

Chapter VIII

Survival Benefit-Based Deceased-Donor Liver Allocation

Overview

- Currently, patients awaiting deceased-donor liver transplantation are prioritized by medical urgency. Specifically, wait-listed chronic liver failure patients are sequenced in decreasing order of Model for End-stage Liver Disease (MELD) score.
- To maximize lifetime gained through liver transplantation, post-transplant survival should be considered in prioritizing liver waiting list candidates. We evaluate a survival benefit based system for allocating deceased-donor livers to chronic liver failure patients.
- Under the proposed system, at the time of offer, the transplant survival benefit score would be computed for each patient active on the waiting list. The proposed score is based on the difference in five-year mean lifetime (with vs without a liver transplant) and accounts for patient and donor characteristics.
- There is great overlap in the distribution of benefit scores across MELD categories, since waiting list mortality is significantly affected by several factors.
- Simulation results indicate that over 2,000 life-years would be saved per year, if benefit-based allocation was implemented.
- The shortage of donor livers increases the need to maximize the life-saving capacity of procured livers.
- Allocation of deceased-donor livers to chronic liver failure patients would be improved by prioritizing patients by transplant survival benefit.

Introduction

In all areas of medicine, it is essential to determine whether or not a patient will benefit from a given treatment. In the case of organ failure, such questions are perhaps even more important since the preferred treatment (organ transplantation) is not available for all patients. In certain cases, even if there were a sufficient number of donor organs, certain types of patients would be better off not receiving a transplant (1, 2) since their waiting list mortality is not sufficiently high to offset the high post- and peri-operative mortality, which is now well-established in the literature (3).

There are at least three possible bases for organ allocation: medical urgency; utility; and transplant benefit. Typically, the prioritization of patients active on the waiting list on a particular date depends strongly on the allocation scheme. Under a medical urgency-based allocation system, patients with worse waiting list outcomes are given higher priority for transplantation. Conversely, a utility-based system would assign priority in accordance with expected post-transplant outcomes. An allocation scheme based on transplant benefit considers both waiting list and post-transplant outcomes. For example, a patient's priority for transplantation could be based on the contrast between two settings: (i) the patient receives the allocated organ (ii) the patient receives no organ.

Each of the allocation schemes mentioned in the preceding paragraph (urgency, utility, benefit) has advantages and disadvantages. For concreteness, suppose that the only outcome considered is mortality. An urgency-based system succeeds in assigning donor organs to patients who are most likely to die on the waiting list. However, this approach may be at the expense of utility since patients at the greatest risk of waiting list death may also be the patients with the highest post-transplant mortality. One can envision an extreme case where medical urgency-based allocation does not result in any fewer deaths, but merely shifts mortality from the pre- to post-transplant side. Conversely, a utility-based allocation system would ensure that transplanted organs are received by patients with lowest post-transplant mortality. However, patients with the best post-transplant outcomes may also have the best waiting list outcomes. In an extreme case, an ordering based on utility could also result in transplantation having no effect on the mortality experience of the patient population, since the low death rate faced by the low-risk patients is merely traded for a low post-transplant death rate. This is different from the extreme scenario we described for an urgency-based system where a high death rate on the waiting list is traded for a high post-transplant death rate. In both cases, however, the lifetime experienced by the patient population is equal to that in the absence of access to transplantation.

A survival benefit based allocation system seeks to minimize mortality to the patient population as a whole by prioritizing patients based on their lifetime gained due to transplantation. To see that maximizing survival benefit results in minimizing patient population mortality, consider the fact that every patient is guaranteed at least their waiting list lifetime. In the absence of transplantation, all patients would experience their waiting list lifetime, and that alone. Suppose that one donor organ is available for transplantation and to be allocated. If that one organ is transplanted, the recipient receives his/her waiting list lifetime, plus any gain in lifetime attributable to the transplant (the

transplant survival benefit). Therefore, allocating the one available organ to the patient with the largest difference in post-transplant and waiting list lifetime (i.e., the greatest transplant survival benefit) will minimize mortality for the patient population as a whole.

In Table VIII-1, we illustrate the last concept from the preceding paragraph: that allocating an available organ to the patient with the greatest transplant benefit maximizes the total life-years lived by the patient population. In this simplified setting, there are three patients on the waiting list at the time a donor organ is to be allocated. The columns in Table VIII-1 are (left to right) ID: patient identification number; WL: predicted waiting list lifetime (i.e., lifetime if no transplant is received), LT: predicted post-transplant lifetime (with the organ to be allocated); B: transplant benefit, computed as $LT - WL$; years lived, summed across all three patients, with row x representing the setting wherein the organ is allocated to patient with $ID = x$. If an urgency based allocation system was in place, the organ would be allocated to patient #2, who has the lowest expected waiting list lifetime. Under a utility-based system, patient #1 would receive the transplant, since that patient is predicted to live the longest with the organ. Under a benefit-based allocation system, the organ would go to patient #3, who is predicted to have neither the greatest post-transplant lifetime, nor the lowest waiting list lifetime, but the greatest difference between the two.

Table VIII-1. Organ Allocation Example Computing Benefit to Patient Population

ID #	Future Lifetime (<i>patient</i>)			Total Future Lifetime (<i>patient population</i>) if assigned to candidate (ID #)
	WL	LT _x	B = LT _x -WL	
1	7	10	3	$17 = 10+2+5 = 3+(7+2+5)$
2	2	3	1	$15 = 7+3+5 = 1+(7+2+5)$
3	5	9	4	$18 = 7+2+9 = 4+(7+2+5)$

B = transplant benefit
 ID = patient identification number
 LT_x = predicted post transplant lifetime
 WL = predicted waiting list lifetime

Continuing with our examination of Table VIII-1, if indeed the organ was allocated to patient #3, as it would be under a survival benefit based allocation scheme, the lifetime lived by the patient population as a whole would equal 18 years; i.e., 7 (patient #1, who remains on the waiting list) plus 2 (patient #2, also left on the waiting list) plus 9 (post-transplant lifetime of patient #3). The population lifetime calculation can be done in a more transparent way as follows. In terms of total lifetime, the worst that could happen is

that the organ is not allocated, in which case all three patients remain on the waiting list. If the organ is allocated, the additional lifetime (i.e., transplant benefit) will be experienced by only the patient who receives the organ. Consider the left-most calculation in row 3 of Table VIII-1. If the organ is allocated to patient #3, each patient is predicted to receive their waiting list lifetime (7+2+5), and patient #3 is predicted to receive their WL lifetime, plus the gain in lifetime due to the transplant (4 years). Since, irrespective of to whom the organ is allocated, each patient is predicted to receive at least their waiting list lifetime, the maximum gain to the patient population as a whole will occur if the patient with the greatest benefit score receives the organ. Although this example considers a very simple scenario, the main ideas extend to more general settings.

Currently in the United States, deceased-donor livers are allocated based primarily on medical urgency. Specifically, acute liver failure patients (Status 1) are given top priority, while chronic end-stage liver disease patients are sequenced on the liver waiting list in decreasing order of Model for End-stage Liver Disease (MELD) score (4, 5, 6). In February 2002, the MELD system replaced the Status system, which was based largely on the Child-Turcotte-Pugh (CTP) score (7, 8). Both the CTP and MELD systems are based on medical urgency, since both utilize scores which are intended primarily to reflect waiting list mortality. Although not a transplant benefit based system (since post-transplant outcomes are not considered), the CTP system represented a great improvement over a system that did not prioritize waiting list patients based on their characteristics such as, for example, a system that ranks patients based on waiting time. Several articles have compared the MELD and CTP scores (9). Almost always, in instances where the analysis found a significant difference between the two scores, MELD was found to more accurately predict waiting list mortality (6). In many cases, however, MELD and CTP scores were not significantly different as predictors of waiting list mortality, due perhaps to inadequate sample size. However, even if one believed the two scores to be equally predictive of waiting list mortality, a system based on MELD would better achieve the objective of urgency-based allocation, since MELD has a finer gradation of risk. Ties are essentially broken by waiting time under either system, meaning that ranks based on the CTP score, which produces more ties, would be more influenced by waiting time. Despite its initial appeal as being equitable, allocation by waiting time identifies patients who have already survived the longest on the waiting list and, in some cases, selects patients for transplantation who need the organ the least. In sum, one would expect that CTP-based allocation is much closer to optimal than allocation by waiting time and that allocation by MELD constitutes a further considerable improvement.

Although the MELD system has proven effective, it was not designed to reflect post-transplant survival. The persistent shortage of donor livers increases the pressure to make the best possible use of those available, which implies that in addition to urgency, utility be considered. As such, the Organ Procurement and Transplantation Network (OPTN) Liver and Intestine Committee is currently evaluating a transplant survival benefit based system of allocating deceased-donor livers to chronic end-stage liver disease patients. We must emphasize that the development of a benefit-based allocation system is a work in progress. This chapter represents the current state of the proposal at the time of its writing. In terms of evaluating and testing the proposed system, much work remains

before implementation can occur. Several important decisions have yet to be made, and some which have been made are subject to modification.

The remainder of this chapter is organized as follows. In *Measuring Transplant Survival Benefit*, we discuss the quantification of transplant survival benefit and describe the currently proposed benefit score. The post-transplant and waiting list survival models are discussed in *Post-Transplant Survival Model* and *Waiting List Survival Model*, respectively. We evaluate the proposed transplant benefit score in *Analysis Of Proposed Liver Transplant Survival Benefit Score*, including comparisons to the MELD and various other scores. We evaluate the implications of benefit based allocation through micro-simulation in *Evaluation of Benefit-Based Allocation Via Simulation*. A Discussion concludes the chapter.

Measuring Transplant Survival Benefit

We quantify the liver transplant survival benefit for a given candidate as that candidate's five-year mean lifetime with a transplant (specifically, with the organ to be allocated) minus his/her five-year mean lifetime without a transplant. Thus, each time an organ is to be allocated, the transplant benefit score would be computed for each chronic liver failure patient active on the waiting list. In this sense, the scores are patient- and organ-specific. After computing each of their benefit scores, all active patients would then be sequenced in decreasing order of benefit score. In principle, the calculation of the benefit score is straightforward. For any donor/patient combination, a predicted post-transplant survival curve is available (described in *Post-Transplant Survival Model*), as is a predicted waiting list survival curve (*Waiting List Survival Model*). In each case, the predicted future lifetime is the area under the survival curve out to five years, while the benefit score is then the difference between those two predictions. For example, if the area under the first 5 years of a patient's post-transplant survival curve equals 3.5, then that patient is expected to live 3.5 of the next 5 years with the transplant. If the patient's benefit score equals 1.5, then it is predicted that out of the next 5 years, the patient will live an extra 1.5 years with the transplant compared to the scenario where the patient receives no transplant. That is, the area between the post-transplant and waiting list survival curves (both followed out to the five-year point) equals 1.5 years. The calculation is truncated at five years for two reasons. First, the available data provided five years worth of pertinent follow-up. Second, lifetime distributions are often skewed far to the right (i.e., the histogram has a long right tail) and the unrestricted mean would be too heavily influenced by this tail.

In *Post-Transplant Survival Model* and *Waiting List Survival Model* we describe the post-transplant and waiting list survival models, respectively. Before doing so, it is useful to briefly compare our proposed approach to estimating transplant benefit to others in the literature. Several authors have quantified the survival benefit of liver transplantation (1, 10) and kidney transplantation (1, 2, 11, 12). Each of these works used a single Cox regression model with "transplant" coded as a binary indicator. That is, one model applies to both the waiting list and the post-transplant deaths, with the regression parameter corresponding to the transplant 0/1 indicator used to quantify the covariate-adjusted

survival benefit of transplantation. As mentioned previously, we use separate models for waiting list and post-transplant survival, similar in spirit to the Lung Allocation Score (13) and in the proposed Life Years From Transplant (LYFT) score to be used in deceased-donor kidney allocation (14).

Having discussed our proposed transplant survival benefit metric, we describe in the next two sections the post-transplant and waiting list models used in its calculation. Each step in the process of building and evaluating the survival models was carried out in consultation with the OPTN Liver and Intestine Committee, as well as SRTR clinicians, surgeons and biostatisticians.

Post-Transplant Survival Model

The transplant study population included patients who received a deceased-donor liver transplant between September 1, 2001 and December 31, 2007. Follow-up began (time 0) at the date of transplant and ended at the earliest of death, re-transplant or loss to follow-up. The death event was considered to be the earliest of death and liver re-transplant. While there is no question any lifetime (post-transplant or otherwise) ends upon death, whether post-transplant lifetime ends at re-transplant requires more careful thought. Basically, one organ is allocated at a time, and of interest is a potential recipient's predicted lifetime with that particular organ. Therefore, in the post-transplant modeling, if a patient was re-transplanted, their lifetime with the original organ was considered to have ended.

Post-transplant survival was modeled using Cox regression (15). Although several regression models are available for survival analysis, the Cox proportional hazards model dominates the biomedical literature. The original paper in which the Cox model was proposed (15) is among the most cited papers in science. Covariate selection was a non-automated form of backward elimination. Specifically, we started the model-building process by fitting a model that contained every covariate suspected of affecting patient or graft survival. This set of covariates consisted of all covariates included in the Program-Specific Report (PSR) models (used to compare center-specific mortality to the covariate-adjusted national average), and/or included in previous SRTR analyses of liver transplant mortality or survival benefit, as well as various additional covariates suggested by members of the SRTR or the OPTN Liver Committee.

The final model consisted of recipient factors: creatinine, albumin, sodium, age, diagnosis, diabetes, dialysis, hospitalization status, previous liver transplant, mechanical support, portal vein thrombosis, previous abdominal surgery, hepatitis C; donor factors: age, cause of death, donation after cardiac death; and transplant factors (i.e., donor/recipient factors): cold ischemia time, and whether or not the transplant represented a regional or national share. In Table VIII-2, we list the recipient factors included in the post-transplant mortality model. There is a baseline death rate that applies to all patients, and the hazard ratios (HR) listed in Table VIII-2 are the multipliers of this baseline death rate. For example, a patient with a previous liver transplant (HR=1.60) has a death rate 60% greater than a patient who would be receiving their first liver transplant.

Table VIII-2. Recipient Characteristics Used in Post-transplant Mean Lifetime Prediction

Covariate	Hazard Ratio	P
Age – 18	1.01	0.47
(Age – 18) × I(age > 18)	0.99	0.33
(Age – 55) × I(age > 55)	1.03	<0.0001
Diagnosis = cholestatic	0.80	0.0002
Diagnosis = non-cholestatic	0.85	0.0002
Diagnosis = acute hepatic necrosis	0.86	0.10
Diagnosis = malignancy	1.20	0.17
Diagnosis = biliary atresia	0.69	0.007
Diagnosis = HCV	1.25	<0.0001
Hepatocellular carcinoma	1.15	<0.0001
Diabetes	1.16	<0.0001
Dialysis (at last WL status update)	1.17	0.005
Hospitalized: ICU	1.34	<0.0001
Hospitalized: not ICU	1.16	<0.0001
Previous liver transplant	1.60	<0.0001
Mechanical support	1.54	<0.0001
Previous abdominal surgery	1.13	<0.0001
Portal vein thrombosis	1.32	<0.0001
Creatinine: 4 th quartile	1.19	<0.0001
Albumin: 1 st quartile	1.15	<0.0001
Growth failure	1.34	0.04

I(A) = applies to patients with characteristic A
HCV = hepatitis C
ICU = intensive care unit
WL = waiting list

Donor factors included in the benefit score calculation are listed in Table VIII-3. Note that the "Donor age ≥ 60 and recipient HCV" entry in Table VIII-3 (fifth row) represents an interaction term. Specifically, from the fourth row, having a donor age ≥ 60 confers a 44% increase in risk, since the hazard ratio is estimated at HR=1.44. If, in addition, the recipient is HCV+, then donor age ≥ 60 results in a greater than twofold increase in post-transplant mortality risk, since the hazard ratio is given by $HR=1.44 \times 1.41=2.03$.

Table VIII-3. Donor Characteristics Used in Post-transplant Mean Lifetime Prediction

Covariate	Hazard Ratio	P
Donor age: < 18	0.89	0.03
Donor age: 40-49	1.16	<0.0001
Donor age: 50-59	1.35	<0.0001
Donor age: ≥ 60	1.44	<0.0001
Donor age: ≥ 60 and recipient HCV	1.41	<0.0001
Donor race: African American	1.14	0.0002
Donor race: Hispanic	1.15	0.0002
Cause of death: anoxia	1.07	0.07
Cause of death: cerebrovascular accident	1.12	0.0002
Donation after cardiac death	1.78	<0.0001

HCV = hepatitis C

Outputs from the final post-transplant Cox model include the hazard ratios corresponding to each covariate, and the baseline survival. The latter can be interpreted as the survival curve for a recipient/donor combination whose characteristics are described by the reference levels of each covariate. Combining the baseline survival (which applies to all patients and donors) and the HRs, a predicted survival curve can be constructed for any patient/donor combination. Five-year post-transplant life expectancy can then be computed as the area under the survival curve, out to five years.

To evaluate the fit and predictive ability of the post-transplant survival model, we computed the index of concordance (or C statistic) based on all patients transplanted in the calendar year 2005. The C statistic is an estimator for the percentage of times that the model correctly predicts which of two patients will die first. To compute the C statistic, the denominator equals the number of pairs of patients where the ordering of the death times is observed. The numerator is the subset of the denominator where the ordering of the death times observed is concordant with that predicted by the model. If $C=1$, then the model is estimated to perfectly predict the first of any two patients to die. If $C=0.5$, the model predicts the first of two patients to die as well as would the toss of a coin. For the post-transplant model, $C=0.63$, indicating satisfactory albeit not exceptional predictive ability. To cross-validate the post-transplant survival model, we randomly split the data set, fitted a Cox model to one half of the patients and computed the C statistic on the other half. We repeated this exercise 10 times, and the average C statistic equaled 0.63, indicating that our internal validation did not appear to overstate the ability of the model to predict post-transplant survival for future patients.

Waiting List Survival Model

To model waiting list survival, the study population consisted of 10 cross-sections of patients. The cross-section dates being May 1 and November 1 of each of calendar years 2002 to 2006. To be included in a particular cross-section (e.g., 11/01/2003), a patient would have to be alive and active on the waiting list as of 11/01/2003. Status 1 patients were excluded. We formed the study population using cross-sections since, when used for allocation, the model will be applied to cross-sections of patients (those active on the waiting list on a given date), as opposed to cohorts of patients. The survival time was defined as time since cross-section, with time previously survived on the waiting list included as a covariate in the model. After being included in a cross-section, patients were censored at the earliest of loss to follow-up or receipt of a liver transplant. To clarify, being active on the waiting list was a requirement to be included in a cross-section. However, subsequent de-activation would not be treated as a censoring event.

As mentioned in the preceding paragraph, patients were censored upon receipt of a liver transplant. Under the MELD system, patients at the highest risk of waiting list death also generally have the highest transplant rate. The potential bias was corrected through inverse probability of censoring weighting (16, 17), a well established method to overcome dependent censoring. Specifically, a time-dependent weight is applied to each cross-section patient, with the weight being equal to the inverse of the probability that the patient remains untransplanted. Basically, patients with a higher (lower) probability of receiving a liver transplant are assigned a higher (lower) weight, to balance out the fact that relatively less (more) follow-up on such patients is actually observed. To compute the weights, we need a time to transplant model which contains all covariates in the waiting list survival model, plus time-dependent MELD and organ procurement organization.

The statistical methodology and other technical issues surrounding our approach to modeling waiting list mortality will be described in a separate report. The underlying ideas are given in some detail in the literature (12, 18). Several SRTR (2, 10) and other analyses (19) have been carried out using closely related methods. Related methodologic material can be found in Liang and Zeger (20), Wei et al. (21) and Zheng and Heagerty (22).

Cox regression was used to model waiting list survival. Covariate selection proceeded in a manner similar to that employed for the post-transplant survival model (Post Transplant Survival Model). Patients were classified based on their most recent measurement as of the cross-section date for lab measures such as the MELD components, albumin and serum sodium. This is appropriate since, when the score is computed in practice, only current and previous lab measurements will be known. Covariates were also defined for the slopes of each patient's prior bilirubin, creatinine, international normalized ratio (INR) and albumin values. Patient characteristics in the final waiting list model included creatinine, bilirubin, INR, albumin, sodium, age, body mass index, previous time on waiting list, diagnosis, diabetes, dialysis, medical condition at listing, and prior history of malignancy. Table VIII-4 lists the patient characteristics included in the waiting list lifetime prediction.

To evaluate its discriminatory ability, we computed the C statistic for the waiting list survival model. Based on the cross-section of patients active on the waiting list on May 1, 2004, $C=0.74$, indicating fairly good predictive ability. The cross-validation of the waiting list survival model proceeded analogously to that described earlier for the post-transplant model; i.e., randomly splitting the data in half, then fitting the model to one half and computing the C statistic on the other half. Averaging over 10 random splits, the C statistic equaled 0.74.

Table VIII-4. Characteristics Used in Waiting List Mean Future Lifetime Prediction

Covariate	Hazard Ratio	P
Previous time on WL: 6 months - 1yr	0.94	0.0007
Previous time on WL: 1 – 2 yrs	0.87	<0.0001
Previous time on WL: 2 – 3 yrs	0.83	<0.0001
Previous time on WL: 3 – 4 yrs	0.78	0.0002
Previous time on WL: 4 – 5 yrs	0.66	0.0005
Previous time on WL: > 5 yrs	0.47	0.03
BMI: 30 - 35	0.87	0.005
BMI: > 35	0.90	0.02
Creatinine*	2.81	<0.0001
Bilirubin*	1.91	<0.0001
INR*	2.93	<0.0001
Albumin	0.47	<0.0001
Sodium: ≤ 131	1.89	<0.0001
Sodium: 132 - 137	1.22	<0.0001
HCC	1.51	<0.0001
Diagnosis = HCV	1.15	0.0006
Diabetes	1.26	<0.0001
Prior malignancy	1.24	0.036
Growth failure	2.11	0.016
Age – 18	1.031	0.08
(Age – 70) \times I(age > 70)	1.045	0.18
Slope: creatinine	1.095	0.09
Slope: bilirubin	1.032	0.26
Slope: albumin	1.042	0.07

BMI = body mass index

HCC = hepatocellular carcinoma

HCV = hepatitis C

I(A) = applies to patients with characteristic A

INR = international normalized ratio

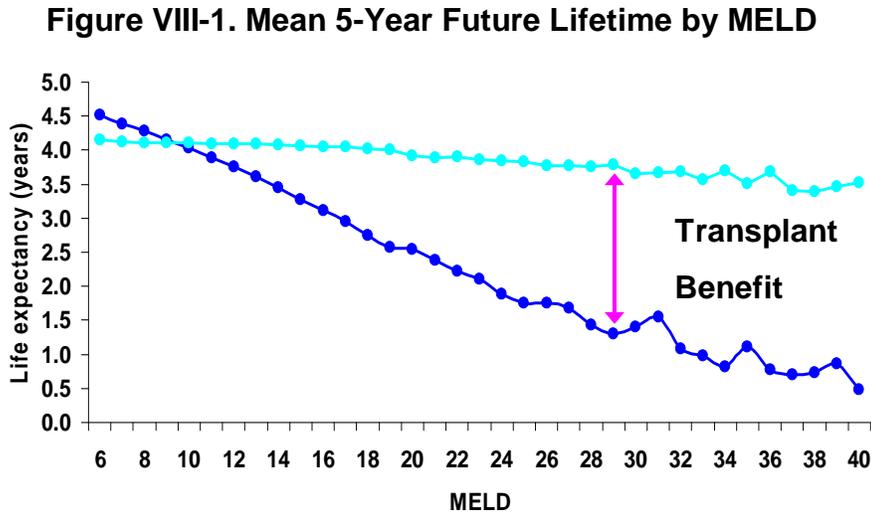
WL = waiting list

* coded as $\log(1+x)$

Analysis of Proposed Liver Transplant Survival Benefit Score

Our objective in this section is to examine the proposed liver transplant benefit score, and to compare it to the MELD score. We used the 10 cross-sections of patients already selected as the study population for the waiting list survival model. Within each cross-section, for each selected patient, we used the most recent lab MELD score and computed the proposed transplant survival benefit based on a typical liver donor; specifically, a donor with characteristics at the reference level for all categorical donor factors and approximately equal to the median of all continuous donor factors.

In Figure VIII-1, we plot mean five-year predicted future waiting list lifetime by MELD, with patients grouped by individual MELD score. It is clear that predicted waiting list lifetime strongly decreases as MELD increases, which would be expected since MELD is a very strong predictor of waiting list mortality. Mean predicted 5-year future waiting list lifetime equals approximately 4.5 years for patients with a MELD score of 6, meaning that on average, a patient with a MELD of 6 would be expected to live 4.5 of the next 5-years. In contrast, mean 5-year waiting list lifetime is only 0.5 years for patients on the waiting list with a MELD score of 40.



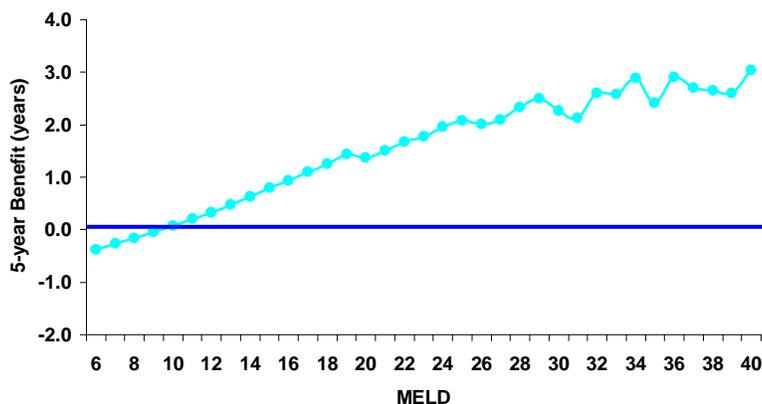
Source: SRTR Analysis. Data as of August 2008.

Also plotted in Figure VIII-1 is mean five-year post-transplant lifetime. Post-transplant life expectancy decreases as MELD increases, although the strength of the decrease is much less than for waiting list lifetime. Patients with a MELD score of 6 are expected to live an average of 4.1 years out of the next five, if they receive a liver transplant; this is an average of 0.4 years less than if they remain on the waiting list. Patients with such a low MELD score do not, on average, benefit from liver transplantation because their

waiting list mortality is too low to offset the high post- and peri-operative mortality risk associated with liver transplantation.

Average five-year transplant benefit is the distance between the two lines in Figure VIII-1. We plot average survival benefit by integer MELD score in Figure VIII-2. Average survival benefit increases steadily as MELD increases; consistent with the fact that as MELD increases, mean waiting list lifetime decreases at a much greater rate than mean post-transplant lifetime. On average, patients with MELD less than 10 have negative benefit scores, indicating reduced lifetime post-transplant, based on a five-year time horizon. It is estimated that patients with a MELD score of 40 gain an average of 3.0 (out of a possible five) future years if they receive a liver transplant.

Figure VIII-2. Mean 5-Year Transplant Benefit by MELD

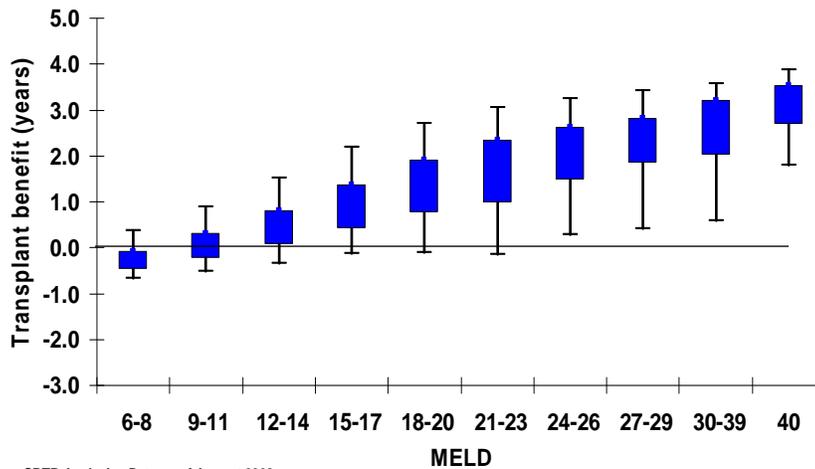


Source: SRTR Analysis. Data as of August 2008.

Of note, Figures VIII-1 and VIII-2 are based on averages. Overlooking this fact could lead one to believe that allocating livers by survival benefit would essentially amount to allocating by MELD, since the trends appear to be in the same direction. However, it is possible that patient X could have a much greater MELD score than patient Y, yet have a lower benefit score. True, the MELD score consists of three very strong predictors of waiting list mortality. However, recent evidence suggests that the MELD components are not weighted optimally in the MELD formula (23). Moreover, one of the MELD components, creatinine, is also a very strong predictor of post-transplant mortality (24). Since creatinine predicts both waiting list and post-transplant mortality, its effect on the benefit score is of less magnitude than one might expect. Moreover, as identified in Measuring Transplant Survival Benefit and Post-Transplant Survival Model, several factors in addition to the MELD components predict waiting list and/or post-transplant mortality. As a result, there is considerable variability in the distribution of the benefit scores at any MELD score, as evidenced by the box and whisker plots in Figure VIII-3.

Note that the boxes in Figure VIII-3 contain the middle 50% of the data (spanning the 25th and 75th percentiles), while the whiskers contain the middle 90%. The pattern in the averages observed in Figure VIII-2 are consistent with the trend repeated in the boxes in Figure VIII-3. However, what is much more prominent is the degree of overlap; not just among the distributions in adjacent MELD categories, but among MELD scores three and four categories apart. Although, MELD 6-8 patients do not, on average, benefit from liver transplantation (Figure VIII-3), approximately 20% of such patients do have positive benefit scores (Figure VIII-3). Similarly, there are patients with MELD ≥ 21 with negative benefit scores (Figure VIII-3) although, on average (Figure VIII-2) the benefit is quite strong in this high MELD subgroup.

Figure VIII-3. Transplant Benefit by MELD Box Plots



Source: SRTR Analysis. Data as of August 2008.

We computed the rank correlation between the proposed transplant benefit score and various other scores (Table VIII-5). The rank correlation (also known as the Spearman correlation) equals the more commonly used correlation coefficient (also known as the Pearson correlation) computed on the ranks, rather than the actual scores. The rank correlation is bounded by -1 and 1. A value of 1 (-1) indicates that as one score increases, the other increases (decreases); values close to 0 indicate no correspondence between the two scores. The rank correlation between the proposed benefit score and MELD score is 0.67, which appears to be consistent with Figure VIII-3. Although the benefit and MELD scores are related, it is clear that one score is not duplicating the other judging by the overlap in benefit scores among MELD categories. Based on the rank correlation being only 0.67, and the overlap in Figure VIII-3, it appears that the ranks of waiting list patients would be altered considerably under a benefit-based allocation system. A rank correlation of 0.67 is perhaps closer to 1 than expected, since MELD considers waiting list mortality while the benefit score captures both pre- and post-transplant mortality. However, the rank correlation between predicted five-year waiting list lifetime (the

waiting list component of the benefit calculation) and MELD was only -0.72. This is further away from -1 than one might anticipate given that, in this case, the rank correlation is being computed on two quantities intended to measure urgency. The fact that the rank correlation between our five-year predicted waiting list lifetime and MELD is not closer to -1 reflects the fact that several factors in addition to MELD affect waiting list mortality.

Table VIII-5. Rank Correlation Among Scores

Scores	Rank Correlation
Transplant benefit score, MELD score	0.67
Predicted 5-year waiting list lifetime, MELD score	-0.72

In Table VIII-6, we further explore the ideas from Table VIII-5. The rank correlation between predicted five-year waiting list lifetime and the benefit score was calculated to be -0.89. The closeness of this particular rank correlation to -1 (which would indicate that the benefit score could be based on predicted waiting list lifetime alone) is consistent with the fact that covariate-adjusted post-transplant death rates are substantially less than (approximately one fifth) those on the waiting list (1).

Table VIII-6. Rank Correlation Between Transplant Benefit and Various Other Scores

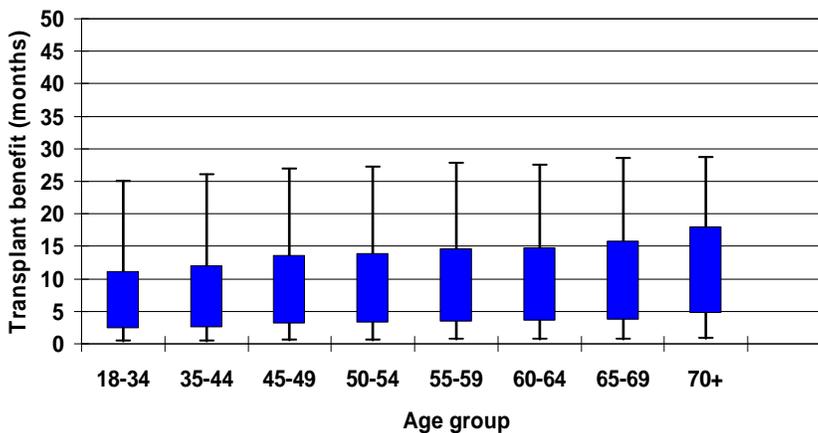
Score	Patients Included	Rank Correlation
Predicted waiting list lifetime	All	-0.89
Predicted waiting list lifetime	Benefit score > 0	-0.90
Benefit score, using MELD components only	All	0.68
Benefit score, using MELD components only	Benefit score > 0	0.61

As such, there appears to be less variation in post-transplant outcomes, since five-year survival on the waiting list is much less than that of post-transplant. If we restrict attention to patients with positive transplant benefit scores (those patients much more likely to obtain offers, under a benefit-based allocation system), the rank correlation between the score and predicted waiting list lifetime equals -0.90 and, hence is approximately equal to that based on all patients.

To obtain a summary measure of the importance of factors other than the MELD components, we computed the rank correlation between the proposed benefit score and a benefit score based on the MELD components alone: 0.68 (Table VIII-6). When restricted to patients with positive (proposed) benefit scores, this correlation decreased to 0.61 (Table VIII-6). In either case, the message is that the MELD components strongly influence but certainly do not dominate the proposed benefit score.

As a follow-up to Figure VIII-3, in Figure VIII-4 we display box and whisker plots for benefit scores by age group. Based on Figure VIII-4, it appears that the proposed benefit score is not strongly influenced by age, judging by the apparent similarity of the distributions across age groups. This may seem strange, since age covariates were selected for inclusion in both the waiting list and post-transplant models. Two things are important, in this regard. First, since age predicts both pre- and post-transplant survival; it would have a stronger influence on the benefit score if it predicted one and not the other, or if it predicted one much more strongly than the other. Second, although age is predictive of mortality, it is much less predictive than several factors, such as the MELD components and albumin.

**Figure VIII-4. Benefit Scores by Age
Patients with Benefit > 0**



Source: SRTR Analysis. Data as of August 2008.

We sought to rank each patient covariate in the benefit score in terms of relative importance. To do so, we took the cross-section of patients active on the waiting list on May 1, 2004 and computed the benefit score for each patient using a randomly selected donor from calendar year 2004. We then fitted a linear regression model, with the benefit score serving as the response variate and each patient covariate serving as the predictor variates. We judged the importance of each covariate by the percentage of variation in the benefit scores it explained. The results of this exercise are listed in Table VIII-7. Each row can be interpreted as the contribution of that covariate, after factoring out the contribution of all covariates listed above.

Table VIII-7. Relative Importance of Covariates to Benefit Score

Covariate	% Score Explained	Cumulative % Explained
Albumin	53.1 %	53.1 %
Bilirubin	14.7	70.8
Donor age	7.7	78.5
Creatinine	5.3	83.9
Recipient age	4.8	88.6
INR	2.9	91.1
DCD	1.8	93.3
Previous liver transplant	1.8	95.1
Diagnosis	1.7	96.8
Remaining, combined	3.4	100.0

DCD = donation after cardiac death
 INR = international normalized ratio

Note that the ordering was based on a sequence of linear regression models to determine the most important covariate (which happened to be albumin), followed by the second most important covariate (after factoring out the contribution of albumin); and so on. Albumin accounted for over nearly half (53%) of the variation in benefit scores. After factoring out the contribution of albumin, bilirubin accounted for an additional 15% of the variation. The next most important factor was donor age (8%). The only remaining covariate which accounted for greater than 5% of the variation in the benefit scores was creatinine (5.3%). Note that recipient age accounted for less than 5% of the variation in scores. Together, the seven factors listed in Table VIII-7 explained almost 97% of the variation in the benefit scores.

The proposed transplant benefit score is based on a five-year timeline, with the truncation point largely chosen based on data availability. To assess the sensitivity of the patient rankings to the five-year truncation time, we computed rank correlations between the proposed score and those based on 1-, 3-, and 10-year lifetimes (Table VIII-8). Survival models for the 10-year life expectancies were based on extrapolations which assumed that the average hazard during years 5-10 was equal to that during year 4-5. As shown in the top row of the matrix in Table VIII-8, the 1-year scores have decreasing rank correlation with the remaining time horizons, with the correlation decreasing as the time horizon is extended. The 3-, 5- and 10-year scores have pair-wise rank correlations very close to 1, indicating that the ordering of patients appears to be quite insensitive to the time horizon, provided the truncation point is three or more years.

Table VIII-8. Rank Correlations Between Benefit Scores Using Different Truncation Points

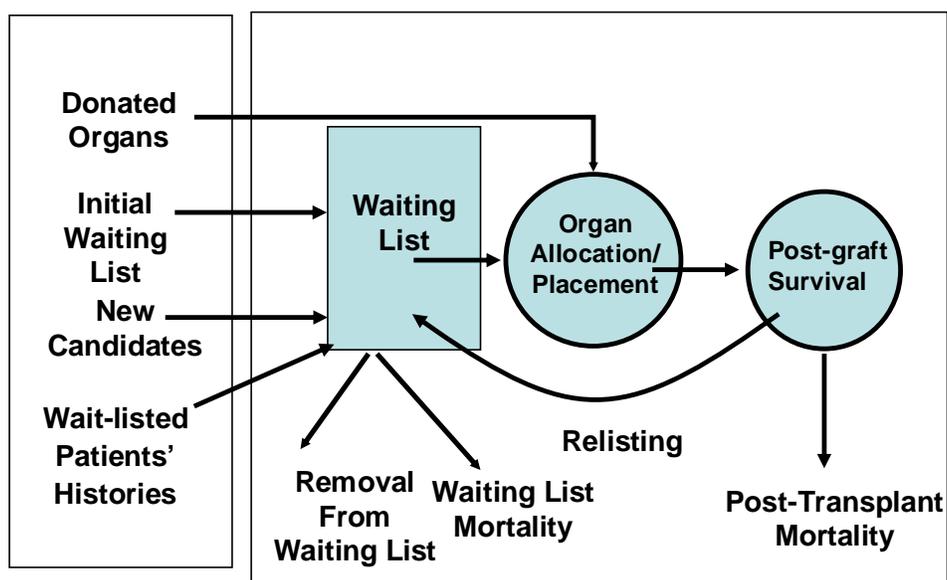
RANK Correlation	1-year	3-year	5-year	10-year*
1-year	1	0.96	0.93	0.91
3-year		1	0.995	0.990
5-year			1	0.997
10-year				1

*Based on extrapolation.

Evaluation of Benefit-Based Allocation Via Simulation

We evaluated the performance of the proposed liver transplant survival benefit score using the Liver Simulated Allocation Model (LSAM) (25). LSAM is part of a family of simulated allocation models (SAMs) developed by the SRTR. Similar models exist for kidney/pancreas allocation (KPSAM) and the thoracic organs (TSAM). The SAMs have been used previously many times to assess the future effect of proposed changes in allocation policy. A schematic diagram of the general flow of data and event processing by LSAM is displayed in Figure VIII-5. The model starts with the actual candidates on the waiting list; in our case, on January 1, 2006. The model runs for one full calendar year, processing new wait-listings (i.e., initial listings, re-listings), transplants, deaths and removals from the waiting list for reasons other than death and transplant.

Figure VIII-5. Liver Simulated Allocation Model (LSAM) Event Processing



Source: SRTR 2008

The primary role of LSAM is to evaluate the changes in experience over the calendar year due to changes in the allocation system. Through LSAM, one can assess the impact (e.g., number of deaths) on one year's worth of experience of changing the allocation rules, which are programmed directly into LSAM for each run. Essentially, a waiting list lifetime (i.e., lifetime, in the absence of transplant) is simulated for each patient. This lifetime without transplant contains status history updates, inactive time and possibly removal. Deceased-donor livers arrive into the system and are allocated based on

whatever rules are programmed into a particular LSAM run. Each organ is offered to patients active on the waiting list, and each offer is accepted or rejected with a probability that depends on the patient and donor characteristics. Post-transplant experience is generated for patients who are transplanted, including death or graft failure and subsequent re-listing. Further details regarding LSAM are described in Thompson et al (25).

We evaluated three allocation schemes. The first is the set of rules currently in place at the time this chapter is being written. Note that the Share15 rule is incorporated into the current allocation scheme. The second allocation system also features MELD/PELD based allocation, but is based on regional sharing for all MELD scores; i.e., local is essentially eliminated as the basis for allocation, with region then being the first geographic level of offer. There are two reasons that 'total' regional sharing is of interest. First, it is currently being considered by the OPTN Liver and Intestine Committee. Second, the Share15 rule has no obvious analog under survival benefit based allocation, meaning that comparisons between regional sharing based allocation and the proposed benefit-based system are easier to interpret than comparisons between the current and benefit based system. The third allocation system evaluated is transplant benefit, again with regional sharing at all MELD scores.

Results of the LSAM modeling are averaged over 10 iterations. In Table VIII-9, we list the number of deaths, by allocation system, with the numbers in parentheses representing the difference between that system and the system in the column to the immediate left. Based on one calendar year of experience, benefit based allocation would result in 83 fewer waiting list deaths, 6 fewer post-transplant deaths and 13 fewer post-removal deaths, for a net saving of 102 lives. Naturally, a regional sharing and transplant benefit allocation systems would result in transplants to different sets of patients.

Table VIII-9. Liver Simulated Allocation Model (LSAM): Number of Deaths by Allocation System

Deaths	Current System	Regional Sharing	Transplant Benefit
Waiting list	1,660	1,602 (-58)	1,519 (-83)
Post-transplant	609	607 (-2)	601 (-6)
Post-removal	407	397 (-10)	384 (-13)
Total	2,675	2,606 (-69)	2,504 (-102)

As such, in Table VIII-10 we estimate the number of liver transplant-attributable life-years saved under each allocation system. Under a MELD/PELD based regional sharing system, we estimate that there would be 6,273 liver transplants and that the mean benefit score would be 1.63 years. Taking the product of these two numbers, there would be $6,273 \times 1.63 = 10,225$ life years saved. Under survival benefit based allocation, there are projected to be 80 fewer transplants, but the mean benefit score is 0.38 years greater than that under a regional sharing system. The result is an additional 2,223 life years. To better appreciate this calculation, recall that a patient's benefit score represents their predicted gain in lifetime (over the next five years) due to receiving a transplant. The total lifetime gained calculation applies the patient/organ-specific predicted gain in five-year lifetime to patients who actually receive a transplant. Naturally, patients who are not transplanted are predicted to receive no transplant survival benefit.

Table VIII-10. Liver Simulated Allocation Model (LSAM): Life-Years Saved by Allocation System

	Current System	Regional Sharing	Transplant Benefit
Mean benefit score at transplant	1.56	1.63 (+0.07)	2.01 (+0.38)
Number of transplants	6,330	6,273 (-27)	6,193 (-80)
Life-years saved	9,875	10,225 (+350)	12,448 (+2,223)

Discussion

In this chapter we describe and evaluate a proposed score currently being considered to serve as the basis for a transplant survival benefit based deceased-donor liver allocation system for chronic liver failure patients. The score is currently being considered by the OPTN Liver and Intestine Committee and may undergo modification prior to its field implementation. The proposed score is based on models of a single end-point, mortality; hence the phrase 'survival benefit'. Mortality has always been considered the most important outcome among organ failure patients. It may be possible to develop a benefit score that incorporates morbidity or quality of life measures. It is unclear how such a score should be constructed and, perhaps most importantly, it appears that reliable pertinent data are currently unavailable for outcomes other than death.

This chapter represents a natural extension of previous work on liver transplant survival benefit from the SRTR (1, 10). Merion et al (1) previously demonstrated that the mortality reduction due to liver transplantation increases with increasing MELD. Subsequently, Schaubel et al (10) demonstrated that liver transplant benefit depends not only on MELD, but on the donor quality, as quantified by the Donor Risk Index (DRI) (26). Rather than using only MELD and DRI, the current chapter proposes a method for computing liver transplant benefit using all pertinent patient and donor characteristics. While the work of Merion et al (1) and Schaubel et al (10) compared average patients, our current proposal focuses on calculations at the patient level. For example, Merion et al (1) indicated that, on average, patients with low MELD do not obtain survival benefit from liver transplantation. This is consistent with Figure VIII-3, which indicates that the average benefit score is negative at low MELD scores. However, Figure VIII-3 also indicates that there are patients at low MELD scores that do benefit from a liver transplant; likely not the patients who receive high DRI livers, judging by our previous work.

We estimate that a substantial number of life-years would be saved if transplant survival benefit based allocation were implemented. Comparing a MELD/PELD and benefit based allocations systems with regional sharing, we estimate that over 2,000 life-years would be saved based on one calendar year's worth of experience and only considering the first five years post-transplant. It should be noted that our calculation vastly underestimates the gain in life-years by the patient population in at least two important ways. First, only one calendar year of liver transplants was simulated by LSAM. Second, the benefit score predicts the gain in lifetime over the next five years, as opposed to the total gain in lifetime.

With respect to model evaluation, the C statistic reflects the ability of the model to correctly rank patients in terms of death rate. For the waiting list model, $C=0.74$ indicating that, among pairs of patients, the model correctly ranked the death times approximately three times as often as it did so incorrectly. The corresponding result was less encouraging for the post-transplant model, with $C=0.63$. It is possible that the most important determinants of post-transplant survival are not known at the time of transplant; e.g., events that occur in the post- and peri-operative period. Additionally, since both patient and donor factors are important to post-transplant outcomes, unlike

waiting list survival, there are two dimensions upon which to misclassify liver transplant recipients. For example, although variables like diabetes, mechanical support and previous abdominal surgery are pre-defined by the OPTN, the classifications for such variables are broad and somewhat subjective. It is thus possible that heterogeneity may exist across centers with respect to the definitions actually applied. The post-transplant model is under continued investigation, with the objective of obtaining a model with better discriminatory power using the observed data.

An important concern is how the switch from an allocation system based on MELD/PELD to a system based on a benefit score would affect pediatric patients. Under the MELD/PELD system, patients aged less than 12 years are given a considerable advantage over adults in terms of priority. In particular, a patient with PELD x faces considerably lower waiting list mortality than a patient with a MELD of x . The advantage given to pediatric patients is implicit, in the sense that no modification is made to convert the MELD score into a PELD score for patients age <12 . Instead, MELD and PELD scores are computed using different formulae, knowing that MELD more accurately reflects waiting list mortality risk than PELD. Given that the advantage given to pediatric patients appears to be well accepted by the liver transplant community, it would presumably be desirable for the advantage to patients age <12 to be preserved under benefit-based allocation. Currently, the same benefit score is computed for pediatric and adult patients, with no explicit modification built in to advantage patients age <12 . One possibility is to add a certain constant (e.g., 0.5 years) to the calculated benefit score for pediatric patients. A second possibility is to down weight the role of waiting list life expectancy, which is known to be quite high among pediatric patients. For example, if the benefit score is calculated as $B = LT - WL$ for adult patients (where LT and WL represent post-transplant and waiting list life expectancy, respectively), then we could set $B^* = LT - 0.75 \times WL$ for pediatric patients. Preliminary LSAM results indicate that the percentage of deceased-donor kidneys transplanted to pediatric patients would not change under benefit-based allocation (data not shown). The most likely reason for the stability across allocation systems, with respect to the percentage of transplants allocated to pediatric patients, is the strong impact of donor-to-recipient size-matching and prioritization to pediatric patients of organs from pediatric donors.

Another important consideration in the transition to benefit-based allocation is the issue of exception scores. Currently, patients may apply to their regional review board to have their allocation MELD score increased. If granted, such patients will be prioritized based on their 'exception' (as opposed to their calculated) MELD score. The intention is that exception scores be granted in cases where the patient's calculated MELD score is known to understate their true waiting list mortality risk. The benefit score is intended to explicitly quantify the impact of each patient characteristic which affects the amount of additional lifetime a liver transplant would provide. Therefore, exception scores should play a greatly reduced role under a benefit-based allocation system, since most factors which, under a MELD/PELD based system, may have prompted an exception score application, would already be accounted for in the benefit score calculation. That is, patients with a certain condition would still get a boost in prioritization, provided that such a boost was consistent with the benefit score, which is evidence-based. It is possible that exception scores would still be granted under a benefit based system; e.g.,

for conditions known to affect waiting list survival, but which are too rare to reliably incorporate into the survival models used to build the allocation score.

Currently, the most frequently occurring exception is that for hepatocellular carcinoma (HCC), although exceptions are granted for many other conditions. As mentioned previously, HCC is one of the waiting list model covariates. Based on our waiting list model, it is estimated that HCC patients face a significant 51% increase in covariate-adjusted waiting list mortality. Since HCC does not significantly affect post-transplant survival and is therefore not included in the post-transplant life expectancy computation, HCC patients would be given an advantage under benefit-based allocation, relative to patients without HCC. This is not to imply that the increased priority offered to HCC patients would be greater under a benefit- (vs MELD) based allocation system. In fact, the opposite may be true, particularly since the currently applied boost to a MELD exception score of 22 for HCC patients is likely indefensible empirically. This issue is currently under investigation, again using LSAM modeling.

References

- 1.) Merion RM, Schaubel DE, Dykstra DM, Freeman RB, Port FK, Wolfe RA. The survival benefit of liver transplantation. *Am J Transplant* 2005; 5(2):307-13.
- 2.) Miles CD, Schaubel DE, Jia X, Ojo AO, Port FK, Rao PS. Mortality experience in recipients undergoing repeat transplantation with expanded criteria donor and non-ECD deceased-donor kidneys. *Am J Transplant* 2007; 7(5):1140-7.
- 3.) Wolfe RA, Ashby VB, Milford EL, Ojo AO, Ettenger RE, Agodoa LY, et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *N Eng J Med* 1999; 341(23): 1725-30.
- 4.) 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(4): 864-71.
- 5.) Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; 33(2): 464-70.
- 6.) Wiesner RH, McDiarmid SV, Kamath PS, Edwards EB, Malinchoc M, Kremers WK, et al. MELD and PELD: application of survival models to liver allocation. *Liver Transpl* 2001; 7 (7): 567-80.
- 7.) Child CG, Turcotte JG. Surgery and portal hypertension. In: The liver and portal hypertension. Edited by CG Child. Philadelphia: Saunders 1964:50-64.
- 8.) Pugh RNH, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the esophagus in bleeding oesophageal varices. *Br J Surg* 1973;60:648-52.
- 9.) Cholongitas E, Marelli L, Shusang V, Senzolo M, Rolles K, Patch D, Burroughs AK. A systematic review of the performance of the model for end-stage liver disease (MELD) in the setting of liver transplantation. *Liver Transpl* 2006 Jul;12(7):1049-61.
- 10.) 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(2):419-25.
- 11.) Ojo AO, Hanson JA, Meier-Kriesche H, Okechukwu CN, Wolfe RA, Leichtman AB, et al. Survival in recipients of marginal cadaveric donor kidneys compared with other recipients and wait-listed transplant candidates. *J Am Soc Nephrol* 2001; 12(3): 589-97.
- 12.) Schaubel DE, Wolfe RA, Port FK. A sequential stratification method for estimating the effect of a time-dependent experimental treatment in observational studies. *Biometrics* 2006; 62 (3):910-7.

- 13.) Egan TM, Murray S, Bustami RT, Shearon TH, McCullough KP, Edwards LB, et al. Development of the new lung allocation system in the United States. *Am J Transplant* 2006; 6(5 Pt 2):1212-27.
- 14.) Wolfe RA, McCullough KP, Schaubel DE, Kalbfleisch JD, Murray S, Stegall MD, et al. Calculating life years from transplant (LYFT): methods for kidney and kidney-pancreas candidates. *Am J Transplant* 2008; Apr;8(4 Pt 2):997-1011.
- 15.) Cox DR. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B (Methodological)*. 1972, 34 (2), 187-220.
- 16.) Robins JM, Rotnitzky A. Recovery of Information and Adjustment for Dependent Censoring using Surrogate Markers. AIDS Epidemiology-Methodological Issues, Proceedings of the Biopharmaceutical Section. In: eds. N. Jewell, N., K. Dietz and V. Farewell. American Statistical Association, 1992: 24-33.
- 17.) Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; 56(3): 779-88.
- 18.) Schaubel DE, Wolfe RA, Sima CS, Merion RM. Estimating the effect of a time-dependent treatment in by levels of an internal time-dependent covariate. in press: *Journal of the American Statistical Association*.
- 19.) Berg CL, Gillespie BW, Merion RM, Brown RS Jr, Abecassis MM, Trotter JF, et al. Improvement in survival associated with adult-to-adult living donor liver transplantation. *Gastroenterology* 2007; 133(6):1806-13.
- 20.) Liang KY, Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, 1986, 73(1), 13-22.
- 21.) Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 1989; 84, (n408):1065-1073.
- 22.) Zheng Y, Heagerty PJ. Partly Conditional Survival Models for Longitudinal Data. *Biometrics* 2005; 61, 379-391.
- 23.) Sharma P, Schaubel DE, Sima CS, Merion RM, Lok A.S.F. Re-weighting the Model for End-Stage Liver Disease score components. *Gastroenterology* 2008; 135(5): 1575-1581.
- 24.) Nair S, Verma S, Thuluvath PJ. Pretransplant renal function predicts survival in patients undergoing orthotopic liver transplantation. *Hepatology* 2002; 35:1179-85.
- 25.) Thompson D, Waisanen L, Wolfe R, Merion RM, McCullough K, Rodgers A. Simulating the allocation of organs for transplantation. *Health Care Manag Sci* 2004 Nov;7(4):331-8.

26. Feng S, Goodrich NP, Bragg-Gresham JL, Dykstra DM, Punch JD, DeRoy MA. Characteristics associated with liver graft failure: the concept of a donor risk index. *Am J Transplant* 2006; 6:783-790.