DYNAMIC ALLOCATION OF KIDNEYS TO CANDIDATES ON THE TRANSPLANT WAITING LIST

STEFANOS A. ZENIOS

Graduate School of Business, Stanford University, Stanford, California 94305, stefzen@leland.stanford.edu

GLENN M. CHERTOW

Division of Nephrology, University of California San Francisco, San Francisco, California 94143

LAWRENCE M. WEIN

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, lwein@mit.edu

The crux of the kidney allocation problem is the trade-off between clinical efficiency and equity. We consider a dynamic resource allocation problem with the tri-criteria objective of maximizing the quality-adjusted life expectancy of transplant candidates (clinical efficiency) and minimizing two measures of inequity: a linear function of the likelihood of transplantation of the various types of patients, and a quadratic function that quantifies the differences in mean waiting times across patient types. The dynamic status of patients is modeled by a set of linear differential equations, and an approximate analysis of the optimal control problem yields a dynamic index policy. We construct a large-scale simulation model using data from over 30,000 transplants, and the simulation results demonstrate that, relative to the organ allocation policy currently employed in the United States, the dynamic index policy increases the quality-adjusted life expectancy and reduces the mean waiting time until transplantation for all six demographic groups (two sexes, races, and age groups) under consideration.

K idney transplantation is the treatment of choice for patients suffering from end stage renal disease (ESRD), also known as chronic kidney failure. However, the supply of cadaveric kidneys for transplantation is not enough to meet the increasing demand; e.g., at the end of 1996 the waiting list for kidney transplants had 34,550 registered candidates, but the total number of cadaveric transplants performed during that year was only 7833 (UNOS 1997). Unfortunately, repeated attempts to increase the supply of organs have been unsuccessful. This organ shortage exacerbates the trade-off between clinical efficiency and equity that lies at the heart of the kidney allocation problem. The goal of this paper is to shed light on and better quantify this trade-off in an attempt to assist policy makers in the formidable task of allocating cadaveric kidneys to potential transplant recipients.

In the United States, the distribution of organs to transplant candidates is controlled by the United Network of Organ Sharing (UNOS). This organization has developed a point system that dictates who receives a donated kidney when it becomes available; a description of ESRD and the UNOS policy are given in §1. This allocation scheme, which attempts to match tissue types of donors and recipients, is still the subject of heated debate, even though it has been in use for more than seven years and has been revised at least twice. Several empirical studies (Kasiske et al. 1991, Sanfilippo et al. 1992, Ellison et al. 1993, Gaston et al. 1993, Gaylin et al. 1993) demonstrate that this allocation policy

generates inequities; most notably, African-Americans experience longer waiting times than other candidates. In addition, a report by the Office of Inspector General (OIG 1991) concludes that the current allocation scheme fails to meet the expectations of the general public.

In an attempt to understand the limitations of the UNOS scheme, medical researchers have assessed how tissue type matching affects the efficiency and fairness of an allocation policy. For example, Ghjertson et al. (1991) and Held et al. (1994) estimate that an allocation algorithm based solely on tissue matching would enhance (relative to the UNOS policy) graft survival (*graft* is the medical term for a transplanted organ) by 5% over a 10-year period and 2% over a 5-year period, respectively. However, much of this enhancement would be achieved by patients receiving a perfectly matched kidney, and such an algorithm would reduce access to transplantation for African-Americans.

Wujciak and Opelz (1994) and Yuan et al. (1994) attempt to clarify these issues by developing simulation models that compare different allocation policies. While Wujciak and Opelz conclude that it is possible to achieve considerable efficiency improvements without generating additional inequity, Yuan et al. conclude that efficiency and equity are conflicting and argue that devising an acceptable allocation policy requires explicit value judgements about efficiency and equity. However, neither simulation model allows candidates to exit the waiting list (due to death) and the former model does not permit new arrivals to the waiting list.

Medical researchers have also investigated other factors that affect the clinical efficiency of organ allocations. Gaston et al. (1996) and Miles et al. (1996) conclude that the relative kidney size of the donor and recipient has little impact on graft survival. Chertow et al. (1996) conclude that nonimmunological factors, such as donor and patient age, gender, and race, are predictive of graft failure.

Several researchers have also attempted to clarify the ethical dilemmas that underlie the organ allocation problem. Bailey (1988) discusses the problem in the general context of rationing of expensive health care resources. More recently, the Council on Ethical and Judicial Affairs of the American Medical Association (AMA 1995) has proposed ethically acceptable criteria for organ allocation. These criteria include the likelihood of benefit, the duration of benefit, the urgency of need, and the improvement in quality of life.

In addition to the medical literature, the operations research literature also includes several studies that address some aspects of the organ allocation problem. One of the first papers in this area is by Ruth et al. (1985), who present a simulation model for the waiting list in Michigan. Righter (1989) formulates the organ allocation problem as a stochastic assignment problem (Derman et al. 1972) and develops properties of the optimal policy. David and Yechiali (1985, 1990, 1995) and David (1995) study several sequential decision problems that are motivated by organ transplantation, from the perspectives of both a potential recipient (1985) and a centralized decision maker. In recent studies that combine analytical and empirical research, Ahn and Hornberger (1996) and Hornberger and Ahn (1997) develop kidney acceptance policies for potential recipients that explicitly incorporate patient preferences and demonstrate that some patients can afford to be selective when making transplantation decisions. Pritsker (1998) describes a large-scale simulation model for the liver allocation system that is used by UNOS to compare alternative liver allocation policies.

In summary, the efficiency-equity trade-off that is at the root of the kidney allocation problem is not well understood. To gain a better understanding of this trade-off, we use an analytical model to develop a new set of policies, and we use a simulation model to compare these policies to existing policies. Specifically, we develop a fluid model that provides a stylistic representation of the organ allocation problem in §2, and we formulate an objective function that captures both efficiency and equity criteria in §3. In §4, we utilize the theory of optimal control to develop a heuristic dynamic index policy that attempts to maximize the objective. Finally, we compare an empirically calibrated version of the heuristic policy to the first-come first-transplanted policy and to the one used by UNOS. This comparison employs a detailed simulation model that utilizes the most current transplant data; the simulation model is described in §5, and the simulation results are reported and discussed in §6. Concluding remarks appear in §7.

1. BACKGROUND ON ESRD

This section contains a brief description of ESRD and the organ allocation scheme that is currently used in the United States. ESRD is a fatal disease unless treated with dialysis or kidney transplantation. Patients undergoing dialysis suffer the inconvenience of visiting a dialysis center for at least 12 hours per week (Allen and Chapman 1994). Although transplant recipients run the risk of graft rejection, transplantation appears to be a superior form of treatment because it enables patients to resume normal life activities. This risk can be reduced by receiving an organ from either a living-related donor or a carefully *matched* cadaveric donor; we focus on cadaveric donors in this paper because there are no allocation decisions associated with living donors.

Matching cadaveric donors to recipients is a two-step process. The first step involves comparing the *tissue type* of the donor to the tissue type of the recipient. Specifically, the tissue type, also known as HLA type, is a combination of six proteins: two of type A, two of type B, and two of type DR. Empirical and clinical evidence shows that when the donor and recipient share all six proteins in common (zero mismatches), the risk of graft rejection is minimized. However, the risk increases with the number of mismatches, so allocating an organ to the recipient with the smallest number of tissue type *mismatches* reduces the chance of graft failure; the number of mismatches gives the number of donor HLA proteins that are absent in the recipient. The second step of the process involves a blood test that determines whether the candidate exhibits antibodies to the proteins of the donor. (In addition, candidates must have a blood type that is compatible with the donor.) Because patients who test positive (also known as *positive-crossmatch* or *presensitized* patients) are at a high risk of acute graft rejection (Allen and Chapman, 1994, Ghjertson et al. 1991, Chertow et al. 1996), transplantation is performed only on candidates with a negative test (*negative-crossmatch* or *nonpresensitized* patients).

As previously mentioned, the United Network of Organ Sharing governs the allocation of organs to transplant candidates in the United States. UNOS coordinates the activities of 72 organ procurement organizations (OPO) that operate in distinct geographic regions and are responsible for procuring all the organs donated in their region and allocating them to candidates registered on their transplant waiting list. When an organ becomes available, the OPO prioritizes all the blood-compatible transplant candidates using a point system devised by UNOS (or an UNOS-approved variant of this system). The point system (see Table 1) gives priority points based on the total number of tissue matches. To compensate candidates with rare tissue types, the policy also awards points based on the waiting time and the rank in the waiting list; consequently, candidates do not stay on the waiting list indefinitely. Finally, the system allocates priority points to candidates with high *panel reactive antibodies* (pra); the pra of a candidate is the probability that the candidate will crossmatch positive with a randomly selected donor. The pra points ensure that a golden (but rare)

Table 1. The UNOS point system after (prior to) July 31, 1995.

Category	Points
Waiting Time	1 (0.5) points for each full year in the waiting list
Rank in the Waiting List	1 (1) point for the longest waiting candidate; fractions of points are assigned proportionately to all other candidates.
Tissue Mismatches	∞ (10) points for no mismatches 7 (7) points for 0 B or DR mismatches 0 (6) points for 0 A or B mismatches 5 (3) points for 1 B or DR mismatches 2 (2) points for 2 B or DR mismatches 0 (1) point for 3 B or DR mismatches
Panel Reactive Antibodies	4 (4) points for PRA > 80%
Pediatric Candidates	4 (2) points when age < 11 (5) years 3 (1) points when 11 (5) years < age < 18 (10) years

opportunity of a negative crossmatch will not be missed by those candidates.

Once the candidates are prioritized, the OPO offers the kidney to the top priority candidate. If the candidate is readily available (a procured kidney is immediately frozen and needs to be transplanted quickly; between 1987 and 1991, 75% of organs were transplanted within 31 hours) and crossmatches negative, the transplantation is performed. Otherwise, the organ is offered to the next candidate, and the procedure is repeated until an available candidate who crossmatches negative is found. Under this allocation scheme, organs are transplanted locally, i.e., into candidates that are registered on the waiting list of the OPO that procured the organ. The only exception is when a zero-mismatched candidate exists somewhere else in the country. In that case, UNOS demands that the kidney be offered to that candidate and in exchange, the OPO of that candidate offers the next kidney it procures to the OPO that provided the zero-mismatched kidney.

2. THE FLUID MODEL

In this section we construct a continuous time, continuous space, deterministic model that provides a stylistic representation of the organ allocation process and tracks the dynamics of the ESRD population over time. The model divides the ESRD population into K distinct categories, or *classes*, based on demographic (age, gender, race), immunological (blood type, tissue type, pra) and physiological characteristics (height, weight). However, other relevant characteristics (e.g., health status) can also be incorporated into the class division. The model also divides the donor population into J classes, based again on demographic, immunological, and physiological characteristics. Without loss of generality we assume that patients of class $k = 1, \ldots, K_W$ are registered on the waiting list, and patients of class $k = K_W + 1, \ldots, K$ have a functioning graft.

The state of the system at time t is described by the K-dimensional column vector $x(t) = (x_1(t), \ldots, x_K(t))'$ (where primes denote transposes), which gives the number of patients in each class. Transplant candidates of class $k = 1, \ldots, K_W$ join the waiting list at rate λ_k^+ per unit time. These patients depart from the waiting list via death, which occurs at rate μ_k per unit time for class k patients, or organ transplantation. Organs of class $j = 1, \ldots, J$ arrive at rate λ_j^- per unit time; we assume that the demand-to-supply ratio $\rho = \sum_{s=1}^{K_W} \lambda_s^+ / \sum_{j=1}^{J} \lambda_j^-$ is greater than one. A fraction $v_{jk}(t)$ of class j organs is allocated to transplant candidates of class k; thus, $v_{jk}(t)$ is a control variable and $u_{jk}(t) = \lambda_j^- v_{jk}(t)$ is the instantaneous transplantation rate of class j kidneys into class k candidates.

When a class $j = 1, \ldots, J$ kidney is transplanted into a class $k = 1, \ldots, K_W$ candidate, the class k candidate leaves the waiting list and becomes a patient of class $c(k, j) \in \{K_W + 1, \ldots, K\}$. Patients of class $c(k, j)$ depart this class at rate $\mu_{c(k,j)}$ per unit time; a fraction $q_{c(k,j)} \in [0, 1)$ of these patients experience graft failure and rejoin the waiting list as patients of class k, and the remaining fraction exit the system due to death.

The dynamics of the system state are described by the ordinary linear differential equations:

$$\frac{d}{dt}x_k(t) = \lambda_k^+ - \mu_k x_k(t) - \sum_{j=1}^{J} u_{jk}(t)$$

$$+ \sum_{j=1}^{J} q_{c(k,j)}\mu_{c(k,j)}x_{c(k,j)}(t); \qquad k = 1, \ldots, K_W,$$

$$(1)$$

and

$$\frac{d}{dt}x_k(t) = \sum_{j=1}^{J} \sum_{i=1}^{K_W} u_{ji}(t)1_{\{c(i,j)=k\}} - \mu_k x_k(t);$$

$$k = K_W + 1, \ldots, K. \qquad (2)$$

However, our model can be expressed more compactly using matrix notation. Let $e = (1, \ldots, 1)'$ be the K_W-dimensional unit vector, $\lambda^- = (\lambda_1^-, \ldots, \lambda_J^-)'$ and $\lambda^+ = (\lambda_1^+, \ldots, \lambda_K^+)'$, where $\lambda_{K_W+1}^+ = \cdots = \lambda_K^+ = 0$. In addition, let $u_{j\cdot}(t) = (u_{j1}(t), \ldots, u_{jK_W}(t))'$ and $u(t) = (u_1.(t)' | \cdots | u_J.(t)')'$. Define the matrices $A \in \mathfrak{R}^{K \times K}$ and $B \in \mathfrak{R}^{K \times JK_W}$ by

$$
A_{ki} = \begin{cases} -\mu_k, & \text{if } k = i; \\ q_i \mu_i, & \text{if } i > K_W, \ k \leqslant K_W \text{ and } i = c(k, j) \\ & \quad \text{for some } j \in \{1, \ldots, J\}; \\ 0 & \text{otherwise,} \end{cases} \tag{3}
$$

$$
B_{ki} = \begin{cases} -1, & \text{if } i \bmod K_W = k \text{ and } k \leqslant K_W; \\ 1, & \text{if } k = c \ (i \bmod K_W, \lceil i/K_W \rceil) \text{ and} \\ & \quad i \bmod K_W \leqslant K_W; \\ 0 & \text{otherwise,} \end{cases} \tag{4}
$$

where $i \bmod K_W$ gives the remainder from the division of i by K_W, and $\lceil i/K_W \rceil$ gives the smallest integer greater than or equal to i/K_W (these operators appear because the vector $u(t)$ lists the control variables $u_{jk}(t)$ in lexicographic order). Finally, define $D \in \mathfrak{R}^{J \times JK_W}$ satisfying

$$
D = \begin{pmatrix} e' & 0 & \ldots & 0 \\ 0 & e' & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & e' \end{pmatrix}. \tag{5}
$$

Then the system state equations are given by

$$
\frac{d}{dt} x(t) = \lambda^+ - Ax(t) + Bu(t), \tag{6}
$$

and are subject to the state constraints

$$
x(t) \geqslant 0. \tag{7}
$$

The organ allocation rates $u(t)$ must satisfy the constraints

$$
Du(t) \leqslant \lambda^-, \tag{8}
$$

$$
u(t) \geqslant 0. \tag{9}
$$

Model (6)–(9) provides a stylized representation of the organ allocation system that sacrifices accuracy for tractability. Specifically, it ignores three aspects of the allocation process: crossmatching between donor and recipient, unavailability of some recipients, and organ sharing between OPOs. In addition, the model assumes that the system evolution is deterministic, the patient and organ survivor curves are exponential and independent of age; and following graft rejection, patients rejoin the patient class they occupied before transplantation. Zenios (1996) analyzes a more general version of the fluid model that allows mixed Erlang survivor curves and a nonmemoryless graft failure process.

Another limitation of the model relates to the definition of the patient and organ classes. To capture the fine interaction between patients and donors, we assume in §6 that the number of classes is approximately 10^7, but the number of patients during the planning horizon is roughly 30,000. This clearly represents a significant deviation from the traditional

application of multiclass fluid models, where a multitude of individuals are classified into a modest number of classes. Furthermore, it creates some technical problems because in reality the state vector is a very long list of zeroes and ones.

Despite these limitations, the overall structure of the model captures the major driving forces behind the equity-efficiency trade-off and the first-order dynamics of the organ allocation process (because organ demand is greater than organ supply). As with other fluid models of single-resource queueing systems, we hypothesize that an analysis of the fluid model will result in a robust and effective allocation policy (e.g., Avram et al. 1995), even though the fluid model is too crude to be used as a reliable performance analysis tool. Hence, we develop a much more realistic simulation model in §5 to assess the impact of various policies. The simulation results in §6, which avoid the simplifying assumptions described above, confirm that the derived policies are indeed effective and robust.

3. THE OBJECTIVE FUNCTION

In this section we construct a tri-criteria objective function for the fluid model. This objective reflects the perspective of a central decision maker who makes organ allocation decisions based on a combination of efficiency and equity criteria. Both criteria are captured using simple analytical metrics, and the objective is assumed to be additive in these metrics.

3.1. Efficiency

Following an accepted convention in medical decision making research, we measure clinical efficiency using *quality-adjusted life years* (QALY); the reader is referred to Gold et al. (1996) for a detailed introduction to QALY and their applications to societal decision making. This measure assumes that the centralized decision maker assigns a *quality of life* (QOL) score h_k to each patient class $k = 1, \ldots, K$, and the QALY over a finite time horizon T is the total number of life years multiplied by the QOL scores,

$$
\int_0^T \sum_{k=1}^K h_k x_k(t) \, dt. \tag{10}
$$

QOL scores aggregate all distinct elements of a health state into a single number that reflects the desirability of that state, and they are scaled so that zero represents death and one represents perfect health. A common practice when utilizing QALY is to include discounting to account for the perception that a current year of life is more valuable than a future one. Introducing discounting into Equation (10) involves a trivial, but algebraically cumbersome, extension of the undiscounted problem. Therefore, for clarity of exposition we focus on the undiscounted problem.

Traditionally, the QALY provide the basis for clinical decision making at the individual level: an individual assigns subjective QOL scores to all possible health states and then

ranks alternative clinical strategies based on the total QALY. But contrast, our approach represents a significant departure from the traditional use of QALY and implies that the central decision maker can use QALY to rank alternative allocation policies. Although there are problems associated with this use of QALY (see Richardson and Nord 1997), recent research developments have demonstrated the utility of this approach (see Gold et al. 1996, Loomes and McKenzie 1989).

Two major assumptions underlie our use of QALY in the context of the organ allocation problem. First, to ensure the validity of objective (10) in the context of the organ allocation policy, one must assume that one year of healthy life always has the same weight regardless of who enjoys the benefits of this life year and when (see Loomes and McKenzie 1989). A consequence of this assumption is that the QOL scores should not depend on the age, gender, and race of the patients in each class, nor should they depend on the duration of the stay in each class; rather, they should only depend on the health state associated with each class. Second, although the abstract formulation (10) does not specify how to estimate the QOL scores for the different health states, we assume that the aggregate QOL scores are statistical averages of QOL scores elicited from a random sample of patients in different patient classes. Although this mechanism may create serious biases, it appears to be a reasonable starting point (see Torrence 1986).

3.2. Equity

According to Webster's English Dictionary, *equity* means *justice according to natural law or right*. To mathematically formalize this definition, one must formally define *justice* in terms of system outcomes. This would entail identifying a reference point that represents the natural law of justice, and defining equity metrics that measure deviations of the system outcomes from this reference point. Because justice is an elusive concept in our setting, we introduce two reference points for justice and propose two equity metrics that are based on two distinct system outcomes; more sophisticated reference points and outcomes exist but are not discussed here.

The first reference point adopts an *absolute equity* viewpoint and starts with the premise that the most equitable policy is one that completely eliminates outcome discrepancies across the various patient groups. The second reference point adopts a *relative equity* viewpoint and starts with the premise that the first-come first-transplanted (FCFT) policy is the most equitable policy. The relative equity approach assesses the discrepancies across patient groups (with respect to various system outcomes) generated by a policy *by comparing them to* the corresponding discrepancies produced by the FCFT policy: The closer a policy's discrepancies coincide with those of FCFT, the more equitable the policy. By contrast, in the absolute equity viewpoint, a policy's equity is assessed solely by the magnitude of its between-group discrepancies.

While the absolute equity viewpoint has some appeal, equating outcomes across all groups of patients can lead to policies that may be perceived as untenable; for example, the differential mortality rate between 10-year olds and 70-year olds would require giving preferential treatment to 10-year olds in order to equalize the waiting times for both groups. In contrast, because FCFT is viewed as the most socially just policy in many queueing systems (Larson 1987), it complements the concept of absolute equity and avoids some of the difficulties associated with it. Furthermore, FCFT is a natural reference point in our setting because it is widely considered to be a fair organ allocation policy (OIG 1991) and is actually employed by at least one major OPO (the OPO in the San Francisco Bay area).

Both the absolute and relative equity viewpoints provide an aggregate approach to equity and thus fail to capture the impact of different allocation policies on individual patients. As such, these viewpoints may lead to policies that are socially undesirable when examined from the perspective of individuals instead of patient groups. To overcome this limitation, one must expand the two viewpoints to consider discrepancies both *between* and *within demographic groups*; the within-group discrepancies can capture the impact of policies on individuals. Zenios et al. (1999) discuss this issue in detail and illustrate the trade-off between within-group and between-group discrepancies. The present paper focuses on between-group discrepancies because of the controversy surrounding demographic-based imbalances for various groups of transplant recipients.

The absolute equity and relative equity viewpoints differ in our fluid and simulation models because FCFT does not eliminate discrepancies between patient groups; this is because of screens for blood compatibility and presensitization, the mismatch of tissue type and blood type across demographic groups, unequal demand-to-supply ratios of different groups, and different mortality rates across groups. Our simulation results in §6 report between-group discrepancies for FCFT and other policies, which allows one to assess inequity under either of these viewpoints. However, for analytical tractability, our objective function is formulated from the absolute equity viewpoint.

There are a variety of outcomes that could be used as a basis for assessing compliance with the two reference points. Although QALY is a natural candidate (and our simulation results in §6 report QALYs for different groups), our analysis focuses on two complementary queueing-based outcomes that have received considerable attention (e.g., OIG 1991, Sanfilippo et al. 1992, Gaston et al. 1996): the mean *waiting time until transplantation* and the *likelihood of transplantation*. Both outcomes are necessary, and consideration of one of them in isolation may generate an inequitable organ allocation system. The first outcome, which focuses on a factor of psychological importance that underlies society's perceptions about fairness in queueing systems (see Larson 1987), can induce absolute equity in the waiting times but may do so at the expense of absolute equity in the likelihood of transplantation. To

illustrate this, consider what happens when the transplant candidates can be separated into a low-mortality group and a high-mortality group. In this case, one can enforce absolute equity in the waiting times by assigning a disproportionately higher fraction of organs to the low-mortality group; see Zenios (1999) for details. However, this will induce absolute inequity in the likelihood of transplantation. On the other hand, consideration of the likelihood of transplantation can alleviate this problem, but clearly it cannot induce absolute equity in the waiting times. Section 6 (Table 4) highlights the distinction between the two outcomes and the absolute and relative equity viewpoints; it shows that the FCFT policy is effective at reducing differences in the waiting times until transplantation but generates large differences in the likelihood of transplantation across groups.

In the remainder of this section, we mathematically formulate two equity metrics that capture deviations of the waiting time until transplantation and the likelihood of transplantation from the absolute equity viewpoint. Although our mathematical formulation of equity focuses on differences among classes, our simulation results in §6 concentrate on differences across gender, race, and age, which constitute sets of classes.

Deviations of the waiting time until transplantation from the absolute equity viewpoint can be captured by the variance of the waiting time until transplantation. Furthermore, this variance can be decomposed into a between-class component and a within-class component. Based on our earlier discussion, we focus on the between-class component, which is equal to the weighted sum of the pairwise square difference of the steady-state waiting time until transplantation for each patient class (the weight of each pairwise square difference is equal to the product of the arrival rates of the two classes being compared). While it would be more natural to consider the square root of the weighted sum, doing so leads to intractability. Because the pairwise square difference terms cannot be formulated in terms of the state variables of the fluid model, we use the following cruder measure, which will be referred to as the *waiting time inequity*:

$$\frac{1}{2} \int_0^T \sum_{k=1}^{K_W} \sum_{i=1}^{K_W} \lambda_k(t, u(t)) \lambda_i(t, u(t))$$

$$\cdot \left(\frac{x_k(t)}{\lambda_k(t, u(t))} - \frac{x_i(t)}{\lambda_i(t, u(t))} \right)^2 dt, \quad (11)$$

where $\lambda(t, u(t)) = (\lambda_1(t, u(t)), \ldots, \lambda_{K_W}(t, u(t)))$ denotes the instantaneous arrival rate into class k under allocation policy $u(t)$, and is given by

$$\lambda_k(t, u(t)) = \lambda_k^+ + \sum_{j=1}^{J} q_{c(k,j)} \mu_{c(k,j)} x_{c(k,j)}(t)$$

for $k = 1, \ldots, K_W$. $\quad (12)$

If we define $R(t, u(t)) \in \mathfrak{R}^{K \times K}$ such that

$$[R(t, u(t))]_{k,l}$$

$$= \begin{cases} -1, & \text{if } k \neq l \text{ and } 1 \leqslant k, l \leqslant K_W; \\ \dfrac{\sum_{s=1}^{K_W} \lambda_s(t, u(t))}{\lambda_k(t, u(t))} - 1, & \text{if } k = l \text{ and } 1 \leqslant k, l \leqslant K_W; \\ 0, & \text{otherwise,} \end{cases}$$

$$(13)$$

then (11) becomes

$$\int_0^T x(t)' R(t, u(t)) x(t) \, dt. \quad (14)$$

If the waiting time inequity can be maintained at zero, then (by Little's law) each candidate class would have the same steady-state mean waiting time. Although the waiting time of a class is different than its waiting time until transplantation (the former includes patients who die while on the waiting list), these two quantities are closely related, and Zenios (1999) demonstrates that (14) captures the first-order effect of equalizing the mean waiting time until transplantation across classes.

The second equity measure focuses on equalizing the likelihood of transplantation. Consider the quantity $\int_0^T \sum_{j=1}^J u_{jk}(t) \, dt / (\lambda_k^+ T)$, which converges (as $T \to \infty$) to the percentage of class k candidates who receive transplantation. If we define the matrix $\tilde{D} \in \mathfrak{R}^{K_W \times K_W J}$, where

$$\tilde{D}_{ki} = \begin{cases} 1 & \text{if } i \bmod K_W = k; \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

then the vector of likelihoods of transplantation is given by $\int_0^T \tilde{D} u(t) \, dt / (\lambda^+ T)$. Because a quadratic function similar to (11) leads to a mathematically intractable optimal control problem, we take an alternative approach where the likelihood of transplantation equity is captured by the service level constraints

$$\left(\frac{1}{\rho} - \varepsilon \right) \hat{\lambda}^+ T \leqslant \int_0^T \tilde{D} u(t) \, dt \leqslant \left(\frac{1}{\rho} + \varepsilon \right) \hat{\lambda}^+ T; \quad (16)$$

where $\hat{\lambda}^+ = (\lambda_1^+, \ldots, \lambda_{K_W}^+)'$. These constraints state that the likelihood of transplantation for each class should be approximately equal to the average likelihood of transplantation $1/\rho$, but deviations of magnitude ε can be tolerated.

Unfortunately, the constrained problem is not analytically tractable either, and so we resort to the Lagrangian version of (16), by inserting into the objective function the expression

$$\int_0^T \gamma' \tilde{D} u(t) \, dt. \quad (17)$$

Although we have no way of knowing a *priori* what values of the Lagrange multipliers $\gamma = (\gamma_1, \ldots, \gamma_{K_W})'$ correspond to what values of the service level bounds in (16), in our computational results we vary γ to generate a set of policies

with varying degrees of restrictiveness (for various sets of patients) with respect to the likelihood of transplantation.

In the next section, we formulate and analyze a control problem that incorporates the waiting time inequity measure (14) and the likelihood of transplantation cost (17).

4. ANALYSIS OF THE CONTROL PROBLEM

We define the control problem in §4.1, analyze the problem and derive a heuristic index policy in §4.2, and simplify the policy in §4.3.

4.1. The Control Problem

We combine the three objectives (10), (14), and (17) and the fluid model to obtain the following optimal control problem: Choose the allocation rates $u(t)$ to maximize the tri-criteria objective

$$\int_0^T (\beta h'x(t) - (1 - \beta)x(t)'R(t,u(t))x(t) + \gamma'\tilde{D}u(t))\,dt,$$
(18)

subject to (6)–(9), where $\beta \in [0,1]$ and $h = (h_1,\ldots,h_K)'$.

4.2. A Dynamic Index Policy

Unfortunately, this control problem does not appear to admit a closed-form solution. Because this problem is of very high dimension and because our intention is to develop an allocation scheme that might be attractive to health policy makers, we focus on deriving a suboptimal policy that is easy to implement and describe. We employ three approximations in this subsection; these approximations, in conjunction with one iteration of the policy improvement algorithm, allow us to construct a closed-form heuristic index policy.

A key approximation. The biggest stumbling block in analyzing this problem is that the matrix $R(t,u(t))$ is a function of both time and the policy. Hence, our first approximation replaces this function by a fixed matrix R. The natural matrix to use is $R(\infty,\bar{u})$, where $\bar{u} = (\bar{u}_{11},\ldots,\bar{u}_{JK_W})'$ is the optimal equilibrium policy; this replacement would be somewhat accurate if the optimally controlled system operated near the optimal equilibrium pair $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_K)'$ and $\bar{u}$ (see Horwood and Whittle 1988 for an analysis that makes this assumption). However, the optimal equilibrium pair is the solution to a huge (e.g., the matrix A is at least $10^7 \times 10^7$) nonlinear program, and its derivation is not practical. Hence, we replace $R(t,u(t))$ by $R = R(\infty,u^F)$, where $u^F = (u_{11}^F,\ldots,u_{JK_W}^F)'$ are the equilibrium allocation rates under the FCFT policy. (We use the FCFT policy to enforce consistency between the absolute equity viewpoint reflected in (18) and the relative equity viewpoint captured by FCFT. We also considered other alternatives for constructing R, but none performed as well in our numerical study.) The fixed matrix R is computed as follows. The rates u^F are estimated from the detailed simulation model described in §5. Then we substitute u^F

into the steady-state version of (6) to find the unique fixed point, $x^F = (x_1^F,\ldots,x_k^F)' = A^{-1}(\lambda^+ + Bu^F)$. The fixed point x^F is then substituted into the steady-state version of (12), which is $\lambda_k(\infty,u^F) = \lambda_k^+ + \Sigma_{j=1}^J q_{c(k,j)}\mu_{c(k,j)}x_{c(k,j)}^F$ for $k = 1,\ldots,K_W$. Finally, $\lambda_k(\infty,u^F)$ is substituted into the matrix $R(t,u(t))$ in (13) to obtain R. Under this key approximation, the objective function can be re-expressed as

$$\text{maximize} \quad \int_0^T (\beta h'x(t) - (1 - \beta)x(t)'Rx(t) + \gamma'\tilde{D}u(t))\,dt.$$
(19)

Problem (6)–(9) and (19) are linear quadratic control problems with state and control constraints and can also be viewed as an infinite-dimensional quadratic programming problem. This problem is mathematically difficult and little is known about the structure of the optimal solution (Hartl et al. 1995).

An iteration of the policy improvement algorithm. Our basic approach is to start with the equilibrium FCFT policy u^F and perform one step of the policy improvement algorithm. Let $V(x,t)$ denote the value function (as given by the objective in (19)) from time t to time T under the optimal policy, given that the state of the system at time t is x; for brevity of notation, we suppress the dependence of V on its arguments. Then Bellman's dynamic programming optimality equation is

$$-\frac{\partial V}{\partial t} = \beta h'x - (1 - \beta)x'Rx$$
$$+ \max_{u \in \Omega}\{(\nabla_x V)'(\lambda^+ - Ax + Bu) + \gamma'\tilde{D}u\}, \quad (20)$$

where Ω is the set of admissible allocation policies that satisfy constraints (7)–(9). Our approach is as follows. We solve the linear differential equations (6) to obtain the trajectory $x(t)$ under the equilibrium FCFT allocation policy (i.e., $u_{jk}(t) = u^F$ for all t), where u^F is calculated from the simulation model; to avoid complications arising from the boundary conditions $x(t) \geqslant 0$, we make our second assumption that the trajectory $x(t)$ is in the interior of the state space. This assumption is questionable in that the number of classes is much greater than the total number of customers (see the end of §2), but it is required for analytical tractability. To obtain the state trajectory under policy u^F, we solve the boundary value problem $dx(\tau)/d\tau = \lambda^+ - Ax(\tau) + Bu^F$, $x(t) = x$. Elementary results from the theory of ordinary differential equations show that

$$x(\tau) = e^{-A(\tau-t)}(x - x(\infty)) + A^{-1}(\lambda^+ + Bu^F)$$
for $\tau \in [t,T]$. (21)

By (19), the approximate value function under the equilibrium FCFT policy is

$$V^F = \int_t^T [\beta h'x(\tau) - (1 - \beta)x(\tau)'Rx(\tau) + \gamma'\tilde{D}u(\tau)]\,d\tau,$$
(22)

where $x(\tau)$ is defined in (21). This integration is performed in Zenios et al. (1997, Appendix B). Our last step is to perform one iteration of the policy improvement algorithm; i.e., we obtain $\nabla_x V^F$, substitute it into the right side of (20) and perform the minimization assuming that $x > 0$. The expression for $\nabla_x V^F$ is simplified by assuming that the time horizon T is very large (i.e., in years) in relation to the time scale of the system dynamics, which change on a daily or weekly basis (patients arrive and leave the waiting list daily). Therefore, our third approximation sets $T = \infty$ and $t = 0$ in $\nabla_x V^F$, which generates a policy that is independent of the time horizon. If we define $\pi(x) = (\pi_1(x), \dots, \pi_K(x))$ to be equal to $(\nabla_x V^s)'$, then

$$\pi(x) = \beta h' A^{-1} - (1 - \beta)(x^F)' R A^{-1} - (1 - \beta)(x - x^F)' \bar{R},$$

(23)

where the matrix $\bar{R}$ is defined in Zenios et al. (1997, Appendix B).

We can now substitute the partial derivatives (23) into the right side of (20) and perform the maximization. Because the function to be optimized is linear in the controls and (8)–(9) are knapsack constraints, the solution generates a *dynamic index policy* (Gittins 1989). If we define the indices

$$G_{jk} = \pi_{c(k,j)}(x(t)) - \pi_k(x(t)) + \gamma_k,$$

(24)

then at time t the proposed policy allocates all organs of class j to the transplant candidate class k with the highest index $G_{jk}(t)$. By (24), the cost γ_k of the likelihood of transplantation behaves as a *subsidy* for class k patients. Because the quantity $\pi_K(x(t))$ represents the marginal increase in $\int_t^T [\beta h' x(\tau) - (1-\beta)x(\tau)' R x(\tau)] d\tau$ if one additional candidate of class k is allowed to join the system at time t, the index $G_{jk}(t)$ equals the subsidy γ_k plus the marginal increase in $\int_t^T [\beta h' x(\tau) - (1 - \beta)x(\tau)' R x(\tau)] d\tau$ if an organ of class j is transplanted into a candidate of class k.

If the expression for V^F was exact, then the derived policy would be better than the equilibrium FCFT policy u^F. However, our expression is not exact (it assumes that $x(t) > 0$ and $T \to \infty$), and we cannot draw this conclusion. Nevertheless, this approach (performing one iteration of the policy improvement algorithm using approximate values for V) has been used to design dynamic call acceptance/rejection protocols for queueing network models of telephone traffic (Ott and Krishnan 1985, Key 1990) and develop dynamic multidrug therapies for patients infected with HIV (Wein et al. 1997).

4.3. Policy Simplification

The following proposition gives a closed-form expression for the dynamic indices $G_{jk}(t)$. The proof of this result involves cumbersome algebraic manipulations and is omitted.

PROPOSITION 1. *The dynamic index in (24) is given by* $G_{jk}(t) = \beta G_{jk}^1(t) + (1 - \beta) G_{jk}^0(t) + \gamma_k$, *where*

$$G_{jk}^1(t) = \frac{h_{c(k,j)}}{\mu_{c(k,j)}} + q_{c(k,j)} \frac{h_k}{\mu_k} - \frac{h_k}{\mu_k},$$

(25)

and

$$G_{jk}^0(t) = \sum_{i=1}^{K_W} \left[\frac{\lambda_i(\infty, u^F)}{\mu_k}(1 - q_{c(k,j)}) \right.$$
$$\left. \cdot \left(\frac{x_k^F}{\lambda_k(\infty, u^F)} - \frac{x_i^F}{\lambda_i(\infty, u^F)} \right) \right]$$
$$+ \sum_{i=1}^{K_W} \left[\frac{\lambda_i(\infty, u^F)}{2\mu_k} \left(1 - \frac{q_{c(k,j)}\mu_{c(k,j)}}{\mu_k + \mu_{c(k,j)}} \right) \right.$$
$$\cdot \left(\frac{x_k(t)}{\lambda_k(\infty, u^F)} - \frac{x_k^F}{\lambda_k(\infty, u^F)} \right)$$
$$- \frac{\lambda_i(\infty, u^F)}{(\mu_k + \mu_i)} \left(1 - \frac{q_{c(k,j)}\mu_{c(k,j)}}{\mu_i + \mu_{c(k,j)}} \right)$$
$$\left. \cdot \left(\frac{x_i(t)}{\lambda_i(\infty, u^F)} - \frac{x_i^F}{\lambda_i(\infty, u^F)} \right) \right]$$
$$+ \sum_{i=1}^{K_W} \sum_{l=1}^{J} \left[\frac{q_{c(k,l)}\mu_{c(k,l)}\lambda_i(\infty, u^F)}{2\mu_k(\mu_k + \mu_{c(k,l)})} \right.$$
$$\cdot \left(1 - \frac{q_{c(k,j)}\mu_{c(k,j)}(2\mu_k + \mu_{c(k,l)} + \mu_{c(k,j)})}{(\mu_k + \mu_{c(k,j)})(\mu_{c(k,j)} + \mu_{c(k,l)})} \right)$$
$$\cdot \left(\frac{x_{c(k,l)}(t)}{\lambda_k(\infty, u^F)} - \frac{x_{c(k,l)}^F}{\lambda_k(\infty, u^F)} \right)$$
$$- \frac{q_{c(i,l)}\mu_{c(i,l)}\lambda_i(\infty, u^F)}{(\mu_k + \mu_i)(\mu_k + \mu_{c(i,l)})}$$
$$\cdot \left(1 - \frac{q_{c(k,j)}\mu_{c(k,j)}(\mu_k + \mu_i + \mu_{c(i,l)} + \mu_{c(k,j)})}{(\mu_i + \mu_{c(k,j)})(\mu_{c(k,j)} + \mu_{c(i,l)})} \right)$$
$$\left. \cdot \left(\frac{x_{c(i,l)}(t)}{\lambda_i(\infty, u^F)} - \frac{x_{c(i,l)}^F}{\lambda_i(\infty, u^F)} \right) \right].$$

(26)

Proposition 1 reveals that the organs are allocated based on a weighted combination of the efficiency index $G_{jk}^1(t)$, the equity index $G_{jk}^0(t)$ and the subsidy γ_k. Notice that $h_{c(k,j)}/\mu_{c(k,j)} + q_{c(k,j)}h_k/\mu_k$ in (25) gives the quality-adjusted life expectancy (QALE) for a class k patient with a class j transplant (assuming that if the transplant fails, the patient will not receive additional transplants), and h_k/μ_k gives the QALE for a class k patient without transplantation. Hence, the efficiency index $G_{jk}^1(t)$ gives the increase in the QALE for a class k transplant candidate who receives a transplant of class j.

The equity index decomposes into three components, which are given by the three summations in Equation (26). The first component does not depend on the state of the system and is calculated using the equilibrium FCFT pair (x^F, u^F). When the state $x(t)$ is close to the equilibrium state x^F under the equilibrium FCFT policy, the second and

third components of Equation (26) are near zero, and the heuristic policy assigns priorities based largely on a static priority rule. The second and third components of $G_{jk}^0(t)$ attempt to maintain $x(t)$ close to x^F. The second component assigns highest priority to the candidate class with the waiting list size that most exceeds the waiting list size of the state x^F. The third component anticipates the influx of retransplantation candidates into the waiting list by measuring the total number of patients of each class that have a functioning graft, and then assigns highest priority to the class that expects the highest influx. Although our objective (14) adopts the absolute equity viewpoint, our approximation of $R(t, u(t))$ by $R(\infty, u^F)$ has led us to a policy that is similar in spirit to the relative equity viewpoint.

Two features of expression (26) may prevent its adoption by UNOS. First, it utilizes the number of functioning grafts of each class ($x_k(t)$ for $k = K_W + 1, \ldots, K$), and this information is not readily available in practice. Second, the expression for $G_{jk}(t)$ depends on the waiting list for candidate classes $l \neq k$. This represents a substantial deviation from the current UNOS policy, where the priority score for each candidate is computed using only the profile of that particular candidate. Thus, to make expression (26) more attractive, we replace $x_i(t)$ by the equilibrium FCFT value x_i^F for $i \neq k$ in (26), so that the index $G_{jk}(t)$ depends only on $x_k(t)$. This substitution eliminates the third component in (26), which was an order of magnitude smaller than the first two components in our exploratory numerical investigations, and part of the second component, and yields

$$\tilde{G}_{jk}(t) = \beta \left(\frac{h_{c(k,j)}}{\mu_{c(k,j)}} + q_{c(k,j)} \frac{h_k}{\mu_k} - \frac{h_k}{\mu_k} \right)$$

$$+ (1 - \beta) \left[\sum_{i=1}^{K_W} \left(\frac{x_k^F}{\lambda_k(\infty, u^F)} - \frac{x_i^F}{\lambda_i(\infty, u^F)} \right) \right.$$

$$\cdot \frac{\lambda_i(\infty, u^F)}{\mu_k} (1 - q_{c(k,j)})$$

$$+ \frac{\sum_{l=1}^{K_W} \lambda_l(\infty, u^F)}{\mu_k} \left(1 - \frac{q_{c(k,j)} \mu_{c(k,j)}}{\mu_{c(k,j)} + \mu_k} \right)$$

$$\left. \cdot \left(\frac{x_k(t)}{\lambda_k(\infty, u^F)} - \frac{x_k^F}{\lambda_k(\infty, u^F)} \right) \right] + \gamma_k.$$

$$(27)$$

The equity part of expression (27) reveals that higher priority is assigned to candidates with longer life expectancy and lower graft failure rates. Candidates with long life expectancy are expected to stay on the waiting list for a longer amount of time and thus contribute more to the total waiting time inequity. Awarding priority to candidates with low graft failure rates decreases the demand for retransplantation, thereby reducing the disparity between supply and demand, which is the driving force behind the efficiency-equity trade-off.
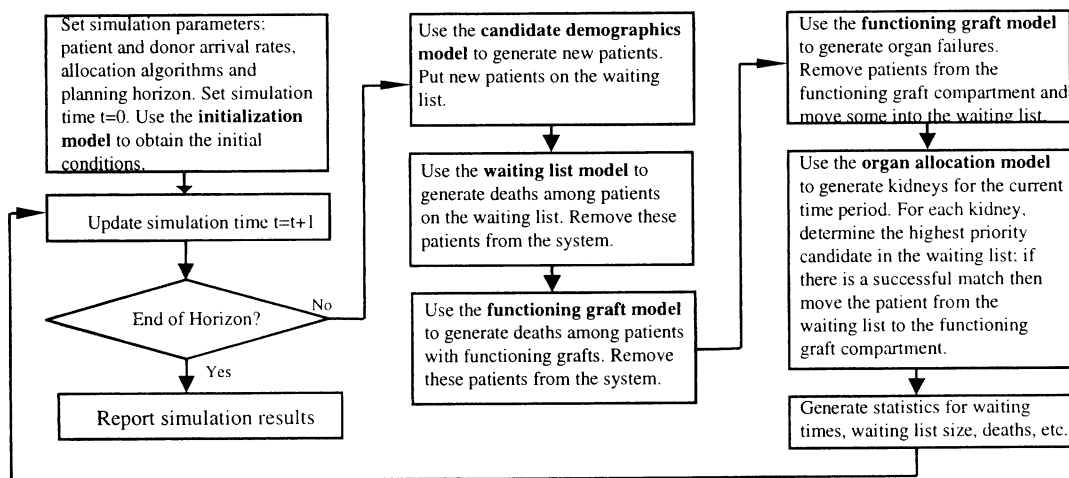
5. THE SIMULATION MODEL

To assess the performance of our proposed kidney allocation policies, we constructed a modular simulation model that mimics the operation of a single OPO. In the following subsections we provide a brief overview of the model and a detailed description of the component of the model that simulates the graft failure process. A more detailed description of the model is provided in Zenios et al. (1997), and the validation of the model is described in Zenios et al. (1998). The limitations of the model are discussed in §7.

5.1. An Overview of the Model

The simulation model is modular and contains five main models (see Figure 1): the candidate demographics model, the waiting list model, the organ allocation model, the functioning graft model, and the initialization model. The candidate demographics model generates new candidates for the transplant waiting list. These candidates remain in the waiting list model until they either exit the system by death or receive an organ from the organ allocation model and move into the functioning graft model. Patients remain in the functioning graft model until either they exit the system by death or the graft fails, in which case they can either rejoin the waiting list or exit the system. The initial population in the system is generated by the initialization model.

The state of the model consists of two vectors: one that maintains the characteristics of the waiting list population, and one that maintains the characteristics of the patients in the functioning graft model. The characteristics recorded by the model are gender, race (African-American, Caucasian, all non-African-Americans are treated as Caucasian in this paper), age (each year between 20 and 90; pediatric candidates are excluded from the model because they make up only 7% of the candidates and are more than twice as likely as adults to receive living-related transplants), tissue type, pra ($<60\%$, $>60\%$), prior transplants (none, one or more), body surface area (bsa), blood type, time of entry into the system, and time of entry into the current sector of the model. In addition, the functioning graft model records information about the graft for each transplant recipient; this information takes the form of a single variable extracted from the proportional hazards model for the graft survival process (see §5.2). The system state is updated monthly (this frequency provides accuracy and computational efficiency) according to the sequence of events presented in the flowchart in Figure 1.

The main inputs for the simulation model are the patient and donor arrival rates, the distributions for the patient and donor characteristics, the pre-transplantation and post-transplantation mortality rates, the graft failure rates, the QOL scores; see Appendix A for a summary of these inputs. These inputs were estimated using data from the UNOS (UNOS 1995), the United States Renal Data System (USRDS 1995), and the New England Organ Bank, as well as data reported in the published literature. The estimation of all the inputs except the graft failure rates (§5.2) involved

Figure 1. Flowchart for the simulation model.

Set simulation parameters: patient and donor arrival rates, allocation algorithms and planning horizon. Set simulation time t=0. Use the **initialization model** to obtain the initial conditions.

Update simulation time t=t+1

End of Horizon? — No

Yes

Report simulation results

Use the **candidate demographics model** to generate new patients. Put new patients on the waiting list.

Use the **waiting list model** to generate deaths among patients on the waiting list. Remove these patients from the system.

Use the **functioning graft model** to generate deaths among patients with functioning grafts. Remove these patients from the system.

Use the **functioning graft model** to generate organ failures. Remove patients from the functioning graft compartment and move some into the waiting list.

Use the **organ allocation model** to generate kidneys for the current time period. For each kidney, determine the highest priority candidate in the waiting list: if there is a successful match then move the patient from the waiting list to the functioning graft compartment.

Generate statistics for waiting times, waiting list size, deaths, etc.

straightforward statistical analysis, and thus the details are omitted.

The model considers two main scenarios for the patient and donor arrival rates. The first scenario, which is referred to as the typical OPO, represents an OPO that covers 1/72 of the national waiting list (recall that there are 72 OPOs in the United States). It assumes that the patient arrival rate at year t is $\lambda^+(t) = 142.90 + 4.48t$ patients per year ($t = 0$ refers to 1995), and the donor arrival rate is 57.09 donors per year (each donor constitutes two identical organs). On the other hand, the second scenario considers a congested OPO that faces a severe shortage of organs, and it assumes that the patient arrival rate at year t is $\lambda^+(t) = 642.74 + 20.16t$ patients per year and the donor arrival rate is 169.0 donors per year (these parameters were obtained from the New England Organ Bank). In addition, the model assigns two QOL scores: 0.60 for transplant candidates and 0.75 for transplant recipients with a functioning graft. These QOL scores reflect the experience of ESRD patients surveyed in the published literature (Deniston et al. 1989, Hornberger et al. 1991).

We conclude this subsection by highlighting the similarities and differences between the simulation model and the fluid model. In general, while the simulation model provides a very fine micro-representation of the organ allocation process, the fluid model provides an aggregate macro-representation. However, the fluid model adopts the following simplifications relative to the simulation model: (a) Rather than maintain a list of all candidates and recipients, it simply classifies the candidates and recipients according to their characteristics, and keeps track of the total number of candidates and recipients in each class; (b) instead of modeling the discrete stochastic flow of candidates and donors, it assumes a continuous and deterministic flow that is driven by a homogeneous inflow of donors and patients, and a constant-hazard mortality and graft failure process; (c) rather than allowing for different graft failure rates between first and second time recipients, it assumes that the failure rates are not affected by the number of transplantations; (d) instead of allowing a positive crossmatch, it assumes that crossmatching is always negative; (e) instead of allowing candidates to be unavailable, it assumes that candidates are always available for transplantation; (f) rather than allowing organ sharing between OPOs, it assumes no such sharing.

5.2. The Functioning Graft Model

The functioning graft model monitors the health of individuals after transplantation. Individuals can exit this model via death or graft failure, and these two processes are assumed to operate independently of each other. The mortality is simulated using the mortality rates presented in Appendix A.

The graft failure rates are modeled by a proportional hazards model with a nonparametric baseline; see Cox and Oakes (1984). The model hypothesizes that the hazard function for a patient with a known covariate vector x is $h(t) = h_0(t)e^{\theta' x}$, where θ is the (unknown) vector of regression coefficients, and $h_0(t)$ is the (unknown) baseline hazard function; the hazard function $h(t)$ gives the instantaneous conditional probability that the organ will fail during $[t, t + \Delta t)$ given that it did not fail before time t. The vector of covariates summarizes all risk factors that are known to predict the short-term and long-term graft survival, and is listed in Table 2. The exponent $e^{\theta' x}$ is referred to as the *prognostic index*.

Proportional hazards models have been used to study risk factors associated with graft failure (Chertow et al. 1996) and to develop prognostic models for graft failures (Van Houwelingen and Thorogood 1995). In fact, we employ the covariates that were found to have the greatest contribution to the overall graft survival in Chertow et al.

The data used for the analysis are extracted from the UNOS Public-Use Data Set (UNOS 1995). The data set is divided into two subsets: the training set, which is used to calibrate the model (sample size = 23,538); and the validation set, which is used to validate the model

Table 2. Results from fitting a proportional hazard model to the training set.

Factor	Categories	Regression Coefficient	Conditional Relative Risk
Donor and recipient sex	male to male or male to female or female to female	0 (baseline)	1
	female to male	0.114 ± 0.059	1.120
Recipient race	Non-African-American	0 (baseline)	1
	African-American	0.421 ± 0.055	1.523
Donor race	Non-African-American	0 (baseline)	1
	African-American	0.165 ± 0.082	1.180
Recipient Age	0–10	0 (baseline)	1
	10–20	0.071 ± 0.203	1.073
	20–30	−0.185 ± 0.190	0.831
	30–40	−0.280 ± 0.187	0.756
	40–50	−0.362 ± 0.189	0.696
	50–60	−0.435 ± 0.100	0.647
	60–70	−0.488 ± 0.205	0.614
	70–80	−0.277 ± 0.361	0.758
Donor Age	0–10	0 (baseline)	1
	10–20	−0.467 ± 0.097	0.627
	20–30	−0.502 ± 0.097	0.605
	30–40	−0.359 ± 0.101	0.698
	40–50	−0.217 ± 0.103	0.805
	50–60	−0.015 ± 0.107	0.985
	60–70	0.181 ± 0.137	1.200
	70–80	−0.403 ± 0.659	0.668
Peak pra	presensitized	0 (baseline)	1
	non-presensitized	−0.384 ± 0.069	0.681
Body surface area	<1.60	0 (baseline)	1
	1.60–1.80	0.071 ± 0.079	1.073
	1.80–2.00	0.104 ± 0.080	1.110
	2.00–2.20	0.199 ± 0.093	1.220
	>2.20	0.353 ± 0.130	1.422
Previous transplants	0	0 (baseline)	1
	>0	0.253 ± 0.065	1.288
HLA-A mismatches	0	0 (baseline)	1
	1	0.092 ± 0.103	1.096
	2	0.122 ± 0.105	1.129
HLA-B mismatches	0	0 (baseline)	1
	1	0.190 ± 0.115	1.209
	2	0.264 ± 0.114	1.302
HLA-DR mismatches	0	0 (baseline)	1
	1	0.099 ± 0.095	1.104
	2	0.250 ± 0.996	1.283

(sample size = 7,713). Patients who were lost to follow-up, died with a functioning graft, or had a functioning graft on the last day of follow-up (December 1991) were assumed to contribute a right-censored observation. Patients with one or more missing covariates were included in the analysis, and the missing covariates were treated as new covariates.

The regression coefficients θ and the nonparametric baseline hazard function $h_0(t)$ were obtained using the S-Plus system. The results are presented in Table 2 and are discussed in the context of our simulation results in §6.1.

We perform three tests to validate the proportional hazards model. In the first test, we compare the distribution

Figure 2. Validation of the graft survival model. Panel (a) gives the Kaplan–Meier survivor curve (solid line) and the mean survivor curve predicted by our model (dotted line). Panel (b) gives the Kaplan–Meier survivor curve (solid lines) and the mean survivor curve predicted by our model (dotted line) for four groups: $-0.5410 \leqslant$ prognostic index $< -0.3052, -0.3052 \leqslant$ prognostic index $< -0.0141, -0.0141 \leqslant$ prognostic index < 0.14141, and prognostic index > 0.1414.
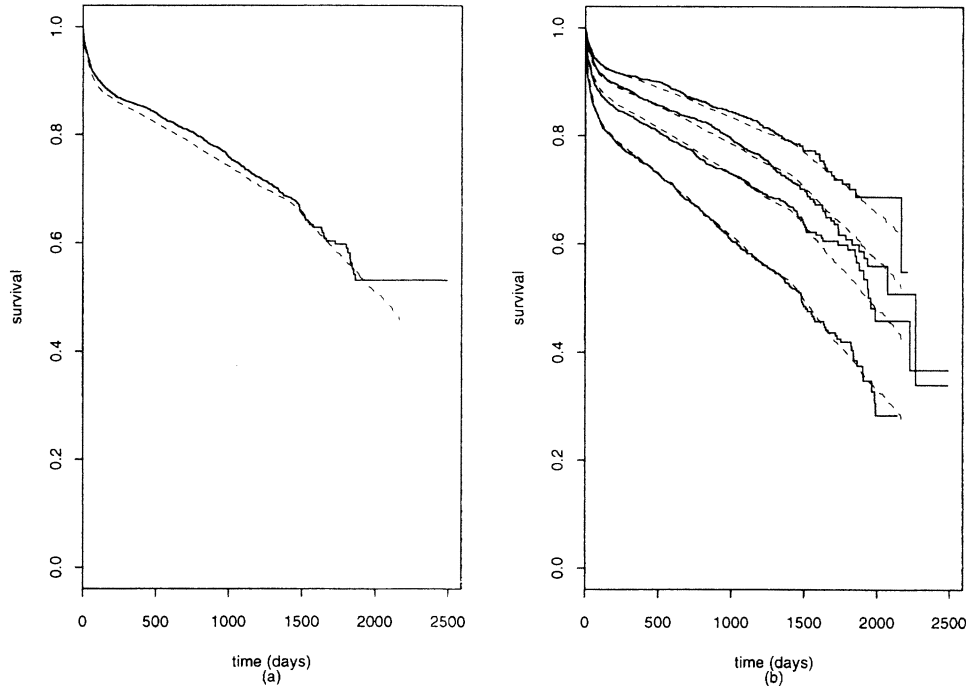
of the prognostic indices in the validation set to the prognostic indices in the training set. The two distributions are virtually indistinguishable. In the second test, we fit the validation data set to a proportional hazards model with a single covariate, the prognostic index. The resulting regression coefficient is 1.0 ± 0.058, indicating that the underlying model is correct. The final test compares the Kaplan-Meier survivor curve (which is a nonparametric unbiased estimate of the underlying survivor curve) of the validation set to the mean survivor curve predicted by the graft survivor model; see Figure 2(a). If the proposed model is correct, the mean survivor curve is an unbiased estimator of the exact survivor curve. The test also categorizes the prognostic index for the validation set into four groups (each group contains about 25% of the data), and for each group, computes the Kaplan-Meier survivor curve and compares it to the mean survivor curve predicted by the proportional hazards model; see Figure 2(b). The figures show that the mean survivor functions predict with precision the Kaplan-Meier curves. These three tests establish the validity of the proportional hazards model.

Because the graft survival process can be summarized very succinctly by the prognostic index $e^{\theta'x}$, it is not necessary to record detailed information about the graft in the functioning graft model. Rather, as mentioned in §5.1, it suffices to summarize all important information using the prognostic index.

6. SIMULATION RESULTS

In this section, we use the simulation model of §5 to study the kidney allocation problem. The main results for the typical OPO are presented and discussed in §6.1, and a variety of other policies and scenarios are considered in §6.2.

6.1. The Typical OPO

This subsection describes the performance of three policies (FCFT, UNOS, SEEP (β, γ)) for the typical OPO.

Experimental design. For each policy and scenario considered, the OPO was simulated for a period of 10 years (1995–2004); longer horizons were deemed inappropriate because the field of transplantation evolves at a very rapid pace, and predictions that extend to more than 10 years would be questionable. Confidence intervals of 95% for the performance measures were obtained by performing 40 independent runs of the simulation model, starting from the same initial conditions. For the proposed SEEP (β, γ), we set $\gamma = 0$ and considered several values of $\beta \in [0, 1]$. Then we set $\beta = 1$ (i.e., maximum efficiency) and varied the subsidy γ for African-Americans, because they have experienced long waiting times until transplantation. More details about the experimental design and the implementation of the SEEP policy are provided in Zenios et al. (1997).

Table 3. Efficiency performance (with 95% confidence intervals) of different allocation policies in the typical OPO.

Policy	QALY	Mean WTT
FCFT	32.64 ± 0.12	27.90 ± 0.15
UNOS	32.72 ± 0.13	25.41 ± 0.34
SEEP (1.00, 0)	34.20 ± 0.12	8.67 ± 0.12
SEEP (0.83, 0)	33.82 ± 0.12	21.94 ± 0.31
SEEP (0.62, 0)	33.15 ± 0.13	26.40 ± 0.31
SEEP (0.50, 0)	32.98 ± 0.11	26.81 ± 0.34
SEEP (0.32, 0)	32.73 ± 0.13	26.92 ± 0.30
SEEP (0.00, 0)	32.57 ± 0.12	26.75 ± 0.28
SEEP (1.00, 1.6)	33.86 ± 0.13	10.61 ± 0.15
SEEP (1.00, 1.8)	33.56 ± 0.17	10.81 ± 0.16

Results. The results are presented in Tables 3 and 4. For 10 policies (FCFT, UNOS, and 8 SEEP (β, γ)s), Table 3 presents two clinical efficiency measures: the quality-adjusted life years (all results are reported in months and truncated at the end of the 10-year simulation period) per transplant candidate and the mean waiting time until transplantation (WTT); we do not report the aggregate likelihood of transplantation because it is unaffected by the allocation policy (it is dictated by the demand-to-supply ratio). For our three model outputs (QALY, mean waiting time until transplantation and likelihood of transplantation) and these 10 policies, Table 4 presents the difference in the mean output between the two genders (G), between the two races (R), and between patients above and below 50 years old (A). This table allows us to assess nine inequities from both the absolute equity and relative (to FCFT) equity viewpoints: gender, race, and age inequity for three performance measures. These differences are more informative than the pairwise square difference terms used in our mathematical objective function because they reveal the *sign*, not just the magnitude, of the difference between a pair of numbers.

If policy A has a larger absolute value of an entry (in a particular column) in Table 4 than policy B, then we say that policy A has a larger *absolute* inequity (for that particular column) than policy B. If the absolute value of an entry

(in a particular column) for non-FCFT policy A minus the corresponding FCFT entry is greater than the absolute value of the entry for non-FCFT policy B minus the FCFT entry, then we say that policy A has a larger *relative* inequity (for that particular column) than policy B. We also abbreviate the three performance measure inequities by QALYI, WTTI, and LTI and will usually preface them with two descriptors: absolute or relative, and gender, race or age.

Discussion. The following seven observations can be extracted from our numerical results.

(1) *Summary of demographic statistics*: Many of the numerical results are driven by the historical demographic statistics found in Appendix A and the parameters of the proportional hazards model in Table 2. Before discussing the results, we summarize the salient characteristics of the data in Appendix A. African-Americans comprise 29.8% of the transplant candidates and 11.2% of the donors. Females comprise 39.0% of the candidates and 36.7% of the donors. Females make up 43.0% of African-American candidates, 37.3% of Caucasian candidates, 36.6% of African-American donors, and 36.7% of Caucasian donors. Male African-American candidates are younger (40.4% are less than 50) than the other three gender-race pairs (about 25.0% are less than 50). Female candidates have 8.3% higher average pra than males, and African-Americans have 10.1% higher average pra than Caucasians. Age-specific mortality rates are higher for dialysis patients than for transplant recipients. The mortality rates for dialysis patients and for transplant recipients increase with age and are higher for males than females. African-Americans have lower mortality rates on the waiting list than Caucasians but have higher mortality rates with transplants. Also, the *differential mortality rates*, which are the age-specific mortality rates for dialysis patients minus the age-specific mortality rates for transplant recipients, are significantly higher (e.g., 60% higher for 45–50 year olds) for Caucasians than African-Americans. African-American candidates are 16.1% less likely than Caucasians to have type A blood and 10.6% more likely to have type B blood. Finally, the

Table 4. Absolute Inequity Outcomes. (The columns denoted G give the difference Females *minus* Males, the columns R give the difference African-Americans *minus* Caucasians, and the columns A give the difference older than 50 *minus* younger than 50.)

Policy	QALYI			WTTI			LTI		
	G	R	A	G	R	A	G	R	A
FCFT	1.08	0.63	−19.15	0.22	3.11	−1.02	−0.72	12.68	−44.24
UNOS	1.50	0.62	−18.47	−1.80	5.42	−2.49	4.04	8.56	−46.02
SEEP (1.00, 0)	1.78	−5.09	−16.03	−1.33	20.56	−6.67	3.14	−94.97	13.82
SEEP (0.83, 0)	2.02	−2.66	−18.86	−2.25	32.09	−13.10	6.35	−41.61	−25.05
SEEP (0.62, 0)	1.74	−0.28	−20.14	−1.09	11.87	−2.89	3.85	−1.50	−47.77
SEEP (0.50, 0)	1.29	0.38	−20.80	−0.53	7.02	−0.00	1.66	9.34	−54.76
SEEP (0.32, 0)	1.13	0.89	−20.85	−0.17	3.72	2.14	2.65	18.77	−60.95
SEEP (0.00, 0)	1.09	1.31	−20.83	0.05	0.81	11.91	1.49	26.07	−66.35
SEEP (1.00, 1.6)	2.08	−0.84	−16.50	−2.14	3.95	−10.48	8.75	−16.39	20.93
SEEP (1.00, 1.8)	2.15	−0.10	−15.16	−2.41	2.24	−11.20	6.18	0.62	25.28

frequencies for the different tissue type proteins differ considerably between the two races (see Zenios et al. 1997, Appendix C). For example, protein 53 appears in 15% of African Americans but in only 0.8% of Caucasians.

(2) *Equity performance of FCFT.* To understand how the facts in observation (1) and the parameter estimates in Table 2 affect the results, we begin with a discussion of the absolute equity performance measures under the FCFT policy in Table 4. Recall that the FCFT policy is subject to the screens for presensitization and blood type compatibility. With regards to gender, females are less likely to receive a transplant than males because they have higher pra levels. Their pra levels and their lower mortality rates on the waiting list lead to slightly longer waits until transplantation. Despite females being less likely to receive a transplant, the low mortality rates for female dialysis patients and transplant recipients allow them to experience more QALY than males.

With regards to race, African-Americans have longer waiting times until transplantation for three reasons: They have higher pra levels, lower mortality rates as dialysis patients, and a blood type mismatch with Caucasians coupled with a higher demand-to-supply ratio than Caucasians. Nevertheless, African-Americans are more likely to receive a transplant than Caucasians because they are younger and have lower mortality rates on the waiting list. This higher likelihood of transplantation allows African-Americans to experience more QALY than Caucasians, even though they have higher graft failure rates and higher mortality rates as transplant recipients.

Turning to age equity, dialysis patients over 50 have higher mortality rates than younger patients, and hence experience smaller waits until transplantation (as predicted by the results in Zenios 1999) and a much smaller likelihood of transplantation. The smaller likelihood of transplantation, coupled with the higher mortality rates on the waiting list and with functioning grafts, causes patients over 50 to have much fewer QALY than patients under 50.

(3) *FCFT vs. UNOS.* Relative to FCFT, the UNOS policy favors females and African-Americans via its pra points but puts African-Americans at a disadvantage via its tissue matching points. Surprisingly, these two policies have approximately the same QALY in Table 3: The mean value for UNOS is only 2.4 days higher than for FCFT. Most of the efficiency gained by good tissue matching is offset by awarding priority to presensitized patients (under UNOS, the mean waiting time is 14.71 months for presensitized patients and 32.49 months for nonpresensitized patients), who have higher graft failure rates on average.

FCFT has the highest mean waiting time until transplantation among the 10 policies under consideration. This is because any policy that gives priority to a set of patients should generate a waiting time distribution that has fatter left and right tails than the distribution generated by FCFT. Moreover, the waiting time until transplantation distribution comprises the left most portion of the waiting time distribution, because patients who die while on the waiting list essentially truncate much of the right tail of the waiting time distribution.

With regard to inequity, UNOS's pra points reduce the waiting time until transplantation for females (which increases the absolute gender WTTI relative to FCFT) and gives females a higher likelihood of transplantation (which reduces the absolute gender LTI) and an increased QALY (which increases the absolute gender QALYI). The effect of UNOS's tissue matching points appears to dominate the effect of its pra points for African-Americans; relative to FCFT, UNOS generates longer waiting times until transplantation for African-Americans (and an increase in the absolute race WTTI). Although UNOS's tissue matching points reduce African-Americans' likelihood of transplantation (and the absolute race LTI), their lower mortality rates on the waiting list still allow them to be more apt to receive a transplant than Caucasians. There is no significant difference in QALY for African-Americans under the two policies: The improved tissue matching provided by UNOS compensates for the fewer African-American transplants undertaken by this policy. Because African-American candidates are younger on average than Caucasians, UNOS offers a slight advantage to older patients relative to FCFT.

(4) *The efficiency-equity trade-off.* The efficiency-equity trade-off is epitomized by comparing the FCFT policy, which represents the gold standard for equity under the relative equity viewpoint, and the SEEP(1,0), which represents the gold standard for efficiency. Relative to FCFT, SEEP(1,0) is disadvantageous to African-Americans, presensitized patients (and hence females and African-Americans), and retransplantation candidates and favors good tissue type matches, older recipients (because of their low graft failure rates), and female-to-female allocations; see Table 2. SEEP(1,0) increases the QALY per patient by 1.56 months compared to FCFT. This increase is equivalent to 7596 graft years per 100,000 recipient years, which is 84.4% of the estimated improvement achieved by the introduction of immunosuppressive drugs in the late 1970s (Opelz and Wujciak 1995). SEEP(1,0) reduces the mean waiting time until transplantation by 19.2 months. This dramatic reduction is achieved at the expense of low priority candidates, who are forced to spend prolonged periods of time on the waiting list and then die before receiving a transplant (however, among the 10 policies tested, SEEP(1,0) minimizes the number of deaths on the waiting list and the mean waiting time of all candidates).

With regard to gender equity, the female-to-female effect in the SEEP(1,0) dominates the disadvantage females experience via the pra points and the slightly higher proportion of African-American candidates who are female. Hence, relative to FCFT, females experience shorter waiting times until transplantation (which increases the absolute gender WTTI), a higher likelihood of transplantation (which reduces the absolute gender LTI), and more QALY (which increases the absolute gender QALYI). African-Americans are at a distinct disadvantage under the SEEP(1,0) policy because they have high graft failure rates, high pra levels, rare tissue types

combined with a high demand-to-supply ratio, and small differential mortality rates. Consequently, they incur a large increase in waiting times until transplantation, are much less apt to receive a transplant (95.3% of all organs are allocated to Caucasians), and incur a reduction in QALY, all of which increase the absolute race inequity measures. The low graft failure rate of older patients and the fact that African-Americans are younger lead the SEEP(1,0) policy to favor older patients, who experience relatively shorter waits until transplantation (which increases the absolute age WTTI), a much higher likelihood of transplantation (which reduces the absolute age LTI), and more QALY (which reduces the absolute age QALYI).

(5) *SEEP*(β, 0)*s*. Tables 3 and 4 show that as β decreases from 1 to 0, clinical efficiency (i.e., QALY) decreases and most of the (absolute and relative) inequities decrease. Although none of the SEEP(β, 0)s achieve good clinical efficiency and low absolute race WTTI and LTI, the SEEP(0.5, 0) increases the QALY (relative to UNOS) *for all six patient groups* (females, males, African-Americans, Caucasians, over 50, under 50).

The absolute WTTIs do not exhibit monotonic behavior with respect to β: They increase until β reaches 0.83, and then decrease. This occurs because as β is reduced from its maximum value of unity, the SEEP(β, 0) policy starts to allocate an increasingly larger fraction of the organs to some low priority candidates with very long waiting times, and these candidates cause the peak in the WTTIs. As β decreases even further, the waiting time differences between the low- and high-priority candidates become less significant and the WTTIs starts to decrease again. Although the numbers are not reported here, the waiting time inequities and the WTTI measures are quite similar when $\beta < 0.83$, which is reassuring in view of our substitution of the former for the latter in §3.

(6) *Subsidized policies*. Policies SEEP(1,1.6) and SEEP(1,1.8) demonstrate that increases in the QALY per candidate need not always be achieved at the expense of equity between races, SEEP(1,1.8) has a higher QALY and lower absolute race WTTI and LTI than UNOS and FCFT. Relative to SEEP(1,0), the subsidized policies provide superior health outcomes for African-Americans but lead to a reduction in QALY for Caucasians and an increase in the (absolute and relative) age and gender inequity measures. However, relative to UNOS, the SEEP(1,1.6) policy increases the QALY and reduces the mean waiting time until transplantation for all six patient groups.

(7) *Efficient frontier*. We used the data in Tables 3 and 4 to generate nine efficiency-equity trade-off plots (that are not displayed here), one for each of the columns of inequity numbers in Table 4. Each plot contains 10 points, one for each of the policies in these two tables, and we plot the QALY versus the inequity measure. These plots allow us to identify policies that are on the efficient frontier of efficiency and equity (i.e., there are no other policies that dominate on both dimensions). From the absolute equity viewpoint, 28 of the 90 points (i.e., policies) on the 9 plots are on an efficient

frontier: Of these, 19 are SEEP(β,0)s, 7 are SEEP(1, γ)s, 2 are FCFT (gender QALYI and LTI) and none are UNOS. Of the 38 points that are on an efficient frontier from the relative equity viewpoint, 23 are SEEP(β,0)s, 9 are FCFT (by definition of relative equity), 5 are SEEP(1, γ)s, and 2 are UNOS (race QALYI and age LTI). In summary, the SEEP(β, γ)s constitute most of the efficient frontiers, and UNOS is far from these frontiers with respect to about half of the (absolute and relative) inequity measures.

6.2. Further Results

In this subsection we consider a variety of different policies and several other scenarios and perform a sensitivity analysis.

Patient characteristics and clinical efficiency. To investigate how the various patient characteristics that are utilized by the SEEP(1,0) affect the QALY per candidate, we consider several new policies that assign priorities using the regression coefficients of the proportional hazards model in §5.7. Seven policies are considered, and each policy assigns priorities using only one of the following patient characteristics: age, gender, race, tissue, type, prior transplants, body surface area (bsa), and pra. For example, the policy that utilizes age ignores any patient characteristic other than age and assigns highest priority to candidates of age 60–70 (because they have the lowest regression coefficient in Table 2 and thus the lowest graft rejection rate). As a reference point, we also consider a policy that assigns priorities using all seven patient characteristics.

The results are described in Table 5. The far-right column gives the increase in QALY compared to UNOS as a fraction of the QALY increase achieved by SEEP(1,0); i.e., [QALY(policy) − QALY(UNOS)]/[QALY(SEEP(1,0)) − QALY(UNOS)]. The race-based policy achieves the largest increase (75.0%) and the bsa-based policy the smallest (18.92%). Furthermore, characteristics that are not used by the current UNOS system, such as prior transplants, are almost as important as the characteristics that are used by the current system, such as tissue types. In addition, the policy that utilizes all seven characteristics slightly outperforms the SEEP(1,0) but the difference is not statistically significant (statistical significance corresponds to about a

Table 5. The impact of various patient characteristics on clinical efficiency.

Patient Characteristics	Quality Adjusted Life Months	% of Improvement Achieved
all	34.28 ± 0.15	105.40
race	33.83 ± 0.13	75.00
tissue type	33.58 ± 0.12	58.08
prior transplants	33.36 ± 0.16	43.24
pra	33.32 ± 0.15	40.54
age	33.18 ± 0.12	31.08
gender	33.04 ± 0.13	21.62
bsa	33.00 ± 0.16	18.92

10% difference in the far-right column of Table 5). This has important practical implications as it suggests that simple policies based on the prognostic index can achieve most, if not all, of the efficiency improvement achieved by the theoretically justified SEEP. The effects of the various characteristics in Table 5 are not additive, but a detailed study of their interactions is beyond the scope of this paper.

Attempted improvement of SEEP. Motivated by the structural results in Derman et al. (1972) and the statistical results in Chertow et al. (1996), we altered the SEEP(1,0) policy by disallowing extreme disparities in the age of donors and recipients: Candidates under 35 could not receive an organ from donors over 65, and candidates over 65 could not receive an organ from donors under 35. This change led to a *reduction* in efficiency relative to SEEP(1,0).

Attempted improvements of the UNOS policy. Because FCFT (which underlies the waiting time points of the UNOS policy) is nearly as efficient as UNOS, the only possible improvements *within the current structure* of the UNOS policy are to alter the number of points awarded to tissue matching and presensitization (recall that our model omits pediatric patients). To examine the first possibility, we analyze a variant of UNOS where priorities are assigned using only the tissue matching portion of the policy. The simulation results show that the QALY per patient for this policy is 33.5 ± 0.15 months, which is nearly identical to the QALY per patient for the tissue type-based policy in Table 5. Hence, no further improvements can be achieved with this portion of the policy.

Observation (3) in §6.1 motivates us to consider a variant of UNOS where the priority points for presensitized patients are decreased from four to one. This variant increases the QALY by 0.41 months and increases the mean waiting time until transplantation by 1.01 months. It also decreases six of the nine absolute inequity measures (all except age QALYI, race WTTI, and age WTTI) and six of the nine relative inequity measures (the gender and age inequities). Overall, African-Americans, prior transplant recipients (these patients get priority under UNOS because they have higher pra levels on average) and females are slightly worse off under this variant of the UNOS policy.

The congested OPO. The large variation in the demand-to-supply ratio among the various OPOs leads to significant disparities in performance across OPOs. The simulation results for the congested OPO, which has a 34.2% higher demand-to-supply ratio than the typical OPO, confirm that the demand-to-supply ratio underlies the equity-efficiency trade-off: Compared to the typical OPO, the efficiency measures decrease considerably, and all the absolute inequity measures (except for age LTI) increase. Averaging over the eight policies that were tested (the 10 policies in Table 3 with the exception of SEEP(1, 1.8) and SEEP(0.62, 0)), we find that the QALY decreases by 3.81 months (this loss is roughly twice the estimated gain achieved by immuno-

suppressive drugs) and the mean waiting time until transplantation increases by 15.64 months (a 70.6% increase). A comparison of FCFT and SEEP(1,0) for the congested and typical OPOs shows that as the demand-to-supply ratio increases, the efficiency-equity trade-off becomes more severe; i.e., increases in the QALY per patient are achieved at the expense of larger increases in all the absolute inequity measures (except for age LTI).

Effect of organ supply. To explore the impact of organ shortage on the performance of our policies, we simulated FCFT, UNOS, SEEP(1,0) and SEEP(0.5,0), assuming that the organ supply increases by 20%. The results show that increasing the organ supply improves the clinical efficiency and decreases all the absolute inequity measures. Averaging over the four policies, we find that the QALY per candidate increases by 1.03 months (which is 66% of the improvement achieved by SEEP(1,0) in the typical OPO), and the mean waiting time until transplantation decreases by 1.42 months.

Effect of QOL Scores. Because the QOL scores represent a subjective evaluation of different health states, we now evaluate the sensitivity of our results to these scores. We performed additional simulation runs for the SEEP(1,0) and SEEP(0.5,0) using the typical OPO and QOL scores of 0.5 and 0.7 (the original value was 0.6) for waiting list patients (QOL scores for transplant recipients remained at 0.75). The results did not produce statistically significant differences in the waiting times until transplantation or the (unadjusted) life years per patient. This suggests that the relative ranking of the indices (27) is not sensitive to the actual QOL scores, and patients who would receive an organ based on one set of QOL scores would most likely receive the same organ under a different set of scores (provided that the QOL for transplantation is sufficiently higher that the QOL for the waiting list).

7. CONCLUDING REMARKS

The efficiency-equity trade-off in kidney transplantation can be alleviated by decreasing the demand-to-supply ratio (in total and within specific demographic groups), reducing demographic- and nondemographic-based differences in graft survival rate, and employing an organ allocation policy that explicitly addresses this trade-off. The first approach requires changes in the legal system or the attitudes or behavior of the general public, and the second approach entails considerable advances in clinical research. Because these changes are not expected to be achieved in the foreseeable future, this paper has analyzed the third approach.

This trade-off is very complex, and there are a variety of ways to measure performance, segment the ESRD population into groups, and assess equity. We focus on three metrics that span the basic issues: quality-adjusted life years, waiting time until transplantation, and likelihood of transplantation. We report on nine types of equities, which are generated by comparing these three quantities across

gender, race (African-Americans and Caucasians), and age (patients above and below 50). In addition, we believe that the problem's complexity demands that inequity be assessed relative to a benchmark that represents perfect equity. We discuss two simple and natural benchmarks that lead to two approaches for assessing equity: an absolute equity viewpoint, which takes as its gold standard a hypothetical policy that equalizes all three performance measures across all the demographic groups; and a relative equity viewpoint, which uses the first-come first-transplanted policy as its gold standard. These two viewpoints lead to different conclusions: FCFT does not treat different demographic groups identically with regard to our three performance measures because it is subject to pre-screening for presensitization and blood compatibility and because of demographic-based differences in survival rates of dialysis patients and demand-to-supply ratios. Most notably, African-Americans experience longer waits until transplantation but are more likely to receive a transplant, and patients over 50 are much less likely to receive a transplant and have less QALY than patients under 50.

We have developed a mathematical model that describes the first-order dynamics of the ESRD population and superimposed on it an objective function that maximizes QALY and minimizes inequity in waiting times and inequity in likelihood of transplantation. Our analysis, which includes a number of simplifying approximations, eventually leads to a closed-form heuristic dynamic index policy. This policy assigns priorities using a mixture of efficiency points and equity points that are similar to the points used by UNOS, but it uses (in a nonobvious way) a larger set of historical information about the patient and donor characteristics and their relationship to clinical efficiency, and employs subsidies that can be used to reduce disparities between specific groups.

We have constructed (and validated in Zenios et al. 1998) a large-scale simulation model to test an empirically calibrated version of our index policy. Although the simulation results demonstrate that this index policy outperforms FCFT and the UNOS policy, we have attempted neither to compare the performance measures generated by the large-scale simulation model and the simplifying fluid model nor to assess the suboptimality (within the fluid model) of our dynamic index policy.

A surprising result from our simulation study is that the policy currently used by UNOS is not appreciably more efficient (in terms of quality-adjusted life years per patient) than the FCFT policy; hence, if one views QALY as the primary efficiency measure and adopts the relative inequity viewpoint, then FCFT is preferable to UNOS. In contrast, the most efficient policy within our class of proposed policies (i.e., there is no weight given to equity and no subsidy points) achieves an increase in QALY per patient (relative to FCFT over a 10-year period) that is 84.4% of the estimated improvement achieved by immuno-suppressive drugs and is roughly comparable to what would be achieved by a 30% increase in the supply of donated organs. Much of the effi-

ciency gains are achieved by employing demographic (e.g., favoring female-to-female transplants, penalizing African-American candidates) and nondemographic (e.g., penalizing re-transplantations) factors that are not included in the UNOS policy.

However, African-Americans fare poorly under this policy on all three performance measures. Nevertheless, our proposed policy can reduce these inequities by increasing the weight given to equity or by providing a subsidy to African-Americans. Tables 3 and 4 show that our class of policies generate efficient frontiers on the efficiency-equity trade-off plots that dominate the UNOS policy for most equity measures; however, fine tuning of both the equity weight and the subsidies is required to generate a policy that dominates the UNOS policy on all nine equity measures simultaneously.

Rather than seek policies that are on the efficient frontier of efficiency and equity, another approach is to find policies that simultaneously improve the health outcomes of all patient groups (even though such policies may increase discrepancies across groups). One of our proposed policies (SEEP(1, 1.6)) increases the QALY and decreases the mean waiting time until transplantation for all six patient groups (females, males, African-Americans, Caucasians, over 50 years old, under 50) relative to the UNOS policy.

Our simulation model has several limitations, and therefore our simulation results should be regarded with some degree of caution. As noted by Paltiel (1997), the kidney allocation process is actually a two-step process, where a centralized point system offers an organ to the top candidate and then the candidate decides whether or not to accept the offer. Although our assumption that the candidate acceptance probability is independent of the centralized allocation policy is simplistic, it is also *conservative*: Because our proposed policies generate more efficient policies than UNOS and because well-informed candidates are less likely to turn down a well-matched kidney, our assumption underestimates the relative efficiency improvements gained by our proposed policy.

A second weakness relates to the timing of the waiting list registration. The choice of when to register a patient on the waiting list remains ambiguous, and patients who register prematurely can gain an unfair advantage. Furthermore, in the absence of uniform criteria for waiting list registrations, there is always the potential for abuse. Currently, there are no uniform criteria, and thus the possibility of abuse represents a real concern. Unfortunately, our model fails to capture the effect of such abuses.

The remaining weaknesses of our model relate to the data. Lack of data about new waiting list registrations prevents us from developing direct estimates about the demographics of new transplant candidates and their mortality rates. We assumed that all deaths with graft failures were a result of competing risks, but an analysis of historical data is needed to determine how many deaths were a result of graft failures; anecdotal evidence suggests that the vast majority of deaths with graft failure are indeed a result of competing risks. Also, the data utilized by the proportional hazards model in

§5.2 provide little information about the nonimmunological characteristics of transplant donors (e.g., kidney size, pre-existing conditions). Incorporation of these additional factors is expected to produce a more accurate model and more effective allocation policies. In addition, the regression coefficients in the proportional hazards model for recipients and donors over 70 years old are imprecise, but this uncertainty should be resolved as more recent data becomes available. Although the recipient race coefficient in the proportional hazards model is known with great precision, it would be remiss to use this quantity before gaining a better understanding of the reasons behind its large magnitude; possible contributors (aside from the variables in the proportional hazards model) to African-Americans' high graft failure rate include their reduced access to immunosuppressive drugs (Kasiske et al. 1991), the high error rate in tissue type identification (Opelz et al. 1993), inter-OPO variation (UNOS 1995), and their long waiting times until transplantation (in addition, Kerman et al. 1992 show that African-Americans' high pra levels may be due in part to pre-transplant blood transfusions from Caucasians). Finally, most of the simulation results presented here are for a hypothetical "typical OPO" and should not be extrapolated to any of the OPOs that are currently operating in the United States.

In summary, the kidney allocation problem is extremely difficult for two reasons. First, equity is a complex and multifaceted concept that incorporates a variety of performance metrics and patient groups. Second, system performance is driven by the demographic and nondemographic factors affecting graft failure rates, demographic-based survival rates of dialysis patients and transplant recipients, and the demographic and nondemographic composition (e.g., tissue type and blood type) of the waiting list and donor pool; see Table 2 and observation (1) in §6.1 for a summary of these factors.

Devising an effective and fair allocation policy is an arduous task that involves difficult choices. This paper has attempted to clarify the choices involved and demonstrate how and reveal why different policies affect different groups of patients. Although any policy will inevitably be disadvantageous to some groups of patients, we have illustrated that it is possible, by employing a large set of demographic and nondemographic information, to develop policies that simultaneously improve the health outcomes of the six demographic groups of patients considered here. Clearly, the final decision on an organ allocation policy cannot be reached by relying solely on the type of models and results presented in this paper. Rather, we hope that our analytical and empirical results enrich the decision making process by providing a systematic framework for generating and comparing various policies.

In our view, the development of a kidney allocation scheme requires policy makers to answer three difficult questions. First, what is the relative importance of the various equity metrics (e.g., waiting time until transplantation vs. likelihood of transplantation vs. QALY, gender vs. race vs. age, within-group vs. between-group)? Second, how is equity assessed (e.g., is it measured with respect to the performance of the FCFT policy, with respect to having perfectly balanced performance across all demographic groups, or with respect to a more sophisticated benchmark)? Third, what is the relative importance of efficiency and equity? Ultimately, any recommendation for revising the current allocation policy will only be successful if it meets the stringent requirements of public approval.

APPENDIX A. SUMMARY OF BASELINE PARAMETERS FOR THE SIMULATION MODEL

Variable	Density	Source
Candidate Demographics		
Fraction of first time transplant candidates by gender and race[1]	0.128 0.262 0.170 0.441	a
Fraction of first time transplant candidates by age given gender and race[2]	0.020 0.031 0.042 0.046 0.053 0.069 0.087 0.116 0.144 0.156 0.114 0.075 0.034 0.015 0.021 0.035 0.040 0.044 0.048 0.053 0.068 0.091 0.131 0.158 0.137 0.105 0.053 0.019 0.024 0.040 0.061 0.091 0.097 0.091 0.095 0.101 0.112 0.109 0.082 0.057 0.027 0.012 0.015 0.027 0.039 0.046 0.055 0.058 0.065 0.079 0.111 0.151 0.150 0.119 0.062 0.023	a
Fraction of first time transplant candidates by blood type given race[3]	0.251 0.210 0.035 0.505 0.442 0.104 0.028 0.456	b
Fraction of first time transplant candidates by pra given gender and race[4]	0.674 0.326 0.784 0.216 0.768 0.232 0.855 0.145	b
Donor Demographics		
Fraction of donors by gender and race[1]	0.025 0.230 0.071 0.561	b
Fraction of donors by age given gender and race[5]	0.074 0.240 0.257 0.151 0.134 0.099 0.044 0.002 0.086 0.224 0.234 0.168 0.147 0.102 0.036 0.002 0.069 0.212 0.252 0.177 0.149 0.106 0.033 0.002 0.068 0.223 0.246 0.174 0.144 0.107 0.036 0.002	b
Fraction of donors by blood type given race[3]	0.251 0.210 0.035 0.505 0.442 0.104 0.028 0.456	b
Distribution of Body Surface Area[6]	$\log(\text{BSA}) \sim$ $N(-0.420 + 0.121\ l(\text{male donor}) + 0.693\ l(\text{age}=11\text{–}20) + 0.881\ l(\text{age}=21\text{–}30)$ $+0.921\ l(\text{age}=31\text{–}40) + 0.948\ l(\text{age}=41\text{–}50) + 0.952\ l(\text{age}=51\text{–}60)$ $+l(\text{age}=61\text{–}70) + l(\text{age}=71\text{–}80), 0.1471)$	
Mortality Rates		
Annual mortality rates for dialysis patients by age, gender and race[7]	0.060 0.085 0.080 0.097 0.105 0.108 0.134 0.145 0.179 0.230 0.271 0.336 0.394 0.055 0.071 0.106 0.113 0.127 0.156 0.175 0.216 0.262 0.312 0.366 0.430 0.498 0.056 0.088 0.108 0.123 0.116 0.121 0.137 0.154 0.191 0.241 0.303 0.364 0.407 0.051 0.073 0.102 0.125 0.148 0.161 0.196 0.242 0.290 0.324 0.388 0.447 0.553	a
Annual mortality rates for transplant recipients by age, gender and race[7]	0.011 0.022 0.017 0.026 0.027 0.030 0.046 0.046 0.071 0.073 0.151 0.151 0.151 0.006 0.009 0.013 0.019 0.026 0.033 0.032 0.039 0.043 0.075 0.086 0.062 0.062 0.012 0.019 0.019 0.019 0.034 0.050 0.057 0.060 0.095 0.127 0.070 0.102 0.102 0.011 0.010 0.016 0.021 0.030 0.038 0.047 0.047 0.074 0.086 0.100 0.094 0.242	a
Static Waiting List		
Presensitized African-Americans	3844	c
Nonpresensitized African-Americans	9889	
Presensitized Caucasians	2471	
Nonpresensitized Caucasians	11251	

[1] The two rows refer to the two sexes (F,M), and the two columns to the two races (African-American, Caucasian).

[2] The four rows refer to (Female African-American, Female Caucasian, Male African-American, Male Caucasian), and the 14 columns refer to the age ranges (20−24, 25−29, 30−34, 35−39, 40−44, 45−49, 50−54, 55−59, 60−64, 65−69, 70−74, 75−79, 80−84, 85+).

[3] The two rows refer to (African-American, Caucasian), and the four columns to blood types (A, B, AB, O).

[4] The two rows refer to pra (<60%, >60%) and the four columns refer to (Female African-American, Female Caucasian, Male African-American, Male Caucasian).

[5] The four rows refer to (Female African-American, Female Caucasian, Male African-American, Male Caucasian), and the 8 columns refer to the age ranges (0−10, 11−20, 21−30, 31−40, 41−50, 51−60, 61−70, 71−80).

[6] This model assumes that the body surface area for each donor is generated from a log normal distribution. The mean of this distribution depends on the donor's age and gender. The standard deviation is 0.1471. The model provides an excellent fit to the data in the UNOS Public-Use Data Set (see Zenios, 1996).

[7] The four rows refer to (Female African-American, Female Caucasian, Male African-American, Male Caucasian), and the 13 columns refer to the age ranges (20−24, 25−29, 30−34, 35−39, 40−44, 45−49, 50−54, 55−59, 60−64, 65−69, 70−74, 75−79, 80−84).

a USRDS Annual Report 1995.

b UNOS Public-Use Data Set.

c Zenios (1996).

ACKNOWLEDGMENT

REFERENCES

Ahn, J. H., J. C. Hornberger. 1996. Involving patients in the cadaveric kidney transplant allocation process: a decision-theoretic perspective. *Management Sci.* **42** 629–641.

Allen, R. D. M., J. R. Chapman. 1994. *A Manual of Renal Transplantation.* Edward Arnold, London.

American Medical Association. 1995. Ethical considerations in the allocation of organs and other scarce medical resources among patients. *Archives of Internal Medicine* **155** 29–40.

Avram F., D. Bertsimas, M. Ricard. 1995. Fluid models of sequencing problems in open queueing networks: an optimal control approach. In *Stochastic Networks.* F. P. Kelly, R. Williams (eds.), volume 71 of *Proceedings of the IMA,* Springer-Verlag, New York, 199–234.

Bailey, M. A. 1988. Economics issues in organ substitution technology. In *Organ Substitution Technology: Ethical, Legal and Public Policy Issues.* D. Mathieu (ed.), Westview Press, Boulder, CO.

Chertow, G. M., E. L. Milford, H. S. Mackenzie, B. M. Brenner. 1996. Antigen independent determinants of cadaveric renal allograft failure. *J. Amer. Medical Assoc.* **276** 1732–1736.

Cox, D. R., D. Oakes. 1984. *Analysis of Survival Data.* Chapman and Hall, London.

David, I. 1995. A sequential assignment match process with general renewal arrival times. *Probab. Engrg. Info. Sci.* **9** 475–492.

——, U. Yechiali. 1985. A time-dependent stopping problem with application to live organ transplants. *Oper. Res.* **33** 491–504.

——, ——. 1990. Sequential assignment match processes with arrivals of candidates and offers. *Probab. Engrg. Info. Sci.* **4** 413–430.

——, ——. 1995. One-attribute sequential assignment match processes in discrete time. *Oper. Res.* **43** 879–884.

Deniston, O. L., P. Carpentier-Alting, J. Kneisley, V. M. Hawthorne, F. K. Port. 1989. Assessment of quality of life in end-stage renal disease. *Health Services Res.* **24** 555–578.

Derman, C., G. J. Lieberman, S. M. Ross. 1972. A sequential stochastic assignment problem. *Management Sci.* **18** 349–355.

Ellison, M. D., T. J. Breen, T. G. Guo, P. R. G. Cunningham, O. P. Daily. 1993. Blacks and whites on the UNOS renal waiting list: waiting times and patient demographics compared. *Transplantation Proc.* **25** 2462–2466.

Gaston, R. S. et al. 1993. Racial equity in renal transplantation. *J. Amer. Medical Assoc.* **270** 1352–1356.

——, S. L. Hudson, B. A. Julian, D. A. Laskow, M. H. Deiehoi. 1996. Impact of donor/recipient size matching on outcomes in renal transplantation. *Transplantation* **61** 383–388.

Gaylin, D. S., P. J. Held, F. K. Port, L. G. Hunsicker, R. A. Wolfe. 1993. The impact of comorbid and sociodemographic factors on access to renal transplantation. *J. Amer. Medical Assoc.* **269** 603–608.

Ghjertson, D. W., P. I. Terasaki, B. S. Takemoto, M. R. Mickey, 1991. National allocation of cadaveric kidneys by HLA matching. *New England. J. Medicine* **324** 1032–1036.

Gittins, J. C. 1989. *Multi-Armed Bandit Allocation Indices.* Wiley, New York.

Gold, M. R., J. E. Siegel, L. B. Russell, M. C. Weinstein. 1996. *Cost-Effectiveness in Health and Medicine.* Oxford University Press, New York.

Hartl R. F., S. R. Sethi, R. G. Vickson. 1995. A survey of the maximum principles for optimal control problems with state constrains. *SIAM Rev.* **37** 181–218.

Held, P. J., B. K. Kahan, L. G. Hunsicker, D. Liska, R. A. Wolfe. 1994. The impact of HLA mismatches on the survival of first cadaveric kidney transplants. *New England J. Medicine* **331** 765–770.

Hornberger, J., D. Redelmeir, J. Petersen. 1991. Variability among methods to assess patient's well-being and consequent cost-effectiveness analysis. *J. Clinical Epidemiology* **45** 505–512.

——, J. H. Ahn. 1997. Deciding eligibility for transplantation when a donor kidney becomes available. *Medical Decision Making* **17** 160–170.

Horwood, J. W., P. Whittle. 1986. Optimal control in the neighborhood of an optimal equilibrium with examples from fishery models. *IMA J. Math. Appl. Medical Biology* **3** 129–142.

Kasiske, B. L., J. F. Neylan III, R. I. Riggio, G. M. Danovitch, L. Kahana. 1991. The effect of race on access and outcome in transplantation. *New England J. Medicine* **324** 302–307.

Kerman, R. H., P. M. Kimball, C. T. Van Buren, R. M. Lewis, D. Cavazos. 1992. Influence of race on crossmatch outcome and recipient eligibility for transplantation. *Transplantation* **53** 64–67.

Key, P. B. 1990. Optimal control and trunk reservation in loss networks. *Probab. Engrg. Info. Sci.* **4** 203–242.

Loomes, G., L. McKenzie. 1989. The use of QALYs in health care decision making. *Soc. Sci. Medicine* **28** 299–308.

Larson, R. C. 1987. Perspectives on queues: social justice and the psychology of queueing. *Oper. Res.* **35** 895–905.

Miles, A. M. V., N. Sumrani, S. John, M. A. Markell, D. A. Distant. 1996. The effect of kidney size on cadaveric renal allograft outcome. *Transplantation* **61** 894–897.

OIG. 1991. *The Distribution of Organs for Transplantations: Expectations and Practices.* U.S. Department of Health and Human Services, Publication OE1-01-89-00550.

Opelz, G., T. Wujciak. 1995. Cadaveric kidneys should be allocated according to the HLA match. *Transplantation Proc.* **27** 93–99.

——, ——, V. Schwartz, D. Back, J. Mytilincos. 1993. Collaborative transplant study analysis of graft survival in blacks. *Transplantation Proc.* **25** 2443–2445.

Ott, T. J., K. R. Krishnan. 1985. State dependent routing of telephone traffic in the use of separable routing schemes. In *Proceedings of the 11th International Teletraffic Congress.* M. Akiyama (ed.), Elsevier, Amsterdam.

Paltiel, A. D. 1997. Organ allocation and the secretary problem. *Medical Decision Making* **17** 231–232.

Pritsker, A. B. 1998. Life and death decisions: organ transplantation allocation policy analysis. *OR/MS Today* **25**(4) 22–28.

Richardson, J., E. Nord. 1997. The importance of perspective in the measurement of quality-adjusted life years. *Medical Decision Making* **17** 33–41.

Righter, R. 1989. A resource allocation problem in a random environment. *Oper. Res.* **37** 329–338.

Ruth, R. J., L. Wyszewianski, G. Herline. 1985. Kidney transplantation: a simulation model for examining demand and supply. *Management Sci.* **31** 515–526.

Sanfilippo, F. P., W. K. Vaughn. T. G. Peters, C. F. Shield III, P. L. Adams. 1992. Factors affecting the waiting time of cadaveric kidney transplant candidates in the United States. *J. Amer. Medical Assoc.* **267** 247–252.

Torrence, G. W. 1986. Measurement of health state utilities for economic appraisal. *J. Health Econom.* **5** 1–30.

UNOS. 1995. *UNOS Public-Use Data Tape.* UNOS, VA.

——, 1997. Facts and Statistics about Transplantation. World Wide Web, http://204.127.237.11:80/stats.htm.

USRDS. 1995. *1995 Annual Report.* World Wide Web, http://www.med.umich/usrds/.

Van Houwelingen, H. C., J. Thorogood. 1995. Construction, validation and updating of a prognostic model for kidney graft survival. *Statistics in Medicine* **14** 1999–2008.

Wein, L. M., S. A. Zenios, M. A. Nowak. 1997. Dynamic multidrug therapies for HIV: a control theoretic approach. *J. Theoretical Biology* **185** 15–29.

Wujciak, T., G. Opelz. 1993. Computer analysis of cadaver kidney allocation procedures. *Transplantation* **55** 516–521.

Yuan, Y., A. Gafni, J. D. Russell, D. Ludwin. 1994. Development of a central matching system for the allocation of cadaveric kidneys: a simulation of clinical effectiveness. *Medical Decision Making* **14** 124–136.

Zenios, S. A. 1996. Health care applications of optimal control theory. Unpublished Ph.D. Thesis, Operations Research Center, M.I.T, Cambridge, MA.

——. 1999. Modeling the transplant waiting list: a queueing model with reneging. *Queueing Systems* **31** 239–251.

——, G. M. Chertow, L. M. Wein. 1997. Dynamic allocation of kidneys to candidates on the transplant waiting list. Graduate School of Business, Stanford University, Research Paper Series. No. 1429.

——, ——, ——. 1998. Evidence-based organ allocation. Graduate School of Business, Stanford University, Research Paper Series. No. 1552.

——, ——, ——. 1999. Evidence-based organ allocation. *Amer. J. Medicine* **107** 52–61.