TRANSPLANT INTERNATIONAL

Transplant International ISSN 0934-0874

ORIGINAL ARTICLE

Comparison of seven liver allocation models with respect to lives saved among patients on the liver transplant waiting list

Laurence S. Magder,[1] Arie Regev[2] and Ayse L. Mindikoglu[3]

1 Division of Biostatistics and Bioinformatics, Department of Epidemiology and Public Health, University of Maryland, Baltimore, MD, USA
2 Division of Gastroenterology and Hepatology, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA
3 Division of Gastroenterology and Hepatology, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

Keywords
end-stage liver disease, liver transplantation, model for end-stage liver disease, transplant policy.

Correspondence
Laurence Magder PhD, MPH, Division of Biostatistics and Bioinformatics, Department of Epidemiology and Public Health, University of Maryland, 660 W. Redwood Street, Baltimore, MD 21201-1596, USA. Tel.: 410 706 3253; fax: 410 706 8548; e-mail: lmagder@epi.umaryland.edu

Conflicts of interest
The authors have no conflicts of interest.

Received: 5 July 2011
Revision requested: 1 August 2011
Accepted: 5 January 2012
Published online: 2 February 2012

doi:10.1111/j.1432-2277.2012.01431.x

Summary

The patients with end-stage liver disease (ESLD) on the liver transplant waiting list are prioritized for transplant based on the model for end-stage liver disease (MELD) score. We developed and used an innovative approach to compare MELD to six proposed alternatives with respect to waiting list mortality. Our analysis was based on United Network for Organ Sharing data of patients with ESLD on the waiting list between January 2006 and June 2009. We compared six allocation models to MELD. Two models were based on reweighting the variables used by MELD: an ''updated'' MELD, and ReFit MELD. Four models also included serum sodium: MESO, MeldNa, UKELD, and ReFit MELDNa. We estimated that UKELD and the updated MELD would result in significantly fewer lives saved. There were no significant differences between the other models. Our new approach can supplement standard methods to provide insight into the relative performance of liver allocation models in reducing waiting list mortality. Our analysis suggests that UKELD and the updated MELD score would not be optimal for reducing waiting list mortality in the United States.

Introduction

In 2002, in an effort to reduce liver transplant waiting list mortality, the Organ Procurement and Transplantation Network (OPTN) adopted the model for end-stage liver disease (MELD) score to prioritize adult patients with end-stage liver disease (ESLD) on the liver transplant waiting list [1]. The MELD score is a modified version of a score developed by Malinchoc *et al.* [2] based on the survival experience of 231 patients undergoing transjugular intrahepatic portosystemic shunt and was subsequently shown to be a good predictor of 3-month mortality in diverse subpopulations of patients with liver disease [3,4]. However, in developing the new allocation policy based on the MELD score, the United Network for Organ Sharing (UNOS)/OPTN committee recognized that the allocation policy should be a ''fluid, constantly changing system, whereby new data and experience are constantly reanalyzed and incorporated into the system where appropriate'' [1].

Since then, in the spirit of this comment, a number of research teams have sought to develop other models that could better predict waiting list mortality. Huo *et al.* [5] proposed a model that consisted of the ratio of the MELD score and serum sodium (MESO). Using OPTN data regarding 6 769 on the US waiting lists in 2005, Kim *et al.* [6] developed a model based on MELD, serum sodium and their interaction (MELDNa). Sharma *et al.* [7] developed an updated MELD using the same variables

© 2012 The Authors

as in the MELD score, but reweighting them based on an analysis of data on 38 899 patients on the liver transplant waiting list from 2001 to 2006 (which we will refer to as updated MELD). Barber *et al.* [8] used 1 103 patients on the transplant waiting lists in the United Kingdom to derive an allocation model based on the MELD variables (serum creatinine, bilirubin, and INR), and serum sodium. Currently, this model is used in UK to prioritize patients. Most recently Leise *et al.* [9] developed a two models based on 14 214 patients on the US waiting lists, one with the original three MELD variables (reFit MELD; 9), and one including sodium (reFit MELDNA; 9).

Most evaluations of liver allocation risk models compare their performance on historical waiting list mortality data using a ''c-statistic'' [10]. A c-statistic is calculated as the area under a Receiver Operator Curve (ROC), which is a plot of sensitivity versus (1-specificity) at each possible cut-point of the quantitative predictor [10]. C-statistics take values from 0 to 1; values closer to 1 indicate better performance of the predictor [10]. Given one person with the event (e.g., death) and one person without the event, the c-statistic can be shown to equal the probability that the person with the event has a higher value of the quantitative predictor [11]. Although the c-statistic is widely used, its direct clinical implications in this context are difficult to decipher.

In this article, we utilized an alternative method to use historical data to compare liver allocation models. It is based on the number of lives that would have been saved over a period of time had additional donor livers been available. Using this method, we compared the seven models described elsewhere with respect to lives saved on the transplant list over a period of time.

Methods

Our approach is based on the following question: What if 10% more livers were available and were allocated based on a specific risk score? How many lives would be saved? We can estimate that number directly from historical waiting list data by hypothetically allocating 10% more livers and seeing how many of those who received a hypothetical liver in reality died on the waiting list. We performed this data analytic experiment separately for seven different allocation scores, and compared the scores with respect to number of lives saved.

STAR datasets

Our approach was designed to apply to historical liver transplant waiting list files provided by UNOS. These files contain one record for every patient registered on the waiting list in the US since 1987. Each record indicates whether the patient was ever removed from the waiting list, and if so, the reason why they were removed. Reasons included death, transplant, condition deteriorated so that transplant was not indicated, and condition improved.

In addition, we have obtained a supplemental file from UNOS with all the patients' serial laboratory values, whereas they were on the waiting list. This includes repeated measures of serum creatinine, serum sodium, bilirubin, and international normalized ratio (INR). For serum creatinine, the number of measures per person varied from 1 to 102, and the median number of measures per person was 4. The median interval between two measures from the same person was 33 days (range: 1–1 258, 99th percentile = 377). Similar numbers of measures were available for the other laboratory measures. This file also indicates time periods during which the person was ''temporarily inactive'', i.e., time periods when the patient was not a candidate for a liver transplant.

Previously, as we described [12], to implement our approach, we reformatted the waiting list dataset into a person-month dataset that consisted of one record for each month that each person was on the waiting list.

Study population

Our analysis was based on UNOS data as of August 25, 2009. We included the patients with ESLD whose priority for liver transplantation was based on the MELD score and who were listed between January 1, 2006 and June 30, 2009. We excluded those under 18 years of age, patients listed as status 1A and exceptional cases (including those with hepatocellular cancer) whose priority for transplant was not based strictly on MELD score. We also excluded waiting list time during which the values of creatinine, bilirubin, sodium, or INR were unknown. Excluding these periods with unknown values resulted in exclusion of 2% of the waiting list time.

Liver allocation models evaluated

We compared seven different models, including MELD [1,2], MESO [5], MELDNa [6], updated MELD [7], UKELD [8], ReFit MELD [9], and ReFit MELDNa [9] that were previously described in the literature. Their formulae are shown in Table 1. In addition, we compared the performance of each model to what would have happened if length of time on the waiting list had been used as the only factor determining liver allocation.

Estimating the number of lives saved

An ideal way to compare risk models would be to estimate the number of patients who would die on the

© 2012 The Authors

Table 1. Liver allocation models.

Allocation models	Formula*	Details
MELD (1,2)	9.57 $\log_e$ Cr + 3.78 $\log_e$ (Bil) + 11.2 $\log_e$ (INR) + 6.43	Values of Cr greater than 4.0 are set to 4.0. Values of Cr, Bil or INR below 1.0 are set to 1.0. Those on dialysis are given a Cr of 4.0. The score is rounded to the nearest integer and capped at 40
MESO Index (5)	MELD/Na (mEq/l) × 10	MELD is calculated as above, but it is not rounded or capped at 40
MELDNa (6)	MELD – Na – [0.025(MELD)(140-Na)] + 140	MELD is calculated as in the original formula (rounded, capped at 40). Values of Na below 125 are set to 125 and values of Na above 140 are set to 140. After calculation, the score is rounded to the nearest integer
Updated MELD (7)	1.266 $\log_e$ (Cr + 1) + 0.939 $\log_e$ (Bil + 1) + 1.658 $\log_e$ (INR + 1)	Values of Cr, Bil and INR are not changed, irrespective of their values, or whether the patient is on dialysis
UKELD (8)	1.485 $\log_e$ (Cr) + 3.13 $\log_e$ (Bil) + 5.395 $\log_e$ (INR) – 81.565 $\log_e$ (Na) + 435	Values of Cr (μmol/l) greater than 400 are set to 400. Values of Cr, Bil or INR below 1.0 are set to 1.0. Na is truncated to be between 112 and 150. The score is rounded to the nearest integer. For those on dialysis, Cr is given the value found when the patients were not on dialysis†
ReFit MELD (9)	8.485 $\log_e$ (Cr) + 4.082 $\log_e$ (Bil) + 10.671 $\log_e$ (INR) + 7.432	Values of Cr greater than 3.0 are set to 3.0. Values of Cr below 0.8 are set to 0.8. Values of Bil or INR below 1.0 are set to 1.0. Values of INR above 3.0 are set to 3.0. Those on dialysis are given a Cr of 3.0. The score is rounded to the nearest integer
ReFit MELD NA (9)	6.792 $\log_e$ (Cr) + 4.258 $\log_e$ (Bil) + 8.29 $\log_e$ (INR) + 0.652(140-Na) –0.194 (140-Na) $\log_e$ (BiliCC)	Values of Cr, Bil, and INR are defined as for ReFit MELD. Values of sodium below 125 are set to 125 and values of sodium above 140 are set to 140. BiliCC is the same as Bil, however, values over 20 are set to 20. The score is rounded to the nearest integer

*In all the formulae except UKELD, Cr refers to serum creatinine (mg/dl), Bil to total bilirubin (mg/dl), INR to prothrombin time-international normalized ratio, and Na to serum sodium (mmol/l). For UKELD, Cr refers to serum creatinine (μmol/l), Bil to total bilirubin (μmol/l).
†As we did not know the patient's nondialysis creatinine value, those on dialysis were given the maximum value of Cr (400 μmol/l).

waiting list if allocation was based on each model and then compare the results. However, with historical waiting list data, where allocation was already performed (and based on the MELD score), it is difficult to make that comparison because, for patients transplanted based on the MELD score, we do not know what would have happened to them if they had not been given a transplant. To make that comparison, it would be necessary to develop a predictive model to obtain the best guess regarding what would have happened to these patients. Developing such a model requires making numerous modeling assumptions.

However, it is possible to use the historical waiting list data to estimate the number of lives that would have been saved if additional livers had been available without modeling the outcomes for transplanted patients. Our approach was as follows: Assume 25 additional livers were available in the first month of the study period. Hypothetically, ''allocate'' these livers to those active on the waiting list with the highest values of the risk score that month (not including those who actually received a transplant that month). Look ahead in time, to see if any of those who received a hypothetical liver died later on the waiting list during the study period. If a person died, that can be viewed as a ''life saved'' during the study period by the availability of the additional liver and allocation based on that model. This is repeated for each month in the study period, assuming that in each month 25 additional livers are available, and determining the number of lives saved.

Using this approach, we estimated the number of lives saved for each candidate liver allocation model, and then compared the results. We chose to consider 25 additional livers (rather than a different number) because we thought it was large enough so that conclusions about the relative performance of the scores could be reached and small enough so that it represented only a 10% increase in the number of livers available. This gave us more confidence that the relative performance of the models on these additional livers would be similar to their relative performance they had been used to allocate the actually available livers. In a secondary analysis, we calculated the

© 2012 The Authors

number of lives saved based on the assumption that 50 (rather than 25) additional livers were available each month.

One complication is that in the current allocation system in the US, livers are not allocated to the person with the highest MELD score in the country. Instead, they are allocated to the person (if any) with the highest MELD score in the Donation Service Area (DSA), where the patient is listed. There are 59 DSA in the US [13]. It was not possible for us to allocate the hypothetical livers to patients with the highest score in each DSA because DSA information was not included in the standard transplant analysis and research (STAR) file. To approximate this system of allocation, therefore, we randomly divided the patients into five hypothetical DSA's for each of the 11 regions, and allocated the hypothetical livers based on those with the highest risk score in each hypothetical DSA.

Another complication is the fact that sometimes, a patient who is hypothetically allocated a liver in our experiment will, in reality receive a transplant in a later month. In these cases, we assumed that the transplant received in a later month would be available for another patient. It was then hypothetically allocated to another patient, along with the 25 additional hypothetical livers in that later month.

Often, because of acute health problems or other reasons, patients on the waiting list are considered to be not suitable for liver transplant at a point in time. These patients are listed as ''temporarily inactive'' in the hope that their condition would improve, and they will be restored to active status. While they are temporarily inactive, they were not eligible for a transplant. In our hypothetical allocation, we also did not allocate livers to patients, whereas they were temporarily inactive.

If a patient was removed from the waiting list as a result of deteriorating condition, we treated that as a ''death'' in calculating the number of lives saved. This is consistent with the approach used by other analyses of waiting list mortality (e.g., Fink *et al.*; 14).

Statistical analysis

Statistical analysis was performed using SAS software, Version 9.2 (Cary, NC, USA; 15). To assess the statistical significance of differences in estimated numbers of lives saved, we used McNemar's test [16].

Results

There were 27 473 patients with ESLD who were on the liver transplant waiting list at any point in time during the study period from January 2006 through June 2009. This corresponded to 370 082 person months on the waiting list. The characteristics of the patients and results of each month are shown in Table 2. Sixty-four percent of the patients were male, and almost half (45%) were between 50 and 59 years of age. On average, there were 253 transplants per month, therefore the addition of 25 hypothetical donated livers represents only a 10% increase in the number of transplants.

Table 3 shows the estimated number of lives saved during the study period by allocating 25 additional livers each month using each allocation model. The number of lives saved basing allocation on waiting list time was significantly lower than basing allocation on any of the seven models ($P < 0.0001$). There were no significant differences in the number of lives saved between MELD, MESO, ReFit MELDNa, MELDNa, and ReFit MELD, although of these, MELD performed the poorest, and ReFit MELD performed the best. UKELD resulted in significantly fewer lives saved than the other scores, and updated MELD score performed significantly worse than the best performing score.

In secondary analyses, we assumed 50 additional livers were available each month (Table 4). This analysis

Table 2. Liver transplant waiting list experience from January 2006 to June 2009.

Characteristics of patients with ESLD on the liver transplant waiting list	
Number of patients	27 473
Sex	
Male	17 530 (64%)
Female	9 953 (36%)
Age	
18–39	1 614 (6%)
40–49	4 986 (19%)
50–59	11 513 (45%)
60–69	6 851 (27%)
70–79	855 (3%)
Characteristics of person months	
Number of person months	370 082
Outcomes each month	
Number (%) surviving the month, and staying on the list	353 413 (96%)
Number (%) dying	3 650 (1%)
Number (%) leaving list due to becoming medically unsuitable	1 790 (0.5%)
Number (%) transplanted	10 635 (3%)
Number (%) leaving list due condition improved	594 (0.2%)
Average number of patients on the waiting list per month	8 415
Average deaths (or removal for medical unsuitability) per month	130
Average number of transplants per month	253

© 2012 The Authors

Table 3. Estimated number of lives saved if 25 additional livers were available each month and the additional livers were allocated by each model (From January 2006 through June 2009).

Allocation model	Number of lives saved	*P*-values for pairwise comparisons
Time on the waiting list	144	$P < 0.0001$ (compared with each of the other scores)
UKELD (8)	788	$P = 0.022$ compared with updated MELD $P < 0.001$ compared with each of the other scores
Updated MELD (7)	846	$P = 0.25$ compared with MELD, $P = 0.67$ compared with Refitmeldna $P = 0.072$ compared with MESO $P = 0.067$ compared with ReFitMELDNA $P = 0.054$ (compared with MELDNa) $P = 0.021$ compared with ReFit MELD
MELD (1,2)	867	$P > 0.17$ compared with each of the scores below on this table
MESO Index (5)	879	$P > 0.58$ compared with each of the scores on this table below
ReFit MELDNa (9)	882	$P > 0.73$ compared with each of the scores below on this table
MELDNa (6)	884	$P = 0.86$ compared with ReFit MELD
ReFit MELD (9)	887	

Table 4. Estimated number of lives saved if 50 additional livers were available each month and the additional livers were allocated by each model (Starting from January 1, 2006).

Allocation model	Number of lives saved	*P*-values for pairwise comparisons
Time on the waiting list	307	$P < 0.0001$ (compared with each other model on this table)
UKELD (8)	1360	$P = 0.086$ compared with Updated-MELD, and $P < 0.002$ compared with each other score below
Updated MELD (7)	1407	$P = 0.029$ compared with MELD, and $P < 0.002$ compared with each other scores below
MELD (1,2)	1449	$P = 0.11$ compared with MESO $P = 0.095$ compared with ReFit MELD $P = 0.043$ compared with ReFit MELDNa $P = 0.033$ compared with MELDNa
MESO Index (5)	1471	$P > 0.32$ for each score below
ReFit MELD (9)	1473	$P > 0.45$ compared with each score below
ReFit MELDNa (9)	1488	
MELDNa (6)	1488	

resulted in generally similar results, although MELD performed significantly worse than the best performing scores.

For comparison, we also calculated c-statistics to quantitate the degree to which each monthly score predicted death in that month. In order from best to worst, the results were MELDNa, ($c = 0.835$), ReFit MELDNa ($c = 0.834$), MESO ($c = 0.826$), ReFit MELD ($c = 0.824$), MELD ($c = 0.822$), updated MELD ($c = 0.816$), and UKELD ($c = 0.805$).

Discussion

On the basis of current liver transplantation policy, donor organs are allocated to patients with the highest risk of death, which is estimated using the MELD scoring system. Although the MELD score is widely used as a liver allocation system, a few other allocation models have recently been proposed in an attempt to improve the prediction of survival and enhance liver allocation among candidates for liver transplantation [5–7].

A widely used approach for comparing liver allocation models is the c-statistic. The c-statistic is a quantification of the performance of a quantitative predictor of a binary outcome, integrating over all possible cut-points of the quantitative predictor. However, in the context of liver transplantation, where relatively few livers are available, the performance of a risk model should be based on high cut-offs; specifically, on how well-relatively high risk scores predict, which the patients are at imminent risk of death. Another drawback to the c-statistic is that it is difficult to interpret the epidemiologic implications of a difference in c-statistics.

Another approach to comparing risk models is to use the simulation program ''Liver Simulated Allocation Models'' (LSAM) created by the US Scientific Registry of Transplant Recipients [17,18]. Given input on a series of waiting list patients and their risk for various outcomes, and a set of allocation policies, this program uses Monte Carlo methods to simulate outcomes [19]. The series of waiting list patients and their outcomes can be derived from historical experience. However, because some of the patients who were allocated transplants in reality might not receive a transplant in a specific simulation, some method must be used to provide information regarding what would have happened to that patient had they not received a transplant. This can be performed based on modeling or matching with other patients. To our knowledge, there is no published comparison of liver allocation models based with respect to waiting list mortality based on LSAM.

© 2012 The Authors

As an alternative, we developed a way of comparing risk models using historical data that does not require modeling assumptions or matching. The result of our approach is interpretable as the number of lives saved had additional livers been available. In the study reported herein, we used this approach to evaluate previously proposed risk models. However, our approach might also be useful in the development of new risk models, by providing a metric to optimize in the development of a new model.

There are several caveats to the use of our method. First, it is based on the assumption that if a person who died on the waiting list had actually been given a liver earlier, they would not have died. This assumes that the death was as a result of primary liver disease. If, for example, the cause of death while on the waiting list was unrelated to liver disease then this person would have died even if the person had received a transplant. To the extent that waiting list deaths were not related to the primary liver disease, our method overestimates the number of lives saved.

A second caveat is that, whereas our method provides direct insight into how an allocation score performs had additional livers been available, this is not necessarily the same as how it would perform as the primary allocation model for all livers. We chose to compare the models on additional hypothetical livers because to compare them using the available livers we would have had to speculate regarding what would have happened to individuals who actually received a transplant, had they not received a transplant. We believe the performance of a model on a set of additional livers that constitutes only a 10% increase in the number of livers (i.e., 25 per month) is a good indication of how it would perform as the primary allocation model.

A third caveat is that our method does not take into consideration the effect that allocation of additional livers based on a specific risk score will have on the structure of the waiting list. The median score will tend to decline, and this effect will be cumulative over time, resulting in reduced mortality. If this effect varies by allocation score, it could affect the relative number of lives saved by the different scores over an extended period of time.

Our approach compares models strictly with respect to waiting list mortality, without consideration of post-transplant survival probabilities of different types of patients. This policy is in accordance with the mandate of the Department of Health and Human Services as stated in the ''final rule'' [20]. In contrast, it might be argued that post-transplant mortality risk should be considered in prioritizing the patients for liver transplantation [21,22]. This would reduce the probability that scarce livers would be provided to the patients who would die soon after transplantation, and maximize the utility of the available livers. In fact, one of the reasons cited for capping MELD at 40 was to reduce the likelihood of futile transplants [1]. Recently, Schaubel *et al.* [23] proposed an interesting allocation scheme that takes into account both waiting list mortality and post-transplant mortality, in an effort to maximize lifetime gained through liver transplantation.

Even if the policy changes and post-transplant mortality is taken into consideration in the future liver allocation policy, the approach that we used to compare models with respect to waiting list mortality would be useful. This is because a transplant-benefit score such as the one proposed by Schaubel *et al.* [23] is based, in part, on a model to predict waiting list mortality.

In summary, we used a novel approach to quantify the performance of liver transplant priority models. While our approach is based on some strong assumptions, the assumptions required are different than those required by programs such as LSAM, or statistics such as the c-statistic. Thus, our approach can supplement the other approaches and provide additional insight into the relative performance of allocation models. On the basis of our approach, some recently proposed models were found to perform somewhat better than the MELD score, and significantly better than the UKELD score.

Authorship

LSM: developed and implemented the method, and contributed to writing the paper. AR: contributed to developing the method and writing the paper. ALM: contributed to developing the method, obtaining the data, and writing the paper.

Funding

The work was supported in part by grant no. 1 K23 DK089008-01 from the National Institutes of Health (NIH) National Institute of Diabetes and Digestive and Kidney Diseases (to Ayse L. Mindikoglu, MD, MPH) and its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the NIH.

In using UNOS data, this work was supported in part by Health Resources and Services Administration contract 234-2005-370011C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

© 2012 The Authors

Acknowledgements

We thank Jean-Pierre Raufman, MD (Professor of Medicine, Division of Gastroenterology and Hepatology, Department of Medicine, University of Maryland School of Medicine) and Charles D. Howell, MD (Professor of Medicine, Division of Gastroenterology and Hepatology, Department of Medicine, University of Maryland School of Medicine) for reviewing our manuscript.

References

1. Freeman RB Jr, Wiesner RH, Harper A, *et al.* The new liver allocation system: moving toward evidence-based transplantation policy. *Liver Transpl* 2002; **8**: 851.
2. 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.
3. Kamath PS, Wiesner RH, Malinchoc M, *et al.* A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; **33**: 464.
4. Wiesner R, Edwards E, Freeman R, *et al.* Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003; **124**: 91.
5. Huo TI, Wang YW, Yang YY, *et al.* Model for end-stage liver disease score to serum sodium ratio index as a prognostic predictor and its correlation with portal pressure in patients with liver cirrhosis. *Liver Int* 2007; **27**: 498.
6. Kim WR, Biggins SW, Kremers WK, *et al.* Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med* 2008; **359**: 1018.
7. Sharma P, Schaubel DE, Sima CS, Merion RM, Lok AS. Re-weighting the model for end-stage liver disease score components. *Gastroenterology* 2008; **135**: 1575.
8. Barber K, Madden S, Allen J, Collett D, Neuberger J, Gimson A. Elective liver transplant list mortality: development of a United kingdom end-stage liver disease score. *Transplantation* 2011; **92**: 469.
9. Leise MD, Kim WR, Kremers WK, Larson JJ, Benson JT, Therneau TM. A revised model for end-stage liver disease optimizes prediction of mortality among patients awaiting liver transplantation. *Gastroenterology* 2011; **140**: 1952.
10. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 2007; **115**: 654.
11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**: 29.
12. Mindikoglu AL, Regev A, Seliger SL, Magder LS. Gender disparity in liver transplant waiting-list mortality: the importance of kidney function. *Liver Transpl* 2010; **16**: 1147.
13. Technical Methods for Organ Procurement Organizations (OPO) Specific Reports. Scientific registry of transplant recipients. Available at: http://www.ustransplant.org/csr/current/opo-report.aspx#loc_this_opo. Accessed on October 24, 2010.
14. Fink MA, Berry SR, Gow PJ, *et al.* Risk factors for liver transplantation waiting list mortality. *J Gastroenterol Hepatol* 2007; **22**: 119.
15. SAS Software. *SAS Software, Version 9.2*. Copyright 2002–2008. http://www.Sas.Com/. Cary, NC: SAS Institute Inc.
16. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947; **12**: 153.
17. Thompson D, Waisanen L, Wolfe R, Merion RM, McCullough K, Rodgers A. Simulating the allocation of organs for transplantation. *Health Care Manag Sci* 2004; **7**: 331.
18. Simulated allocation models list. Scientific registry of transplant recipients. Available at: http://www.ustransplant.org/sam/Default.aspx. Accessed on October 22, 2010.
19. Metropolis N, Ulam S. The Monte Carlo method. *J Am Stat Assoc* 1949; **44**: 335.
20. Organ Procurement and Transplantation Network – HRSA. Final rule with comment period. *Fed Regist* 1998; **63**: 16296.
21. Merion RM, Schaubel DE, Dykstra DM, Freeman RB, Port FK, Wolfe RA. The survival benefit of liver transplantation. *Am J Transplant* 2005; **5**: 307.
22. Weismuller TJ, Fikatas P, Schmidt J, *et al.* Multicentric evaluation of model for end-stage liver disease-based allocation and survival after liver transplantation in Germany – limitations of the 'sickest first'-concept. *Transpl Int* 2011; **24**: 91.
23. Schaubel DE, Guidinger MK, Biggins SW, *et al.* Survival benefit-based deceased-donor liver allocation. *Am J Transplant* 2009; **9**: 970.

© 2012 The Authors