The Optimal Timing of Living-Donor Liver Transplantation

Oguzhan Alagoz
Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, alagoz@ie.pitt.edu

Lisa M. Maillart
Weatherhead School of Management, Case Western Reserve University, Cleveland, Ohio 44106, lisa.maillart@case.edu

Andrew J. Schaefer
Departments of Industrial Engineering and Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, schaefer@ie.pitt.edu

Mark S. Roberts
Division of General Internal Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, roberts@upmc.edu

Living donors are a significant and increasing source of livers for transplantation, mainly because of the insufficient supply of cadaveric organs. We consider the problem of optimally timing a living-donor liver transplant to maximize the patient’s total reward, such as quality-adjusted life expectancy. We formulate a Markov decision process (MDP) model in which the state of the process is described by patient health. We derive structural properties of the MDP model, including a set of intuitive conditions that ensure the existence of a control-limit optimal policy. We use clinical data in our computational experiments, which show that the optimal policy is typically of control-limit type.

Key words: medical decision making; Markov decision processes; control-limit policy; organ transplantation; liver transplantation; service operations

History: Accepted by Linda V. Green, public sector applications; received September 2003. This paper was with the authors 3½ months for 2 revisions.
Medical decision-making research. The QALD measure is based on the assumption that the patient assigns a quality score between zero and one to each health state (Gold et al. 1996). A quality score of one corresponds to perfect health, whereas a quality score of zero corresponds to death.

Although the financial costs of liver transplantation are significant, we do not consider these costs or perform a cost-effectiveness analysis, because our decision maker is the patient/physician, and the financial costs of the liver transplant are rarely incurred by the patient (Health-Alliance 2004, Mayo Clinic 2004). Furthermore, because cost-effectiveness in health care is typically used as a metric for group- or population-based policies (Eisenberg 1989, Weinstein and Stason 1977), as opposed to individual patient choices, it is not applicable in this situation.

We assume that the donor is indifferent to the timing of transplantation and that the quality of the donated organ is fixed. We assume that there is a finite number of health states, and that a complete ordering over the health states exists. We also assume that the patient is either ineligible or has decided not to receive cadaveric organ offers. We do not consider the risk for the donor in the decision process. Although we assume that the decision maker is both indifferent to the timing of the resolution of uncertainty and risk neutral, we recognize that these assumptions do not necessarily hold (Chew and Ho 1994). Considering these and other patient preference issues is left for future research.

The optimal solution to this problem may appear to be straightforward, i.e., the patient should have the transplant immediately once the living donor comes forward. Consider, however, that there are two components of a patient’s total reward—pretransplant reward and posttransplant reward—and recall that the overall objective is to maximize the total discounted reward rather than maximizing one of these components. When transplantation occurs, the patient’s pretransplant life ends and the posttransplant life begins. Therefore, if transplantation occurs at an early stage of the disease, the patient may maximize her posttransplant reward but not her total reward.

To the best of our knowledge, there are no other studies that consider the optimal timing of living-donor liver transplantation. Several researchers do, however, examine how to allocate cadaveric organs (particularly kidneys) to patients such that society’s benefit is maximized (Righter 1989; David and Yechiali 1995; Zenios et al. 1999, 2000). Several others investigate the problem of when to accept a cadaveric organ offer such that the patient’s benefit is maximized (David and Yechiali 1985, Ahn and Hornberger 1996, Hornberger and Ahn 1997). Most of these studies focus on kidneys and do not explicitly model donor transplantation was almost exclusively used in the pediatric population, where a parent of a child with end-stage liver disease would donate a lobe in order to save their child’s life. However, as demonstrated in Table 1, a large component of the rise in living-donor transplants has been from adult to adult. Bolstered by the success of adult-to-adult living-donor transplantation from Japan, where there are social prohibitions to cadaveric transplantation (Nishizaki et al. 2002), adult-to-adult living-donor transplantation has spread rapidly in the United States as waiting-list death rates have risen, and waiting time for organs has skyrocketed, with a median waiting time of 770 days in 1998.

There are several potential advantages of using a living donor: The organ is usually of higher quality, there is no waiting time on the list for the patient, the time spent by the organ outside the body—cold ischemia time—is essentially zero, and the time of the transplantation can be selected rather than dictated by a cadaveric donation (IALOD 2002). This timing decision is the focus of this paper. In addition to these benefits, a complete preoperative evaluation of the donated liver is possible, which may increase the success of the operation (Hashikura et al. 2002). Moreover, the posttransplant quality of life is generally higher for living-donor recipients than for cadaveric liver recipients (Trotter et al. 2002).

The purpose of this study is to determine the optimal timing of living-donor liver transplantation. That is, we seek a policy describing those health states in which the living-donor liver transplantation should occur, and those where waiting is the optimal action. In current practice, the patient, the transplant surgeon, and/or the physician responsible for the care of the patient are the decision makers (IALOD 2002, UNOS 2004a), and we assume that their objective is to maximize the patient’s total expected discounted reward. The model presented in this paper is general enough to handle various reward function definitions. Possible definitions include total discounted expected life days and total discounted quality-adjusted life days (QALD) of the patient, a common measure in

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients waiting</th>
<th>Deceased donors</th>
<th>Adults</th>
<th>Pediatrics</th>
<th>Living donors</th>
<th>Adults</th>
<th>Pediatrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>7,398</td>
<td>4,017</td>
<td>3,547</td>
<td>470</td>
<td>60</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>1997</td>
<td>9,527</td>
<td>4,101</td>
<td>3,613</td>
<td>488</td>
<td>84</td>
<td>3</td>
<td>81</td>
</tr>
<tr>
<td>1998</td>
<td>11,908</td>
<td>4,416</td>
<td>3,896</td>
<td>520</td>
<td>87</td>
<td>24</td>
<td>63</td>
</tr>
<tr>
<td>1999</td>
<td>14,445</td>
<td>4,845</td>
<td>4,040</td>
<td>445</td>
<td>230</td>
<td>140</td>
<td>90</td>
</tr>
<tr>
<td>2000</td>
<td>16,874</td>
<td>4,582</td>
<td>4,116</td>
<td>464</td>
<td>380</td>
<td>272</td>
<td>108</td>
</tr>
<tr>
<td>2001</td>
<td>18,214</td>
<td>4,665</td>
<td>4,175</td>
<td>490</td>
<td>515</td>
<td>408</td>
<td>107</td>
</tr>
</tbody>
</table>

Table 1 U.S. Liver Transplant and Waiting-List Data Between 1996 and 2001
patient health. Su and Zenios (2002) study the kidney-allocation problem, considering both society’s and the individual patient’s perspective, but do not consider the effects of the dynamic behavior of patient health on the decision process. Howard (2002) models the problem of when to refuse a cadaveric liver, but does not provide any solutions to this decision model. Instead, he provides statistical evidence that explains why a transplant surgeon may reject a cadaveric liver offer.

The remainder of this paper is organized as follows. In §2, we present a Markov decision process (MDP) model of the problem. We derive several structural properties of this MDP model and its optimal policy in §3. In §4, we present and discuss computational results. We draw some conclusions and discuss future research directions in §5.

2. Model Formulation

We formulate a discrete-time, infinite-horizon, discounted MDP model of this problem. The transition probabilities and the reward function are assumed to be stationary. The notation used in the model is as follows.

$h$: Patient health. Note that $h$ is typically described by a vector of lab values. Because we assume a complete ordering of the health states, we model $h$ as a scalar without loss of generality.

$S$: State space, $S = \{1, \ldots, H + 1\}$ where $H + 1$ represents death.

$P(h' \mid h)$: Probability that the patient will be in state $h'$ at time $t + 1$ given that she is in state $h$ at time $t$ and the liver is not transplanted at time $t$.

$P$: Transition probability matrix, i.e., $P = [P(h' \mid h)]$, $h, h' \in S$.

$a(h)$: Action taken at state $h$, to be chosen from $\{\text{Transplant}(T), \text{Wait}(W)\}$.

$r(h, T)$: Total expected discounted posttransplant reward when patient health is $h$ at the time of the transplant. Note that $r(h, T)$ is also a function of the liver quality and patient type, i.e., gender and blood type. However, because these factors are assumed to be fixed, we suppress this dependency for notational convenience.

$r(h, W)$: Expected intermediate reward accrued in the current time period when patient health is $h$ and she chooses to wait.

$V(h)$: Maximum total expected discounted reward that the patient can attain when her current health is $h$.

$\lambda$: Discount rate, $0 \leq \lambda \leq 1$.

Note that $r(h, T)$ accounts for the possibility of death during the transplant operation. Because in practice patients often need retransplantation due to a number of severe posttransplant complications, we incorporate the risk and the reward of retransplantation into $r(h, T)$.

Figure 1 shows the state-transition diagram of the MDP. The decision maker can take one of two actions at state $h$, namely, “Transplant” or “Wait for...
one more decision epoch." If the patient chooses “Transplant” in health state $h$, she receives a reward of $r(h, T)$, quits the process, and moves to absorbing state “Transplant” with probability one. If the patient chooses to “Wait” in health state $h$, then she receives an intermediate reward of $r(h, W)$ and moves to health state $h' \in S$ with probability $P(h' \mid h)$. The optimal solution to this problem can be obtained by solving the following set of recursive equations (Puterman 1994):

\[
V(h) = \max \left\{ r(h, T), r(h, W) + \lambda \sum_{h' \in S} P(h' \mid h)V(h'), \right\}, \\
h = 1, \ldots, H + 1. \tag{1}
\]

It can be shown that there exists a stationary optimal policy for the discounted as well as the undiscounted case, because there is no reward associated with remaining in the absorbing states (Puterman 1994).

### 3. Structural Properties

In this section, we derive some structural properties of the living-donor model (LDM) given by (1). Some of these properties provide closed-form solutions to the LDM under certain assumptions on the reward function and the transition probability matrix. Our main result establishes sufficient conditions that ensure the existence of an optimal control-limit policy. A control-limit policy is of the simple form: Choose the “Transplant” action and “Accept” the organ if and only if the observed health state is one of the states $j, j + 1, \ldots, H$, for some state $j$, called the control limit (Barlow and Proschan 1965). If it is known that the optimal policy is of the control-limit type, the problem can typically be solved more efficiently. Such policies are also easier to implement (Puterman 1994).

The following additional assumptions are made throughout:

**Assumption 1 (As1).** The function $r(h, T)$ is nonincreasing. That is, as the patient gets sicker, her posttransplant discounted expected reward does not increase.

**Assumption 2 (As2).** The function $r(h, W)$ is nonincreasing. That is, the expected intermediate reward that the patient accrues is nonincreasing in $h$.

Let $a^*(h)$ be the optimal action in state $h$. Theorem 1 presents sufficient conditions under which it is optimal to choose “Transplant” for all health states.

**Theorem 1.** If $P$ is upper triangular, i.e., the patient health never improves, and

\[
[1 - \lambda P(h \mid h)]r(h, T) \geq r(h, W) + \lambda r(h + 1, T)[1 - P(h \mid h)]
\]

for $h = 1, \ldots, H$, \tag{2}

then $a^*(s) = “T”$ for all $s \in S$.

Both the upper-triangularity assumption on $P$ and Condition (2) in Theorem 1 are very restrictive. In fact, our computational tests in §4 indicate that transplanting right away is typically suboptimal for the LDM. For this reason, we next explore conditions under which the optimal policy may not be “Transplant” for all states but still has an appealing structure.

In many areas of application, such as maintenance optimization (Pierskalla and Voelker 1976, Valdez-Flores and Feldman 1989, Chen and Feldman 1997), inventory theory (Harrison and Taksar 1983), and queueing (Weiss 1979), authors derive sufficient conditions to ensure the existence of an optimal control-limit policy. Most assume special structure on the transition probability matrix and/or the reward function. Below, we define some concepts that are used to specify these special structures. Interested readers should refer to Barlow and Proschan (1965), Derman (1962; 1963a, b), and Pierskalla and Voelker (1976) for more details.

**Definition 1 (Barlow and Proschan 1965).** (a) A discrete distribution $\{p_i\}_{i=0}^{\infty}$ is IFR (increasing failure rate) if $p_k / \sum_{i=k}^{\infty} p_i$ is nondecreasing in $k = 0, 1, 2, \ldots$.

(b) A Markov chain is said to be IFR if its rows are in increasing stochastic order, that is,

\[
z(i) = \sum_{j=h}^{H+1} P(j \mid i)
\]

is nondecreasing in $i$ for all $h = 1, \ldots, H + 1$.

This definition is equivalent to the well-known notion of stochastic dominance and may be viewed intuitively as follows: The sicker the patient, the more probable the patient will become even sicker.

Control-limit policies are studied for similar models (Derman 1962, 1963a, b; Barlow and Proschan 1965; Rosenfield 1976), but because the structure of our model does not match the structure of these similar models, new conditions need to be derived. For example, the IFR assumption is not sufficient to guarantee the existence of a control-limit optimal policy.

In the following theorem, we consider the case in which the transition probability matrix is IFR and show that the patient’s total discounted expected reward is nonincreasing in $h$.

**Theorem 2.** If $P$ is IFR, then for $s = 1, \ldots, H$, $V(s) \geq V(s + 1)$.

The following theorem, the main result of this section, gives a set of conditions sufficient to guarantee the existence of a control-limit optimal policy.

**Theorem 3.** Let $P$ be an IFR matrix and suppose $P$ and $r(h, T)$ satisfy the following conditions:

\[
\sum_{k=j}^{H} P(k \mid h) \leq \sum_{k=j}^{H} P(k \mid h + 1), \quad \text{for } j = h + 1, \ldots, H \text{ and } \tag{4}
\]

\[
h = 1, \ldots, H,
\]

for $h = 1, \ldots, H$. \tag{4}
and
\[ \frac{r(h, T) - r(h+1, T)}{r(h, T)} \leq \lambda[P(H+1|h+1) - P(H+1|h)], \]
for \( h = 1, \ldots, H - 1 \).  \( \text{(5)} \)

Then there exists an optimal control-limit policy. In other words, there exists a state \( j \) such that \( a^*(1) = a^*(2) = \cdots = a^*(j - 1) = \text{\"W\"} \) and \( a^*(j) = a^*(j + 1) = \cdots = a^*(H) = \text{\"T.\"} \).

In Theorem 3, \( 4 \) implies that the sicker the patient is, the more likely it is that she will move to sicker health states. Note that \( 4 \) and Definition 1 have similar interpretations, but \( 4 \) is neither a consequence of, nor sufficient to establish, Definition 1. Condition \( 5 \) on the reward function has an intuitive explanation. Namely, as the patient gets sicker, the reduction in the benefit of waiting is greater than the reduction in the benefit of performing the transplant.

Theorem 4 addresses the relationship between the optimal policies for two patients with different disease progression rates, but equivalent reward functions. We show that, given both patients have control-limit optimal policies, if Patient 1 deteriorates faster than Patient 2, then Patient 1 has a lower optimal control limit. First, we describe a plausible relationship between the transition probability matrices for two different diseases. Namely, if one of the transition probability matrices has a faster deterioration rate than another, then we say that transition probability matrix is dominated by the other. In medical terms, this condition may arise when the progression of two diseases are different.

**Definition 2.** Let \( P = [P(j | i)] \), \( i, j = 1, \ldots, n \) and \( Q = [Q(j | i)] \), \( i, j = 1, \ldots, n \) be two transition probability matrices. We say \( P \) dominates \( Q \),
\[ P \succeq Q, \text{ if } \sum_{j=1}^{n} P(j | i) \leq \sum_{j=1}^{n} Q(j | i), \quad 1 \leq i, k \leq n. \]

Consider the implications of this definition by letting the random variables \( X(h) \) and \( Y(h) \) be the time to death starting from health state \( h \) under transition probability matrices \( P \) and \( Q \) if no transplant is performed, respectively. It can easily be shown that Definition 2 implies ordinal stochastic dominance between \( X(h) \) and \( Y(h) \), i.e., \( X(h) \succeq_a Y(h) \) for \( h \in S \).

**Theorem 4.** Let \( \Pi_1 \) and \( \Pi_2 \) be two problem instances that satisfy the conditions of Theorem 3 so that the optimal policies for \( \Pi_1 \) and \( \Pi_2 \) are both control-limit optimal policies with control limits \( j_1 \) and \( j_2 \), respectively. Let \( P_1 \) and \( P_2 \) be the transition probability matrices of \( \Pi_1 \) and \( \Pi_2 \), respectively. If \( \Pi_1 \) and \( \Pi_2 \) have the same reward functions, \( r(h, T) \) and \( r(h, W) \), and \( P_1 \succeq P_2 \), then \( j_1 \geq j_2 \).

Note that Theorem 4 holds even if the conditions of Theorem 3 do not hold, as long as both problems have control-limit optimal policies and nonincreasing optimal value functions and the same reward functions.

**Corollary 5.** Let \( \Pi_1 \) and \( \Pi_2 \) be two problem instances with nonincreasing optimal value functions and the optimal policies for \( \Pi_1 \) and \( \Pi_2 \) both be control-limit optimal policies with control limits \( j_1 \) and \( j_2 \). Let \( P_1 \) and \( P_2 \) be the transition probability matrices of \( \Pi_1 \) and \( \Pi_2 \), respectively. If \( \Pi_1 \) and \( \Pi_2 \) have the same reward functions, \( r(h, T) \) and \( r(h, W) \), and \( P_1 \succeq P_2 \), then \( j_1 \geq j_2 \).

### 4. Computational Results

#### 4.1. Data Sources

We use clinical data to estimate transition probabilities, intermediate rewards, and terminal rewards. The data are from two sources. The first is a data set from UNOS that covers 28,717 patients nationwide (UNOS 2004b). UNOS is the organization that administers the organ procurement and allocation system in the United States. The second is a data set from The Thomas E. Starzl Transplantation Institute at the University of Pittsburgh Medical Center (UPMC), one of the largest liver transplant centers in the world. This data set covers 3,009 patients from the UPMC transplant center with greater clinical detail than the UNOS data set.

It is necessary to consider both data sets for several reasons. While the UNOS data set covers every transplant candidate nationwide, it is quite shallow and does not contain some important items. For instance, no clinical laboratory values or patient blood types are included for most patients. The UPMC data set is more detailed and complete, including all pertinent clinical information about listed patients and donor organs, but only provides data for a single transplant center.

#### 4.2. Estimating Parameters

We use the Model for End-stage Liver Disease (MELD) scores to represent patient health. The MELD score, first introduced by Malinchoc et al. (2000) to assess the short-term prognosis of patients with liver cirrhosis (Malinchoc et al. 2000, Wiesner et al. 2001), is a function of several laboratory values (total bilirubin, creatinine, and prothrombin time) that are measures of the liver disease. UNOS currently uses MELD scores to assess the medical urgency of liver patients (UNOS 2004a). MELD scores are restricted to integers ranging from 6 to 40, where 40 is the sickest. In this study, we use this same range to measure a patient’s health status but, because of the sparsity of the data, we aggregate the scores into groups of two or three.

The data are classified into five disease groups based on the underlying etiology (cause) of end-stage
liver disease. We estimate the transition probabilities for each disease group separately because the progression of liver disease is highly disease dependent (Dienstag and Isselbacher 2001, Podolsky and Isselbacher 2001). For each group, we estimate the pretransplant transition probabilities between health states using the Natural History Model (NHM) of Alagoz et al. (2002). The NHM, an empirical stochastic model, employs cubic spline functions to estimate incomplete observations and uses the resulting data sets to predict values of the disease covariates of a particular patient at time \( t+1 \) given known values at time \( t \).

There are two types of rewards in the MDP model: pretransplant and posttransplant reward. Because we are unaware of any existing data on quality-adjusted rewards for MELD scores, we use total discounted life expectancy in days rather than total discounted QALD for \( r(h,W) \) and \( r(h,T) \) in our computational tests.

If the patient chooses to “Wait,” the patient accrues one day as the intermediate reward, i.e., \( r(h,W) = 1 \), \( \forall h \in S \). If the patient chooses the “Transplant” option, then she receives a posttransplant reward that is equal to the expected life days of the patient, given her health status at the time of the transplant and the liver quality. To estimate the expected posttransplant life days of the patient, given her MELD score at the time of transplant and liver quality, we use the Cox proportional hazards model (Cox 1972) of Roberts et al. (2004). The inputs of this model include the donor characteristics and the clinical characteristics of the patient at the time of the transplant.

### 4.3. Numerical Examples

First consider a 60-year-old female patient with primary biliary cirrhosis who has blood type A. Figure 2 depicts the optimal transplant/wait policy for this patient with 10 potential donors. Table 2 contains the characteristics of these donors. We used these characteristics because the model of Roberts et al. (2004) identifies these factors to be most influential on the posttransplant survival rates. We apply a 0.99 annual discount rate (daily \( \lambda = 0.999972 \)). Recall that as the MELD score increases, the patient gets sicker. In Figure 2, the organs are ordered from highest to lowest quality, i.e., Organ 1 is the “best” and Organ 10 is the “worst” organ. The quality of a liver for a given patient is determined by the posttransplant survival model of Roberts et al. (2004). As Figure 2 shows, the optimal policy varies for different liver qualities. If the only liver available to the patient is Organ 1, then the optimal policy for this particular patient is as follows: “Wait” until the MELD score rises over 14, and then accept the liver and have the transplant.

We tested a total of 840 instances and, while none strictly satisfied the conditions of Theorem 3, all of the optimal policies were of control-limit type. To quantify the magnitude of the violations of the conditions of Theorem 3, we define the following metrics:

The maximum violation of the IFR assumption:

\[
\epsilon_1 = \max_{j,h} \left\{ \max \left\{ 0, \sum_{k=j}^{H+1} [P(k \mid h+1) - P(k \mid h)] \right\} \right\}
\]

for \( j = 1, \ldots, H + 1 \) and \( h = 1, \ldots, H - 1 \).

The maximum violation of Condition (4):

\[
\epsilon_2 = \max_{j,h} \left\{ \max \left\{ 0, \sum_{k=j}^{H} [P(k \mid h+1) - P(k \mid h)] \right\} \right\}
\]

for \( j = h+1, \ldots, H \) and \( h = 1, \ldots, H - 1 \).

The maximum violation of Condition (5):

\[
\epsilon_3 = \max_h \left\{ \max \left\{ 0, \lambda \max \left\{ 0, P(H+1 \mid h+1) - P(H+1 \mid h) \right\} \right\} \right\}
\]

\[
\frac{r(h,T) - r(h+1,T)}{r(h,T)} \right\} \right\}
\]

for \( h = 1, \ldots, H - 1 \).

### Table 2: Donor Characteristics

<table>
<thead>
<tr>
<th>Organ no.</th>
<th>Age</th>
<th>Sex</th>
<th>Blood type</th>
<th>Donor white(^1)</th>
<th>CMVGR(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Female</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>Female</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>Female</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>Male</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>Male</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>Male</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>Male</td>
<td>A</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>52</td>
<td>Female</td>
<td>O</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>60</td>
<td>Male</td>
<td>O</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>Male</td>
<td>O</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

\(^1\) Indicates whether the donor is white or not.
\(^2\) Indicates whether the donor has cytomegalovirus (CMVGR) or not.
Table 3  Maximum Violations of Conditions of Theorem 3

<table>
<thead>
<tr>
<th></th>
<th>Disease 1</th>
<th>Disease 2</th>
<th>Disease 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_1$</td>
<td>0.0234</td>
<td>0.0021</td>
<td>0.0189</td>
</tr>
<tr>
<td>$\epsilon_2$</td>
<td>0.0038</td>
<td>0.0012</td>
<td>0.0189</td>
</tr>
<tr>
<td>$\epsilon_3$</td>
<td>0.1052</td>
<td>0.1089</td>
<td>0.1195</td>
</tr>
</tbody>
</table>

Note that the values of $\epsilon_1$ and $\epsilon_2$ depend only on the etiology of the liver disease, whereas $\epsilon_3$ depends on the patient type, donor organ quality, and the discount rate because $r(h, T)$ is a function of these three factors. Table 3 reports the maximum $\epsilon_1$, $\epsilon_2$, and $\epsilon_3$ values obtained with an annual discount rate of 0.99. The values for $\epsilon_1$ and $\epsilon_2$ are very small, whereas $\epsilon_3$ values are relatively larger.

As discussed earlier, the policies in Figure 2 are examples of control-limit policies. The control limit for this patient is a MELD score of 16 when the potential donors are of organ types 7, 8, or 9, and rises to a MELD score of 20 when there is a very low-quality liver such as Organ 10. In general, as the liver quality drops, the control limit is nondecreasing and thus the “Wait” region becomes larger. This result is intuitive: As the patient’s posttransplant life expectancy drops while her pretransplant life expectancy remains the same, she chooses to wait until she reaches a sicker state.

Disease and patient type affect the optimal policies. Figure 3 shows the optimal policy for seven patient types, where patient type is differentiated by characteristics not including the MELD score, 10 levels of liver quality, and three disease groups, again with an annual discount rate of 0.99. Table 4 shows the static characteristics of the patient types that are used in generating Figure 3. In Figure 3, patients are in decreasing order with respect to their posttransplant expected life days for a given liver. As the quality of the patient characteristics drops, the “Wait” region becomes larger. Intuitively, this result makes sense because, as the quality of patient characteristics drops, the ratio of pretransplant survival to posttransplant survival rate increases. In general, our findings indicate that Disease Group 2, which includes Hepatitis B and C, has the highest control limits for the same patients and livers.

5. Summary and Future Work
We consider the problem of optimally timing living-donor liver transplantation. We formulate an MDP model of this decision problem and solve numerical examples using clinical data. To the best of our knowledge, this study is the first to formulate the living-donor organ transplantation problem by explicitly modeling patient health and the first to solve numerical instances of the problem using clinical data.

We derive structural properties of the LDM, including conditions that guarantee the existence of a control-limit policy. We also establish sufficient conditions for some of the intuitive results seen in our com-
computational experiments. For instance, if one disease causes a faster deterioration in patient health than another, and yet results in identical posttransplant life expectancy, then the control limit for this disease is less than or equal to that for the other.

In all of our computational tests, the optimal policy is of control-limit type. In some of the examples, when the liver quality is very low, it is optimal for the patient to choose never to have the transplant. This implies that measuring an allocation system based on total number of transplants may not fully capture the patient’s perspective. Our computational experiments also show that there are significant differences in the optimal policies for identical patients in different disease groups. There are two possible sources of this variation: differences in the disease progression and differences in the posttransplant survival rate of the diseases.

Future research will consider the problem of accepting/refusing cadaveric liver(s) offered for transplantation. The cadaveric liver case is more complicated than the living-donor case, because there is uncertainty regarding the future availability of livers. Furthermore, there are multiple organ types in the cadaveric-donor problem, whereas in the living-donor problem there is a single organ type available to the patient. In the cadaveric-donor problem, the composition of the waiting list and the rank of the patient on the waiting list have a significant effect on the transplant/wait decision. Clearly, a patient at the “top” of the list will be more selective than a patient near the “bottom” of the list. Moreover, the waiting list is dynamic. Capturing these complexities requires a different state definition and different transition probability structures. Future work will combine the two models into a single decision model in which the patient has both a living-donor and the possibility of a cadaveric donor(s).

Lastly, these decision models could be generalized to address the problem of optimally timing other living-donor organ transplantsations, given the following components: a proxy for the patient health, the transition probabilities between patient health states, a measure for organ quality, and a reward function for a given organ and patient health characteristics. If there are additional treatment options (such as dialysis in the kidney transplantation), then the action space could be expanded to include new actions.

Acknowledgments
This research was supported by National Science Foundation grants DMI-0223084 and DMI-0222661, and partially supported by Agency for Healthcare Research and Quality grant R01-HS09694. The authors thank Cindy Bryce, John Fung, Jennifer Kreke, and Steven Shechter, as well as two anonymous referees for their suggestions and insight.

Appendix

Proof of Theorem 1. The theorem holds for \( s = H \) because

\[
V(H) = \max \{ r(H, T), r(H, W) + \lambda(1 - P(H \mid H))r(H + 1, T) + \lambda P(H \mid H)V(H) \}.
\]

Because \( r(H + 1, T) = 0 \) and

\[
r(H, T) \geq r(H, W)/(1 - \lambda P(H \mid H)), \quad V(H) = r(H, T),
\]

and \( a^*(H) = “T.” \) Assume that \( a^*(s) = “T” \) for \( s = h, \ldots, H - 1 \). Then \( V(s) = r(s, T) \) for \( s = h, \ldots, H \). Now

\[
V(h - 1) = \max \left\{ r(h - 1, T), r(h - 1, W) \right\} + \lambda \sum_{h'} P(h' \mid h - 1)V(h'), \quad \text{and}
\]

\[
r(h - 1, W) + \lambda \sum_{h'} P(h' \mid h - 1)V(h') \leq r(h - 1, W) + \lambda \sum_{h' = h - 1} P(h' \mid h - 1)r(h, T) + \lambda P(h - 1 \mid h - 1)V(h - 1) \tag{A1}
\]

\[
= r(h - 1, W) + \lambda r(h, T)[1 - P(h - 1 \mid h - 1)] + \lambda P(h - 1 \mid h - 1)V(h - 1), \tag{A2}
\]

where (A1) follows from the induction hypothesis, \( (AsI) \), and because \( P \) is upper triangular. So, if \( a^*(h - 1) \) is uniquely “W,” then (A2) is equivalent to

\[
V(h - 1)[1 - \lambda P(h - 1 \mid h - 1)] \
\leq r(h - 1, W) + \lambda r(h, T)[1 - P(h - 1 \mid h - 1)] \
\leq r(h - 1, T)[1 - \lambda P(h - 1 \mid h - 1)],
\]

where the last inequality follows from (2). Therefore, \( V(h - 1) \leq r(h - 1, T) \) or \( a^*(h - 1) = “T” \) which is a contradiction from which the result follows. \( \square \)

Before proving our main results, we give two inequalities that hold for IFR matrices and nonincreasing functions.
LEMMA 1. Let $P$ be an IFR transition probability matrix and $V(h)$ be a nonincreasing function. Then the following hold:

(a) $\sum_{h' \geq h} [P(h' \mid h) - P(h' \mid h + 1)]V(h')$

\[ \geq \sum_{h' \geq h} [P(h' \mid h) - P(h' \mid h + 1)]V(h'). \]

(b) $\sum_{h' > h} [P(h' \mid h) - P(h' \mid h + 1)]V(h')$

\[ \geq \sum_{h' > h} [P(h' \mid h) - P(h' \mid h + 1)]V(h + 1). \]

**Proof of Lemma 1.** (a) First note that the IFR assumption requires that $\sum_{i=1}^{\infty} P(i \mid h) \geq \sum_{i=1}^{\infty} P(i \mid h + 1)$ for any $h' \in S$. Now,

\[ \sum_{h' \geq h} [P(h' \mid h) - P(h' \mid h + 1)]V(h') \]

\[ = [P(1 \mid h) - P(1 \mid h + 1)]V(1) + \sum_{h' = 2}^{h} [P(h' \mid h) - P(h' \mid h + 1)]V(h') \]

\[ \geq [P(1 \mid h) - P(1 \mid h + 1)]V(2) + \sum_{h' = 2}^{h} [P(h' \mid h) - P(h' \mid h + 1)]V(h') \]

\[ = [P(1 \mid h) + P(2 \mid h) - P(1 \mid h + 1) - P(2 \mid h + 1)]V(2) + \sum_{h' = 3}^{h} [P(h' \mid h) - P(h' \mid h + 1)]V(h') \]

\[ \geq [P(1 \mid h) + P(2 \mid h) - P(1 \mid h + 1) - P(2 \mid h + 1)]V(3) + \sum_{h' = 3}^{h} [P(h' \mid h) - P(h' \mid h + 1)]V(h'), \] (A3)

where (A3) follows because $P(1 \mid h) \geq P(1 \mid h + 1)$ and $V(1) \geq V(2)$. We obtain the second inequality by recasting (A3), and the last inequality holds because $P(1 \mid h) + P(2 \mid h) \geq P(1 \mid h + 1) + P(2 \mid h + 1)$ and $V(2) \geq V(3)$. The result follows if we apply the same procedure to states 3 through $h$.

(b) The proof is similar to the proof of Part (a) and is omitted.\hfill $\square$

**Proof of Theorem 2.** This theorem is a direct result of the infinite-horizon version of Lemma 3.9.4 in Topkis (1998) or, equivalently, Theorem 4.7.3 in Puterman (1994).\hfill $\square$

**Proof of Theorem 3.** First note that the monotonicity result of Theorem 2 holds. If we show that for some $h$, $a^*(h) = \text{"T"}$ implies $a^*(h + 1) = \text{"T,"}$ then the result follows. Assume that the converse is true. In other words, for some $h$, $a^*(h) = \text{"T"}$ but $a^*(h + 1)$ is uniquely "W." In this case,

\[ r(h, T) \geq r(h, W) + \lambda \sum_{h \leq s} P(h' \mid h)V(h') \]

\[ r(h + 1, T) < r(h + 1, W) + \lambda \sum_{h \leq s} P(h' \mid h + 1)V(h'). \]

It is obvious that,

\[ r(h, T) - r(h + 1, T) > r(h, W) - r(h + 1, W) \]

\[ + \lambda \left[ \sum_{h \leq h} P(h' \mid h)V(h') + \sum_{h \geq h} P(h' \mid h)V(h') \right] - \lambda \left[ \sum_{h \leq h} P(h' \mid h + 1)V(h') + \sum_{h \geq h} P(h' \mid h + 1)V(h') \right], \]

\[ \geq \lambda \left[ \sum_{h \leq h} P(h' \mid h)V(h') + \sum_{h \geq h} P(h' \mid h)V(h') \right] - \lambda \left[ \sum_{h \leq h} P(h' \mid h + 1)V(h') + \sum_{h \geq h} P(h' \mid h + 1)V(h') \right], \]

because of (A2) and $V(H + 1) = 0$. The last inequality can also be rewritten as

\[ r(h, T) - r(h + 1, T) > \lambda \left[ \sum_{h \leq h} [P(h' \mid h) - P(h' \mid h + 1)]V(h') \right] + \sum_{h \geq h} [P(h' \mid h) - P(h' \mid h + 1)]V(h'), \] (A4)

From Theorem 2 we know that $V(h)$ is nonincreasing in $h$. Therefore, the result of Lemma 1 applies to this problem and we can replace each $V(h')$ with $V(h)$ without violating (A4). Similarly, as a result of Lemma 1 and (4), we can replace each $V(h')$ with $V(h + 1)$. We then obtain the following:

\[ r(h, T) - r(h + 1, T) \]

\[ > \lambda \left[ \left( 1 - \sum_{h' = h+1}^{h} P(h' \mid h) - P(H + 1 \mid h) \right) \right] V(h) \]

\[ - \left( 1 - \sum_{h' = h+1}^{h} P(h' \mid h + 1) - P(H + 1 \mid h + 1) \right) \right] V(h) \]

\[ + \sum_{h' = h+1}^{H} [P(h' \mid h) - P(h' \mid h + 1)]V(h + 1), \]

\[ = \lambda \left[ \sum_{h' = h+1}^{H} [P(h' \mid h) - P(h' \mid h + 1)]V(h + 1) \right] \]

\[ + \lambda \left( P(H + 1 \mid h + 1) - P(H + 1 \mid h) \right)V(h). \]

From (4), $\sum_{h' = h+1}^{H} [P(h' \mid h + 1) - P(h' \mid h)]$ is nonnegative, so we can drop the first term and rewrite the last inequality as follows:

\[ r(h, T) - r(h + 1, T) \]

\[ > \lambda \left[ P(H + 1 \mid h + 1) - P(H + 1 \mid h) \right] V(h). \] (A5)

Using (A5) and (5) we obtain

\[ \lambda [P(H + 1 \mid h + 1) - P(H + 1 \mid h)]V(h) \]

\[ < \lambda [P(H + 1 \mid h + 1) - P(H + 1 \mid h)]r(h, T), \]

which is equivalent to assuming that $V(h) < r(h, T)$, which means there exists a contradiction. Therefore, $a^*(h + 1) = \text{"T"}$ also holds, from which the result follows.\hfill $\square$
LEMMA 2. Let $P$ and $Q$ be $n \times n$ transition probability matrices where $P \geq Q$. Furthermore, let $V(j)$ be a monotonically non-increasing function in $j$. Then for any $i$, the following are true:

(a) $\sum_{j \neq i} [P(j \mid i) - Q(j \mid i)] V(j) \geq \sum_{j \neq i} [P(j \mid i) - Q(j \mid i)] V(i)$.

(b) $\sum_{j \neq i} [P(j \mid i) - Q(j \mid i)] V(j) \geq \sum_{j \neq i} [P(j \mid i) - Q(j \mid i)] V(i + 1)$.

The proof of Lemma 2 is very similar to that of Lemma 1 and is omitted.

PROOF OF THEOREM 4. Suppose that we solve the two problems simultaneously using the value-iteration algorithm. We first show that starting with a value of 0 for all states in both problems, at the end of each iteration of the algorithm, the value function of $\Pi_1$ will be greater than or equal to the value function of $\Pi_2$ for each health state. Let $V_i(h)$ be the value function of the state $h$ of problem $i$ at the end of iteration $j$. We start with 0 for both problems. In this case,

$$V_1^1(h) = V_2^1(h) = \max \{r(h, T), r(h, W)\}, \quad h = 1, \ldots, H$$

as shown in Theorem 2. Therefore, the result holds for the base case.

Now, assume that $V_1^j(h) \geq V_2^j(h), h = 1, \ldots, H$, holds for iterations 2, $\ldots$, $n$. Then we want to show that $V_1^{n+1}(h) \geq V_2^{n+1}(h), h = 1, \ldots, H$. If for any state $h$, $V_2^{n+1}(h) = r(h, T)$, then the result immediately follows because $V_1^{n+1}(h) \geq r(h, T)$. Otherwise, the application of the value-iteration algorithm results in the following:

$$V_1^{n+1}(h) \geq r(h, W) + \lambda \sum_{h' \geq h} P_1(h' \mid h) V_1^n(h')$$

and

$$V_2^{n+1}(h) = r(h, W) + \lambda \sum_{h' \geq h} P_2(h' \mid h) V_2^n(h')$$

We easily obtain the following:

$$V_1^{n+1}(h) - V_2^{n+1}(h) \geq \lambda \left( \sum_{h' \geq h} [P_1(h' \mid h) - P_2(h' \mid h)] V_1^n(h') \right) + \lambda \left( \sum_{h' > h} [P_1(h' \mid h) - P_2(h' \mid h)] V_2^n(h') \right)$$

where (A6) follows from the induction assumption and (A7) is obtained by simply rearranging terms. Inequality (A8) holds because $P_1 \geq P_2$ and the monotonicity of the value function imply that $V_2^n(h)$ can be replaced with $V_1^n(h)$, and $V_2^{n+1}(h)$ can be replaced with $V_1^{n+1}(h) + 1$, without violating the inequality as a result of Lemma 2. The first inequality in (A9) follows from rearranging the terms in (A8), and the second part of (A9) follows because $P_1 \geq P_2$ and the value function is monotonic.

Because the value function for $\Pi_1$ is always greater than or equal to that of $\Pi_2$ at each iteration of the value-iteration algorithm, the optimal value function of $\Pi_1$ will always be greater than or equal to that of $\Pi_2$. Hence, if for state $j_1$, $a_1(j_1) = "T"$ in the first problem, because $r(j_1) \leq V_2(j_1) \leq V_1(j_1) = r(j_1)$, $a_2(j_1) = "T"$ always holds in the second problem. As a result, because we have a control-limit optimal policy in both problems, the control limit in the first problem will always be greater than or equal to the control limit in the second problem, from which the result follows. □

References


