19 (Table 2).

Model predicting the risk of dropout in patients with HCC

Clinically and/or statistically significant variables were combined to predict the risk of dropout among HCC candidates in the US training set. Variables included MELD, number of HCCs, maximum HCC size, and alpha-fetoprotein level. The resulting MHCC model weighted by the regression parameters was as follows:

Table 1. Demographics.

	Training set US		Validation set US		Validation set UK	
	No HCC (*N* = 43528)	HCC (*N* = 5498)	No HCC (*N* = 16291)	HCC (*N* = 4184)	No HCC (*N* = 1292)	HCC (*N* = 489)
Age (years)	56.3 (6.5)	58.0 (6.6)	57.1 (6.4)	58.9 (6.1)	56.5 (6.6)	57.9 (6.1)
Women	15373 (35.3%)	1145 (20.8%)	5949 (36.5%)	931 (22.3%)	492 (38.1%)	97 (19.8%)
Diagnosis						
HCV	17108 (39.3%)	2911 (68.9%)	6223 (38.2%)	2503 (59.8%)	213 (16.5%)	220 (45.0%)
HBV	1093 (2.5%)	168 (4.0%)	326 (2.0%)	235 (5.6%)	16 (1.2%)	55 (11.2%)
Alcohol	7980 (18.3%)	458 (10.8%)	3335 (20.5%)	285 (6.8%)	449 (34.8%)	100 (20.4%)
NASH	2916 (6.7%)	169 (4.0%)	2018 (12.4%)	201 (4.8%)	112 (8.7%)	30 (6.1%)
Hemoch.	331 (0.8%)	264 (6.2%)	113 (0.7%)	23 (0.5%)	7 (0.5%)	14 (2.9%)
Other	14100 (32.4%)	258 (6.1%)	4276 (26.2%)	937 (22.4%)	495 (38.3%)	70 (14.3%)
MELD	17.7 (8.3)	12.5 (6.2)	18.7 (8.6)	12.0 (5.8)	16.2 (5.3)	10.9 (3.9)
Number of HCCs						
Two or more		1414 (32.8%)		1020 (29.6%)		165 (33.7%)
md		1186		734		0
Tumor size						
>1 cm		3973 (92.2%)		3400 (98.9%)		471 (96.3%)
md		1187		734		0
Alpha-fetoprotein						
>400		452 (10.7%)		198 (5.7%)		31 (6.3%)
md		1254		731		0

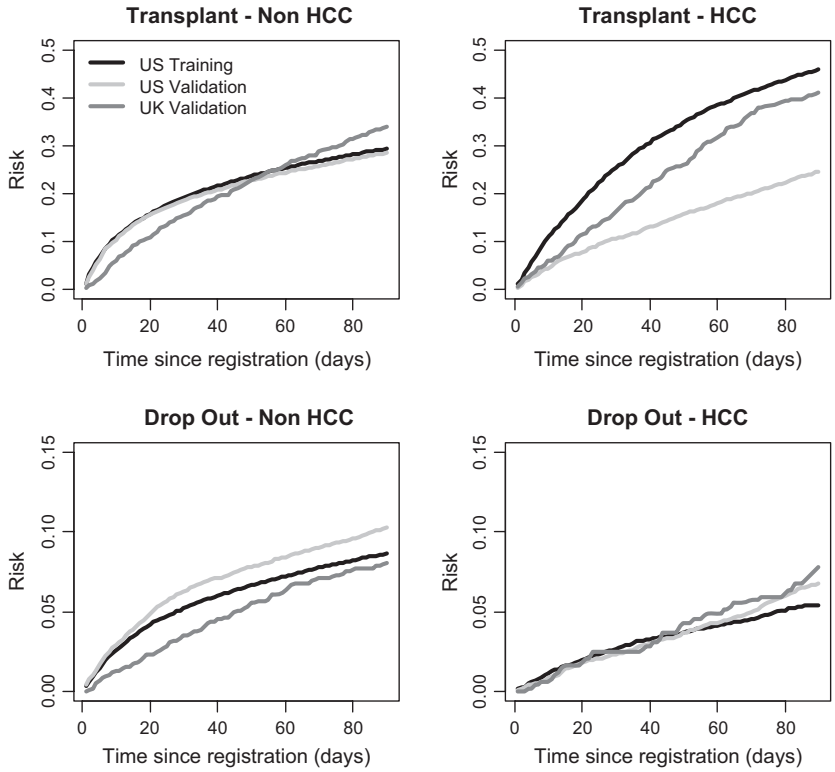

Figure 1 Time-dependent probability of dropout and of transplantation in the HCC and non-HCC groups for the US training, US validation, and UK validation sets.

Table 2. Competitive risk assessment of variables predicting dropout 3 months after listing in the HCC groups.

	Training set US		Validation set US		Validation set UK	
	Dropout		Dropout		Dropout	
HCC patients	HR [95% CI]	*P*	HR [95% CI]	*P*	HR [95% CI]	*P*
MELD						
per unit	1.19 [1.16;1.23]	<0.001	1.14 [1.12; 1.16]	<0.001	1.16 [1.09; 1.24]	<0.001
Number of HCCs						
1	Reference		Reference		Reference	
≥2	1.71 [1.21; 2.41]	0.002	1.01 [0.82; 1.24]	0.94	0.95 [0.48; 1.87]	0.88
HCC size						
≤1 cm	Reference		Reference		Reference	
>1 cm	6.84 [1.69; 27.69]	0.007	2.03 [0.76; 5.43]	0.16	*	
Alpha-fetoprotein						
≤400	Reference		Reference		Reference	
>400	1.70 [1.21; 2.41]	0.03	2.75 [2.08; 3.63]	<0.001	1.13 [0.26; 4.80]	0.87
	Dropout		Dropout		Dropout	
Patients with non-HCC	HR [CI 95%]	*P*	HR [CI 95%]	*P*	HR [CI 95%]	*P*
MELD per unit	1.17 [1.17; 1.17]	<0.001	1.19 [1.18; 1.19]	<0.001	1.17 [1.17; 1.18]	<0.001

*Tumor size (>1) could not be introduced in the model because the associated HR was close to infinite.

$$\text{MHCC} = 0.17 * \text{MELD} + 0.53 \text{ if HCC Nb} \geq 2 + 1.92 * \text{ if HCC size} > 1\text{cm} + 0.53 \text{ if AFP} > 400$$

The MHCC was assessed by category, and a linear regression was established to model the risk of dropout (denoted by P) estimated by the multistate model as a function of the MHCC:

$$\text{Logit}(P) = -8.79 + 1.21 * \text{MHCC}$$

$$R^2 = 0.91$$

A similar regression was established in the non-HCC training set:

$$\text{Logit}(P) = -4.63 + 0.11 * \text{MELD}$$

$$R^2 = 0.97$$

As a final step, both HCC and non-HCC models were combined assuming identical risks of dropout in patients with HCC and non-HCC, and a new deMELD model was designed, where the smallest possible score was 6 and the largest 40:

$$\text{New deMELD} : -37.8 + 1.9 * \text{MELD} + 5.9 \text{ if HCC Nb} \geq 2 + 21.2 \text{ if HCC size} > 1\text{cm} + 5.9 \text{ if AFP} > 400$$

External validations of the model

The risk of dropout was increasing according to MELD in all non-HCC groups (Table 3). The 3-month risks of dropout were lower than 5% for patients with MELDs ≤15. Conversely, it was higher than 20% in patients with MELDs >25 in both US sets or with MELDs >20 in the UK sets.

Because the aim of this study was to obtain similar risks of dropout for patients with HCC and non-HCC to manage equitably both groups on a same waiting list, the risks of dropout linked to MELD, new deMELD, mixed new deMELD, and exception MELD in the patients with HCC were compared to those linked to MELD in the patients without HCC in the various sets (Table 3). The currently used model based on exception MELD points was the least accurate in predicting the risk of dropout. This statement was especially true for patients with exception MELDs between 21 and 25, demonstrating too low risks of dropout, close to those observed for patients with calculated MELDs 11–15 (Table 3). As a result, the risk of dropout was not progressing in a harmonious fashion according to the exception MELDs.

Conversely, the risks of dropout linked to the new deMELD, mixed new deMELD, and MELD were increasing harmoniously in the patients with HCC of the various sets, with only a lower risk of dropout for patients with the highest deMELDs (>25 in the US sets). Both followed similar shapes as MELD in the non-HCC groups (Table 3).

C-indices

The exception MELD model demonstrated the lowest C-indices in the various HCC groups (52.7% to 56.6%, Table 4). They were significantly lower by 9.6–19.4 points than the C-indices observed with the use of the new deMELD, which were between 66.2% and 73.7%. There were no significant differences between the new deMELD, the mixed new deMELD, and the MELD C-indices in the patients with HCC.

The C-indices of MELD were between 72.9% and 77.9% in the patients without HCC, and were not statistically different from the C-indices of new deMELD, mixed new deMELD, and MELD in the patients with HCC in the training and UK sets.

Score distributions

In addition to observing the risk of dropout, the score distributions were also assessed between groups. Exception MELD presented a skewed distribution of patients with a large number of them between 21 and 25 (83% of all HCC candidates in the US training set, 77% in the US validation set, and 78% in the UK validation). The resulting mean exception MELDs between 19.4 and 20.2 were higher than the mean MELDs in the HCC sets ($P < 0.001$ in all sets) and the mean MELDs in the non-HCC sets (16.1–18.7, all with $P \leq 0.03$). With this exception point system, patients with HCC put a significant pressure on the waiting list, with easier access to liver grafts.

The mean new deMELDs were between 9.2 and 9.8, and the mean HCC MELDs between 10.9 and 11.2. As observed by the histograms (Fig. 2), 68.8% (US training), 72.2% (US validation), and 69.7% (UK validation) of the patients demonstrated a new deMELD lower than their calculated MELD. Some demonstrated higher new deMELD, which reflects the presence of a large/aggressive HCC. These deMELD and MELD scores were lower than the MELD in the non-HCC sets ($P < 0.001$ in all sets).

Discussion

The proposed new deMELD can predict the risk of dropout in HCC candidates according to HCC characteristics and MELD. This new score shows a similar dropout distribution as MELD in patients without HCC and is more accurate than the exception MELD score in predicting dropout.

The present study is based on three patient populations from the US and the UK, which demonstrate significant

Table 3. Observed risks of dropout according to categories of MELD, new deMELD, mixed deMELD, and exception MELD (as per OPTN/UNOS rules). Mixed deMELD was defined as the best score between MELD and new deMELD for each patient.

		Patients without HCC		Patients with HCC							
		MELD		MELD		new deMELD		Modified new deMELD		Exception MELD	
		N	% dropout	*N*	% dropout	*N*	% dropout	*N*	% dropout	*N*	% dropout
US Training	All	43528	8.7 [8.4; 8.9]	4244	3.2 [2.7; 3.7]	4244	3.2 [2.7; 3.7]	4244	3.2 [2.7; 3.7]	4244	3.2 [2.7; 3.7]
	6–10	7724	2.0 [1.6; 2.3]	2210	1.4 [0.9; 1.8]	2986	1.6 [1.1; 2.0]	2173	1.3 [0.8; 1.8]	403	1.5 [0.3; 2.8]
	11–15	14062	3.4 [3.1; 3.7]	1425	3.3 [2.4; 4.3]	561	4.2 [2.5; 5.9]	1267	2.6 [1.7; 3.5]	226	2.7 [0.6; 4.8]
	16–20	9538	6.9 [6.4; 7.4]	471	8.5 [5.9; 11.0]	391	6.0 [3.6; 8.4]	493	6.4 [4.2; 8.6]	63	9.6 [2.3; 16.9]
	21–25	5016	12.6 [11.6; 13.5]	114	10.1 [4.4; 15.8]	134	12.9 [7.2; 18.6]	136	12.8 [7.1; 18.5]	3528	3.1 [2.5; 3.7]
	26 +	7188	25.8 [24.8; 26.9]	24	29.2 [11.0; 47.4]	172	14.3 [9.0; 19.6]	175	14.1 [8.8; 19.3]	24	29.2 [11; 47.4]
US Validation	All	16291	10.3 [9.8; 10.7]	3450	5.1 [4.3; 5.9]	3450	5.1 [4.3; 5.9]	3450	5.1 [4.3; 5.9]	3450	5.1 [4.3; 5.9]
	6–10	2356	1.8 [1.2; 2.4]	1856	3.1 [2.3; 4.0]	2516	3.3 [2.5; 4.0]	1856	3.1 [2.3; 4.0]	456	1.3 [0.2; 2.5]
	11–15	4946	2.7 [2.2; 3.1]	1155	5.2 [3.8; 6.5]	436	4.9 [2.8; 7.1]	1137	4.7 [3.4; 6.0]	258	6.3 [3.1; 9.6]
	16–20	3663	7.6 [6.7; 8.5]	351	10.2 [6.8; 13.6]	312	10.9 [7.2; 14.5]	327	10.2 [6.7; 13.7]	70	9.0 [1.4; 16.5]
	21–25	2153	15.0 [13.4; 16.6]	68	21.6 [11.2; 32.0]	92	18.3 [9.6; 27.0]	63	16.6 [6.7; 26.6]	2646	5.3 [4.4; 6.2]
	26 +	3173	26.9 [25.3; 28.5]	20	35.0 [14.1; 55.9]	94	22.2 [13.6; 30.8]	67	29.7 [18.5; 40.8]	20	35.0 [14.1; 55.9]
UK	All	1292	8.1 [6.6; 9.6]	489	7.8 [5.4; 10.2]	489	7.8 [5.4; 10.2]	489	7.8 [5.4; 10.2]	489	7.8 [5.4; 10.2]
	6–10	141	4.4 [1.0; 7.9]	271	4.8 [2.3; 7.4]	358	4.8 [2.6; 7.0]	265	4.9 [2.3; 7.5]	58	10.3 [2.5; 18.2]
	11–15	508	3.0 [1.5; 4.4]	159	7.6 [3.5; 11.7]	67	11.9 [4.2; 19.7]	149	7.4 [3.2; 11.7]	29	0.0
	16–20	421	8.4 [5.7; 11.0]	49	22.4 [10.8; 34.1]	37	16.2 [4.3; 28.1]	47	14.9 [4.7; 25.1]	19	15.8 [−0.6; 32.2]
	21 +	222	21.7 [16.3; 27.1]	10	20.0 [0.0; 44.8]	27	25.9 [9.4; 42.5]	28	25.0 [9.0; 41.0]	383	7.6 [4.9; 10.3]

Table 4. C-indices of MELD, new deMELD, and exception MELD (as per OPTN/UNOS rules). Estimates (expressed in percentages) and 95% confidence intervals are shown. Mixed deMELD was defined as the best score between MELD and new deMELD for each patient.

C-indices (en%)	US training set	US validation set	UK validation set
Patients without HCC			
MELD	76.7 (76.0 to 77.5)	77.9 (76.8 to 79.0)	72.9 (67.8 to 77.6)
Patients with HCC			
MELD	72.8 (68.2 to 77.2)	66.3 (61.6 to 70.7)	65.7 (56.0 to 75.0)
deMELD	73.7 (69.3 to 78.0)	66.2 (61.4 to 70.7)	66.3 (57.0 to 75.5)
MELD exception	54.4 (50.9 to 57.9)	56.6 (53.2 to 59.7)	52.7 (48.3 to 59.4)
Mixed deMELD	73.8 (69.3 to 78.1)	67.0 (62.2 to 71.4)	65.2 (55.3 to 74.5)
Difference in patients with HCC			
deMELD – MELD	0.9 (−1.2 to 3.0)	−0.1 (−2.3 to 2.1)	0.6 (−3.6 to 4.5)
deMELD – MELD exception	19.4 (14.5 to 24.2)	9.6 (4.5 to 15.0)	13.5 (2.0 to 23.4)
Mixed deMELD – MELD	1.0 (0.2 to 1.9)	0.7 (0.0 to 1.5)	−0.5 (−1.3 to 0.2)
Difference in patients with HCC versus without HCC			
MELD non -HCC – MELD HCC	3.9 (−0.5 to 8.6)	11.7 (7.1 to 16.6)	7.1 (−3.6 to 18.2)
MELD non -HCC – deMELD HCC	2.9 (−1.3 to 7.5)	11.8 (7.1 to 16.6)	6.5 (−4.3 to 17.2)
MELD non -HCC – MELD exception	22.3 (18.7 to 25.9)	21.5 (18.1 to 24.9)	20.0 (12.1 to 26.7)
MELD non -HCC – Mixed MELD	2.9 (−1.6 to 7.6)	11.0 (6.4 to 15.8)	7.6 (−3.2 to 18.7)

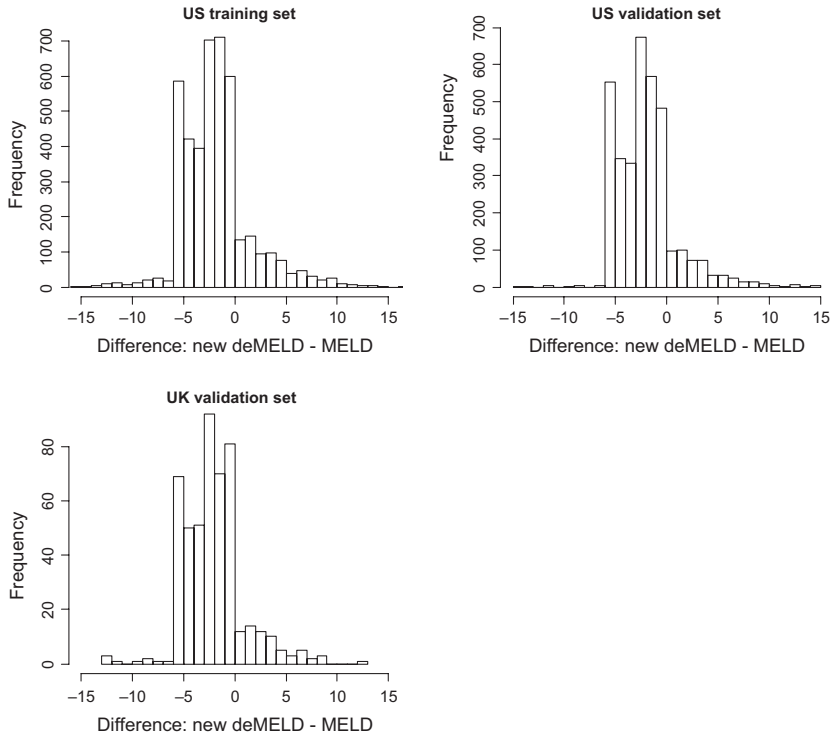

Figure 2 Distributions (histograms) of the difference between new deMELD and calculated MELD in patients with HCC for the US training, US validation, and UK validation sets.

demographic and management differences. The US training set originates from an earlier era with a smaller proportion of patients listed with HCC (11.2%), and the UK set includes more patients with alcohol-induced liver disease and with lower calculated MELDs. In addition, the risk of dropout in patients with HCC was not predicted by the same variables between patient sets, especially regarding HCC number and AFP. Overall, these observations suggest that the epidemiology of liver disease varies between sets, and that policies are different between countries and have evolved over time. They should be viewed as a way to test the proposed score externally, in different patient populations and in more "extreme" conditions, where a validation is more difficult to obtain.

Despite the observed differences between patient sets, a new deMELD model could be designed based on the US training set. Unlike a previously described score, it was only based on HCC characteristics (HCC size, HCC number, and AFP) and MELD, and could determine the risk of dropout along a continuous and harmonious scale [5]. As a result, patients with more aggressive (high AFP) and/or more advanced (large and numerous) HCCs can be transplanted earlier. The score should be updated at least every 3 months while on the list. The response to a local HCC treatment would decrease the deMELD in parallel to the decreased risk of dropout [13,14]. To illustrate, a patient with a 2 cm HCC fully treated by ablation would have a lower new deMELD (with an HCC size of 0 cm). Of note, this strategy implies that all centers apply similarly aggressive policies for HCC management, and this may need a regulation for a mandatory wait-list locoregional treatment.

Of note, the use of the new deMELD may lead to the transplantation of patients with more aggressive HCCs, who may be at increased risk of post-transplant recurrence [15]. Such a risk may be controlled by a required minimum 3–6-month waiting time prior to listing, as suggested after HCC downstaging prior to transplantation, and should be closely monitored if the model is implemented [5,16,17].

The new deMELD has been explored on external validation cohorts, and demonstrated similar patterns in the US and the UK. Its prediction of dropout in the patients with HCC was similar to the one of MELD in the patients without HCC. In addition, the new deMELD was superior to the use of exception MELD for the prediction of dropout, as exception MELD is associated with an unfairly low dropout rate for patients with 21–25 points.

Of note, the use of the proposed new deMELD would decrease the average number of points given to patients with HCC (average scores: 9.2–9.8 points), and 68.8% to 72.2% of candidates would have deMELDs lower than their calculated MELDs. This observation is based on clinical reasons: the risk of dropout in patients with HCC from the training set was modulated by the chance of being transplanted (we use a competitive risk model), and not only by the presence of cirrhosis and HCC. As shown in Fig. 1, the chance of being transplanted was higher in HCC than in non-HCC candidates in the HCC training set, and conversely the risk of dropout was lower in HCC than in non-HCC candidates. This observation explains at least in part the trend for a lower risk of dropout in the new deMELD model. Second, the proposed model may also reflect differences in the type (and the prognosis) linked to cirrhosis in patients with and without HCC, non-HCC, but not HCC ones, being primarily listed for cirrhosis problems. A number of problems linked to cirrhosis (including encephalopathy and ascites) are commonly underestimated by MELD, which again proportionally increased the risk of dropout in non-HCC MELD groups. In other words, a MELD 15 cirrhosis may not have a similar risk of dropout in the two groups of patients, further explaining the MELD and new deMELD differences. In patients with a new deMELD lower than MELD, the calculated MELD score may be used instead (as already implemented), to take into account the risk of dropout linked to liver failure. This model has been explored with the mixed new deMELD.

The present study is based on registry data, where management and reporting biases are possible. To illustrate, patient management may vary between centers, and some transplant candidates with HCC could have potentially been primarily considered for resection in some other centers. In addition, we acknowledge that some events beyond 3 months are missed by the proposed model, although most of them are captured as shown in Fig. 1.

The currently used models, including MELD, are all based on a liver graft allocation constructed on the risk of dropout. While dropout should be considered as a narrow endpoint, as it does not take into account the probability of recurrence and overall results of transplantation, we accept this limitation as reflecting the current practice of wait-list management. Another allocation scheme could be based on transplant benefit, transplanting the candidates with the longest expected graft survival first [18–20]. Although difficult to build, transplant benefit models could be explored using a similar design as the one described herein.

Overall, the proposed new deMELD allows to estimate the risk of dropout in HCC candidates according to HCC characteristics and MELD. It provides a dynamic and more accurate assessment of dropout than the use of exception MELD points. In addition, the new deMELD shows a similar distribution as MELD for patients with non-HCC, and both scores could be used in combination for the management of patients with and without HCC on a common waiting list.

Authorship

CT, PM, GM and CC: designed research/study. CT and CC: performed research/study and analyzed data. CT, PM, TB, PM, GM and CC: wrote and reviewed the paper.

Funding

This study was supported by grants from the Swiss National Science Foundation, the Artères Foundation, the Astellas European Foundation, and the Boninchi Foundation. Christian Toso was supported by a Professorship from the Swiss National Science Foundation (PP00P3_139021).

Acknowledgements

The authors are grateful to Prof. Thomas Perneger for his advice regarding the study. Some of the data reported here have been supplied by the Minneapolis Medical Research Foundation (MMRF) as the contractor for the Scientific Registry of Transplant Recipients (SRTR). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government. The authors thank the NHS Blood and Transplant and the LSAWP for allowing access to the UK data. CC had full access to all of the data provided by the SRTR and the NHSBT, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

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.
2. Wiesner R, Edwards E, Freeman R, *et al.* Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003; **124**: 91.
3. Freeman RB, Edwards EB, Harper AM. Waiting list removal rates among patients with chronic and malignant liver diseases. *Am J Transplant* 2006; **6**: 1416.
4. Washburn K, Edwards E, Harper A, Freeman R. Hepatocellular carcinoma patients are advantaged in the current liver transplant allocation system. *Am J Transplant* 2010; **10**: 1643.
5. Toso C, Dupuis-Lozeron E, Majno P, *et al.* A model for dropout assessment of candidates with or without hepatocellular carcinoma on a common liver transplant waiting list. *Hepatology* 2012; **56**: 149.
6. De Giorgio M, Vezzoli S, Cohen E, *et al.* Prediction of progression-free survival in patients presenting with hepatocellular carcinoma within the Milan criteria. *Liver Transplant* 2010; **16**: 503.
7. Kim WR, Stock PG, Smith JM, *et al.* OPTN/SRTR 2011 Annual Data Report: liver. *Am J Transplant* 2013; **13**(Suppl 1): 73.
8. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; **26**: 2389.
9. Allignol A, Schumacher M, Beyersmann J. *Empirical transition matrix of multistate models: The etm package*. 38 ed, 2011.
10. Wolbers M, Blanche P, Koller MT, Witteman JC, Gerds TA. Concordance for prognostic models with competing risks. *Biostatistics* 2014 [Epub ahead of print].
11. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *JASA* 1999; **94**: 496.
12. Toso C, Asthana S, Bigam DL, Shapiro AM, Kneteman NM. Reassessing selection criteria prior to liver transplantation for hepatocellular carcinoma utilizing the Scientific Registry of Transplant Recipients database. *Hepatology* 2009; **49**: 832.
13. Merani S, Majno P, Kneteman NM, Berney T, Morel P, Mentha G, *et al.* The impact of waiting list alpha-fetoprotein changes on the outcome of liver transplant for hepatocellular carcinoma. *J Hepatol* 2011; **55**: 814.
14. Cescon M, Cucchetti A, Ravaioli M, Pinna AD. Hepatocellular carcinoma locoregional therapies for patients in the waiting list. Impact on transplantability and recurrence rate. *J Hepatol* 2013; **58**: 609.
15. Duvoux C, Roudot-Thoraval F, Decaens T, Pessione F, Badran H, Piardi T, *et al.* Liver transplantation for hepatocellular carcinoma: a model including alpha-fetoprotein improves the performance of Milan criteria. *Gastroenterology* 2012; **143**: 986.
16. Roberts JP, Venook A, Kerlan R, Yao F. Hepatocellular carcinoma: Ablate and wait versus rapid transplantation. *Liver Transplant* 2010; **16**: 925.
17. Toso C, Mentha G, Kneteman NM, Majno P. The place of downstaging for hepatocellular carcinoma. *J Hepatol* 2010; **52**: 930.
18. Ravaioli M, Grazi GL, Dazzi A, Bertuzzo V, Ercolani G, Cescon M, *et al.* Survival benefit after liver transplantation: a single European center experience. *Transplantation* 2009; **88**: 826.
19. Schaubel DE, Sima CS, Goodrich NP, Feng S, Merion RM. The survival benefit of deceased donor liver transplantation as a function of candidate disease severity and donor quality. *Am J Transplant* 2008; **8**: 419.
20. Schaubel DE, Guidinger MK, Biggins SW, Kalbfleisch JD, Pomfret EA, Sharma P, *et al.* Survival benefit-based deceased-donor liver allocation. *Am J Transplant* 2009; **9**: 970.