Impact of MELD on short-term and long-term outcome following liver transplantation: a European perspective
Evi Nagler\textsuperscript{a}, Hans Van Vlierberghe\textsuperscript{a}, Isabelle Colle\textsuperscript{a}, Roberto Troisi\textsuperscript{b} and Bernard de Hemptinne\textsuperscript{b}

Introduction The Model for End-Stage Liver Disease (MELD) has been found to accurately predict pre-transplant mortality and is a valuable system for ranking patients in need of liver transplantation. Its association with post-transplant outcome, however, remains unclear.

Materials and methods We retrospectively studied 121 adult patients who were transplanted for non-fulminant liver failure between January 1991 and December 2001. MELD scores were calculated taking variables as close as possible prior to liver transplantation. Patients were stratified into two or three groups using different cut-off values of the MELD score.

Results Indications for liver transplantation were mainly alcoholic liver disease (47.1\%) or hepatitis C virus (19.0\%). Gender distribution was male 62\% vs female 38\%. Mean age was 54 years ± 10 years. Mean MELD score was 16 ± 6. Follow-up time was 5.4 years (range, 1.6–12.3 years). The use of different MELD cut-off levels yielded no difference in survival at different time points.

Conclusion Higher MELD scores did not have a negative impact on patient and graft survival following OLT. Since MELD is good at identifying those urgently in need of liver transplantation and high MELD scores do not appear to have an influence on long-term outcome, use of MELD in liver allocation seems warranted.

Keywords: allocation system, liver transplantation, Child–Pugh score, MELD score

Departments of \textsuperscript{a}Gastroenterology and Hepatology and \textsuperscript{b}Hepatobiliary Surgery and Liver Transplantation, Ghent University Hospital, Belgium.

Correspondence to Dr Hans Van Vlierberghe, Department of Gastroenterology and Hepatology, Ghent University Hospital, 1K12 IE, De Pintelaan 185, 9000 Gent, Belgium.
Tel: +32 9 240 2371; fax: +32 9 240 2674; e-mail: hans.vanvlierberghe@ugent.be

Accepted 4 May 2005
In addition, it is not sufficient to know the impact of MELD on short-term survival. Since transplantation no longer prolongs life by merely a few years, but offers patients a close to normal life expectancy, the question that inevitably rises is whether the use of MELD scores has a significant impact on long-term outcome after transplantation. The goal of this study subsequently was to verify whether higher MELD scores were associated with worse long-term outcomes.

Materials and methods

Patient population
We conducted a retrospective cohort analysis, using data collected by a file study. In the period from January 1991 until December 2001, 416 patients underwent liver transplantation at the Ghent University Hospital. We selected adult patients undergoing primary orthotopic liver transplantation (OLT) of a cadaver liver for chronic liver failure. We excluded patients in sequential fashion according to the study design. Paediatric patients were excluded ($n = 58$), as were patients undergoing re-transplantation or heterotopic liver transplantation ($n = 48$), patients receiving a graft from a living donor ($n = 20$), and all individuals for whom fulminant liver failure was the indication to liver transplantation ($n = 35$). For reasons of follow-up, all non-Belgians were excluded from this study ($n = 116$). Of the remaining 139 files, 18 lacked crucial information for retrospective MELD calculation. After all exclusions, 121 patients remained for further analysis.

Criteria for liver transplantation
The current allocation policy defines three categories of disease severity in chronic liver disease based on the CTP score: T2, T3 and T4. T1 is a status reserved for those suffering from fulminant liver failure. Patients with uncomplicated cirrhosis are given status T4, whereas with a CTP score less than 10, a patient will receive status T3. A CTP score of 10 is required to be eligible for status T2, which is also the status required for listing. T2 patients are sorted according to their days in urgent T2 since their most recent placement in T2, i.e., the longest waiting suitable T2 patient is ranked first. Patients are considered urgent T2 when they have one of the five defined complications: i.e., hepatic encephalopathy, variceal bleeding, hypertensive gastropathy, hepatorenal syndrome and spontaneous bacterial peritonitis. Patients are granted urgency T2 status for 28 days. After this period, each T2 patient has to be re-evaluated [13].

Definitions
Primary end points were death and graft failure. Patient survival was defined as time from initial transplantation till date of death or last follow-up. Graft survival was defined as time from transplantation until date of death, re-transplantation or last follow-up. Because of the small number of patients, initially no distinction was made between dying or graft failure. Both were interpreted as being the same end point. In a second analysis, a distinction was made between death and graft failure as primary end points and both were treated separately.

MELD score calculation
The MELD score was calculated in accordance with the UNOS guidelines. MELD scores were calculated using variables taken as close as possible prior to liver transplantation. Values for creatinine (mg/dl), INR and total bilirubin (mg/dl) smaller than 1.0 were equalled to 1 to avoid negative logarithms. Creatinine values greater than 4 were made equal to 4. The following formula was used to calculate MELD scores:

$$\text{MELD score} = 10^*\left[0.957^*\ln(\text{creatinine}) + 0.378^*\ln(\text{bilirubin}) + 1.12*\ln(\text{INR}) + 0.643\right].$$

The MELD scores were calculated with and without correction for a pretransplant diagnosis of hepatocellular carcinoma (HCC), as prescribed by the current UNOS guidelines (one lesion $< 1.9$ cm, MELD score $= 20$; one lesion $> 2$ and $< 5$ cm or two to three lesions $< 3$ cm, MELD score $= 24$ unless calculated MELD score is higher) [14].

Statistical analysis
All data were entered using an EXCEL spreadsheet program (MS Office 97, MicroSoft Corporation) and analysed further with the statistical software package SPSS (SPSS Inc., version 11.5).

Patient survival was analysed at 1 to 12 years post-transplantation with intervals of 1 year, using adequate follow-up as the inclusion criterion: for example, at 5 years following transplantation only those cases with a minimal follow-up of 5 years were analysed.

We used the Mann–Whitney $U$ test to compare MELD and CTP scores between those who survived, died or were re-transplanted.

For a first study, the study group was stratified into three groups according to the MELD score: $< 15$, 15–25 and $> 25$. These three categories correlate largely with, respectively, the UNOS status 3, 2B and 2A. This stratification was chosen following the results by Onaca et al. [8] who investigated the short-term survival using MELD as a prioritization tool.

For the second investigation, the population was divided five times into two groups according to a cut-off value of 15, 18, 20, 25 and 30. The cut-off value of 18 was chosen...
following the results of Fernandez-Aguilar et al. [12], while all other values were set arbitrarily.

(1) group: MELD < 15
  group: MELD ≥ 15
(2) group: MELD < 18
  group: MELD ≥ 18
(3) group: MELD < 20
  group: MELD ≥ 20
(4) group: MELD < 25
  group: MELD ≥ 25
(5) group: MELD < 30
  group: MELD ≥ 30

Thirdly, we also stratified the individuals according to CTP category A, B and C and a second time according to a cut-off value equal to 10.

(1) group: CTP A
  group: CTP B
  group: CTP C
(2) group: CTP score < 10 (CTP A + B)
  group: CTP score ≥ 10 (CTP C)

Statistical analysis was performed using chi-squared for categorical variables and Fisher’s exact test for 2 × 2 tables to demonstrate a difference between the different groups.

Survival was estimated using the Kaplan–Meier estimator. Log rank was used to demonstrate a difference in survival between the earlier defined subgroups. Death and graft failure were considered as the same end point as well as different ones and thereby compared to survival, together as well as separately.

Significance level for all statistical testing was set at \( \alpha = 0.05 \).

The predictive powers of MELD concerning survival of both graft and patient were determined using receiver operating characteristic (ROC) curve analysis.

**Results**

**Demographic and clinical characteristics**

The study cohort consisted of 121 patients with a mean age of 50.4 years (range 28.2–72.0, standard deviation 9.8). Amongst them, 75 were male, 46 female. Amongst males, the mean age was 54.3 years (range 34.0–71.8, standard deviation 9.2), amongst females 54.5 years (range 28.2–71.9, standard deviation 10.7). The median waiting time from listing to time for transplantation ranged from 3 months in 1991 to 1.5 years in 2001.

The most common causes of end-stage liver failure and their frequencies are summarized in Table 1. Alcoholic liver disease and hepatitis C were the most important aetiologies with, respectively, 47.1% and 19.0%. Patients with hepatocellular malignancy (\( n = 19 \)) were classified according to the underlying liver disease. Seven conditions were categorized under ‘other’ to maintain transparency in tables and graphs. These were polycystosis (\( n = 2 \)), epitheloid haemangioendothelioma (\( n = 1 \)), acute Budd–Chiari (\( n = 1 \)), hepatocytolysis based on vascular insufficiency (\( n = 1 \)), cirrhosis caused by Wilson’s disease (\( n = 1 \)), an isolated HCC (\( n = 1 \)) and finally primary sclerosing cholangitis (\( n = 1 \)). From a pragmatic point of view, in these patients MELD score was calculated. However, data were analysed with and without results from these patients.

**MELD score distribution**

The mean MELD score was 16.1, and the median 15.3 (range 6.5–40.0). The distribution of the MELD score and the number of patients among the different MELD

![Graph showing MELD score distribution](image-url)
score ranges are shown in Figure 1. After correcting for HCC, the mean MELD score was 17.5 (range 6.7–40.0).

**CTP score distribution**
In 16.5% of all patients, the CTP score varied between 5 and 7, making them belong to the Child A category, while 40.5% belonged to Child B and 43.0% to Child C.

**Outcome**
Mean follow-up was 5.4 years (standard deviation 3.2, range 1.7 to 12.3 years, including those who died or who underwent a re-transplant).

**Survival and re-transplantation**
At the end of the study period 71% was still alive, 15% had died and 14% had undergone re-transplantation. Mean 1 year and 5 year survival was, respectively, 83% and 71% (Fig. 2). Causes of death were diverse and are summarized in Table 2. The reasons for re-transplantation are summarized in Table 3.

**Influence of MELD score on survival**
Use of Mann–Whitney U test yielded no statistically significant difference in MELD for the survivors versus those who died or were re-transplanted at any of the time points. When death and graft failure were considered as separate end points, the same result was obtained. Correcting MELD for HCC revealed no significant difference between both study groups. Analysing the data without the aetiologies categorized under ‘other’ did not change these results.

For those patients with a minimal 1 year follow-up Fisher’s exact test became statistically significant only at a cut-off value of the MELD score of 30 (P = 0.02). For the patients with a minimal follow-up of 2 to 12 years, Fisher’s exact test only became statistically significant when survival was compared to graft failure at a cut-off value of the MELD score of 30.

The chi-squared never became statistically significant. Correction for HCC did not change these results, nor did analysing the data without the category of other aetiologies.

The Kaplan–Meier survival curve did not show a difference in outcome between the different groups, except for a MELD cut-off value of 30. In this case the log rank test became statistically significant, with a P value equalling 0.04. Even so, only two cases with a MELD score above 30 were reported, rendering this result fairly meaningless. When considering death and graft failure as different end points, it was comparison of survival versus graft failure that rendered the log rank statistically significant and only at a cut-off value of MELD of 30. Correcting MELD for HCC did not produce a significant log rank test for a single cut-off. Neither did any of the Kaplan–Meier curves point in the direction of a difference in result (Figs 3 and 4). Analysis of the data without the patients ranked under ‘other’ aetiologies had no significant impact on these results.

<table>
<thead>
<tr>
<th>Table 2 Causes of death</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sepsis + MOF</td>
<td>5</td>
<td>26.6</td>
</tr>
<tr>
<td>Other liver related deaths</td>
<td>4</td>
<td>21.1</td>
</tr>
<tr>
<td>Recurrent tumour</td>
<td>3</td>
<td>15.8</td>
</tr>
<tr>
<td>Chronic rejection</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>Pulmonary infection</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>Tumour ORL</td>
<td>1</td>
<td>5.3</td>
</tr>
<tr>
<td>Metastatic epidermoid epithelioma</td>
<td>1</td>
<td>5.3</td>
</tr>
<tr>
<td>Recurrent cryptogenic cirrhosis</td>
<td>1</td>
<td>5.3</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 Causes of re-transplantation</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNF</td>
<td>6</td>
<td>42.9</td>
</tr>
<tr>
<td>Ischaemic necrosis</td>
<td>3</td>
<td>21.4</td>
</tr>
<tr>
<td>Delayed graft failure</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>Art hep stenosis with recurrent cirrhosis</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>Chronic rejection</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>Recurrent hepatitis C</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>Recurrent cryptogenic liver cirrhosis</td>
<td>1</td>
<td>7.1</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Influence of CTP score on survival
Fisher’s exact never yielded a significant result, nor did the chi-squared. The Kaplan–Meier survival curve did not show a difference in outcome between the different groups.

Prognostic capacity of MELD
Following Desai et al. [11], we evaluated MELD’s capability to predict post-transplant outcome using ROC curve analysis. We calculated the c-statistic at 1 to 12 years post-transplantation, adequate follow-up always being the inclusion criterion. The pre-transplant MELD score appeared to be a weak predictor of post-operative outcome, with c-statistics never above 0.61. When corrected for HCC, prognostic capacity appeared even worse with c-statistics never above 0.56. Excluding the patients with aetiologies grouped under ‘other’ revealed a c-statistic of 0.58.

To gain insight into why the predictive value of MELD score in post-transplant survival was so poor, we conducted a Cox analysis using the three MELD score variables. In this analysis, we found that only the INR was significant in predicting post-transplant patient and graft survival ($P = 0.001, \exp(B) = 1.067$). The two remaining variables, total bilirubin ($P = 0.546, \exp(B) = 1.019$) and creatinine ($P = 0.287, \exp(B) = 1.314$) were not significant in this model.

Discussion
The current allocation algorithm for liver transplantation in the USA for patients suffering from chronic liver disease is based on the MELD score. This score will be implemented in the Eurotransplant region in 2005. It replaces a system that uses the CTP score to evaluate the disease severity of transplant candidates. In comparison to the CTP score, the MELD score has certain important advantages. Its calculation, for instance, uses only objective laboratory parameters. Moreover, its discriminating power in the evaluation of disease severity is found to be significantly greater, reducing the role of waiting time to a point were it only influences decision making in patients with equal MELD scores.

Paediatric candidates for liver transplantation were not included in this study. Children do not have the same risk of dying while on the waiting list and a separate risk score was developed in order to evaluate these patients: the Paediatric End-Stage Liver Disease score [2,15]. Patients with acute fulminant liver failure were also excluded from our study. The MELD score was developed for individuals with chronic liver disease, subsequently the attention was focused on this last category. Status 1 remains a separate allocation category, MELD based allocation is solely implemented in chronic liver failure.

Several studies showed the MELD score to be an equivalent, if not a better prognostic index for mortality on the waiting list, in comparison to the CTP classification [2–5]. In the light of the current scarcity in organs, allocation can no longer happen without consideration for the post-operative results. By transplanting the sickest
patients first, fear has risen that allocation according to medical urgency could result in worse post-operative results. The research conducted on this subject delivers no clear answer. In two large studies a significant correlation between short-term survival of graft and acceptor on one hand, and higher values of the MELD score on the other hand, was found [8,9]. Yet others did not see this connection [11,12,16]. Our investigation confirmed that patients with higher pre-transplant MELD scores do not seem to have worse outcomes on short-term after liver transplantation. Up until now, research into the correlation between MELD score and long-term outcome has not been performed.

The results of our study show that higher MELD scores do not correlate with poor long-term outcomes after orthotopic liver transplantation (OLT). At none of the time points (1–12 years post-operatively) a significant difference in pre-operative MELD scores was seen between the survivors and those who died or were re-transplanted. The lack of association between the MELD score on one hand and 1–12 year patient and graft survival on the other hand, remained obvious when arbitrary MELD scores, as suggested by UNOS, were given to an HCC diagnosis. Neither did we manage to identify a cut-off value of the MELD score above which transplantation had significant worse outcomes. This suggests that the MELD classification has no prognostic value in the post-operative patient and graft survival. In other words, MELD is not an adequate prognostic index for survival after OLT in patients with cirrhosis (c-statistic never above 0.7).

The fact that the MELD score is a weak predictor of post-transplant outcome was confirmed by a Cox proportional hazard analysis, using the three MELD laboratory parameters as co-variables. In this model, only INR proved to be an independent predictor of survival. Our findings support the idea that outcome after OLT might not be immediately related to the disease severity at the time of transplantation.

From our data, the fear that MELD based allocation would lead to less efficient use of the scarce donor livers appears unwarranted. In addition, the adjustment of the allocation algorithm seems to have no positive or negative consequences for the results after liver transplantation.

Actually it seems strange that MELD would not have an influence on post-transplant outcome. Several studies have shown that creatinine [17], especially, and also bilirubin and the INR are independent predictors of survival after transplantation [18]. Still it appears that on calculation of the score, the impact disappears. Perhaps the weighting factors need alteration so as to exert an impact on post-transplant outcome. The finding that MELD poorly predicts outcome after transplantation may have been anticipated because prior studies have identified factors in addition to the severity of liver disease that have a major influence on post-transplant survival. Although studies have demonstrated the prognostic value of pre-transplant laboratory values [19,20], other work has shown that variables related to recipient diagnosis [20,21,22], donor organ quality [23–25], age of both donor and recipient [18], surgical procedure [26–28], and early post-operative events [29–31] are at least as important in predicting post-transplant outcome.

A worst case scenario would be an allocation model that accurately identifies patients most likely to die pre-transplant but at the same time those with the greatest risk of post-operative mortality and graft failure.

Such a system would not result in maximal efficiency in terms of life-years saved. Since MELD is highly predictive for pre-transplant mortality, yet has no connection with post-transplant patient or graft survival, it implies that patients with the highest risk of dying while on the waiting list can still enjoy equivalent post-operative survival.

A limitation of our study is its retrospective character. We are talking about a patient population that was transplanted before instalment of MELD. None of our patients was treated under the current MELD system. Our OLT population may change under the MELD system. For example, some of our status 2B patients with high MELD scores would receive a cadaveric liver more quickly today. Conversely, our patients with extensive waiting times but a low MELD score would have undergone transplantation faster under the old UNOS allocation system. In other words, the patients too ill to be transplanted and subsequently removed from the waiting list were not included in the current analysis. This procedure implies elimination of those with expected poor outcomes, thus possibly those with higher MELD scores.

Possibly these findings are not entirely relevant for transplantations performed under a MELD based allocation algorithm. Given the recent instalment of the MELD score in the USA, good prospective studies are not available, making these retrospective studies with all its limitations the only way to evaluate the change in allocation.

Initially, no distinction was made between death and graft failure. Both were considered to be the same end point. Primarily we wanted to investigate whether a correlation existed between the MELD score and long-term success of transplantation as treatment for chronic end-stage liver
disease. The number of deaths and graft failures was too small to make this distinction in the analysis. However, one could argue that the majority of re-transplantation causes were not related to high or low MELD scores, therefore making the distinction between death and graft failure a necessary one. When considered as separate end points, statistical analysis revealed no correlation between high MELD scores and poor survival post-transplantation. Because of the low number of patients, we will however have to wait for larger studies to provide a definite answer.

In our study, the number of patients with long-term follow-up (> 8 years) was rather limited, making further investigation on national and international scale an absolute necessity, before drawing a definite conclusion. Because the study was limited to patients undergoing OLT in one centre, the results might not be representative for the practice in other centres. However, at the 2004 meeting of European Association for Study of the Liver (EASL), similar results were found in other Belgian and European centres [32,33].

The weaknesses of MELD lie in the fact that this system was not developed for its current use in liver allocation. Originally, the MELD score was developed as a predictor for survival in patients undergoing a transjugular intrahepatic portosystemic shunt (TIPS) procedure [34]. With time, the formula was adjusted for prediction of pre- and post-transplantation mortality in end-stage liver disease.

Conclusion
We can conclude that changing the allocation system is a step in the right direction. The system, however, is far from perfect and the waiting list keeps evolving and changing. As long as there remains a shortage of donor organs, the allocation algorithm needs adjusting, improvement and re-evaluation. In the future, further attempts need to be made to develop a system that can be implemented next to MELD to predict post-transplant survival. That way pointless transplantations could be avoided to achieve the maximum value of each organ.

Conflict of interest
None declared.

Authors’ contributions
EN did the analysis and wrote the article. HV had the original idea and wrote the article. IC did the follow up of the patients. RT and BH carried out the surgery.

References
3 McCaughan GW, Strasser SI. To MELD or not to MELD? Hepatology 2001; 34:215–216.


