IISE Transactions

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uiie21

Quantifying the benefits of increasing decision-making frequency for health applications with regular decision epochs

Suyanpeng Zhang, Sze-Chuan Suen, Vinay Sundaram & Cynthia L. Gong

To cite this article: Suyanpeng Zhang, Sze-Chuan Suen, Vinay Sundaram & Cynthia L. Gong (10 Apr 2024): Quantifying the benefits of increasing decision-making frequency for health applications with regular decision epochs, IISE Transactions, DOI: 10.1080/24725854.2024.2321492

To link to this article: https://doi.org/10.1080/24725854.2024.2321492

IISE TRANSACTIONS
https://doi.org/10.1080/24725854.2024.2321492

Quantifying the benefits of increasing decision-making frequency for health applications with regular decision epochs

Suyanpeng Zhang[a], Sze-Chuan Suen[a], Vinay Sundaram[b], and Cynthia L. Gong[c,d]

[a]Daniel J. Epstein Department of Industrial and Systems Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA; [b]Karsh Division of Gastroenterology and Comprehensive Transplant Center, Cedars-Sinai Medical Center, Los Angeles, CA, USA; [c]Fetal and Neonatal Institute, Division of Neonatology, Department of Pediatrics, Children's Hospital Los Angeles, Los Angeles, CA, USA; [d]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

ABSTRACT
Sequential decision-making problems in the context of uncertainty naturally arise in healthcare settings. In general, the frequency at which decisions can be made or changed is determined by physical limitations, such as the frequency of doctor's visits or transplantation offers. Quantifying the benefits of increasing the frequency of decision-making allows us to quantify the value of changing these physical constraints and thus improve the quality of care. In this article, we study the value provided by having additional decision-making opportunities in each epoch. We model this problem using a Markov Decision Process (MDP) framework. We provide structural properties of the optimal policies and quantify the difference in optimal values between MDP problems of different decision-making frequencies. We analyze numerical examples using liver transplantation in high-risk patients and treatment initiation in chronic kidney disease to illustrate our findings.

ARTICLE HISTORY
Received 17 May 2022
Accepted 11 February 2024

KEYWORDS
Health care; dynamic programming; Markov decision process; acute liver failure (ACLF)

1. Introduction

Sequential decision-making problems with fixed decision intervals in the context of uncertainty naturally arise in healthcare settings: monitoring problems, treatment initiation problems, disease testing, and diagnosis frequency, etc. For example, a patient with a chronic illness may require a tailored treatment regimen as their disease progresses over time, or a patient with organ failures may be offered organ transplants of varying quality, and may choose to wait or accept offered organs as their own survival probability declines.

In general, the frequency at which decisions can be made or changed in these contexts is determined by some physical limitation that occurs regularly over time, e.g., the frequency of doctor's visits or transplantation offers. In many healthcare settings, such limitations are often costly and must occur after a discrete interval of time. For instance, Chronic Kidney Disease (CKD) treatment regimen changes can only be made when the patient visits a doctor's office, which may happen at some interval (e.g., weekly, monthly, etc.). These followup frequencies often vary by patient health or disease progression rate, but the optimal frequency for a particular health state may be unknown. Increasing the frequency may provide benefits – catching disease progression sooner and faster modification of treatment plans as the patient's needs change – but may also impose costs to patients. It is therefore critical to carefully determine whether it is net beneficial to have more frequent visits.

This problem also arises in the context of organ transplantation. A patient waiting for organ transplantation may choose to invest in efforts to increase the frequency of receiving organ offers. Such efforts include transferring to hospitals that have shorter waiting periods (UCSF Health, 2023; UW Health, 2023), and multiple-listing (US Dept. of Health & Human Services, 2023). Transplant centers may have different organ offer frequencies and duration until transplant. Multiple listing entails the process of enrolling at two or more transplant hospitals. Candidates located near the donor hospital are typically prioritized over those farther away, so opting for multiple-listing can enhance patients' prospects of receiving a local organ offer and chance of transplantation (US Dept. of Health & Human Services, 2023). For example, people who are multiple-listed for heart transplantation have a shorter average second listing waiting period (126 days) compared with the first listing waiting period (335 days) (Givens *et al.*, 2015). The Organ Procurement & Transplant Network policy also allows patients to transfer primary waiting time to another hospital or switch wait time between programs if multiple-listed. However, multiple-listing usually involves completing additional evaluations for the new hospital and coordination with the insurance provider. Such efforts may be financially costly and time-consuming, and it may be useful to understand the value of increasing the frequency of donated

CONTACT Suyanpeng Zhang suyanpen@usc.edu
Supplemental data for this article can be accessed online at https://doi.org/10.1080/24725854.2024.2321492.

organs to organ recipients to better determine whether the costs associated with such efforts are justified. As these efforts for increasing the frequency of receiving an organ offer are primarily an individual patient's medical decision, we approach the problem from the patient's perspective.

In these contexts, it is important to identify the best times to offer more frequent decision-making opportunities and quantify the associated benefits. This allows for better evaluation of whether the benefits justify the potential costs of creating these additional decision-making opportunities.

1.1. Research question and approach

What is the value in increasing opportunities to make decisions, specifically in the context of stopping problems when decisions can only happen at regular intervals? While many works have studied timing trade-offs, even within the Markov Decision Process (MDP) literature, we here take a novel approach of directly comparing two MDPs—one with more frequent decisions, structured such that the state outcomes are equivalent if the action is to continue (as opposed to stopping). This comparison allows direct quantification of the value of more frequent decisions in addition to the identification of the optimal stopping time (which is the typical motivation in previous literature).

We will explore this problem in the context of the "more-frequent" and the "less-frequent" MDP problems. In scenario 1 ("less-frequent" decision), a decision-maker has an opportunity to "stop" a process at each interval. We contrast this to scenario 2 ("more-frequent" decision), where the decision is made every $1/k$ intervals (where k is an integer). How much more should a policy-maker value these additional opportunities?

We use the same state space, action space, and discount factor for both the more- and less-frequent problems, but the number of epochs in the more-frequent problem is k times that of the less-frequent problem. The transitions are such that, given the same sequence of actions (i.e., the more-frequent problem follows the same policy as in the less-frequent problem for all additional decision-making epochs at each state), both problems generate the same likelihood of ending up in each state. The reward values over a given duration are also equivalent if the same actions are used in both problems. In the more-frequent framework, a per-period cost accounts for all costs associated with additional decision-making opportunities. This problem setup allows us to use these two scenarios to study the benefits of increasing decision-making frequency in stopping problems, all else equal.

We make four main contributions in this study. First, we provide structural results around the valuation of decision-making frequency in MDP stopping problem frameworks. Understanding this valuation allows us to decide how often decisions ought to be made to increase utility. Despite substantial prior literature in the area of discrete-time MDPs, we are not aware of any prior work that has examined this problem rigorously. Secondly, we provide structural results relating to less-frequent and more-frequent problem solutions. This allows us to partially solve one problem when the optimal solution is known for the other, allowing us to translate knowledge from one context to another. Third, we analyze the difference between the optimal values of the two problems and when this quantity is maximized. This novel approach allows us to quantify the benefits of making more frequent decisions. Moreover, this informs us of when it would be more profitable to switch to a more-frequent decision-making framework. Fourth, we provide two numerical examples using liver transplantation among a particularly severely ill patient population and early-stage CKD treatment initiation using empirical data. These examples demonstrate how this framework might be used in diverse healthcare applications and illustrate its applicability in similar problem contexts.

2. Literature review

2.1. MDPs in healthcare applications

MDPs have long been used in the operation research literature for a variety of applications, including inventory management (Giannoccaro and Pontrandolfo, 2002), portfolio management (Bäuerle and Rieder, 2009), production and storage (Arruda and do Val, 2008), and others. There is a deep literature in solving and understanding MDP structure (Puterman, 1994; Givan and Parr, 2001; Topkis, 2011). These works have provided the foundations of many subsequent results on threshold structures of MDP policies, and we will similarly rely on those results here. As in prior literature, we will examine threshold policies and monotonic structure over time and state space, but we will extend this work to examine their implications when comparing more- and less-frequent decision-making frameworks. We point the reader to Sonnenberg and Beck (1993), and Givan and Parr (2001), Schaefer *et al.* (2004), Alagoz *et al.* (2010) for a more complete review of MDPs.

MDPs are also a commonly used tool for healthcare applications, and have been used for applications such as screening (Chhatwal *et al.*, 2010; Alagoz *et al.*, 2013), sequential disease testing (Arruda *et al.*, 2019; Singh *et al.*, 2020), treatment initiation (Shechter *et al.*, 2008; Liu *et al.*, 2017), and organ transplantation (see below). Within this MDP framework, we focus our analysis on finite horizon stopping problems. Stopping problems are commonly used for treatment initiation problems and organ transplantation problems and form an important healthcare policy decision tool. In these and other health-related problems, a finite decision horizon is typically considered. Among the problems mentioned here, several are stopping problems (David and Yechiali, 1985; Ahn and Hornberger, 1996; Alagoz *et al.*, 2004, 2007; Shechter *et al.*, 2008; Chhatwal *et al.*, 2010; Kurt *et al.*, 2011; Alagoz *et al.*, 2013; Liu *et al.*, 2017). Previously, authors have focused on establishing threshold policies over either state or time in an MDP framework. For instance, Alagoz *et al.* (2007) identified an at-most-three-region optimal policy for an infinite-horizon MDP model for liver transplantation. Shechter *et al.* (2008) considered both state thresholds and time thresholds to find the optimal

HIV treatment initiation time. However, to the best of our knowledge, no paper has considered how these threshold policies may change if the frequency of decision-making is changed. In this work, we extend prior analyses by additionally studying this problem and extending threshold properties to provide novel insights into estimating the value of decision-making frequency.

2.2. Epoch sizes in MDPs

There are two main time-related components that impact a decision-making process: the time horizon and epoch size. The former has been studied in prior literature, as exemplified by literature that considers the effect of different lengths of life on decision-making (Ehrlich, 2000; Dybvig and Liu, 2010). The latter has received less attention, although many authors have investigated questions involving epoch intervals in their work, particularly within the reliability literature. For instance, Barlow and Proschan (1975) focus on probabilistic aspects of reliability theory and include discussion of timing problems, and Kuo (2006) used a partially observable MDP (POMDP) in machine maintenance, allowing the intervals between sampling draws to vary. See Wang (2002) for a review of the reliability literature. However, unlike prior work, we do not only focus on *when* an action should be taken, but also the *additional value* generated from having the opportunity to make more frequent decisions. Although we may find that the optimal time to act may be the same, there may be value in having had more chances to change one's decision.

While an alternative would be to use a Continuous-Time MDP (CTMDP) or Semi-MDP (SMDP) model, we focus on a discrete-time formulation in alignment with the majority of the work in clinical and healthcare applications using MDPs, with the hope that this makes our work more generalizable. CTMDP and SMDP frameworks are usually more difficult and more computationally costly to solve than discrete-time MDP models, and this may also contribute towards their relative unpopularity in the healthcare application context.

However, even within the context of discrete-time models in healthcare, the choice of epoch size is not always clear. This has led to prior work on methods to convert between epoch sizes; for example, Chhatwal *et al.* (2016) shows how eigen-decomposition methods can be used for converting transition probability matrices between different lengths of time. We will use this technique to convert transition probabilities and rewards between frequencies in our work. One notable prior work has tangentially addressed the issue of epoch size in an MDP using a variable decision-making frequency model. Alagoz *et al.* (2013) formulated a finite-horizon MDP model (a stopping problem) in breast cancer diagnosis. The goal of the work is to reduce unnecessary follow-ups by considering follow up with different frequencies. Alagoz *et al.* (2013) introduced two non-terminate actions (follow-ups) which may be chosen every 6 and 12 months respectively. This problem introduces the utility of considering different action frequencies when solving for optimal health policies, but does not quantify the benefits of more frequent decision-making, which we do here.

There are also examples of using restless bandits to choose epoch sizes in decision-making problems. For instance, Herlihy *et al.* (2023) develop a restless multi-armed bandit framework for monitoring drug adherence. The doctor can choose to observe the patient's state at each decision epoch, potentially resulting in variable lengths of time between observations. However, this approach cannot compare the exact additional value of more frequent decision-making, as we do in this work by comparing two MDP formulations. In addition, our approach extends the existing MDP literature on organ transplantation, of which there is a rich legacy (David and Yechiali, 1985; Ahn and Hornberger, 1996; Alagoz *et al.*, 2004, 2007; Sandıkçı *et al.*, 2008; Sandıkçı *et al.*, 2013).

2.3. Organ transplantation with stochastic dynamic models

We use the optimal timing of liver transplantation as one of our motivating examples, for which we will also provide a numerical analysis. Prior work has applied MDPs to organ transplantation problems (David and Yechiali, 1985; Ahn and Hornberger, 1996; Sandıkçı *et al.*, 2008; Sandıkçı *et al.*, 2013; Boloori *et al.*, 2020). Although prior work has examined liver transplantation problems for patients with end-stage liver failure under an MDP framework, they have not examined the value of increasing transplant offers. In this article, we determine the value of increasing the frequency at which livers are offered to inform patients of how much cost would be justified in doing so. We consider a particularly vulnerable patient population (acute-on-liver-failure grade 2 or 3, or ACLF2 and ACLF3, patients, who have two, three, or more failed organs), where patients are severely ill and at very high priority for liver transplant, making the offer of more frequent organ offers particularly salient (Mahmud *et al.*, 2020).

2.4. Treatment initiation with stochastic dynamic models

In this work, we also examine when to initiate treatment for early-stage CKD patients. Prior work has studied the optimal time to initiate treatment in the context of stochastic disease progression; for instance, Shechter *et al.* (2008) identified the optimal timing of initiating HIV treatment using an MDP, Kurt *et al.* (2011) studied structural properties of statin initiation for type 2 diabetic patients using an MDP framework, and Liu *et al.* (2017) proposed an MDP framework to find the optimal strategy for treatments considering technology changes. Although the structural properties of the optimal policy have been thoroughly analyzed by many, there has been limited exploration of how changes in the frequency of decision-making can affect the optimal policy and value, which is what we focus on here.

3. Model formulation

We formulate two finite-horizon, discrete-time MDPs (the less-frequent and more-frequent problem). Both MDPs have the same objective, which is to maximize the total expected discounted rewards for a patient. As shown in Figure 1, in the *more-frequent problem*, the decision-maker is able to make $k-1$ additional decisions in each interval compared with the *less-frequent problem*, resulting in k times as many decision-making opportunities in the more-frequent problem. We consider non-stationary transition probabilities and rewards as these are common in health application problems.

The notation used in this article is as follows. The set of health states in both more- and less-frequent problems is the same and denoted by $S = \{1, 2, \ldots,$ post-decision-making state $(|S|-1)$, *death* $(|S|)\}$. We assume there exists an ordering of the states. As typical in many healthcare MDP problems, we will order the states such that state 1 is the healthiest state and the health status in the state j is worse than that of state i if $i < j$. Some states may be absorbing (the death state and post-decision-making state).

We limit our analysis to stopping problems, in which the decision-maker may continue or stop the problem. In our motivating organ transplantation problem, this is equivalent to continuing to wait for a better organ or stopping the decision process by accepting an organ offer. We therefore denote the set of available actions in both more- and less-frequent problems as $A = \{wait\ (w),\ accept\ (a)\}$. If wait is chosen, the patient can remain alive or die before the next decision period, when the process repeats. Once the *accept* decision is made, the patient will permanently enter a post-decision-making state. We allow the action space to vary across states, where only *wait* is allowed in some states ("wait states") whereas both *wait* and *accept* are allowed in all others ("non-wait states"). This allows us to model situations where no decision other than *wait* can be made (e.g., if no liver is offered this period). We use $\tilde{S} \subseteq S$ to denote a set containing all non-wait states.

The total number of decision epochs in the more-frequent problem is denoted by N. We assume N is a multiple of k. $T = \{1, \ldots, N\}$ is the set of possible decision periods for the more-frequent problem. $\tilde{T} = \{1, k+1, 2k+1, \ldots, N\}$ is the set of possible decision periods for the less-frequent problem. We use $\tilde{t}^+$ to represent the decision period after time $\tilde{t}$ in the less-frequent problem, and use $\tilde{t}^-$ to represent the epoch before $\tilde{t}$, i.e., $\tilde{t}^+ = \tilde{t} + k$, $\tilde{t}^- = \tilde{t} - k$.

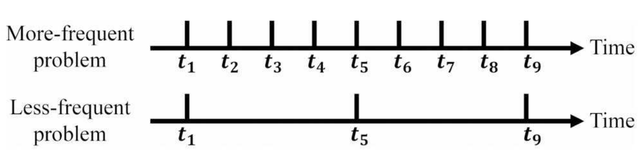

Figure 1. Timeline of the more- and less-frequent problems (using $k=4$ as an example). In the less-frequent problem, the decision-maker can make one decision every four time units at the beginning of each decision epoch. In the more-frequent problem, the decision-maker can make one decision every one time unit at the beginning each decision epoch, resulting in four times as many decisions as the less-frequent problem.

We denote $p_t(d)$ as the transition probability matrix for the more-frequent problem when the decision-maker chooses action $d \in A$ and $t \in T$. In the less-frequent problem, we use $P_{\tilde{t}}(d) = p^k_{\tilde{t}}(d), \tilde{t} \in \tilde{T}$, which is $p_{\tilde{t}}(d)$ multiplied by itself k times, to represent the transition probability matrix for action $d \in A$. This means that the likelihood of being in any state for each $t \in T$ in the more-frequent problem is the same as in the less-frequent problem at those same epochs, provided the same actions were taken. In the more-frequent problem, we assume $p_t(d) = p_{t+m}(d)$ for $m \leq k$ for any $t \in \tilde{T}$. We make this assumption as the transition probabilities typically do not vary much within a short interval (daily, monthly, or yearly). We could relax this assumption using a continuous approximation, e.g., Gompertz functions (Gompertz, 1825; Sonnenberg and Beck, 1993), but we omit this here for simplicity. We use $p_t(i|d)$ to represent the ith row of the matrix $p_t(d)$. Throughout this article, we use $p_{ss',t}(d)$ to denote an element of the matrix, the transition probability from state s to state s' at time t given action d. We discuss how we parameterize the matrix in Section 5. λ denotes the discount factor for the more-frequent problem, $0 \leq \lambda \leq 1$. For the less-frequent problem, the discount factor is λ^k.

The reward earned for the patient at state $s \in S$ and taking action $d \in A$ for $t \in T$ for the more-frequent problem is denoted using $r_t(s, d)$, the health benefits to the patient. We consider two types of rewards: the immediate reward and the lump-sum reward. Once the decision-maker chooses *wait*, the patient will earn immediate reward $r_t(s, w)$ based on s, t and advance to the next decision period. If the decision-maker chooses *accept*, the patient will earn the lump-sum reward $r_t(s, a)$ given s, t and enter the post-decision-making state. We assume that the value of the reward is the same for both t and $t + m$ $(t \in \tilde{T}, m \leq k)$ for both types of rewards in the more-frequent problem. We therefore use $r_t(s, d)$ to represent both $r_t(s, d)$ and $r_{t+m}(s, d)$ for the more-frequent problem $(t \in \tilde{T}, m \leq k)$. We use $\overrightarrow{r}_t(d)$ to denote the vector of rewards for action d for different states at time t. In the less-frequent problem, we use $R_{\tilde{t}}(s, d)$ to denote the reward for the patient at state $s \in S$ and taking action $d \in A$ for $\tilde{t} \in \tilde{T}$. We assume the lump-sum reward of the less-frequent problem is the same as the more-frequent problem at $\tilde{t} \in \tilde{T}$ $(R_{\tilde{t}}(s, a) = r_{\tilde{t}}(s, a))$. Also, we assume the lump-sum reward of wait states is zero. For the immediate reward, the rewards earned at time $\tilde{t}$ in the less-frequent problem should equal the immediate reward earned in the more-frequent problem at time $\tilde{t}$ plus the expected discounted immediate reward in the remainder of that interval $(R_{\tilde{t}}(s, w) = r_{\tilde{t}}(s, w) + \sum_{j=1}^{k-1} \lambda^j p^j_{\tilde{t}}(s|w) \overrightarrow{r}_{\tilde{t}+j}(w) = r_{\tilde{t}}(s, w) + \sum_{j=1}^{k-1} \lambda^j p^j_{\tilde{t}}(s|w) \overrightarrow{r}_{\tilde{t}}(w))$.

C_k represents the per-period costs needed to use the more-frequent decision-making frequency compared with the less-frequent frequency; this value is dependent on k. We assume that C_k is time-homogeneous, and, since it captures costs, is non-positive. C_k only appears in the more-frequent problem, as this cost is not incurred in the less-frequent problem.

Let $v_t(s), V_{\tilde{t}}(s)$ denote the optimal value function of the state $s \in S, t \in T, \tilde{t} \in \tilde{T}$ for the more-frequent and less-frequent problem respectively. At optimality, the following must hold for the more-frequent problem:

$$v_t(s) = \left\{ \begin{array}{ll} C_k + \max\left[r_t(s,a), r_t(s,w) + \lambda p_t(s|w)\overrightarrow{v}_{t+1}\right] & \text{if } s \notin \{\text{wait states, post-decision, death}\} \\ C_k + r_t(s,w) + \lambda p_t(s|w)\overrightarrow{v}_{t+1} & \text{if } s \in \{\text{wait states}\} \\ 0 & \text{if } s \in \{\text{post-decision, death}\} \end{array} \right\}.$$

Similarly, for the less-frequent problem:

$$V_{\tilde{t}}(s) = \left\{ \begin{array}{ll} \max\left[R_{\tilde{t}}(s,a), R_{\tilde{t}}(s,w) + \lambda^k P_{\tilde{t}}(s|w)\overrightarrow{V}_{\tilde{t}^+}\right] & \text{if } s \notin \{\text{wait states, post-decision, death}\} \\ R_{\tilde{t}}(s,w) + \lambda^k P_{\tilde{t}}(s|w)\overrightarrow{V}_{\tilde{t}^+} & \text{if } s \in \{\text{wait states}\} \\ 0 & \text{if } s \in \{\text{post-decision, death}\} \end{array} \right\}.$$

This problem is equivalent to one where the decision-maker may choose to *accept* while in a wait state if the lump-sum reward in that state is smaller than $\min(\lambda v_t(s), 0), \forall t \in T, s \in S$, as *wait* will then always be chosen in wait states (see Lemma 1 in Appendix 1.1). We can set the reward for *accept* in wait states to satisfy this condition for any realistic problem. Therefore, for ease of notation in the remainder of this manuscript, we assume the action space is {*accept*, *wait*} for all states and the lump-sum reward for *accept* is sufficiently small in wait states.

4. Structural properties

4.1. Assumptions

We first make the following reasonable assumptions for the more-frequent problem; we also make analogous assumptions for the less-frequent problem as well (not shown for simplicity). In this and the following sections, many results pertain to a threshold policy. Similar to Bertsekas (2012) and Puterman (1994), we define a threshold policy where, given that the optimal action is *accept* for non-wait state i at time t, the optimal action will also be to *accept* for non-wait states $j > i$ or for time $t' > t$.

Assumption 1. *Rewards are non-increasing over time and non-wait states.*

Assumption 2. *Both P and p have the increasing failure rate property for all non-wait states.*

This means that as a patient progresses to a worse state, then this patient has a higher chance of progressing to an even worse state compared with patients in better health condition states. This is generally true in the healthcare context.

Assumption 3.

$$r_t(s,w) + \sum_{j=1}^{|\tilde{S}|} p_{sj,t}(d)u(j) - r_t(s,a) \geq r_t(\bar{s},w)$$
$$+\sum_{j=1}^{|\tilde{S}|} p_{\bar{s}j,t}(d)u(j) - r_t(\bar{s},a), \ \textit{for any non-increasing}$$
$$u \ \textit{vector over state}, \forall s, \bar{s} \in \tilde{S}, \bar{s} > s, t \in T$$

This means that the reward difference between *wait* and *accept* is non-increasing over states. For instance, the benefit of waiting is higher in a healthier state, as sicker states usually have higher mortality. Similar assumptions are commonly used in sequential decision-making problems in healthcare and have been used in prior work to show there exists a threshold policy over states (see Puterman (1994), page 107, and Chhatwal *et al.* (2010), for example). We will use this assumption for a similar purpose.

Assumption 4. $r_{t-1}(s,w) + \sum_{j=1}^{|\tilde{S}|} p_{sj,t-1}(d)u(j) - r_{t-1}(s,a) \geq r_t(s,w) + \sum_{j=1}^{|\tilde{S}|} p_{sj,t}(d)\bar{u}(j) - r_t(s,a), \forall$ *non-increasing* $u, \bar{u}$ *over state j such that* $u(j) \geq \bar{u}(j) > 0, \forall s \in \tilde{S}, \forall t \in T, \forall d \in A.$

This means that the reward difference between *wait* and *accept* is non-increasing over time. For instance, the benefit of waiting is higher in an earlier decision epoch, as later decision epochs usually have higher mortality. Diseases with increasing mortality risk and progression probabilities satisfy this assumption. This assumption is very common in healthcare problems, as patients in worse health states are more likely to become sicker, and that effect worsens over time. For instance, one disease can lead to complications and comorbidities, as biological systems within the body are linked (e.g., having severe cirrhosis of the liver can lead to liver failure (University of California San Francisco, 2021), but over time this can also lead to other organs failing as well (Wu and Sundaram, 2019)).

Assumption 5. *The probability of entering wait states is zero.*

We focus our attention on non-wait states only, as *wait* will always be preferred in wait states by construction. This assumption allows us to establish a threshold policy and monotonicity over non-wait states. We show numerically that our structural outcomes can hold in non-wait states with violation of this assumption in Section 5.

The above assumptions are important to establish threshold policies and the theoretical results below. Topkis (2011), Puterman (1994) and other works have used supermodularity assumptions to establish control-type policies. We will also establish theoretical results based on similar assumptions.

4.2. Structural observations

Prior work on MDPs shows that our assumptions will generate monotonic (non-increasing) optimal policies in state for both the more-frequent and the less-frequent problem (Puterman, 1994). This means that if *wait* is preferred at non-wait state *s* at any time, then it is also the optimal policy for any non-wait state before *s*; if *accept* is preferred at non-wait state *s*, then it will be for all non-wait states after *s*. Similarly, there also exists monotonic optimal policies in time for both problems, as we show using a similar procedure as in Chhatwal *et al.* (2016) (see Lemma 3 in Appendix). At any non-wait state, if *accept* is preferred at time *t*, then it will also be preferred in any time after *t*; if *wait* is preferred at time *t*, then it is also preferred for all times before *t*. These threshold policy results imply that if it is optimal to stop the problem at one time or state, then it is also optimal to do so at any later time or sicker state.

Proposition 1.

$$v_{\tilde{t}}(s) - V_{\tilde{t}}(s) \geq \sum_{i=0}^{N-\tilde{t}-\frac{N-\tilde{t}}{k}} \lambda^i C_k, \forall \tilde{t} \in \tilde{T}, s \in \tilde{S}.$$

All proofs are provided in the Appendix. Proposition 1 holds without any of the previous assumptions. In the more-frequent problem, the decision-maker has more opportunities to make decisions, so it is intuitive that the optimal value for the more-frequent problem should be no less than the optimal value for the less-frequent problem if the cost of increasing the frequency is not considered. The right-hand-side is the total discounted cost of increasing the decision-making frequency *k* times. We can then think of a decision-maker comparing total costs to this additional benefit when making a decision about which offer frequency to use. This proposition also serves as the basis for discovering relationships between the more-frequent and the less-frequent problem.

Proposition 2. *When accept is the optimal action in the more-frequent problem for non-wait state s, epoch $\tilde{t} \in \tilde{T}$, if $C_k \leq v_{\tilde{t}}(s) - V_{\tilde{t}}(s)$, then accept is also the optimal action for the less-frequent problem for state s at time $\tilde{t}$.*

Proposition 2 helps us understand the relationship between the more-frequent and less-frequent problem in the case where financial costs are less than the additional benefits. This can serve as the foundation when building other interesting structural properties. This proposition does not rely on any assumptions and should hold for any problem (even without any threshold policies over state or time) described in the problem setup. Using this proposition, we can identify properties of the solution of the more-frequent problem by examining the solution for the less-frequent problem via contraposition. For instance, for a non-wait state *s* in a period $\tilde{t}$, if the optimal action in the less-frequent problem is *wait*, then we know the optimal action at state *s* in $\tilde{t}$ must also be *wait* in the more-frequent problem (as it cannot be *accept*, as it would also then be *accept* in the less-frequent problem). If there exists a threshold policy in both the less-frequent and more-frequent problems, the optimal action at $t \leq \tilde{t}$ for the more-frequent problem must also be *wait*. We can then pre-solve part of the optimal strategy for the more-frequent problem by solving the less-frequent problem. Similar logic using this proposition shows that when there exists a threshold policy over time, then it must be the case that whenever the more-frequent and less-frequent problems differ in optimal policy, the less-frequent problem chooses *accept* while the more-frequent problem chooses *wait* for all $\tilde{t} \in \tilde{T}$.

However, note that this proposition does not comment on the optimal actions at time $t \notin \tilde{T}$, and it is possible for the more-frequent problem to choose *accept* at a time $t \notin \tilde{T}$ while the less-frequent problem does not have the opportunity to change from the *wait* action. If so, the less-frequent problem's optimal action at the next opportunity $t \in \tilde{T}$ should be *accept* (provided a threshold policy over time exists in the more-frequent problem). Also, this proposition does not provide a way to check if $v_{\tilde{t}}(s) - V_{\tilde{t}}(s) \geq C_k$. In Theorem 4, we will provide a sufficient condition to ensure $v_{\tilde{t}}(s) - V_{\tilde{t}}(s) \geq C_k$.

Let $D_{\tilde{t}}(s) = v_{\tilde{t}}(s) - V_{\tilde{t}}(s)$ denote the difference in optimal values between the more- and less-frequent problem at time $\tilde{t}$ and state *s*. We next develop properties concerning $D_{\tilde{t}}(s)$.

Theorem 1. *When both problems have different optimal actions and Assumption 4 and Assumption 5 hold, then the difference in the optimal value, $D_{\tilde{t}}(s), s \in \tilde{S}$, is non-increasing in time for all $\tilde{t} \in \tilde{T}$ when the optimal action for the more-frequent problem is to wait. Otherwise, $D_{\tilde{t}}(s)$ is non-decreasing in time for all $\tilde{t} \in \tilde{T}$.*

Theorem 2. *When both problems have different optimal actions, if Assumption 3 and Assumption 5 hold, then the difference in the optimal value, $D_{\tilde{t}}(s), s \in \tilde{S}$, is non-increasing in state for all $s \in \tilde{S}$ when the optimal action for the more-frequent problem is to wait. Otherwise, $D_{\tilde{t}}(s)$, is non-decreasing in state for all $s \in \tilde{S}$.*

Theorem 1 and Theorem 2 state that the difference in the optimal value is always non-increasing over time and over non-wait states if both the more- and less-frequent problems give different optimal actions and more-frequent problem choose to *wait*. This means that the additional benefit we could earn from switching to the more-frequent problems from the less-frequent problem will be non-increasing over both time and non-wait states if both problems do not agree with each other (however, note that when *accept* is the optimal action at time $\tilde{t} \in \tilde{T}$ for both problems then $D_{\tilde{t}}$, the difference between optimal values, is zero). When both the more-frequent and less-frequent problems give different optimal actions and the more-frequent problem prefers *accept*, using Proposition 2, we have $D_{\tilde{t}}(s) \leq C_k$. In this case, using the more-frequent problem costs more than the additional benefits gained.

Let $B(s) \in \tilde{T}$ be the last epoch where the optimal policy for both problems is *wait* for state $s \in \tilde{S}$.

Theorem 3. *$D_{\tilde{t}}(s), s \in \tilde{S}$ is non-decreasing over time and states for $\tilde{t} \in [0, B^+(s)) \cap \tilde{T}, s \in \tilde{S}$, where $B(s)^+ = B(s) + k, B^-(s) = B(s) - k$, if: (a) there is a threshold policy over time and non-wait states; (b) optimal values are non-increasing over time; (c) $D_{B^+(s)}$ is non-decreasing over non-wait state s; (d) $p_{B^-(s)}(i|w)\overrightarrow{D}_{B(s)} \leq p_{B(s)}(i|w)\overrightarrow{D}_{B^+(s)}$, $i \in \tilde{S}$; (e) and Assumption 5 holds.*

Theorem 3 requires that a threshold policy over non-wait states exists. However, note that if there exists a threshold policy only over some states, Theorem 3 may still be true. Condition (c) means that the additional benefit of more-frequent decision-making is non-decreasing if health states are worse. This means a more severely ill patient would gain more from additional decision-making opportunities, which may be realistic as the potential health gains of intervention increase as a patient nears death. Condition (d) means the one-step ahead expected additional benefits of more-frequent decision-making by choosing *wait* is smaller at time $B(s)$ than at time $B(s) + k$. For example, this would mean that the one-step ahead expected additional benefits for an ACLF3 patient aged 50 is less than that of the same individual at the same state at an older age. This may be reasonable as an older individual may have a faster expected rate of health decline and therefore have larger expected one-step ahead expected benefits of more frequent decision-making. When Theorem 1 and Theorem 3 hold, we know the largest additional benefit from switching to the more-frequent problem in each non-wait state will be garnered either in the last time period where both problems choose *wait*, in the first time the optimal policies do not agree, or the last decision epoch.

We now turn to when $D_{\tilde{t}}(s) \geq 0$. Let $\psi_{\tilde{t}}(s) = \max\{r_{\tilde{t}}(s,a), r_{\tilde{t}}(s,w) + \lambda p_{\tilde{t}}(s)\overrightarrow{r}_{\tilde{t}+1}(a)\}$.

Theorem 4. *For $s \in \tilde{S}, \tilde{t} \in \tilde{T}$, if $C_k + \lambda C_k + r_{\tilde{t}}(s,w) + \lambda p_{\tilde{t}}(s|w)\overrightarrow{\psi}_{\tilde{t}} \geq V_{\tilde{t}}(s)$, then $D_{\tilde{t}}(s) \geq 0$.*

Theorem 4 provides a sufficient condition for when $D_{\tilde{t}}(s) \geq 0$. When this theorem holds, the largest additional benefit when comparing the less-frequent problem to the more-frequent problem will either be in the last time both problems choose *wait* or in the first time the optimal policies do not agree. Intuitively, this means that if the benefits of making an additional decision is greater than the costs of doing it, then $D_{\tilde{t}}(s) \geq 0$.

Furthermore, for all time $\tilde{t}$ and state s such that $D_{\tilde{t}}(s) \geq 0$, according to Theorem 4, the optimal action for the more-frequent problem is always *wait*. In other words, we can identify a time threshold t_s (for state s) for the more-frequent problem, such that for any $\tilde{t} \leq t_s$, the optimal action for state s is to *wait*. With this, we only need to focus on solving the periods $(t_s, N]$ for each state s for the more-frequent problem.

These four theorems are important in determining and quantifying the difference in value when comparing the less-frequent problem to the more-frequent problem. In the next section, we show numeric examples to better illustrate our theoretical results.

5. Numerical examples

We provide two numerical examples in this work. The first is on liver transplantation decision-making with $k = 2$, formulated through a partnership with a physician specializing in liver transplantation at Cedars-Sinai Hospital in Los Angeles (Section 5.1). The second example examines treatment initiation for early-stage CKD patients (Section 5.5), with model details and results in the Appendix Section 1.3 due to space constraints.

5.1. Organ transplantation decisions among ACLF patients

Unlike typical prior literature in organ transplantation (Alagoz *et al.*, 2004, 2007), which focuses on general End-Stage Liver Diseases (ESLD) where patients have a relatively lower death probability within a year after entering the ESLD health state, we focus on a cohort of patients diagnosed with acute-on-liver-failure grade-2 and grade-3 (ACLF2 and ACLF3, respectively). ACLF are types of acute liver failure where the patient has two (ACLF2) or three or more (ACLF3) simultaneous Organ Failures (OF) and is therefore in a severe, life-threatening condition. Conventionally, the transplant decision for these patients is made within a week or a month after transplant eligibility to avoid the high likelihood of death (Mahmud *et al.*, 2020; Zhang *et al.*, 2021).

Some livers may have a higher probability of resulting in a successful transplantation, as measured by the Donor Risk Index (DRI), which is a function of age, cause of donor death (if donor is dead), race, donation after cardiac death, partial/split grafts, donor height, donor location, and organ cool time (Feng *et al.*, 2006; Rosenberger *et al.*, 2014). DRI depends on the donor, but not the recipient. Marginal livers (DRI $\geq$ 1.7) are less likely to lead to successful transplantation, whereas optimal livers (DRI < 1.7) are more likely to do so. The 1.7 threshold is well accepted in clinical practice; for instance, it is used in various prior medical literature (Avolio *et al.*, 2008; Croome *et al.*, 2012; Jesudian *et al.*, 2016).

Patient-physician teams may decline an offered organ in hopes of being offered another with better outcome probabilities later. Different health systems have different expected waiting periods (Thuluvath *et al.*, 2018). Our physician partner indicated that in some health systems, ACLF2 and ACLF3 patients are offered a liver for transplantation as often as once per day. We therefore consider a 28-day decision-making framework, and a liver (either marginal or optimal) may be offered to the patient every 2 days with some probability. Given these high-need patients, we forgo the queuing systems often seen in liver transplantation models (e.g., Bandi *et al.* (2019)) as these patients cannot survive a long transplant waitlist, and instead use an MDP stopping problem framework. Although our decision time horizon is short (a month), we include health outcomes one-year post-transplant and lifetime expected outcomes to capture clinically relevant outcomes. The model can also be extended to longer time horizons without structural modifications.

We assume a liver is offered to eligible patients with probability Ω on a bi-daily basis. Conditional on a liver offer, we assume the likelihood of receiving an optimal liver offer is time-invariant. We assume that a liver offered on the last decision epoch must be accepted if the patient chooses to *wait* on all prior decision epochs.

Our objective is to evaluate the outcome of providing more frequent liver offers to eligible patients at specific times/health states, which may improve patients' health at a cost. Additional liver offers may rely on multiple listing or transferring to other transplant centers (Ardekani and Orlowski, 2010; Vagefi *et al.*, 2014; UCSF Health, 2023; US Dept. of Health & Human Services, 2023), and may also depend on a hospital's resources, the price of organ transportation, and other logistical costs. It may also mean additional costs borne by the patient, if they must make additional decisions regarding accepting/rejecting a liver. Identifying the value in increasing the frequency of liver offers is therefore of policy importance; this value must exceed the costs if the increase in offers is to be net beneficial.

We compare the less-frequent problem where an organ is offered with probability Ω every 2 days to the more-frequent problem where a liver is offered with probability ω every day (the more-frequent problem). Ω and ω depend on the number of organ failures suffered by the patient; those with more failures are prioritized over those with fewer (Thuluvath *et al.*, 2018). If a liver is offered, with probability o it will be an optimal liver offer, and we assume this is invariant across patient-types. Thus at each epoch in the more-frequent problem, an ACLF2 or ACLF3 patient may be offered one of three options: a marginal liver ($(1-o)\omega$), an optimal liver ($o\omega$), or no liver at all ($1-\omega$). The patient can decide to accept the offer if a liver is offered. If an optimal liver is offered, there are no benefits to rejection, so it will be accepted. This is confirmed by both clinical experts and model outputs if we allow patients to make decisions when receiving an optimal liver. However, if a marginal liver is offered, the patient may decide whether to accept the liver or not. Accepting a marginal liver will lead to a lower post-transplant quality of life, while rejecting will mean the patient is exposed to mortality risk for at least the time until another liver is offered.

Our goal is to quantify the benefit of increased decision-making frequency. In this scenario, there are benefits accrued from both additional offered livers (ω livers offered per 2 days in expectation in the less-frequent problem and Ω livers offered per day in expectation in the more-frequent problem), as well as additional opportunities to make more decisions as the patient's health state is observed (ACLF3, ACLF2, dead, etc.). We first quantify the total benefit, then analyze the contribution of each separately in Section 5.1.4.

We quantify the benefit of increased quality of life and duration of life (the rewards R) through Quality Adjusted Life Years (QALYs) and the willingness-to-pay threshold. QALYs were first adopted for cost-effectiveness analysis and is now widely used in medical decision-making problems (Weinstein *et al.*, 2009). QALY weights range from zero to one, and they take both quality of life and quantity of life lived into consideration. For example, for a patient with ACLF, the QALY weight for a year of life is 0.4 (Wells *et al.*, 2004) while a perfectly healthy person will have QALY weight of one. Because we need to compare these benefits with financial costs, we convert these QALY rewards to dollars using a conversion factor, $\mathcal{T}$ (the commonly used "Willingness-To-Pay (WTP) threshold") (Weinstein *et al.*, 2009). For example, an accepted WTP per QALY gained in typical cost-effectiveness analyses is \$50,000 per QALY gained (Grosse, 2008); the dollar value of the health benefit is then the product of the QALYs gained and \$50,000 per QALY gained. Thus, $R = \mathcal{T} \times$ QALYs. Our value function at time t, which additionally takes into account the financial costs C_k in the more-frequent problem, is then the *Net Monetary Benefits (NMB)*, as commonly referred to in healthcare decision-making (Trippoli, 2017).

We solve both the more- and less-frequent MDPs with NMB objective values. This allows us to compare the marginal benefits of increasing the frequency of organ donations. We also identify when the difference between the optimal values of the two problems (D) is maximized in each state. We perform this analysis using empirical data and perform sensitivity analyses around uncertain parameters. We examine how D changes over states in our sensitivity analysis to numerically demonstrate Theorems 2 and 3.

5.1.1. Model inputs

We use United Network for Organ Sharing (UNOS) data and values from the medical literature to parameterize both MDPs. We use the likelihood of receiving an organ (ω), the conditional likelihood of receiving an optimal liver given a liver offer (o), death probability (γ), and the probability of improving from a worse to a better health state (ξ) to parameterize the transition matrix $p_{\tilde{t}}(w)$. We use the eigen-decomposition method to calculate $P_{\tilde{t}}(w)$. The time-invariant relative risk (rr) of survival is used to calculate the post-transplant survival probabilities for a marginal organ. The model structure and inputs were validated by clinical experts from the Cedars Sinai Health System. we vary o and

rr in sensitivity analysis as they are uncertain and may vary by transplant center. Critically, we relax the assumption that there is no probability of entering a wait state to numerically evaluate whether this will change outcomes from the theoretically proven results above. See Appendix 1.2.1 for details on model inputs.

5.1.2. Assumptions

We make the following additional assumptions in our numeric experiments while relaxing Assumption 5. First, we assume a ACLF2/3 patient will always accept an optimal liver if it is offered. Second, we stratify ACLF3 patients to ACLF =3OF and ACLF >3OF with the former having exactly three OFs and the latter having more than three OFs. We assume the likelihood of improving from ACLF >3OF to ACLF = 3OF is equivalent to the probability of improving from ACLF = 3OF to ACLF2. Third, we assume that all patients will accept the offered liver at the end of the time horizon if a liver is offered, regardless of organ type.

5.1.3. Base case results

We relax the assumption that the probability of entering wait states is zero in our numerical analyses. Even so, our base case outcomes are consistent with our theoretical results on threshold policies over time and state when a marginal liver is offered, as well as with all propositions and theorems. In the less-frequent problem, for ACLF2, ACLF =3OF, and ACLF >3OF patients, respectively, we find that the optimal policy recommends waiting at most 4 days, 2 days, and 2 days for an optimal liver. We then use these results and Theorem 4 to identify the threshold t_s for all states (the threshold for state s such that for any $\tilde{t} \leq t_s$, $D_{\tilde{t}}(s) \geq 0$ – the optimal action for state s at time $\tilde{t}$ must be *wait*). The thresholds are day 4, 2, and 2 for ACLF2, ACLF = 3 OF, and ACLF > 3 OF patients, respectively. Given the solution from the less-frequent problem, we know the optimal action for the more-frequent problem must be *wait* from day 0-4, 0-2, and 0-2 for ACLF2, ACLF = 3 OF, and ACLF > 3 OF patients, respectively, in the more frequent problem. We then only need to solve for the optimal action on days 4-28, 2-28, and 2-28 for ACLF2, ACLF = 3 OF, and ACLF > 3 OF patients, respectively. After doing so, we find that the optimal policy recommends waiting at most 6 days, 2 days, and 2 days for ACLF2, ACLF = 3 OF, and ACLF > 3 OF patients. As expected, the results from the more-frequent problem recommend a longer wait duration than the less-frequent problem as the more-frequent problem provides an additional offer every 2 days.

We show the difference in optimal values (D) in Figure 2. The largest difference in the optimal value between the more-frequent and less-frequent problems ranges from \$50,730 to \$59,203. For ACLF3 patients (both with 3OF and >3OF), these differences decrease over time until both problems have the same optimal action, when the difference value goes below zero (when both problems' optimal action is *accept*). When D >0, it would be beneficial for the patient

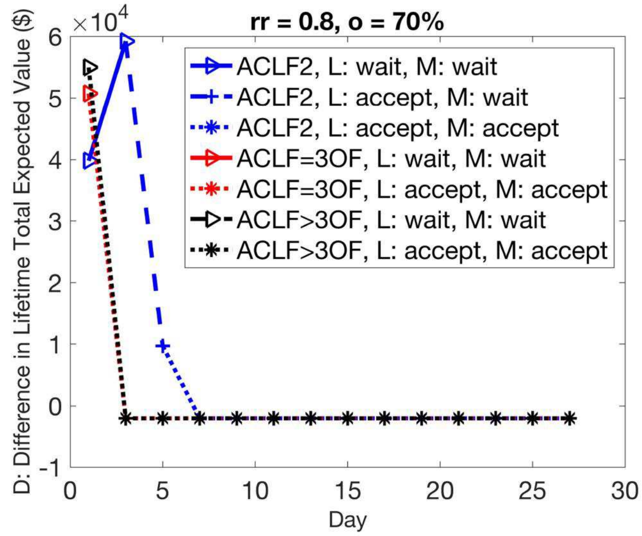

Figure 2. Difference in the expected reward earned over **time** between more-frequent (M) and less-frequent (L) problems for the marginal states with ACLF2, ACLF =3OF, ACLF >3OF. The difference between more-frequent and less-frequent problem drops below \$0 once both problems' optimal actions are *accept*. Triangles denote when the optimal actions for both problems are *wait*, "+" when it is *wait* in M and *acccept* in L, and "*" when the optimal actions for both problems are *accept*.

to choose multiple-listing or transfer to another transplant center for a higher frequency of receiving livers. For example, suppose there was an ACLF2 patient (without an optimal liver offer) who had an opportunity to transfer to a transplant center with average offer frequency of once a day from another transplant center with an average offer frequency of once every 2 days. From Figure 2, we see that this patient would not benefit from the transfer after day 6, so transfers should be made before then.

From Figure 2, we see $D > 0$ even when both problems have the same recommendation, but note that the difference in expected rewards may not be C_k even when the optimal actions for both problems across time are the same. The difference in expected values can vary because the more-frequent problem allows the decision-maker to more closely track the status of the patient's state, leading to a higher expected reward even when the optimal policy is the same. Moreover, the decision-maker is provided additional offers which also leads to a higher expected reward.

We identify when the difference between the optimal values of the two problems (D) is maximized for ACLF2, ACLF =3OF, and ACLF >3OF patients when a marginal liver is offered. We will refer to this epoch as the "time of peak D" and the D value at this time as the "value of peak D." Note that different states have possibly distinct "time of peak D" and "value of peak D." Identifying the difference at the time of peak D is important as it is the maximum per-epoch benefit of switching to the more frequent decision-making framework. In the base case, the time of peak D are day 4, day 2, and day 2 for ACLF2, ACLF =3OF, and ACLF >3OF, respectively. The values of peak D are \$59,203, \$50,730, and \$54,974 for ACLF2, ACLF =3OF, and ACLF >3OF, respectively.

Analyzing how the difference value changes over time helps us determine the time of peak D. The difference increases until the more-frequent and less-frequent problems have different optimal strategies, after which it decreases. When the optimal actions for both problems are *accept*, the difference value becomes negative, as both problems have the same action but the more-frequent problem has costs of additional decision-making opportunities. This insight illustrates a general insight into when the time of peak D occurs: the largest difference in value occurs at either the last decision epoch when the optimal action for both problems is *wait* or the first decision epoch when both problems have different optimal actions. This is because the difference value is non-decreasing when the optimal action for both problems is *wait* and non-increasing when the problems have different optimal actions. For instance, according to Theorem 3, for ACLF2 patients, the difference value D is non-decreasing between day 0 to 4 as both problems recommend to *wait*.

This difference in value function between the two problems represents the additional lifetime expected NMBs for using the more-frequent framework. As C_k increases, D will decrease. For example, for ACLF2 patients using base case parameter values, the peak value of D will decrease by approximately \$1200 if C_k increases by \$500.

We find that twice-as-frequent offers would not be net beneficial at any time if C_k is greater than \$27,209 for ACLF2 patients, \$24,256 for ACLF =3OF patients, or \$27,044 for ACLF >3OF patients. If costs were higher than these values, it would not be worthwhile to pursue increased organ offers even if it resulted in twice as frequent offers. Although our analysis is conservative and very uncertain, this provides a ballpark threshold for costs.

We can also use Theorem 4 to identify an upper bound on the C_k to ensure that twice-as-frequent decisions would be net beneficial. We find these values to be \$21,121, \$20,395, and \$27,030 for ACLF2, ACLF =3OF, and ACLF >3OF patients, respectively. Comparing these values with the actual C_k upper bound values derived numerically above, we find that these values are smaller, as expected (as Theorem 4 only provides sufficient conditions). However, all theoretical values are close to the numerically derived values, showing that this sufficient condition can be practically useful. The gap between actual and Theorem 4-derived C_k is larger for healthier patient, as one-step look ahead approximation is less accurate when patient has longer time to *wait*.

Variation in k. Our framework also allows us to consider situations where the offer frequency is more than doubled—k can by any integer. With a larger k, we observed a longer maximal waiting duration for an optimal liver for patients and a larger value of peak D. For more details, we refer the reader to Appendix 1.2.7.

5.1.4. Sensitivity analyses: Variation in relative risk of mortality and probability of being offered an optimal liver

Changes in difference value changes over time. The value of the relative risk (rr) of post-transplant mortality and the likelihood of receiving an optimal liver (o) depends on hospital characteristics. We vary o between 0.5 *and* 0.7 and rr between 0.7 *and* 0.9 (ranges determined from discussions with clinical experts). Appendix table 3 shows the maximum number of days the model recommends waiting for an optimal liver and the maximum benefits provided by the more-frequent compared to the less-frequent problem.

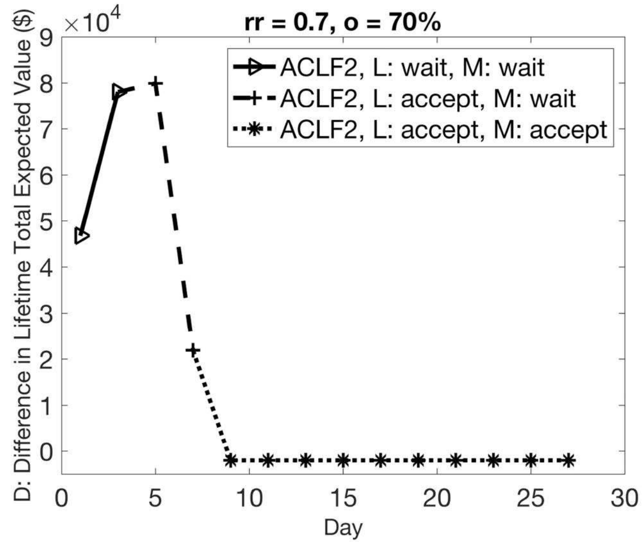

Figure 3. Difference in the expected reward earned over **time** between more-frequent (M) and less-frequent (L) problems for the marginal states with ACLF2. The difference in optimal value is marked triangular when the optimal action for both problems is *wait*, marked "+" when the optimal action for the more-frequent problem is *wait* while the optimal action for less-frequent problem is *accept*, and marked "*" when the optimal action for both problems is *accept*.

We find that the propositions and theorems demonstrated in the base case analysis also hold for cases in the sensitivity analysis: there exists threshold policies over time and states with marginal liver offered for all cases, the more-frequent problem always provides more benefits, and the optimal value function is always non-increasing in time. The value of D is monotone over the appropriate times as defined in Theorem 1 and Theorem 3.

Figure 3 illustrates the outcome described in Theorem 1, which states that when the problems have different optimal actions, the difference in optimal values must be non-increasing over time. We see this from day 6-8 for ACLF2 patients – in this period, the optimal action for the more-frequent problem is *wait* whereas the optimal action for the less-frequent problem is *accept*. Similarly, Theorem 3 states that when the optimal policy for both problems is the same (*wait*), the difference in optimal values must be non-decreasing over time, as seen in days 0-4.

We see substantial variation across both problems in the maximum number of days to wait for an optimal liver across the o and rr ranges evaluated. In all but the last case in Appendix table 3, the results from the more-frequent problem recommend a longer waiting period, as expected. Offering an additional liver each day as well as allowing the decision-maker to track the patient's status more closely makes the *wait* option more beneficial, as there will be more opportunities to be offered an optimal liver in the future

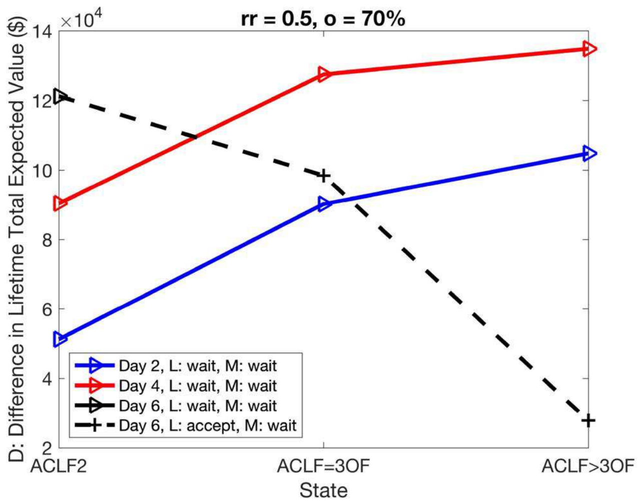

Figure 4. Difference in the expected reward earned over **states** between more-frequent (M) and less-frequent (L) problems for day 2, 4, and 6.

and there will be more opportunities to halt a quickly deteriorating health state.

As shown in Appendix table 3, we notice that both the time of peak D and the cost at that time increases as the relative risk decreases, because the decrease in the relative risk will lead to a larger difference in benefit between an optimal and a marginal liver. The time of peak D also increases as the likelihood of receiving an optimal liver increases. This suggests that the decision-maker would switch from a less-frequent problem to a more-frequent problem later when the likelihood of receiving an optimal liver is larger.

Appendix table 3 also illustrates that the largest difference in value occurs at either the last decision epoch when the optimal action for both problems is *wait* or the first decision epoch when both problems have different optimal actions. For instance, when $o = 70\%$ and $rr = 0.7$ for ACLF = 3OF patients, the maximum D value is achieved at the last decision epoch when the optimal action for both problems is *wait* (day 2). For ACLF2 patients in this case, however, the time of peak D is the first decision epoch with discordant optimal policies. Also, by finding the value differences at peak D for these two cases, we know that the maximum benefit of switching from the less-frequent problem to the more-frequent problem is $79,937 for ACLF2 patients and $68,923 for ACLF =3OF patients.

Changes in difference value over states. Figure 4 shows D when $o = 70\%$ and $rr =$ 0.5. Theorem 2 states that when both problems have different optimal actions, the difference in optimal values must be non-increasing over states. We see this among ACLF = 3OFOF and ACLF >3OF patients on day 6, as the optimal action for the more-frequent problem is *wait* whereas it is *accept* for the less-frequent problem. Similarly, Theorem 3 states that when the optimal policy for both problems is the same (*wait*), the difference in optimal values must be non-decreasing over states, as seen in all states in days 2 and 4, when both more- and less-frequent problems recommend *wait*.

Analyzing how the difference value changes over non-wait states helps us understand how the benefits change over the health state. Generally, in early decision-making periods, when both problems recommend *wait*, patients with more organ failures will benefit more from additional liver offers. Later, the two problems recommend different optimal actions (if offered a marginal liver, the more-frequent problem recommends *wait* whereas the less-frequent problem recommends *accept*) in sicker states while both problems recommend *wait* in healthier states. This shows that the sicker states will gain less benefit from additional decision-making as sicker patients are recommended to *accept* a marginal liver in the less-frequent problem, due to higher mortality probability, and therefore have lower benefits compared with healthier states when both problems recommend *wait*.

Where does the benefit come from? In this example, the benefit of more-frequent decision-making arises from both receiving more liver offers and additional decision-making opportunities. To understand which contributes more towards increasing the value of peak D, we consider a case study for ACLF2 patients, where we fix the probability of receiving a liver offer to be the same over a 2-day interval in both the more- and less-frequent problems. The difference between the benefit accrued in this scenario and the previous base case values will then quantify the benefit associated with receiving more livers. To do this, we set the likelihood of receiving a liver for the less-frequent problem (Ω) to one. Then we set the probability for the more-frequent problem (ω) such that both problems produce one liver offer every 2 days in expectation. Assuming $o = 70\%$ and $rr = 0.7$, we found that the value of peak D is $1490. The base case found $79,937 of additional benefit obtained for ACLF2 patients, so the additional benefit mainly comes from receiving liver offers.

5.2. Treatment initiation for early-stage CKD

Our second numerical example focuses on a monitoring problem for treatment initiation, where physicians can only observe patient health states for updating treatment plans during an office visit. Disease progression progresses stochastically, and the optimal frequency of these visits may vary by health state and time.

In this example, we examine a decision for how frequently to monitor early-stage CKD patients to determine treatment initiation time. CKD patients are categorized into five stages by disease severity, as measured using the estimated Glomerular Filtration Rate (eGFR). According to current guidelines, CKD patients should undergo at least yearly eGFR checkups and initiate Angiotensin-Converting Enzyme (ACE)/Angiotensin Receptor Blocker (ARB) treatment immediately after reaching stage 3 (New *et al.*, 2007; Tahir *et al.*, 2007). However, there is some controversy regarding whether annual checkups are sufficient (Hirano *et al.*, 2019), and considering the low cost associated with ACE/ARB treatment, there is a possibility that early initiation of such treatment could benefit patients (Sharma *et al.*, 2011). The optimal monitoring frequency and time to initiate treatment would depend on the lifetime health benefits and costs (as measured using NMB). However, more frequent office visits for CKD status monitoring incurs additional costs, and more-frequent physician visits would only be beneficial if these costs were offset by the health benefits of catching disease progression earlier.

We therefore use our approach to quantify the relative benefit of more frequent treatment initiation opportunities for early stage (stage 1-2) CKD patients. To do this, we create two equivalent MDPs for patients in stages 1, 2, and 3+. The less-frequent MDP problem uses a 12-month decision-making frequency, whereas the more-frequent assumes 6-month checkups. The available actions in both MDPs are to either initiate treatment or wait. We assume increasing the decision-making frequency incurs a cost (C_k). We use a previously established CKD simulation model to estimate rewards and transition probabilities (Hoerger *et al.*, 2010). Details of the model, solution, and outcomes are given in Appendix Section 1.3.

Our investigation shows that early treatment initiation is necessary and beneficial for stage 1 and 2 CKD patients for some age groups, regardless of the decision-making frequency. Remarkably, the optimal policies derived from both MDPs are identical, indicating that the decision-making frequency does not significantly impact the recommended course of action. Furthermore, the differences in optimal values between the two MDPs fall below zero. This suggests that a higher frequency decision-making framework may not yield substantial benefits, as the progression of chronic diseases like CKD tends to be slow. These results indicate that the current monitoring frequency recommendations are sufficient, despite ongoing controversy to the contrary.

Our findings in this example also validate and strengthen our theoretical results, as all results were consistent with our theoretical conclusions and served to illustrate that they were applicable in a CKD context. These results demonstrate that our approach to quantifying the benefit of additional decision-making opportunities can be extended and applied to various healthcare scenarios, showcasing its generalizability and relevance in different healthcare applications.

6. Conclusions

We examine the value of making more frequent decisions in a discrete-time, finite-horizon MDP stopping problem framework. We quantify the benefits of making more frequent decisions as well as provide useful structural properties that can help decision-makers decide which frequency decision-making model to use when more frequent decisions are costly.

In our theorems, we provide novel insights comparing the value difference between the more- and less-frequent problems. We identified properties of the more- and less-frequent problems to determine when there exists a threshold policy in time and in non-wait states. We additionally found properties of the difference in expected value between these problems and provide sufficient conditions for when more-frequent decision-making would be net beneficial. We provide sufficient conditions under which we can guarantee that the more-frequent decision-making framework would be preferred for some states at some time, and we demonstrate all of our theoretical results numerically in our examples.

We provide two numerical examples to showcase the practical application of our work. The first examines liver transplantation to determine the value of more frequent liver donations. We use data from the UNOS database to parameterize our model for severely ill patients with multiple organ failures (ACLF2/3). Our numerical example shows that the maximum benefit from additional decisions (the value at peak D) is roughly \$60,000 in NMB over the post-transplant lifetime when C_k is \$2000. This benefit, which includes both health outcomes and financial costs, decreases to zero if the per-patient, per-period cost of using the more-frequent framework increases to roughly \$25,000. In sensitivity analyses, we find that this benefit is inversely proportional to the relative risk of mortality, rr, but is relatively invariant with the likelihood of receiving an optimal liver, o.

The policy implications of our work are that (i) we can identify cost-thresholds over which it would not be beneficial to provide even severely ill patients (ACLF2/3) patients with more frequent organ offers, and (ii) even if it is net beneficial to do so, it may not be beneficial to do so for all periods and patient health states—and our analysis framework can identify when it is worth it, thus allowing for targeted offer frequencies. Currently, liver offers already consider patient severity (through MELD scores, etc. (US Dept. of Health & Human Services, 2022)), and given the push for more patient-tailored healthcare, perhaps it would be realistic to push for individualized organ offer frequency.

In a second example, we demonstrate the broad applicability of this framework by considering a treatment initiation problem for CKD patients. This example not only confirms the validity of our theoretical findings, but also offers insights into the importance of early treatment initiation for CKD patients. We found that the value at peak D is negative, indicating that it would not be net beneficial to increase the frequency of early-stage CKD patient monitoring from the recommended annual check-up. This shows how our comparative MDP analysis can contribute towards clinical controversies in monitoring frequency.

We must acknowledge several limitations of this work. We set up the more-frequent problem in a way that the two transition probability matrices for every 2 days are identical if their time indices belong to the same time epoch in the less-frequent problem for simplicity. In reality, the transition probability matrix is changing continuously over time. However, since the change in the transitions within a short time period is negligible, our results should not change substantially in our numerical analysis even had we allowed the transition matrices for t and $t+1$ periods to be different.

We limit our analysis to stopping problems, as these occur naturally in many healthcare settings, but many interesting MDPs also fall outside this category. We only analyze problems where the more-frequent framework allows for k times as many decisions, where k is an integer; this limits the generalizability of our analysis. In our theoretical analysis, we assume that the probability of entering wait states is zero. Future work should focus on incorporating the possibility of transitions into wait states. Our analysis focuses on finite horizon models, and we leave the extension of this work to infinite horizon models for future work.

Our numerical analysis relies on highly uncertain input values, particularly C_k. However, we perform sensitivity analyses and identify upper bounds for C_k above which the cost would no longer justify the additional decision-making opportunities, and we compare these numerical outcomes to our theoretical bounds.

Despite these limitations, we believe that this work provides interesting insights for not only transplantation applications, but also other applications such as monitoring problems with regular decision epochs. This article draws attention to quantifying the benefits of a more-frequent decision-making framework in healthcare settings. We derived structural properties between the more-frequent and the less-frequent problems, and provided a numerical example to show how to make use of these results. These results have implications for a wide variety of MDP stopping problems and provide insight into future work on improving the speed of MDP solution algorithms.

Acknowledgments

The authors would like to acknowledge Yiwen Cao for her help in identifying an appropriate chronic kidney disease application problem for this context, as well as for helping to build and run the simulation code for the second numerical example.

Funding

This material is based upon work supported in part by the National Science Foundation under Grant No. 2237959.

Notes on contributors

Suyanpeng Zhang is a PhD candidate in the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California. He completed his BS in mathematics at Rensselaer Polytechnic Institute and his MS in industrial and operations engineering at the University of Michigan. His research focuses on developing novel and data-driven modeling methods that adapt to challenges in healthcare to facilitate the generation of practical policy recommendations.

Sze-chuan Suen is a WiSE Gabilan Assistant Professor in the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California. She received her PhD in the department of Management Science and Engineering from Stanford University in 2016. Her research interests include developing applied mathematical models to identify epidemiological trends and evaluate health policies to support informed decision-making. Her work in health policy modeling draws from a variety of techniques, including optimization, dynamic systems modeling, Markov decision processes, simulation, cost-effectiveness analysis, and decision analysis. Her previous work has examined the optimal management of tuberculosis, HIV, COVID-19, as well as chronic diseases.

Vinay Sundaram, MD, MSc, FACP, FAASLD, transplant hepatologist at Cedars-Sinai Medical Center, sadly passed away in 2022. During his life, he was a distinguished physician, patient advocate, and scientist. He graduated from the New York University School of Medicine in 2004, completed his residency at the University of Virginia Medical Center in Charlottesville in 2007, and completed fellowships in gastroenterology at the University of Pittsburgh Medical Center (2010) and hepatology at Beth Israel Deaconess Medical Center (Harvard Medical School) in 2011. Dr. Sundaram published over 140 research papers over the course of life, many of which focused on liver transplantation in acute-on-chronic liver failure (ACLF) patients. He was a Steering Committee member of the AASLD ACLF Special Interest Group, and participated in numerous campaigns to raise awareness about ACLF and issues around liver transplantation in these patients.

Cynthia L. Gong, PharmD, PhD is a Senior Research Associate at Curta. She received her PharmD and PhD in health economics from the USC School of Pharmacy, as well as completed a pharmacy residency in pharmacoeconomics at the VA Northern California Healthcare System. She previously served as a research faculty member at the USC Keck School of Medicine, and as Senior Pharmacoeconomics Advisor for Stanford Healthcare, and has published over 30 peer-reviewed articles with a focus primarily on economic evaluation, discrete choice, and health-related quality-of-life.

ORCID

Suyanpeng Zhang http://orcid.org/0000-0002-8957-1293
Sze-Chuan Suen http://orcid.org/0000-0001-9453-5863

Data Availability Statement

UNOS data can be requested from UNOS directly.

References

Ahn, J.-H. and Hornberger, J.C. (1996) Involving patients in the cadaveric kidney transplant allocation process: A decision-theoretic perspective. *Management Science*, **42**(5), 629–641.

Alagoz, O., Chhatwal, J. and Burnside, E.S. (2013) Optimal policies for reducing unnecessary follow-up mammography exams in breast cancer diagnosis. *Decision Analysis*, **10**(3), 200–224.

Alagoz, O., Hsu, H.E., Schaefer, A.J. and Roberts, M.S. (2010) Markov decision processes: A tool for sequential decision making under uncertainty. *Medical Decision Making*, **30**, 474–483.

Alagoz, O., Maillart, L.M., Schaefer, A.J. and Roberts, M.S. (2004) The optimal timing of living-donor liver transplantation. *Management Science*, **50**(10), 1420–1430.

Alagoz, O., Maillart, L.M., Schaefer, A.J. and Roberts, M.S. (2007) Choosing among living-donor and cadaveric livers. *Management Science*, **53**(11), 1702–1715.

Ardekani, M.S. and Orlowski, J.M. (2010) Multiple listing in kidney transplantation. *American Journal of Kidney Diseases*, **55**(4), 717–725.

Arruda, E.F. and do Val, J.B.R. (2008) Stability and optimality of a multi-product production and storage system under demand uncertainty. *European Journal of Operational Research*, **188**(2), 406–427.

Arruda, E.F., Pereira, B.B., Thiers, C.A. and Tura, B.R. (2019) Optimal testing policies for diagnosing patients with intermediary probability of disease. *Artificial Intelligence in Medicine*, **97**, 89–97.

Avolio, A., Siciliano, M., Barbarino, R., Nure, E., Annicchiarico, B., Gasbarrini, A., Agnes, S. and Castagneto, M. (2008) Donor risk index and organ patient index as predictors of graft survival after liver transplantation, in *Transplantation Proceedings*, **40**, 1899–1902.

Bandi, C., Trichakis, N. and Vayanos, P. (2019) Robust multiclass queuing theory for wait time estimation in resource allocation systems. *Management Science*, **65**(1), 152–187.

Barlow, R.E. and Proschan, F. (1975) Statistical theory of reliability and life testing: Probability models. Technical report, Florida State University Tallahassee.
Bäuerle, N. and Rieder, U. (2009) MDP algorithms for portfolio optimization problems in pure jump markets. *Finance and Stochastics*, **13**(4), 591–611.
Bertsekas, D. (2012) *Dynamic Programming and Optimal Control: Volume I*. Athena scientific, Belmont, MA.
Boloori, A., Saghafian, S., Chakkera, H.A. and Cook, C.B. (2020) Data-driven management of post-transplant medications: An ambiguous partially observable Markov decision process approach. *Manufacturing & Service Operations Management*, **22**(5), 1066–1087.
Chhatwal, J., Alagoz, O. and Burnside, E.S. (2010) Optimal breast biopsy decision-making based on mammographic features and demographic factors. *Operations Research*, **58**(6), 1577–1591.
Chhatwal, J., Jayasuriya, S. and Elbasha, E.H. (2016) Changing cycle lengths in state-transition models: Challenges and solutions. *Medical Decision Making*, **36**(8), 952–964.
Croome, K., Marotta, P., Wall, W., Dale, C., Levstik, M., Chandok, N. and Hernandez-Alejandro, R. (2012) Should a lower quality organ go to the least sick patient? Model for end-stage liver disease score and donor risk index as predictors of early allograft dysfunction. *Transplantation Proceedings*, **44**, 1303–1306.
David, I. and Yechiali, U. (1985) A time-dependent stopping problem with application to live organ transplants. *Operations Research*, **33**(3), 491–504.
Dybvig, P.H. and Liu, H. (2010) Lifetime consumption and investment: Retirement and constrained borrowing. *Journal of Economic Theory*, **145**(3), 885–907.
Ehrlich, I. (2000) Uncertain lifetime, life protection, and the value of life saving. *Journal of Health Economics*, **19**(3), 341–367.
Feng, S., Goodrich, N., Bragg-Gresham, J., Dykstra, D., Punch, J., DebRoy, M., Greenstein, S. and Merion, R. (2006) Characteristics associated with liver graft failure: The concept of a donor risk index. *American Journal of Transplantation*, **6**(4), 783–790.
Giannoccaro, I. and Pontrandolfo, P. (2002) Inventory management in supply chains: A reinforcement learning approach. *International Journal of Production Economics*, **78**(2), 153–161.
Givan, B. and Parr, R. (2001) An introduction to Markov decision processes. Available at https://faculty.kfupm.edu.sa/coe/ashraf/RichFilesTeaching/COE101_540/Projects/givan1.pdf (accessed June 6, 2023).
Givens, R.C., Dardas, T., Clerkin, K.J., Restaino, S., Schulze, P.C. and Mancini, D.M. (2015) Outcomes of multiple listing for adult heart transplantation in the United States: Analysis of OPTN data from 2000 to 2013. *JACC: Heart Failure*, **3**(12), 933–941.
Gompertz, B. (1825) XXIV. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, in a letter to Francis Baily, Esq. FRS &c. *Philosophical Transactions of the Royal Society of London* (115), 513–583.
Grosse, S.D. (2008) Assessing cost-effectiveness in healthcare: History of the $50,000 per qaly threshold. *Expert Review of Pharmacoeconomics & Outcomes Research*, **8**(2), 165–178.
Herlihy, C., Prins, A., Srinivasan, A. and Dickerson, J.P. (2023) Planning to fairly allocate: Probabilistic fairness in the restless bandit setting, in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*, Association for Computing Machinery, New York, NY, pp.732–740.
Hirano, K., Kobayashi, D., Kohtani, N., Uemura, Y., Ohashi, Y., Komatsu, Y., Yanagita, M. and Hishida, A. (2019) Optimal follow-up intervals for different stages of chronic kidney disease: A prospective observational study. *Clinical and Experimental Nephrology*, **23**, 613–620.
Hoerger, T.J., Wittenborn, J.S., Segel, J.E., Burrows, N.R., Imai, K., Eggers, P., Pavkov, M.E., Jordan, R., Hailpern, S.M., Schoolwerth, A.C. et al. (2010) A health policy model of CKD: 1. Model construction, assumptions, and validation of health consequences. *American Journal of Kidney Diseases*, **55**(3), 452–462.
Jesudian, A., Desale, S., Julia, J., Landry, E., Maxwell, C., Kallakury, B., Laurin, J. and Shetty, K. (2016) Donor factors including donor risk index predict fibrosis progression, allograft loss, and patient survival following liver transplantation for hepatitis c virus. *Journal of Clinical and Experimental Hepatology*, **6**(2), 109–114.
Kuo, Y. (2006) Optimal adaptive control policy for joint machine maintenance and product quality control. *European Journal of Operational Research*, **171**(2), 586–597.
Kurt, M., Denton, B.T., Schaefer, A.J., Shah, N.D. and Smith, S.A. (2011) The structure of optimal statin initiation policies for patients with type 2 diabetes. *IIE Transactions on Healthcare Systems Engineering*, **1**(1), 49–65.
Liu, S., Brandeau, M.L. and Goldhaber-Fiebert, J.D. (2017, Mar) Optimizing patient treatment decisions in an era of rapid technological advances: The case of hepatitis C treatment. *Health Care Management Science*, **20**(1), 16–32.
Mahmud, N., Hernaez, R., Wu, T. and Sundaram, V. (2020, Sep) Early transplantation in acute on chronic liver failure: Who and when. *Current Hepatology Reports*, **19**(3), 168–173.
New, J., Middleton, R., Klebe, B., Farmer, C., De Lusignan, S., Stevens, P. and O'Donoghue, D. (2007) Assessing the prevalence, monitoring and management of chronic kidney disease in patients with diabetes compared with those without diabetes in general practice. *Diabetic Medicine*, **24**(4), 364–369.
Puterman, M.L. (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (1st ed.). John Wiley & Sons, Inc, New York, NY.
Rosenberger, L.H., Gillen, J.R., Hranjec, T., Stokes, J.B., Brayman, K.L., Kumer, S.C., Schmitt, T.M. and Sawyer, R.G. (2014) Donor risk index predicts graft failure reliably but not post-transplant infections. *Surgical infections*, **15**(2), 94–98.
Sandıkçı, B., Maillart, L.M., Schaefer, A.J., Alagoz, O. and Roberts, M.S. (2008) Estimating the patient's price of privacy in liver transplantation. *Operations Research*, **56**(6), 1393–1410.
Sandıkçı, B., Maillart, L.M., Schaefer, A.J. and Roberts, M.S. (2013) Alleviating the patient's price of privacy through a partially observable waiting list. *Management Science*, **59**(8), 1836–1854.
Schaefer, A.J., Bailey, M.D., Shechter, S.M. and Roberts, M.S. (2004) *Modeling Medical Treatment Using Markov Decision Processes*, Springer, Boston, MA, pp. 593–612.
Sharma, P., Blackburn, R.C., Parke, C.L., McCullough, K., Marks, A. and Black, C. (2011) Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers for adults with early (stage 1 to 3) non-diabetic chronic kidney disease. *Cochrane Database of Systematic Reviews* (10), CD007751.
Shechter, S.M., Bailey, M.D., Schaefer, A.J. and Roberts, M.S. (2008) The optimal time to initiate HIV therapy under ordered health states. *Operations Research*, **56**(1), 20–33.
Singh, R., Liu, F. and Shroff, N.B. (2020) A partially observable MDP approach for sequential testing for infectious diseases such as Covid-19. *ArXiv preprint ArXiv:2007.13023*.
Sonnenberg, F.A. and Beck, J.R. (1993) Markov models in medical decision making: A practical guide. *Medical Decision Making*, **13**(4), 322–338.
Tahir, H., Jackson, L.L. and Warnock, D.G. (2007) Antiproteinuric therapy and fabry nephropathy: Sustained reduction of proteinuria in patients receiving enzyme replacement therapy with agalsidase-β. *Journal of the American Society of Nephrology*, **18**(9), 2609–2617.
Thuluvath, P.J., Thuluvath, A.J., Hanish, S. and Savva, Y. (2018) Liver transplantation in patients with multiple organ failures: Feasibility and outcomes. *Journal of Hepatology*, **69**(5), 1047–1056.
Topkis, D.M. (2011) *Supermodularity and Complementarity*. Princeton University Press, Princeton, NJ.
Trippoli, S. (2017) Incremental cost-effectiveness ratio and net monetary benefit: Current use in pharmacoeconomics and future perspectives. *European Journal of Internal Medicine*, **43**, e36.
UCSF Health (2023) FAQ: Getting a liver transplant. Availible at https://www.ucsfhealth.org/education/faq-getting-a-liver-transplant (accessed June 9, 2023).

University of California San Francisco (2021) End-stage liver disease (ESLD). Availible at https://surgery.ucsf.edu/conditions–procedures/end-stage-liver-disease-(esld).aspx (accessed November 3, 2022).

US Dept. of Health & Human Services (2022) Questions and answers about liver allocation. Availible at https://optn.transplant.hrsa.gov/patients/by-organ/liver/questions-and-answers-about-liver-allocation/ (accessed November 3, 2022).

US Dept. of Health & Human Services (2023) Multiple listing. Availible at https://optn.transplant.hrsa.gov/patients/about-transplantation/multiple-listing/ (accessed June 9, 2023).

UW Health (2023) *Multiple listing with the UW Health transplant center*. Availible at https://patient.uwhealth.org/education/multiple-wait-list-listings (accessed June 9, 2023).

Vagefi, P.A., Feng, S., Dodge, J.L., Markmann, J.F. and Roberts, J.P. (2014) Multiple listings as a reflection of geographic disparity in liver transplantation. *Journal of the American College of Surgeons*, **219**(3), 496–504.

Wang, H. (2002) A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, **139**(3), 469–489.

Weinstein, M.C., Torrance, G. and McGuire, A. (2009) QALYs: The basics. *Value in Health*, **12**, S5–S9.

Wells, C.D., Murrill, W. and Arguedas, M.R. (2004) Comparison of health-related quality of life preferences between physicians and cirrhotic patients: Implications for cost–utility analyses in chronic liver disease. *Digestive Diseases and Sciences*, **49**, 453–458.

Wu, T. and Sundaram, V. (2019) Transplantation for acute-on-chronic liver failure. *Clinical Liver Disease*, **14**(4), 152.

Zhang, S., Suen, S.-C., Gong, C.L., Pham, J., Trebicka, J., Duvoux, C., Klein, A.S., Wu, T., Jalan, R. and Sundaram, V. (2021) Early transplantation maximizes survival in severe acute-on-chronic liver failure: Results of a Markov decision process model. *JHEP Reports*, 100367.