A Clinically Based Discrete-Event Simulation of End-Stage Liver Disease and the Organ Allocation Process

Steven M. Shechter, MS, Cindy L. Bryce, PhD, Oguzhan Alagoz, MS, Jennifer E. Kreke, MS, James E. Stahl, MD, CM, MPH, Andrew J. Schaefer, PhD, Derek C. Angus, MD, MPH, Mark S. Roberts, MD, MPP

Background**. The optimal allocation of scarce donor livers is a contentious health care issue requiring careful analysis. The objective of this article was to design a biologically based discrete-event simulation to test proposed changes in allocation policies. **Methods**. The authors used data from multiple sources to simulate end-stage liver disease and the complex allocation system. To validate the model, they compared simulation output with historical data. **Results**. Simulation outcomes were within 1% to 2% of actual results for measures such as new candidates, donated livers, and transplants by year. The model overestimated the yearly size of the waiting list by 5% in the last year of the simulation and the total number of pretransplant deaths by 10%. **Conclusion**. The authors created a discrete-event simulation model that represents the biology of end-stage liver disease and the health care organization of transplantation in the United States. **Key words:** liver transplantation; discrete-event simulation; simulation modeling; Monte Carlo simulation; organ allocation; patient survival; graft survival; policy analysis. **(Med Decis Making 2005;25:199–209)

Liver transplantation is the only viable therapy that has been demonstrated to enhance the quantity and quality of life for patients with end-stage liver disease (ESLD).[1] Unfortunately, the therapy is severely limited by the scarcity of donated livers, as evidenced by the more than 17,000 patients in the United States awaiting liver transplantation.[2] With this severe shortage of a life-saving resource, policies must be developed to determine the best way to allocate the donated livers.

Since 1984, the United Network for Organ Sharing (UNOS) has overseen the organ-matching process in the United States and has developed allocation policies.[2] Concerns about regional preference and patient prioritization have fueled a debate over the appropriate allocation mechanism. As a result, there have been multiple revisions in UNOS policies over the years.[1]

The debate surrounding allocation policy has occurred without rigorous, biologically based estimates of policy changes on outcomes. Conventional techniques such as randomized controlled trials are impractical for evaluating allocation policies because they would require many years to evaluate. Furthermore, it is unrealistic to divide a national waiting list into multiple lists to test various proposed policy changes. As an alternative to clinical trials, we developed a liver allocation model that uses discrete-event simulation (DES) to estimate the effects of policy

Received 2 January 2004 from the Department of Industrial Engineering (SMS, OA, JEK, AJS, MSR) and the Center for Research on Health Care (CLB, AJS, DCA, MSR), University of Pittsburgh, Pennsylvania; the Section of Decision Sciences and Clinical Systems Modeling, Division of General Internal Medicine, Department of Medicine, University of Pittsburgh School of Medicine, Pennsylvania (CLB, AJS, MSR); MGH–Institute for Technology Assessment, Massachusetts General Hospital, Boston (JES); and the CRISMA Laboratory, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pennsylvania (DCA). This work was supported through grant R01 HS09694 from the Agency for Healthcare Research and Quality and the National Institute of Diabetes and Digestive and Kidney Diseases and partially from grant DMI-0223084 from the National Science Foundation. Parts of this article were presented in abstract form at the 2001, 2002, and 2003 annual meetings of the Society for Medical Decision Making. Revision accepted for publication 25 June 2004.

Address correspondence and reprint requests to Mark S. Roberts, MD, MPP, Center for Research on Health Care, 230 McKee Place, Suite 600, Pittsburgh, PA 15213.

DOI: 10.1177/0272989X04268956

Table 1 Data Sources and Samples Sizes for Various Model Components

Model Component	Sample Size	Source
Disease progression module	1,997	Clinical records of patients awaiting liver transplantation at the University of Pittsburgh Medical Center (UPMC)
Relisted patient characteristics	655	Clinical records of patients relisted for transplant at UPMC
Patient generator	36,651	Candidate registry from the United Network for Organ Sharing (UNOS)
Survival estimation module	17,044	Liver transplant registry data from UNOS
Organ generator	17,044	Donor organ registry from UNOS
Quality-of-life module	95	Prospective evaluation of a cohort of patients awaiting liver transplantation at UPMC

changes. Our goal was to provide a clinically useful tool to test proposed changes to the allocation system prior to possible implementation.

Two earlier models designed to inform liver allocation policy in the United States simulated the listing process, organ availability, and donor-recipient matching.[3,4] Although calibrated with the same data, their results differed primarily because one model included gradual improvement in posttransplant survival rates whereas the other did not. In addition, both used transition probabilities between UNOS priority statuses (prior to 2002, there were 4 statuses that determined urgency level) as the only measure of disease progression. Consequently, the 2 models could explore only changes in allocation policies that maintained the existing prioritization rules.

This article describes the development, calibration, and validation of a biologically based simulation model of ESLD that is integrated with a model of the US liver allocation system. We modeled disease progression independent of the UNOS status assignments and can therefore test virtually any proposed allocation policy. We used our model to simulate the experiences of adult patients (16 years or older) with ESLD who were placed on the waiting list for a liver transplant between 1992 and 1996. We validated the model by comparing the simulation results with key UNOS statistics over the same period, including the median waiting time for a transplant, the length of survival of patients who receive transplants, the number of patients who die while waiting, and the mean time from listing until death.

METHODS

Choice of Modeling Method

DES is an analytic tool developed in industrial engineering that is designed to represent complex stochastic systems that include queues and competition for resources and investigate potential changes in those systems. It is most often used for industrial applications such as testing changes in manufacturing plants and communications systems. However, DES can also inform the potential life-altering consequences associated with health policy and clinical decision making[5] and has, in fact, been used to inform UNOS liver allocation policies.[3]

A model used to test organ allocation policies must allow the interaction between the model components to generate the waiting list and competition for organs. Statistical summary measures from an existing waiting list would not provide a clear understanding of how policy changes may affect outcomes. Instead, the actual composition, size, and characteristics of the waiting list must result from the execution of the model. Furthermore, the model must be able to predict the natural history of ESLD independent of the particular organ allocation prioritization rules that exist at any given time. Because standard decision-modeling techniques such as Markov models cannot represent the generation of queues nor allow for individual patients represented in the model to interact (or compete), we based our model on DES.

Data Sources

Data to calibrate the model came from several sources (Table 1). The simulation is driven by data available from the beginning of 1991 through the end of 1996. We used 36,651 patient records from the candidate listing data that included 17,044 records of patients who received transplants. In addition, because UNOS had only limited patient information between listing, transplant, and death, we used 1,997 patient data records from the University of Pittsburgh Medical Center (UPMC) to obtain detailed longitudinal information on patients' clinical characteristics, which is

required for the pretransplant natural history module. To generate clinical characteristics for patients requiring retransplantation, we used 655 UPMC records of relisted patients. Quality-of-life estimates were obtained from the literature[6] and derived from a prospective study of patients awaiting transplant at UPMC.[7]

Overview of the Model

To allow for modeling flexibility and faster execution times, we built a discrete event, Monte Carlo micro simulation of the liver allocation process in the C programming language. The model has 5 core modules: the patient generator, organ generator, pretransplant natural history, matching algorithm, and posttransplant survival (Figure 1). Users may set parameters to vary the conditions of the simulation, including the allocation matching algorithm, the number of simulation replications, the start year, or the initial random seed.

Patient Generator Module

The patient generator 1) creates patient arrivals to the waiting list; 2) assigns the patients various clinical and demographic attributes such as liver disease category, age, gender, and geographic region (Table 2); and 3) initializes patient-specific variables such as waiting time and quality-adjusted life years.

Patients arrive to the waiting list according to a nonstationary Poisson process with the arrival rate varying by year, as there was a rise in patient listings from year to year. Daily arrival rates for each year were obtained by dividing the annual UNOS listings during the 1992–6 period by 365. We assumed a Poisson process within each year as this is a common approach to modeling arrivals to a system.[8] Consistent with the Poisson process, the interarrival times between patient listings were distributed exponentially with the appropriate yearly rate.

UNOS assigns more than 70 separate diagnoses of ESLD; however, concerns about having adequate sample sizes for each group led us initially to aggregate these into 10 broader disease categories based on discussions with our National Clinical Oversight Committee (see the appendix). For the purpose of obtaining more reliable probability distributions for certain disease characteristics, we further aggregated these to form 5 major disease groups (Table 3). We generated other clinical and demographic attributes based on chi-square tests for independence and on clinical recommendations (data dependencies are shown in Figure 2). Geographic data were based on the 59 organ procurement organizations (OPOs) that formed 11 regions between 1992 and 1996.

Organ Generator Module

The organ generator 1) registers the arrival of each donated liver; 2) determines if the liver is used for adult transplantation, used for pediatric transplantation, or wasted; and 3) generates the characteristics of the liver (Table 2) that will be used in the donor-recipient matching process if the liver is intended for adult transplantation.

As with the patient arrivals, cadaveric livers arrive according to a nonstationary Poisson process, with the arrival rate varying by year, as there was also a rise in donated livers from year to year. Because our donor file consists only of livers that were transplanted into adults, we used UNOS estimates to increase our arrival rate to account for livers intended for pediatric transplant and livers that are wasted because they were not transplanted within a reasonable amount of time. Then, as each liver becomes available, a discrete distribution based on UNOS wastage and usage rates determines if and how the liver will be used. We used statistical methods and clinical expertise to form distributions for the different liver attributes, just as we did with the patient attributes. Donor age and gender are jointly derived from a distribution on 4 categories: young male (younger than 45 years), old male (older than or equal to 45 years), young female, and old female. A continuous age is then obtained for each donated liver by fitting a distribution using Arena© (Rockwell Software, Sewickley, PA) to the various ages within each of the 4 categories and then drawing from that distribution using the inverse transformation technique.[8] Donor race is conditioned on the age and gender category, and blood type is then based on race. Finally, presence of the cytomegalovirus antibody is dependent on donor age. After the liver attributes are assigned, the liver becomes available for donor-recipient matching.

Pretransplant Natural History Module

The natural history module simulates the progression of disease from the time a transplant candidate is placed on the waiting list until the occurrence of either transplantation or death in the absence of transplantation. It does this by periodically updating a patient's health status with respect to clinical variables that predict future survival of ESLD patients.[9,10] Variables updated by the model include 4 laboratory values (prothrombin time and levels of bilirubin, creatinine,

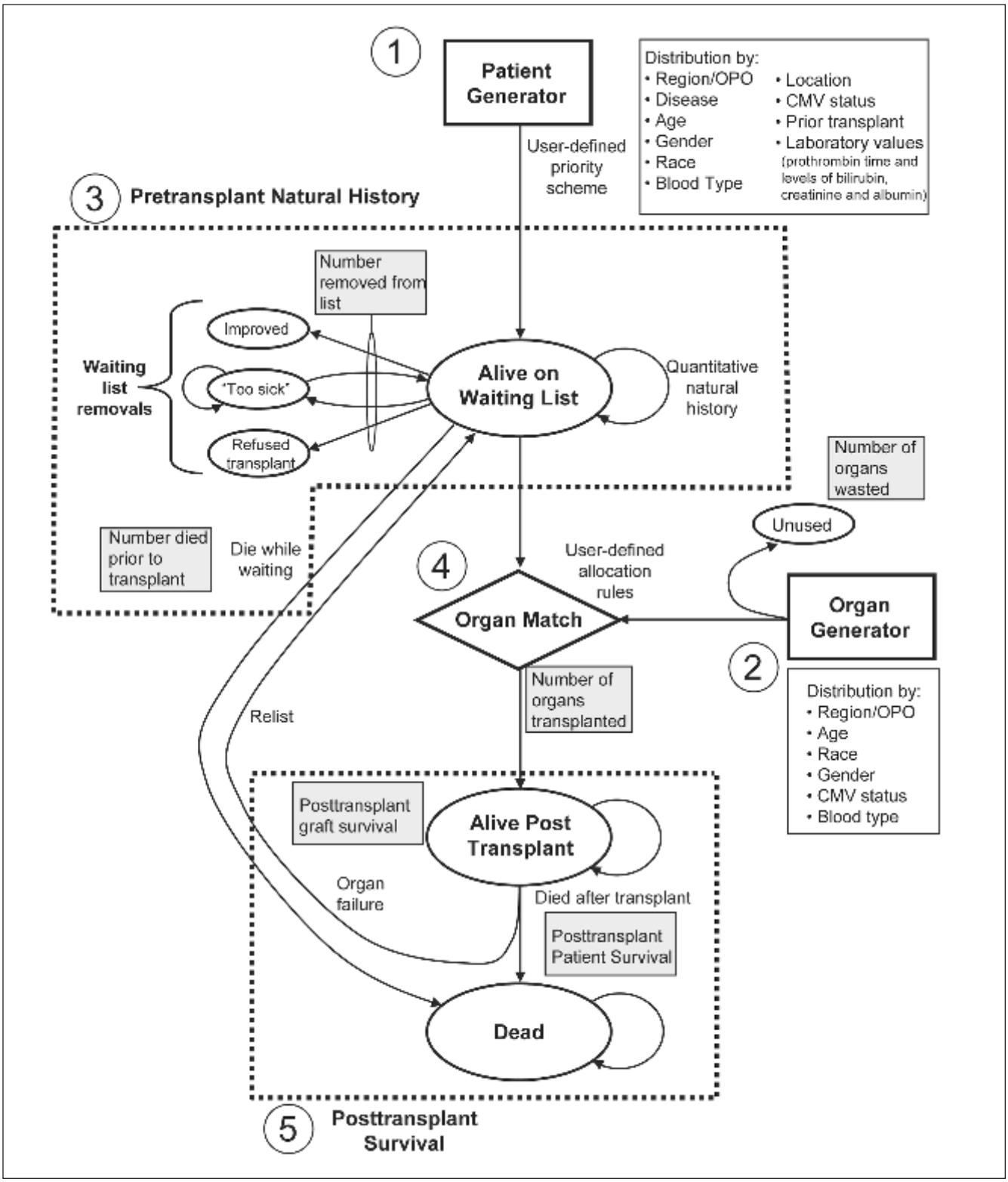

Figure 1 Liver allocation simulation model.

Note: The discrete-event simulation (DES) model has 5 core modules: the patient generator, organ generator, pretransplant natural history, donor-recipient matching algorithm, and posttransplant survival modules. OPO = organ procurement organization; CMV = cytomegalovirus.

Table 2 Study Variables and Data Sources

Variable	Data Source
Demographic characteristics of liver transplant candidate	
Age	UNOS candidate file
Gender	UNOS candidate file
Race	UNOS candidate file
Region and OPO	UNOS candidate file
Clinical characteristics of liver transplant candidate	
Category of end-stage liver disease (10 groups)	UNOS candidate file
Blood type	UNOS candidate file
Cytomegalovirus antibody status	UNOS candidate file
Prothrombin time	UPMC medical records
Bilirubin level	UPMC medical records
Creatinine level	UPMC medical records
Albumin level	UPMC medical records
On mechanical ventilator	UPMC medical records
Alive/dead indicator	UPMC medical records
Presence of encephalopathy	UNOS candidate file
History of nonliver transplant	UNOS candidate file
Location (home, hospital ward, intensive care unit)	UNOS candidate file
Demographic characteristics of liver donor	
Age	UNOS donor file
Gender	UNOS donor file
Race	UNOS donor file
Region and OPO	UNOS donor file
Clinical characteristics of liver donor	
Blood type	UNOS donor file
Cytomegalovirus antibody status	UNOS donor file

Note: UNOS = United Network for Organ Sharing; OPO = organ procurement organization; UPMC = University of Pittsburgh Medical Center.

and albumin) and the presence or absence of encephalopathy. These are the primary time-varying attributes that determine pretherapy health progression and posttherapy outcomes if a donor liver becomes available. Because we assumed that patients out of the hospital were more stable than those in the hospital, our model updates outpatients monthly and inpatients daily.

Although the average progression of disease in a cohort may be well behaved and easily estimated, the progression of disease in an individual is often chaotic. For example, a patient may undergo a slow and steady decline for a long period and then experience an acute exacerbation. To simulate the stochastic progression of disease in individuals, we used laboratory data from the records of transplant candidates at UPMC. Based on these data, we estimated cubic spline functions (continuous curves that fit discrete data) to obtain predicted laboratory values at intervening time points.[9,10] Every historical patient had 4 spline functions, 1 for each of the 4 laboratory values used as predictors of posttransplant survival. Each spline-derived series was decomposed into sequential pairs of laboratory values measured at specified time intervals. This method allows us to estimate daily and monthly changes in lab values regardless of whether measurements were actually taken at those intervals.[10]

When updating laboratory values for a particular simulated patient, the model searches the subset of actual patients with similar lab values (according to a set of nearness criteria), randomly chooses 1 patient from the sample, and returns an indication if the patient is dead by the next period, or, if the patient is alive, it returns the spline-estimated laboratory profile of that patient at the next time period. The appeal of this approach is that the estimated changes in laboratory profiles are proportional to actual changes observed in real data. Common changes occur more frequently, but unusual exacerbations remain possible.

If the profile returned at the next time period indicates that the patient died, the simulated patient is removed from the model, and we increment the deaths-while-waiting statistic. After testing revealed that our natural history data underestimated the number of such deaths, we included a yearly adjustment to the death rate. We did this by first considering the difference between the yearly death probabilities implied by the UNOS data and our simulation output and then randomly removing patients to make up for this difference. For patients out of the hospital, the yearly adjustment to the probability was an increase between 0.01 and 0.02 depending on the year, and for patients in the hospital, the increase was between 0.0002 and 0.0006.

Matching Algorithm Module

The matching algorithm module is flexible: It can represent allocation schemes based on any level of aggregation (from OPO level to national level) and can use any variable tracked in the model to determine a candidate's position within the waiting list. For baseline model validation, we implemented the UNOS allocation rule that existed between 1992 and 1996.[11] This policy grouped patients into 4 statuses and began its search for the most urgent patient, within the same OPO as the harvested liver, whose blood type was compatible with the liver. If no such patient was found at this OPO level, the algorithm expanded its search to

Table 3 Disease Categories

Classification		
5-Group	**10-Group**	**Category**
1	1	Primary biliary cirrhosis: primary biliary cirrhosis.
	2	Primary sclerosing cholangitis: Crohn disease, ulcerative colitis, cholangitis with no bowel disease
	3	Alcoholic liver disease: Laënnec cirrhosis
	4	Autoimmune disorders: cirrhosis (drug or industrial exposure, cryptogenic, idiopathic), chronic autoimmune hepatitis (etiology unknown, postnecrotic)
2	5	Hepatitis C and similar infections: postnecrotic cirrhosis (non-A, non-B; type C; type D; types B and C; types B and D), Laënnec cirrhosis (post-necrotic, other)
	6	Hepatitis B: postnecrotic cirrhosis (HBsAg-positive)
3	7	Acute hepatic failure: acute hepatic necrosis (drug exposure; hepatitis A; hepatitis B, HBsAg-positive; non-A, non-B hepatitis; hepatitis C; hepatitis D; hepatitis B and C; hepatitis B and D; other acute viral infection), etiology unknown
4	8	Cancers: primary liver malignancy (hepatoma, fibrolamellar hepatocellular carcinoma, cholangiocarcinoma, hepatoblastoma, hemangio-endothelioma, hemangiosarcoma, angiosarcoma)
5	9	Metabolic disorders: alpha-1-antitrypsin deficiency, glycogen storage disease type I, glycogen storage disease type II, hemochromatosis, hemosiderosis, hyperlipidemia type II, homozygous hypercholesterolemia, primary oxalosis or oxaluria, hyperoxaluria, tyrosinemia, Wilson's disease or other copper disorder, urea cycle disorder, Crigler-Najjar syndrome, Wolman's disease, protoporphyria, Niemann-Pick disease, abetalipoproteinemia, Gaucher's disease, Rendu-Osler-Weber syndrome, carbamoylphosphate synthase deficiency, amyloidosis, Wiskott-Aldrich syndrome
5	10	Other liver diseases: cirrhosis (post-necrotic hepatitis A), secondary biliary cirrhosis (Caroli disease, choledochal cyst, other), familial cholestasis (Byler disease, other), cholestatic liver disease not listed above, neonatal hepatitis, biliary atresia (extrahepatic biliary atresia, hypoplasia, Alagille syndrome, other), congenital hepatic fibrosis, cystic fibrosis, Budd-Chiari syndrome, benign tumor (hepatic adenoma, polycystic liver disease, other), liver disease induced by total parenteral nutrition or hyperalimentation, graft-versus-host disease, trauma, biliary stricture or stenosis, gastroschisis, idiopathic adult ductopenia, unknown

patients in other OPOs within the same region as the harvested liver. Finally, if still no match was found, it would consider patients anywhere in the United States.

Posttransplant Survival Module

To simulate the possibility of patient death and organ rejection or loss, the posttransplant survival module generates 2 estimates of survival time: one for the patient and another for the organ. If patient survival is shorter than graft survival, the patient dies and is removed from the system. The graft is also removed from the system because organs are never transplanted more than once. If, however, the graft fails before the patient dies, then the patient requires another transplant and is relisted in the simulation with characteristics representative of actual relisted patients at UPMC. In the United States, approximately 11% of transplant recipients undergo transplantation more than once.[12] After generating the clinical characteristics, relisted patients are handled in the same manner as all other patients in the model.

To estimate the survival probability distributions, we estimated disease-specific Cox proportional hazards models[13] for each of 10 disease groups using data from UNOS.[14] With the Cox model, a baseline hazard rate is generated and then adjusted according to the

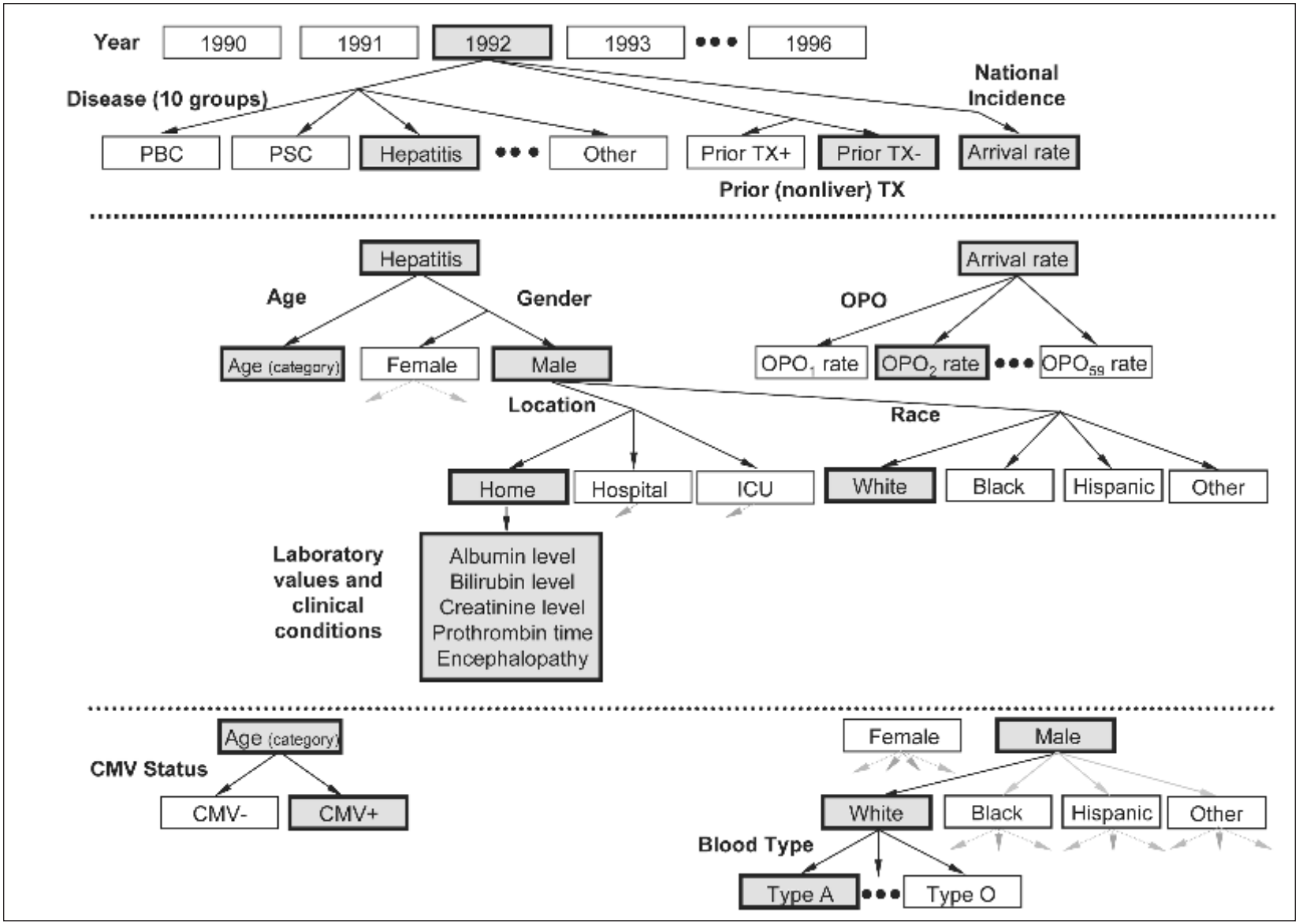

Figure 2 Pattern of dependencies in patient arrival data.

Note: When the simulation generates a patient's attributes, it does so according to a series of conditional relationships determined by clinical oversight and statistical investigation. The 3 strata of dependencies are separated by heavy dotted lines. Groups of variables between the dotted lines are dependent only on the relationships in that particular stratum. For example, the following are dependent on the simulation year: the number of patients listed for a transplant (arrival rate), the distribution of different types of end-stage liver disease, and the rate of listed patients who have had other organ transplants (prior TX+). Once a disease is selected, it determines the distribution of age and gender (e.g., more middle-aged women than men present with primary biliary cirrhosis). The overall yearly national rate of patient arrival is distributed across the 59 organ procurement organizations (OPOs) in proportion to the fraction of all patients listed at each OPO during the years of the study. The combination of gender and disease determines the race and also determines the severity of illness as proxied by location (at home, in a hospital ward, or in an intensive care unit [ICU] at the time of listing). Initial laboratory values are sampled from distributions stratified by disease and location. The age category of the patient determines the cytomegalovirus (CMV) status. Gender and race determine the blood type.

characteristics of a specific patient and specific donated organ. The Cox model is usually used to compare survival between groups or to describe survival as a function of a set of variables. However, it can also be used to create a patient-specific survival curve from which a pseudorandom observation can be generated to obtain the time of a particular event (e.g., patient death or graft failure). One advantage of this approach is that it uses the entire survival distribution, rather than just the mean of the distribution. Alternatives to the Cox model can also be easily incorporated into this module.

Incorporation of Quality of Life

Quality of life is currently incorporated in the model as a characteristic of a particular state of the patient. Based on data from a prospective analysis of 95 patients awaiting liver transplantation at UPMC,[7] we used time tradeoff assessments to attach a utility of 0.73 for pretransplant outpatients, 0.54 for pretransplant inpatients, and 0.83 for posttransplant patients. Because our sample did not include patients in the intensive care unit or on ventilators, we assigned a utility of 0.4 for each of these, based on estimates from the literature.[6] Relative value scale and standard gamble quality-of-life assessments were also collected from the 95 UPMC patients, and we can run the model with those as well.

Variance Reduction

When comparing results between allocation policies, we want to increase our confidence that any system output differences are based on real performance differences as opposed to random sources of variance in the waiting lists. To eliminate some differences in random variation observed by different simulated systems, we used the standard variance-reduction technique of common random numbers.[8] For example, because new patient and liver arrivals were assumed to be independent of the allocation schemes, we used 1 random number stream to generate the same patient arrivals and characteristics across different policies, and we used a 2nd random number stream to generate the same organ arrivals and characteristics across different policies. By eliminating any variance attributable to different arrival characteristics, we should obtain a tighter confidence interval for system output differences.

Because UNOS waiting list data did not contain sufficient information to create the exact waiting list at the beginning of 1992, and because allocation policy changes occur when patients are already waiting under the former policy, we warmed up the system using the same baseline matching process for testing each alternative policy. We chose 1992 as the beginning of our simulation period because it is the earliest year for which UNOS displays complete data on its Web site. Using the baseline allocation policy, we started our simulation with an empty waiting list, used 1991 patient and organ arrival rates and characteristics, and let the simulation run until the number of simulated waiting patients equaled the actual number waiting at the start of 1992. Then, from 1992 through 1996, we applied the organ allocation policy we wanted to test. For example, if we wanted to test the difference in outcomes between a national priority list and the current Model for End-Stage Liver Disease (MELD) policy, we would run 2 simulations. For both of them, we would warm up the simulation using the same allocation policy that existed prior to 1992. Then, to simulate the years 1992 through 1996, we would have one simulation use the national policy and the other use the MELD policy. By using the same patient and organ arrivals and the same allocation method in the warm-up period across policies, we guaranteed that any differences in policies were not attributable to different types of patients waiting at the start of 1992. However, to reflect the fact that we used our simulation to generate the starting wait list, and doing so represents one possible instance of waiting patients, we randomly changed the arrival streams between replications to consider different instances of the starting wait list.

Validation of the Model

Our model underwent various stages of validation. We met numerous times with our National Clinical Oversight Committee to ensure conceptual validity of the model prior to its actual development. Two key components of our model, the pretransplant natural history and posttransplant survival modules, were developed as independent research projects and validated in isolation of the present simulation.[10,14]

Once all of the modules were completed, we validated the model by comparing simulated output measures with the same types of output measures found on the UNOS Web site. Measures included the number of patients awaiting transplant, number of cadaveric liver donors, number of transplants performed, number of deaths while waiting, 1-year graft survival, 1-year patient survival, and median time until transplant. We calculated this last measure using the Kaplan-Meier method for handling censored observations.[15]

Table 4 Validation of Model Results

Outcome Measure	1992	1993	1994	1995	1996
Patients already listed for transplant					
UNOS	1880	2548	5072	6795	
Model mean ($\bar{x}$)	1914	2816	3824	5348	7156
Model std dev (s)	68.0	96.6	104.1	121.9	130.4
Difference (%)	1.8	10.5	7.9	5.4	5.3
New patients listed for transplant					
UNOS	3695	4315	4767	5857	6543
Model mean ($\bar{x}$)	3682	4312	4761	5852	6482
Model std dev (s)	43.9	73.0	66.3	72.1	66.9
Difference (%)	–0.4	–0.1	–0.1	–0.1	–0.9
Cadaveric donors					
UNOS	3334	3764	4093	4335	4463
Model mean ($\bar{x}$)	3358	3826	4046	4475	4631
Model std dev (s)	66.1	75.8	63.6	75.9	71.4
Difference (%)	0.7	1.6	–1.1	3.2	3.8
Transplants performed					
UNOS	2599	2946	3124	3460	3583
Model mean ($\bar{x}$)	2600	2963	3138	3464	3582
Model std dev (s)	57.4	67.4	55.6	64.6	58.9
Difference (%)	0.0	0.6	0.4	0.1	0.0
Deaths while on waiting list					
UNOS	473	514	589	754	919
Model mean ($\bar{x}$)	519	574	678	831	1004
Model std dev (s)	22.2	28.6	34.4	35.4	29.5
Difference (%)	9.7	11.7	15.1	10.2	9.2
Median waiting time (days)					
UNOS	142	193	217	316	NA
Model mean ($\bar{x}$)	125	177	242	343	
Model std dev (s)	12.1	14.0	16.9	16.6	
Difference (%)	–12.0	–8.3	11.5	8.5	
1-year patient survival after transplant					
UNOS	0.82	0.83	0.85	0.84	0.85
Model mean ($\bar{x}$)	0.84	0.84	0.84	0.84	0.83
Model std dev (s)	0.01	0.01	0.01	0.01	0.01
Difference (%)	2.4	1.2	–1.2	0	–2.3
1-year graft survival after transplant					
UNOS	0.75	0.77	0.79	0.80	0.78
Model mean ($\bar{x}$)	0.79	0.77	0.77	0.78	0.78
Model std dev (s)	0.01	0.01	0.01	0.01	0.01
Difference (%)	5.3	0	–2.5	–2.5	0

Note: UNOS = United Network for Organ Sharing; NA = not available. Model results are presented as the mean ($\bar{x}$) and standard deviation (s) of 30 replications performed by the model.

Traditional hypothesis testing is not the appropriate statistical approach for measuring model validity because the null hypothesis that the true system and model parameters are identical is almost always false[8]; with enough simulation replications, any arbitrarily small difference between the model parameter and the actual value would be found to be statistically significant. Rather, one may form confidence intervals around the model estimates to understand how precise those estimates are and then determine if there is a clinically significant difference from actual data—a subjective decision for system experts to make.[8]

Table 5 Comparison of Posttransplant Survival Rates between UNOS Data and Rates Generated by the Model

Outcome Measure	1992	1993	1994	1995	1996
1-year survival					
UNOS	0.84	0.84	0.87	0.86	0.87
Model	0.84	0.84	0.84	0.84	0.83
3-year survival					
UNOS	0.75	0.76	0.78	0.78	0.79
Model	0.77	0.76	0.76	0.76	0.75

Note: UNOS = United Network for Organ Sharing. UNOS survival rates were calculated using Kaplan-Meier plots of survival data for patients aged 16 years or older.

RESULTS

Table 4 compares the actual UNOS data for the years 1992–6 with the averages and standard deviations obtained after 30 replications of our baseline model. For most years, the simulation output closely matches UNOS data for new patients listed, cadaveric donors, transplants, median waiting time for a transplant, and survival rates for patients and organs 1 year after receiving a transplant. Our model overestimated the waiting list in each year and consequently overestimated the deaths while waiting (more patients waiting means more patients are at risk of dying before transplantation).

Table 5 shows 1-year and 3-year posttransplant patient survival rates for the years 1992–6. The UNOS results for these years were calculated from the 1-year Kaplan-Meier analyses of the database for organ recipients aged 16 years or older. The UNOS data demonstrate a slowly improving trend in posttransplant survival, whereas the model output shows a near constant rate of survival. The model currently does not incorporate a time trend in the posttransplant survival analysis, so it will capture only changes in survival suggested by changes in the level of illness of individuals undergoing transplantation.

DISCUSSION

The liver allocation process is a complex stochastic system in which individual patients and livers arrive

with various characteristics, patient health deteriorates while waiting, and liver and patient characteristics affect the length of graft and patient survival. At the individual level, we used mathematical models of disease progression and posttransplant survival to track patients' health. In addition, a predetermined matching algorithm assigns an incoming liver to one of the many patients on the waiting list, and this policy affects overall statistics such as the number of successful transplants, median waiting time for a transplant, and average time from listing until death. At the system level, there is no feasible way to estimate analytically the effects of the implemented allocation policy on the various output measures. Hence, we built a Monte Carlo DES that allows us to estimate these statistics empirically. Our model is clinically based and modular, allowing us to incorporate possible changes to natural history, posttransplant survival, or patient prioritization and allocation policy.

As mentioned in the Results section and shown in Table 4, the baseline model results are close to UNOS data for multiple measures. This gives us confidence that the model can reasonably approximate certain important measures of interest to the transplant community. As a result, one can use our model to test various allocation policies to get some idea of how they will perform. However, we recognize that the current version has several limitations, and by addressing these, we can create a more accurate decision-making tool for policy makers to use in determining allocation rules.

First, the main limitation to developing an empirical distribution of natural history was that no national databases contained the necessary level of clinical detail on natural history, a problem that we describe elsewhere.[9] We used a single institution (UPMC) that treated sicker patients on average, so the results may not be representative of transplant candidates in general. Furthermore, by using a single institution, we lost patients to follow-up at other transplant centers and deaths that occurred at non-UPMC hospitals. Both of these issues caused us to underestimate the number of deaths while waiting for a transplant and required adjustments of the pretransplant death rates in our model to match the UNOS experience.

Second, our matching algorithm discards some donated livers on the basis of estimated wastage rates, but it does not consider specific factors that would influence these rates, such as the liver characteristics, the time it takes for the liver to reach the patient, or the reasons that a patient and physician decline an offered liver. The collection of information on all organ arrivals and detailed data on the factors affecting organ wastage will allow us to model liver arrivals and usage more accurately in the future.

Third, there are reasons other than death for a patient's removal from the UNOS waiting list, including improving condition without a transplant or being too sick for a transplant, which accounted for 3% and 4%, respectively, of 1996 list removals.[2] UNOS also includes a removal category labeled "other" that accounted for 5% of the 1996 removals; however, it is not clear what patients comprise this group.[2] Our data did not include removal information other than deaths, so we used clinical expertise to approximate improved patient conditions in our model. To account for other removal reasons, we randomly removed patients each day according to probabilities implied by the UNOS data. We believe the lack of data to support more detailed modeling of patient list removals relates to our overestimation of patients waiting at the end of each year.

Fourth, our natural history data did not include the actual patient statuses assigned by UNOS. Instead, we mapped each patient's location (in the intensive care unit on a ventilator, in the intensive care unit not on a ventilator, in a hospital ward, at home) to statuses 1 through 4. With longitudinal data that include status assignments, we could derive more accurate status mappings based on location and other clinical factors.

We are currently involved in efforts to address the above limitations, especially those related to data sources. In addition, we plan to improve our model's ability to inform policy decisions about transplantation by adding cost components, enhancing the quality-of-life components, and incorporating cost-effectiveness analyses. The cost component will include the costs of pretransplant and posttransplant care for patients who have different diagnoses of ESLD, along with the costs of transplantation itself.

In summary, we have built a DES model of the liver allocation process in the United States that represents both the stochastic biological progression of ESLD and the deterministic assignment of livers to patients. The model includes sufficient detail to estimate the effects of a wide range of questions regarding liver allocation and policy change.[16] For example, one may use the model to examine differences between the current MELD allocation rules and a national allocation policy with regard to important outcome measures such as total transplants, average waiting time for a transplant, and the average posttransplant survival time. Because our pretransplant natural history runs independently of the UNOS allocation policy in place at the time, one can test virtually any hypothetical policy change with our model.

APPENDIX
Members of the National Clinical Oversight Committee

John Fung, MD, PhD
Associate Professor of Surgery and Transplantation Research Director
University of Pittsburgh
School of Medicine
Pittsburgh, PA 15213

Roger Jenkins, MD
Chief of Transplantation
Beth Israel/Deaconess Medical Center
Harvard Medical School
Boston, MA 02215

John R. Lake, MD
Professor of Medicine and Surgery
University of Minnesota
Minneapolis, MN 55455

Wallis Marsh, MD
Associate Professor of Surgery
University of Pittsburgh
School of Medicine
Pittsburgh, PA 15213

Andre C. Stieber, MD, FACS
Associate Professor of Surgery, Liver Transplantation
Emory University
School of Medicine
Atlanta, GA 30322

Russell H. Wiesner, MD
Medical Director of Liver Transplantation
Mayo Clinic
Rochester, MN 55905

Rowen K. Zetterman, MD, FACP, FACG
Professor, Department of Internal Medicine
University of Nebraska Medical Center
Omaha, NE 68198

REFERENCES

1. Committee on Organ Procurement and Transplantation Policy, Institute of Medicine. Organ Procurement and Transplantation. Washington (DC): National Academy Press; 1999.
2. United Network for Organ Sharing. US transplantation data. Available from: URL: http://www.unos.org/data/default.asp?displayType=usData. Accessed 23 December 2003.
3. Pritzker AAB, Martin DL, Reust JS, et al. Organ transplantation policy evaluation. In: Alexopoulos C, Kang K, Lilegdon WR, Goldsman D, eds. Proceedings of the 1995 Winter Simulation Conference, 3-6 December 1995. Arlington (VA): 1314–23.
4. Consad Research. An analysis of alternative national policies for allocating donor livers for transplantation. Final Report, Consad Research, Pittsburgh (PA); 1995.
5. Klein RW, Dittus RS, Roberts SD, Wilson JR. Simulation modeling and health-care decision making. Med Decis Making. 1993;13(4):347–54.
6. Tsevat J, Cook EF, Green ML, et al. Health values of the seriously ill. SUPPORT Investigators. Ann Intern Med. 1995;122(7):514–20.
7. Bryce CL, Angus DC, Roberts MS, Tsevat J. Health status versus utilities of patients with end-stage liver disease. Qual Life Res. 2004;13:773–82.
8. Law AM, Kelton WD. Simulation Modeling and Analysis. 3rd ed. New York: McGraw-Hill; 2000.
9. Kreke J, Schaefer AJ, Angus DC, Bryce CL, Roberts MS. Incorporating biology into discrete event simulation models of organ allocation. In: Proceedings of the 2002 Winter Simulation Conference. Piscataway (NJ): Institute of Electrical and Electronics Engineers; 2002.
10. Alagoz O, Roberts MS, Bryce CL, Schaefer AJ, Chang J, Angus DC. Incorporating clinical natural history in simulation models: empiric estimates of the progression of end-stage liver disease. Med Decis Making. Under review.
11. United Network for Organ Sharing and the Division of Organ Transplantation, Bureau of Health Resources Development, Health Resources and Services Administration, US Department of Health and Human Services. Annual Report of the U.S. Scientific Registry for Transplant Recipients and the Organ Procurement and Transplantation Network—Transplant Data: 1988-1993. Bethesda (MD): US Department of Health and Human Services; 1994.
12. United Network for Organ Sharing. Annual Report of the United States Scientific Registry for Transplant Recipients and the Organ Procurement and Transplant Network: Transplant Data 1989-1998. Richmond (VA): United Network for Organ Sharing; 1999.
13. Cox DR. Regression models and life tables. J R Stat Soc [Ser B]. 1972;34:187–220.
14. Roberts MS, Angus DC, Bryce CL, Valenta Z, Weissfeld L. Survival after liver transplantation in the United States: a disease-specific analysis of the UNOS database. Liver Transpl. 2004;10(7):886-897.
15. Collett D. Modelling Survival Data in Medical Research. New York: Chapman & Hall; 1994.
16. Roberts MS, Bryce CL, Schaefer A, Shechter S, Alagoz O, Angus D. Optimal organ allocation policy: a biologically-based simulation model [abstract]. Med Decis Making. 2003;23:575.