OPTIMAL TIMING OF LIVING-DONOR LIVER TRANSPLANTATION UNDER RISK-AVERSION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

By
Ümit Emre Köse
July 2016

OPTIMAL TIMING OF LIVING-DONOR LIVER TRANSPLANTATION UNDER RISK-AVERSION
By Ümit Emre Köse
July 2016

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Özlem Çavuş İyigün(Advisor)

Firdevs Ulus

İsmail S. Bakal

Approved for the Graduate School of Engineering and Science:

Levent Onural
Director of the Graduate School

ABSTRACT

OPTIMAL TIMING OF LIVING-DONOR LIVER TRANSPLANTATION UNDER RISK-AVERSION

Ümit Emre Köse
M.S. in Industrial Engineering
Advisor: Özlem Çavuş İyigün
July 2016

Liver transplantation, which can be performed from either living-donors or cadavers, is the only viable treatment for end-stage liver diseases. In this study, we focus on living-donor liver transplantation. The timing of the transplantation from a living-donor is crucial as it affects the quality and the length of the patient's lifetime. The studies in the literature use risk-neutral Markov decision processes (MDPs) to optimize the timing of transplantation. However, in real life, the patients and the physicians are usually risk-averse, therefore, those risk neutral models fail to represent the real behavior. In this study, we model the living-donor liver transplantation problem as a risk-averse MDP. We incorporate risk-aversion into the MDP model using dynamic coherent measures of risk, and in order to be able to reflect varying risk preferences of the decision makers, we use first-order mean-semi-deviation and mean-AVaR as the one-step conditional measures of risk. We obtain optimal policies for patients having cirrhotic diseases or hepatitis B under different risk preferences and organs of different quality. We also measure the sensitivity of the optimal policies to the transition probabilities and to the quality of life. We further perform a simulation study in order to find the distribution of lifetime under the risk-averse optimal policies.

Keywords: Liver transplantation, Markov decision process, dynamic risk measures, coherent risk measures.

ÖZET

YAŞAYAN DONÖRDEN KARACİĞER NAKLİ ZAMANLAMASININ RİSKTEN KAÇINARAK ENİYİLENMESİ

Ümit Emre Köse
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Özlem Çavuş İyigün
Temmuz 2016

Karaciğer nakli, yaşayan donörden veya kadavradan gerçekleştirilebilmekle beraber, son evre karaciğer hastalıkları için tek tedavi yöntemi olarak öne çıkmaktadır. Bu çalışmada, yaşayan donörden nakle odaklanılmıştır. Yaşayan donörden nakilde zamanlama, hastanın nakil öncesi ve sonrası yaşam kalitesini ve uzunluğunu etkilediği için büyük önem arz etmektedir. Literatürdeki çalışmalar, nakil zamanlamasını eniyilemek için, riske duyarsız Markov karar süreçleri (MKS) kullanmaktadır. Halbuki, gerçek hayatta, hastalar ve doktorlar genellikle riskten kaçınmaktadır. Bu çalışmada, yaşayan donörden karaciğer nakli problemi, riskten kaçınan bir MKS modeli ile modellenmiştir. Riskten kaçınma, modele dinamik tutarlı risk ölçütleri kullanılarak dahil edilmiştir ve karar vericilerin değişken risk tercihlerini yansıtabilmek için ortalama-kısmi sapma ve ortalama-riske maruz ortalama değer risk ölçütleri kullanılmıştır. Eniyi politikalar, sirotik hastalıklar veya hepatit B taşıyan hastalarda, farklı risk tercihleri ve organ kaliteleri göz önüne alınarak elde edilmiştir. Eniyi politikaların geçiş olasılıklarına ve yaşam kalitesine duyarlılığı ölçülmüştür. Bunun yanında, riskten kaçınan eniyi politikaların yaşam süresi dağılımını bulmak için bir benzetim çalışması yapılmıştır.

Anahtar sözcükler: Karaciğer nakli, Markov karar süreci, dinamik risk ölçütleri, tutarlı risk ölçütleri.

Acknowledgement

I would like to extend my utmost gratitude to my advisor Asst. Prof. Özlem Çavuş İyigün for her continuous support in my study and research, as well as her motivation, patience and guidance. It has been a privilege for me to conduct my studies as her student and benefit from her immense knowledge and ever-constructive criticism. Her meticulous and rigorous nature has always been an inspiration for me.

I also would like to express my gratitude to Assoc. Prof. Oğuzhan Alagöz for his support, guidance and significant contributions.

I thank Prof. Andrew J. Schaefer for his contributions.

I convey my thanks to Asst. Prof. Firdevs Ulus and Assoc. Prof. İsmail S. Bakal for their precious time to read and review my thesis.

I am also grateful to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting this work under Program 1001 and Project No. 213M442.

Last but not the least, I would like to thank my family - my parents, Harun and Kadem Köse and my brother, Berkay Köse for their ceaseless support, patience and understanding throughout my study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Liver transplantation is the only viable treatment for end-stage liver diseases such as cirrhosis and hepatitis B. There have been 5,950 liver transplantations in the USA in 2015. As of January 15th, 2016, there are 15,212 patients waiting for transplantation. For a liver transplantation, either livers from living-donors or cadaveric livers are used. Although cadaveric transplantations account for the majority of transplantations throughout history, living-donor transplantations have become a considerable proportion of all transplants since it was firstly performed (see Organ Procurement and Transplantation Network (OPTN) data[1]). A cadaveric liver becomes available after the donor's demise and must be used for transplantation immediately. At this point, living-donor transplantation has an advantage over cadaveric transplantation. Transplantation from a living-donor can be scheduled whenever deemed appropriate. Timing of transplantation has to aim maximization of recipient's pre-transplant and post-transplant lifetime as a whole considering the mortality risks associated with the operation.

An important issue in liver transplantation and all medical decision making problems is the risk-aversion of the decision makers. While making their decisions, patients and physicians often tend to be risk-averse. In a study reflecting the risk behavior of patients on a lung surgery decision, it has been found that lung cancer patients often choose to reject surgical treatments due to the risk preferences of

physicians or themselves (Cykert [2]). According to Massin et al. [3], there is a strong correlation between influenza vaccination decision and risk-aversion — the more risk-averse the physicians become, the more they are likely to prescribe vaccination. Additionally, Wakker and Deneffe [4] found that risk-aversion in patients is especially higher in cases where the objective is maximizing lifetime compared to decisions where medical treatment cost is in question. The examples emphasizing the decisiveness of risk-aversion are abundant in the medical literature. Therefore risk-neutral decision making models may yield unrealistic results.

In an end-stage liver disease, there is a certain number of health states which indicate the progression of the disease. End-stage liver diseases progress in a discrete Markovian manner, i.e. there is a discrete probability distribution associated with the health states the patient may be at in the next day and the next health state depends on the current health state of the patient. End-stage liver disease problem is previously modeled as an MDP under risk-neutrality assumption [5, 6, 7]. This study aims to present a model that optimizes timing of living-donor liver transplantation, i.e. finds the optimal health state for transplantation, while considering risk-aversion. That is, the problem of choosing the health state for transplantation has the objective of maximizing risk-adjusted pre-transplant and post-transplant lifetime as a whole. The problem is therefore formulated as a risk-averse undiscounted infinite horizon Markov Decision Process (MDP). Risk-aversion is incorporated into the model using dynamic coherent measures of risk. The reader is referred to the works of Ruszczynski [8] and Cavus and Ruszczynski [9, 10] for the details of risk-averse MDPs with dynamic risk measures. Coherent risk measures (see Artzner et al. [11] for details) have simpler interpretations than the risk representations in utility-based approaches. This study offers the first risk-averse undiscounted infinite horizon MDP dealing with living-donor liver transplantation timing problem. The parameters of the risk-averse MDP model have been estimated using actual clinical data obtained from United Network for Organ Sharing (UNOS) and Thomas E. Starzl Transplantation Institute within the University of Pittsburgh Medical Center (UPMC). Numerical results are obtained using health state transition probability matrices

corresponding to cirrhotic diseases or hepatitis B and post-transplant lifetime distributions for various organ types and patients of different ages. We implement the risk-averse MDP model using the risk-averse value iteration and policy iteration suggested by Cavus and Ruszczynski [9] and we compare the efficiency of these algorithms under the mentioned clinical data. We also obtain results when quality-adjusted life years (QALYs) are used. The results indicate that an optimal policy is indeed a control-limit policy and optimal transplantation decisions shift to earlier stages of the disease as risk-aversion increases. We also observe that as the organ quality or patient age increases, under the same risk preferences, optimal transplantation timing shifts to later stages of the disease. Another finding indicates that incorporating QALY into the model noticeably shifts the control-limits to later stages of the disease, as well as increasing the risk-sensitivity of the results. Since we also observe that disease type is a significant factor determining the risk-averse optimal control-limits, we perform a sensitivity analysis using different transition probability matrices with different progression rates. In order to find the lifetime distributions corresponding to the risk-averse optimal policies and compare them with the risk-neutral optimal policies, we use a simulation model. We see that risk-aversion significantly reduces lifetime standard deviation and pre-transplant mortality rate.

This thesis is structured as follows: In Chapter 2, we mention the previous studies in related fields of research. Next, in Chapter 3, we present a basic knowledge of conditional and dynamic measures of risk and an insight about risk-averse discrete-time undiscounted infinite horizon MDPs which make use of these conditional and dynamic measures of risk. In Chapter 4, we model the liver transplantation problem using such an MDP and we provide the dynamic programming equations necessary to solve the problem. Chapter 5 describes the risk-averse value iteration and policy iteration algorithms which are used to compute the optimal policies. In Chapter 6, we describe the problem data and estimation of problem parameters using these data, then we present the optimal policies. Chapter 6 also contains the sensitivity analysis with respect to the transition probability matrix and the simulation study. Finally, in Chapter 7 we give our concluding remarks.

Chapter 2

Literature Review

In this chapter, studies belonging to a number of research fields closely related to our problem are mentioned. To our best knowledge, the works concentrating on liver transplantation problems are solely based on risk-neutral models. The other studies under the umbrella term "medical decision making" mostly incorporate the risk attitude of the decision makers using utility functions. Although Markov decision processes (MDP) are relatively frequent in modeling medical decision making problems, the Markov decision processes using modern measures of risk, such as coherent measures of risk, are extremely uncommon in medical decision making.

2.1 Liver Transplantation Studies

As far as liver transplantation is considered, to our best knowledge, the previous works have focused on risk-neutral models. The decision making studies on liver transplantation mostly consider one of the two problems - they either suggest an optimal organ allocation policy or they aim to maximize a particular patient's benefit in liver transplantation.

In order to assess the severity of the end-stage liver disease and pre-transplant

mortality, Wiesner [12] develops MELD (Model for End-Stage Liver Disease) score as a function of blood levels such as serum creatinine, total serum bilirubin, international normalized ratio (INR) of prothrombin time and etiology of cirrhosis. This score has a range of 6-40 and higher values indicate a worse health status. Wiesner [12] suggests that an organ allocation policy which prioritizes patients with high MELD scores might decrease the number of patients who die while waiting for transplantation. Alternatively, Schaubel et al. [13] propose a survival-benefit based liver allocation system where the patients having the greatest benefit from a transplantation are prioritized. Using the observational data, this benefit is calculated as the difference of the expected post-transplant lifetime and expected lifetime staying in the waiting list without transplantation. They object MELD score based allocation system, stating that, although it reduces pre-transplant mortality rate, it disregards the expected post-transplant lifetime of a patient. These two studies intend to offer liver allocation policies maximizing the utilization of available livers having two distinct objectives. In a later study, Akan et al. [14] propose a liver allocation system design which has two objectives: minimizing total number of patient deaths while waiting for transplantation and maximizing total quality-adjusted life years (QALYs) of a patient. A quality-adjusted life year is between 0 and 1 years, considering the detrimental effects of the current health state. Usage of QALY facilitates representing real life situtations more accurately. Akan et al. [14] model the waiting list as a multiclass queue – in their study patients can dynamically switch classes where a class corresponds to a health status characterized by MELD score.

Alongside the liver allocation policies, another approach is to view the liver transplantation problem from the patient's perspective. Alagoz et al. [5] consider an end stage liver disease setting where the liver is transplanted from a living donor. They suggest an MDP model and use MELD score of the patient to represent the states of the stochastic process to optimally decide the timing of liver transplantation. The aim of their study is to maximize the expected summation of pre-transplant and post-transplant lifetimes. A following study by Alagoz et al. [6] generalizes this study to include organ availability and quality into the MDP model where cadaveric livers are considered. Alagoz et al. [6] define

the state of the system in terms of patient health and the available organ. At each decision stage, the offered organ is subject to a probability distribution which depends on the current health state. The study by Alagoz et al. [15] synthesizes the studies by Alagoz et al. [5] and Alagoz et al. [6] and offers an MDP model for risk-neutral liver transplantation timing problem while considering the availability of a cadaveric liver. They incorporate a disutility associated with using the living-donor liver instead of using a cadaveric liver. A further study by Sandikci et al. [7] analyzes how the information on waiting list affects the patient's optimal decisions, noting that United Network for Organ Sharing (UNOS) ranks the patients in the waiting list considering the severity of the patient, geography and physiologic compatibility between the donor and the recipient. They propose an MDP model where the states of the system are augmented to contain the ranking of the patient in the waiting list in addition to the patient health and the quality of the available organ. They demonstrate that the information on waiting list has a positive effect on expected lifetime of a patient. However, the models in this series of studies strictly use risk-neutral models and rely on expected lifetime calculations. In most real life problems, risk-aversion plays a crucial part in representing patient preferences.

There are also studies dealing with resource allocation problems, in general, which have possible applications to liver allocation problem. Righter [16] offers such a model where resources arrive according to a Poisson process. These resources are sequentially allocated to activities which have independent random deadlines and the objective is to maximize the expected return – in a medical decision making context, expected return may become expected total lifetime. David [17] addresses the same problem in a continuous time dynamic programming framework. In addition to these studies, some works approach the problem having the patient's lifetime benefit as the objective. David and Yechiali [18] propose a solution to a general optimal stopping problem which has applications in organ transplantation context. They consider when to accept an offer (e.g. organ) where offers arrive according to a random process and there is a random deadline for accepting offers (e.g. death). They firstly show that when the offers arrive at fixed instants, then there exists an optimal, time-dependent, control limit, i.e.

the time to accept the incoming offer. They also show that when the underlying process has a distribution with increasing failure rate and the arrivals form a renewal process, then the optimal control limit is a continuous nonincreasing function of time. They further generalize the model to include arrivals as a non-homogeneous Poisson process and they use it to determine an optimal policy for a kidney transplantation problem based on actual data. However, they emphasize that their model is intended to be used for organ transplantation problems in general. These studies which are intended to be used in organ transplantation problems do not consider risk at any point in their models. In the next section, we proceed to review the known models in medical decision making literature which incorporate risk-aversion.

2.2 General Risk-Averse Medical Decision Making Approaches

In contrast to risk-neutral approaches on liver transplantation, there are studies for general medical decision making problems, which consider the risk-aversion of decision makers. They mostly incorporate risk using utility functions of health states or life years. In a utility function approach, utilities are elicited from decision makers through inquiries about their preferences, and risk-aversion is defined as decreasing marginal utility for increasing reward [4]. The most well known elicitation methods are standard gamble [19, 20] and time trade-off [21] methods. The standard gamble method is used to determine the probability p such that the patient is indifferent between staying at his or her current health state, and the lottery of moving to the perfect health state (with utility 1) with probability p and the worst health state (with utility 0) with probability $(1-p)$. It is conventional to think that p gives the patient's current utility. In the time trade-off method, the patient is asked how many life years of perfect health is equivalent to a predefined number of life years in his or her current health state. Then the ratio of the number of life years in perfect health state and the specific number of life years in the current health state gives the utility of the current

health state. There exists a study by Stiggelbout et al. [22] where these two methods are compared for cancer patients. After the utilities are obtained for certain health states or specific numbers of life years, usually a utility function is fitted on the elicited utilities.

A risk-averse utility function is characterized as a non-decreasing concave function of rewards and usually power utility functions or exponential utility function are used to approximate the elicited utilities. For example, Pliskin et al. [23] determine the utility of health status using the standard gamble method and offer a power utility function of life years. Then they combine these into additive bivariate functional forms. Another exemplary work by Karni [24] builds an axiomatic expected utility model in a medical decision making context where utilities and risk attitudes are obtained by asking the decision makers (patients) about certainty equivalents of different lotteries – certainty equivalent of a lottery corresponds to the quantity making the decision maker indifferent between the certain amount and taking the lottery. In the model offered by Karni [24], a patient is faced with a diagnosis and an action (treatment) needs to be taken. He models the utility of an outcome resulting from a treatment as a function of the financial outcome and the post-treatment health state and adopts an exponential utility function. He states that the risk attitude is outcome dependent but action independent. Karni et al. [25] tests the model by Karni [24] on a problem, using actual data, where pregnant women decide to undergo prenatal diagnosis test and have to choose the physician to administer the test. They state that there is a trade-off between financial cost of hiring a physician and the risk of losing the fetus.

Expected utility theory has been challenged by various works which argue that its basic assumptions might be violated by many decision makers who are risk-averse. Kahneman and Tversky [26] propose Prospect Theory (PT) as a critique to expected utility theory. In their approach, based on the inquiries made with the decision makers, probabilities of the outcomes are transformed into decision weights using a weighting function, e.g. an inverse S-shaped probability weighting function – overweighting small probabilities and underweighting intermediate and high probabilities. An extension to the Prospect Theory is offered as Cumulative

Prospect Theory (CPT) by Tversky and Kahneman [27]. In CPT, cumulative decision weights are used rather than separate decision weights. Bleichrodt and Pinto [28] use weighting functions in a medical decision making context and offer a parameter-free elicitation method for weighting functions where, unlike previous studies, no functional form is assumed beforehand.

Utility functions and similar approaches are used in many areas including medical decision making. However, since they are constructed using elicitation inquiries, it is mostly difficult to interpret the degree of risk-aversion in utility functions. In our work, we model risk-aversion using coherent measures of risk.

2.3 MDPs in Medical Decision Making

Medical treatments often employ sequential decision making where outcomes of the decisions are often stochastic. Hence, there exists an abundant literature on MDPs used in medical decision making context, most of which disregard the fact that patients may prefer to avoid risk in their decisions. Alongside with MDPs, partially observable MDPs (POMDP) are also used in problem formulations where the state of the system is also stochastic. In this section we will review the works so far which make use of MDP models within a medical decision making context.

Schaefer et al. [29] offer a survey paper where they discuss implementation of various types of MDPs in medical treatment problems. One of the earliest works in the area is offered by Lefevre [30] where an epidemic control problem is formulated using a continuous-time MDP aiming to minimize the cost associated with the epidemic by deciding how much people to quarantine and how much medical treatment to allocate. Faissol et al. [31] consider an MDP model to optimally solve treatment and test timing problem for Hepatitis C patients in terms of cost effectiveness. Alterovitz et al. [32] provide a motion plan for visually guided surgical needles where the optimal steering actions are found using an MDP so that the probability that the needle will reach the target is maximized. A recent

study by Kumar and Ghildayal [33] compare different vaccination strategies in terms of cost effectiveness. These works mostly concentrate on problems of physicians or planners in medical decision making. However, most of the other studies consider problems from the patient's perspective. Barry [34] formulates an MDP model for patients with moderate symptoms of benign prostatic hyperplasia in order to optimally decide between transurethral resection surgery and waiting. Ahn and Hornberger [35] try to model a kidney transplantation problem using an MDP where organs can be accepted or rejected based on their quality as they become available. In another study, Shechter et al. [36] provide an MDP model to determine when to commence HIV treatment to maximize expected lifetime considering the detrimental effects and benefits of the treatment. Chhatwal et al. [37] provide an MDP model to decide biopsy timing for potential breast cancer patients. Denton et al. [38] and Kurt et al. [39] give MDP models which can be used to determine optimal statin therapy timing for diabetes patients. Van Arendonk et al. [40] notice the fact that most pediatric kidney transplant receivers need retransplantation afterwards and compare different retransplantation strategies using an MDP model. The works by Alagoz et al. [5] and Sandikci et al. [7] mentioned in Section 2.1 are also other examples for the application of MDPs in healthcare.

Alongside MDP models, there are also examples of POMDP models in medical decision making literature. Hu et al. [41] use a POMDP approach in order to determine a dosage policy to optimize anaesthesia administration of patients. Similarly, Hauskrecht and Fraser [42] use a POMDP to model the treatment of patients with ischemic heart disease. Maillart et al. [43] suggest a POMDP model to assess breast cancer screening policies. Kreke et al. [44] use POMDP and MDP techniques to model pneumonia related sepsis progression and aim to obtain optimal testing and hospital discharge policies. Another POMDP approach is used by Ayer et al. [45] in order to optimize mammography timing for potential breast cancer patients considering personal risk factors. Bennett and Hauser [46] develop a general purpose artificial intelligence framework which makes use of POMDPs and MDPs in order to make clinical decisions.

The works listed above assume that the decision maker is risk-neutral. Aside

from these works, there are a few works that incorporate the risk in MDP formulations in a medical decision making context. Cher et al. [47] incorporate risk-adjusted quality-adjusted life years (RA-QALY) into Barry's [34] MDP model in the same prostatic surgery problem. They use power-utility and exponential utility functions in order to obtain RA-QALY. Tilson and Tilson [48] use exponential utility function in an MDP model for treatment selection in an asymptomatic disease. However, unlike our study, these risk-averse works make use of utility functions as they incorporate risk-aversion into their models. In our study we use coherent measures of risk since it allows us to interpret risk preferences in a more convenient fashion.

2.4 Risk-Averse MDPs in General Context

In this study, we use a risk-averse MDP model using dynamic coherent measures of risk. In the existing literature, risk has been incorporated into MDPs using a number of approaches. A common method is to use exponential or entropic risk measure as the objective criterion [49, 50, 51, 52, 53, 54, 55, 56, 48, 57]. Di Masi and Stettner [58] demonstrate the existence and uniqueness of the solution to the Bellman equation using entropic risk measure in risk-averse MDPs. Another approach in risk-averse MDPs is to minimize the probability of failing to exceed a certain reward threshold [59, 60, 61]. Boda et al. [62] offer an analogous approach for a retirement problem where the probability that the value of a person's assets exceeds a threshold is maximized. Incorporating risk into MDPs is also achieved using mean-risk models. Filar et al. [63] propose a variance penalized MDP in a reward maximization context. Bauerle and Mundt [64] consider another mean-risk model where either average value-at-risk (AVaR) is minimized such that expected value is above a certain threshold or expected value is maximized such that AVaR is below a certain threshold. In the first problem the risk is minimized while keeping the expected return high enough whereas in the latter problem the expected return is maximized while the risk is kept at low levels. Haskell and Jain [65] provide a convex analytic approach for risk-sensitive MDPs including history dependent formulations. They augment the state space to include the history in

order to provide linear programs for AVaR minimization, stochastic dominance constrained problems and chance constrained problems. They also provide approximation methods in order to deal with infinite dimensional linear programs. These approximations use a sequence of finite dimensional linear programs such as aggregation of constraints, relaxations of the aggregate constraints and then an inner approximation of the decision variables. Similarly, Bauerle and Ott [66] and Chow et al. [67] provide models to minimize AVaR; however, minimizing risk without additional constraints might be criticized to have limited applications since in most problems expected value is a concern as well. Ruszczynski [8] uses dynamic coherent measures of risk in finite or discounted infinite horizon MDPs and offers iterative solution methods. Cavus and Ruszczynski [10] provide an extension of Ruszczynski [8] where they consider undiscounted infinite horizon problems and they model simple asset selling and organ transplantation models, without depending on actual data. A following work by Cavus and Ruszczynski [9] suggests computational methods for undiscounted infinite horizon MDPs with coherent dynamic risk measures.

Most of the decision analytic works listed in this section do not make use of coherent measures of risk. Some of the works here minimize the risk disregarding the expected value. In a medical decision making context, a mean-risk model would be much more appropriate since both expected value and the deviation from it can be valued differently by different patients. In this study, we use coherent mean-risk measures. It is important to note that coherent risk measures also allow easy interpretations of risk preferences unlike utility functions. This is especially useful where one intends to build a risk-averse model considering differing risk preferences. There are some works among the listed which do use mean-risk models, however, to our best knowledge, there is no study using an MDP where risk is incorporated using dynamic coherent measures of risk in a medical decision making context.

Chapter 3

Risk-Averse Undiscounted Transient MDPs

The nature of the living donor liver transplantation problem requires sequential decision making where at each period, the decision maker needs to decide between waiting or transplantation based on the current state of the disease. We assume that the disease progresses in a Markovian manner, in other words, the next state of the disease depends on the current state of the disease only, not on the entire history of the process. Therefore, this problem can be modeled as an MDP. We use a risk-averse discrete-time infinite horizon undiscounted transient MDP with dynamic coherent measures of risk (see [9] for details) to model the problem where the decision maker, that is the patient or the physician, is assumed to be risk-averse.

3.1 Conditional and Dynamic Risk Measures

Prior to proceeding to the infinite horizon risk-averse undiscounted MDP model we need to introduce conditional and dynamic measures of risk. These concepts are essential in forming multi-stage risk-averse stochastic programs. Throughout

this study, we will formulate the risk-averse liver transplantation problem as a cost minimization problem where we consider negatives of rewards, i.e. negatives of life days. Hence, we will deal with random costs and cost minimization. Let us have a sequence of random costs $\tilde{Z} = \{Z_1, Z_2, \ldots, Z_T\}$ where Z_t, $t = 1, 2, \ldots, T$ is defined on a probability space $\mathscr{Z}_t = (\Omega, \mathcal{F}_t, P)$ for $t = 1, 2, \ldots, T$ and the sequence adopts a filtration $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_T$. This implies that the uncertainty is resolved as we proceed through the periods. A function $\rho_t : \mathscr{Z}_{t+1} \to \mathscr{Z}_t$ is called a *one-step conditional risk measure*, as defined in [68], if it satisfies the following axioms:

(A1) $\rho_t(\lambda Z + (1-\lambda)W)) \leq \lambda\rho_t(Z) + (1-\lambda)\rho_t(W), \quad \lambda \in (0,1), \forall Z, W \in \mathscr{Z}_{t+1}$

(A2) If $Z \preceq W$, then $\rho_t(Z) \leq \rho_t(W), \quad \forall Z, W \in \mathscr{Z}_{t+1}$

(A3) $\rho_t(Z + W) = Z + \rho_t(W), \quad \forall Z \in \mathscr{Z}_t, W \in \mathscr{Z}_{t+1}$

(A4) $\rho_t(\alpha Z) = \alpha\rho_t(Z), \quad \forall Z \in \mathscr{Z}_{t+1}, \alpha \geq 0.$

The convexity axiom (A1) implies that diversification over multiple random costs yields to lower risk. The expression $Z \preceq W$ in the monotonicity axiom (A2) means that the random cost Z is less than the random cost W under all possible scenarios. Accordingly, the risk of Z is lower than the risk of W. The translation equivariance axiom (A3) is a result of the fact that the sequence of random variables $\tilde{Z}$ is subject to the filtration $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_T$. As uncertainty resolves through the periods $t = 1, 2, \ldots, T$, we observe that, at period t, the uncertainty concerning Z is resolved. Hence it can be moved outside the one-step conditional risk-measure. The positive homogeneity axiom (A4) can be interpreted as the risk concerning the random cost Z remains the same regardless of the currency we wish to express it in. These axioms are conditional extensions of axioms of coherent measures of risk defined by Artzner et al. [11]. We use two examples of such one-step conditional risk measures in our study. When ρ_t is defined as *conditional first-order mean-semi-deviation*, it takes the following form:

$$\rho_t(Z_{t+1}|\mathcal{F}_t) = E[Z_{t+1}|\mathcal{F}_t] + \kappa E\Big[Z_{t+1} - E[Z_{t+1}|\mathcal{F}_t]\Big|\mathcal{F}_t\Big]_+, \quad Z_{t+1} \in \mathscr{Z}_{t+1}. \quad (3.1)$$

In (3.1), $\rho_t(Z_{t+1})$ is the weighted sum of the conditional expectation of Z_{t+1} and expected upper semi-deviation from the conditional mean. Here $\kappa \in [0,1]$ indicates the weight given to the conditional semi-deviation and it can be $\mathcal{F}_t$ measurable (see [69, 70, 71, 68]). A higher κ value implies more risk-aversion.

Another example of one-step conditional risk measure is *mean-AVaR*, which can be defined as follows:

$$\rho_t(Z_{t+1}|\mathcal{F}_t) = \lambda E[Z_{t+1}|\mathcal{F}_t] + (1-\lambda) \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} E[Z_{t+1} - \eta | \mathcal{F}_t]_+ \right\}. \quad (3.2)$$

In the expression above, using $\lambda \in [0,1]$, conditional expectation is weighted with conditional average value-at-risk (AVaR) and α is the parameter of conditional AVaR. Here, both λ and α can be $\mathcal{F}_t$ measurable. It can be seen that high α and low λ values indicate high risk-aversion (see [72, 73, 68, 71]).

Now, we can proceed to define a dynamic risk measure. Let $\mathscr{Z} = \mathscr{Z}_1 \times \mathscr{Z}_2 \times \ldots \times \mathscr{Z}_T$ and $\tilde{Z} = \{Z_1, Z_2, \ldots, Z_T\}$. Having $\rho_t(\cdot)$ as a one-step conditional risk measure as defined above, a dynamic measure of risk (see [8, 68]) can be defined as follows:

$$J_T(\tilde{Z}) = \rho_1(Z_1 + \rho_2(Z_2 + \ldots + \rho_{T-1}(Z_{T-1} + \rho_T(Z_T)) \ldots)), \quad (3.3)$$

The case where $\rho_t(\cdot), t = 1, 2, \ldots, T$ is substituted with the conditional expectation $E[Z_{t+1}|\mathcal{F}_t]$, gives us the expected total value of the random cost sequence $\tilde{Z}$. In the next section, we will use dynamic coherent measures of risk within an MDP and obtain risk-averse undiscounted transient MDP formulations.

3.2 Risk-Averse Undiscounted Transient MDP Formulation

In order to provide a basis of understanding for our model, we give an introduction to infinite horizon risk-averse undiscounted transient MDPs (see the works of Cavus and Ruszczynski [9, 10]). Let us have a finite state space $\mathscr{X}$ and an action

space $\mathscr{U}$. For all $x \in \mathscr{X}$, we have a nonempty action set $U(x) \subset \mathscr{U}$, i.e. the set of actions or controls available at that state. For all $x \in \mathscr{X}$ and $u \in U(x)$, we have a probability measure defined over $\mathscr{X}$. Let $P(y|x,u)$ denote the transition probability from state $x \in \mathscr{X}$ to state $y \in \mathscr{X}$ under the action u in $U(x)$. At state $x \in \mathscr{X}$, taking action $u \in U(x)$ and transitioning to state $y \in \mathscr{X}$ has a cost of $c(x,y,u)$. A *stationary Markov decision process* is defined by its state space $\mathscr{X}$, action space $\mathscr{U}$, action set $U(\cdot)$, transition probabilities $P(\cdot|\cdot,\cdot)$ and cost function $c(\cdot,\cdot,\cdot)$. Since we are dealing with transient MDPs, let us define $\mathscr{X}_A \subset \mathscr{X}$ such that if $x \in \mathscr{X}_A$, x is an absorbing state, i.e. $P(x|x,u) = 1$ for all $x \in \mathscr{X}_A, u \in U(x)$. Any state $x \in \mathscr{X} \setminus \mathscr{X}_A$ is a transient state. Since it is assumed that no more costs are incurred after reaching absorbing state, we have $c(x,x,u) = 0$ for all $x \in \mathscr{X}_A$ and $u \in U(x)$.

Before proceeding to infinite horizon case, let us consider a finite horizon MDP with T periods. Let us define the sequence $x = (x_1, x_2, \ldots, x_T) \in \mathscr{X}^T$ as the history of states. Here x_1 can be considered as the starting state. A policy is a sequence $\Pi = \{\pi_1, \pi_2, \ldots, \pi_T\}$ of decision rules $\pi_t : \mathscr{X} \to \mathscr{U}$ such that $\pi_t(x) \in U(x), \forall x \in \mathscr{X}$ for $t = 1, \ldots, T$. If $\pi_t = \pi$, $\forall t = 1, 2, \ldots, T$ where $\pi : \mathscr{X} \to \mathscr{U}$ and $\pi(x) \in U(x), \forall x \in \mathscr{X}$, then such a policy is called stationary. The constant decision rule π is sufficient to define such a stationary *Markov policy*, hence we will refer to π as the policy hereafter. Using $c(x_t, x_{t+1}, \pi(x_t))$ as Z_t in (3.3), for each such policy, we can define the following *dynamic measure of risk*:

$$\begin{aligned} J_T(\pi, x_1) = \rho_1(c(x_1, x_2, \pi(x_1)) + \rho_2(c(x_2, x_3, \pi(x_2)) + \ldots \\ +\rho_{T-1}(c(x_{T-1}, x_T, \pi(x_{T-1})) + \rho_T(c(x_T, x_{T+1}, \pi(x_T))) \ldots)). \end{aligned} \tag{3.4}$$

Here $\rho_t(\cdot)$ is a one-step conditional risk measure as defined in the previous section. When the dynamic risk measure is extended to infinite horizon, it assumes the following form:

$$J_\infty(\pi, x_1) = \lim_{T \to \infty} J_T(\pi, x_1).$$

Then the infinite horizon problem is to minimize the dynamic measure of risk over possible policies:

$$J^*_\infty(x_1) = \min_\pi \lim_{T \to \infty} J_T(\pi, x_1). \tag{3.5}$$

Ruszczynski [8] states that the difficulty with the above formulation is that $\rho_t(\cdot)$, $t = 1, 2, \ldots, T$ can depend on the entire history of the process. For example, if the multipliers κ, λ and α in formulations (3.1) and (3.2) are allowed to depend on $\{x_1, x_2, \ldots, x_t\}$, we cannot obtain an optimal Markov policy described as above.

Hence, in order to overcome the difficulty in obtaining a stationary optimal policy, Ruszczynski [8] defines the concepts of *Markov risk transition mapping* and *Markov risk measure*. Let $\mathscr{V}$ be the finite dimensional space of all real functions on $\mathscr{X} \times \mathscr{U}$. The dual space of $\mathscr{V}$, $\mathscr{V}'$, can be thought as the space of signed measures on $\mathscr{U} \times \mathscr{X}$. Let $\mathscr{M}$ the set of probability measures in $\mathscr{V}'$. Accordingly, $\sigma : \mathscr{V} \times \mathscr{X} \times \mathscr{M} \to \mathbb{R}$ is called a *Markov risk transition mapping* if for every $x \in \mathscr{X}$, $u \in U(x)$ and for every $P(\cdot|x, \cdot) \in \mathscr{M}$, the function $\phi \mapsto \sigma(\phi, x, P(\cdot|x, u))$ is a coherent measure of risk on $\mathscr{V}$. When first-order mean-semi deviation is used as one-step conditional risk measure, $\sigma(\cdot, \cdot, \cdot)$ takes the following form:

$$\sigma(c(x, \cdot, \pi(x)) + v, x, P(\cdot|x, \pi(x))) = \underbrace{\sum_{y \in \mathscr{X}} \big(c(x, y, \pi(x)) + v(y)\big) P(y|x, \pi(x))}_{\mu}$$
$$+\kappa \sum_{y \in \mathscr{X}} \big(c(x, y, \pi(x)) + v(y) - \mu\big)_+ P(y|x, \pi(x)), \qquad \kappa \in (0, 1). \quad (3.6)$$

In case mean-AVaR is used, we get the following form for $\sigma(\cdot, \cdot, \cdot)$:

$$\sigma(c(x, \cdot, \pi(x)) + v, x, P(\cdot|x, \pi(x))) = \lambda \sum_{y \in \mathscr{X}} \big(c(x, y, \pi(x)) + v(y)\big) P(y|x, \pi(x))$$
$$+(1 - \lambda) \inf_{\eta \in \mathbb{R}} \Big\{\eta + \frac{1}{1 - \alpha} \sum_{y \in \mathscr{X}} \big(c(x, y, \pi(x)) + v(y) - \eta\big)_+ P(y|x, \pi(x))\Big\}. \quad (3.7)$$

Assume that for every $x \in \mathscr{X}$ and $m \in \mathscr{M}$ the risk transition mapping $\sigma(\cdot, \cdot, \cdot)$ is continuous. Then, from Theorem 2.2. of [74], it follows that there is a closed, convex and bounded set $A(x, m) \subset \mathscr{M}$ such that for all $v \in \mathscr{V}$ we have:

$$\sigma(v, x, m) = \max_{\mu \in A(x,m)} \langle v, \mu \rangle. \quad (3.8)$$

The set $A(x, m)$ is previously derived for different risk transition mappings [74]. According to Ruszczynski [8], the set $A(x, m)$ is formulated as follows for

first-order mean-semi-deviation risk measure:

$$\begin{aligned} A(x,m) = \{\mu \in \mathscr{M} : \exists (h \in \mathscr{V}) \\ \mu(u,y) = m(u,y)[1 + h(u,y) - \langle h, m \rangle], \forall (u,y) \in \mathscr{U} \times \mathscr{X}, \\ h(y) \leq \kappa, \forall y \in \mathscr{X}, \\ h \geq 0\}, \end{aligned} \tag{3.9}$$

and for mean-AVaR risk measure, it takes the following form:

$$\begin{aligned} A(x,m) = \{\mu \in \mathscr{M} : \exists z : \mathscr{X} \to \mathbb{R} \quad \text{s.t.} \\ \mu(u,y) = \lambda m(u,y) + (1-\lambda) z(y), \forall (u,y) \in \mathscr{U} \times \mathscr{X} \\ 0 \leq z(y) \leq \frac{1}{1-\alpha} m(u,y), y \in \mathscr{X}\}. \end{aligned} \tag{3.10}$$

A one-step conditional risk measure $\rho_t(\cdot)$ is called a *Markov risk measure* if there exists a risk transition mapping $\sigma_t : \mathscr{V} \times \mathscr{X} \times \mathscr{M} \to \mathbb{R}$ such that for all $v \in \mathscr{V}$ and $u_t \in U(x_t)$ we have:

$$\rho_t(v(x_{t+1})) = \sigma_t(v, x_t, P(\cdot|x_t, u_t)),$$

where the risk at period t is parametrized by x_t, and thus only depends on the past via state x_t.

The problem given in 3.5 can be solved using dynamic programming (see [9, 10]). In the infinite horizon problem, the aim is to minimize $J_\infty(\pi, x_1)$ with respect to π. Dynamic programming equations take the following form:

$$v(x) = \min_\pi \sigma(c(x, \cdot, \pi(x)) + v, x, P(\cdot|x, \pi(x))), \quad x \in \mathscr{X} \backslash \mathscr{X}_A \tag{3.11}$$

$$v(x) = 0, \quad x \in \mathscr{X}_A, \tag{3.12}$$

where $v \in \mathbb{R}^{|\mathscr{X}|}$ is the value function and $\sigma(\cdot,\cdot,\cdot)$ is a Markov risk transition mapping. It is given that if a Markov risk transition mapping satisfies (3.11) and (3.12), then the following also holds:

$$v(x) = \min_\pi J_\infty(\pi, x), \quad x \in \mathscr{X}. \tag{3.13}$$

In the next chapter we model the risk-averse liver transplantation timing problem as a risk-averse undiscounted transient MDP.

Chapter 4

Liver Transplantation: MDP Model

As mentioned in the previous chapter, the living donor liver transplantation problem requires sequential decision making where at each decision stage, the decision maker, patient or physician, needs to decide between waiting or transplantation based on the current state of the disease. End-stage diseases such as cirrhosis and hepatitis B progress in a Markovian manner, i.e. the next health state of the patient depends solely on the current health state of the patient. In the previous study by Alagoz et al. [5] the living-donor liver transplantation problem is modeled using an MDP; however, that study assumes risk-neutrality. We use a similar model where health states are based on MELD scores and days are used as time units; but we assume that the decision makers are risk-averse. Therefore we model the liver transplantation problem as a risk-averse discrete-time infinite horizon undiscounted transient MDP. In accordance with the available clinical data on the progression of the disease, days are used as time units.

4.1 Model Description

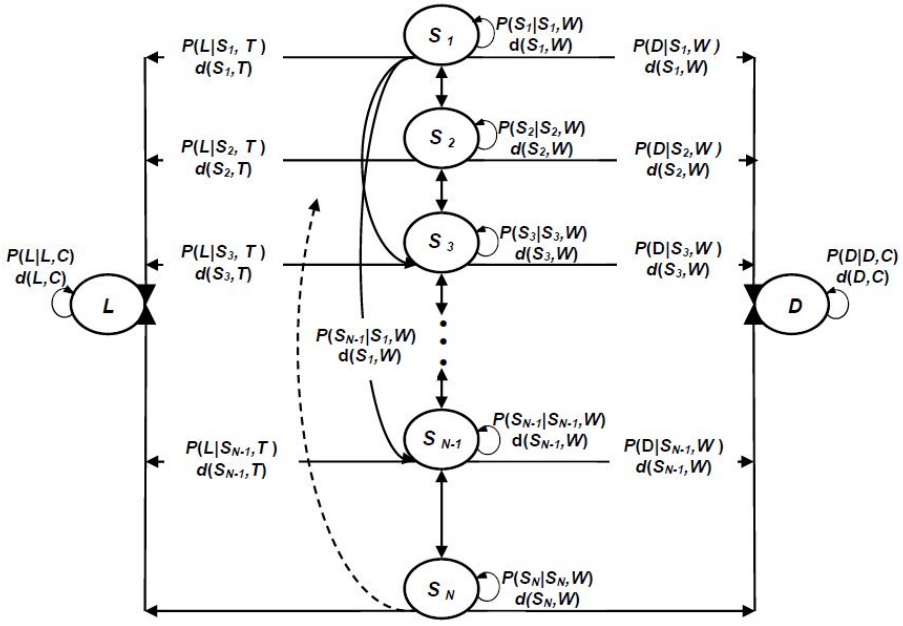
In an end-stage liver disease, health status of patients changes throughout time until the disease progression ends either in liver transplantation or death. According to Wiesner [12], these changes can be monitored using MELD (Model for End Stage Liver Disease) score which is a function of certain blood values such as total bilirubin, creatinine, prothrombin time. These blood values also constitute disease symptoms. MELD score has been previously used to guide liver allocation decisions as an indicator of disease severity [75, 12]. MELD score varies between 6 and 40. Typically, a high MELD score indicates a severe health status. In an MDP model based on patient's health status, it is reasonable to incorporate MELD score into the states of the MDP. Let us have N such health states denoted as $S_1, S_2, \ldots, S_N$ where each health state corresponds to a certain MELD score range, e.g. S_1 for MELD score 6-7, S_2 for MELD score 8-9 and so on. This aggregation of MELD scores in modeling health states is due to the sparsity in the post-transplant survival data for high MELD scores. Let S_1 be the best health state whereas S_N is the worst. Additionally, let D and L correspond to death and post-transplant states respectively, and these two states are considered as absorbing states, i.e. $\mathscr{X}_A = \{L, D\}$. Then the state space is defined as $\mathscr{X} = \{S_1, S_2, \ldots, S_N, D, L\}$ in the infinite horizon MDP model.

For each health state $x \in \mathscr{X}$, $\mathscr{X} = \{S_1, S_2, \ldots, S_N, D, L\}$ there is a set $U(x)$ which denotes the actions (controls) available at state x. For any health state $S_i \in \{S_1, \ldots, S_N\}$, $U(S_i) = \{W, T\}$ where W and T correspond to waiting and transplantation, respectively. When wait action is taken at state $S_i \in \{S_1, \ldots, S_N\}$, the patient moves to health state $S_j \in \{S_1, \ldots, S_N\}$ with probability $P(S_j|S_i, W)$. On the other hand, waiting at state $S_i \in \{S_1, \ldots, S_N\}$ can result in death of the patient with probability $P(D|S_i, W)$. If transplantation is decided at state $S_i \in \{S_1, \ldots, S_N\}$, then the patient moves to the post-transplant state L with certainty, i.e. $P(L|S_i, T) = 1$. Once the states D and L are reached, a decision is no longer required since either the patient dies or the process ends with a transplantation. Hence, we state $U(D) = U(L) = \{C\}$ where C is a formal action which ensures staying at the same state. Since L and D are

absorbing states, $P(L|L,C) = P(D|D,C) = 1$.

As the patient moves through the states before transplantation or death, he collects rewards which correspond to (quality-adjusted) lifetime — days in our case. Waiting at a state $S_i \in \{S_1, \ldots, S_N\}$ has an associated reward $d(S_i, W)$. This reward corresponds to one day if not quality-adjusted, otherwise it is a proportion of a day. Likewise, transplanting at a state $S_i \in \{S_1, \ldots, S_N\}$ has the reward $d(S_i, T)$ and this reward represents the risk (and quality) adjusted lifetime after transplantation. Once the absorbing states L and D have been reached, then no more reward is supposed to be collected. Therefore $d(D, C) = d(L, C) = 0$. State transition diagram of the problem with N health states is illustrated on Figure 4.1.

Figure 4.1: State transition diagram of the problem.



Accordingly, living donor liver transplantation problem is to determine which actions to take, waiting or transplantation, at each health state in order to maximize the risk-adjusted pre-transplant and post-transplant lifetime as a whole. An array of optimal actions for each health state, regardless of the current decision period, constitutes a policy. Needless to say, a policy defined as such would be deterministic and stationary.

In order to be able to define the problem more clearly, let us start with the finite horizon case. Assume we have T periods and a stationary policy π for this finite horizon where $\pi : \mathscr{X} \to \mathscr{U}$ and the system has been in states $x = \{x_t\}_{t=1}^{T}$ throughout the history. Then using (3.4), we could devise the following dynamic risk measure:

$$\begin{aligned} J_T(\pi, x_1) = \rho_1(-d(x_1, \pi(x_1)) + \rho_2(-d(x_2, \pi(x_2)) + \dots \\ +\rho_{T-1}(-d(x_{T-1}, \pi(x_{T-1})) + \rho_T(-d(x_T, \pi(x_T)))\dots)). \end{aligned}$$

In the infinite horizon problem, the aim is to minimize $J_\infty(\pi, x_1)$ with respect to π, where π is a stationary and deterministic policy.

4.2 Dynamic Programming Equations

This problem can be solved using dynamic programming (see [9, 10] for details). Dynamic programming equations take the following form:

$$v(x) = \min_{\pi}\{\sigma(-d(x, \pi(x)) + v, x, P(\cdot|x, \pi(x)))\}, \quad x \in \mathscr{X}\backslash\{L, D\} \quad (4.1)$$

$$v(x) = 0, \quad x \in \{L, D\}. \quad (4.2)$$

The reward parameter $-d(x, \pi(x))$ can move outside the risk-transition mapping since it is deterministic. It is left inside the risk-transition mapping for our convenience and consistence. In (4.1) and (4.2), $v(x)$, $x \in \mathscr{X}$ is the value function and $\sigma(\cdot, \cdot, \cdot)$ is a Markov risk transition mapping. Let us define $v(S_i)$ as the negative maximum risk-averse lifetime that could be obtained starting from health state S_i. Note that since we are in a cost minimization context, we minimize the negative risk-averse lifetime. When first-order mean-semi-deviation is used as the one-step conditional measure of risk, using (4.1), (4.2) and (3.6), the dynamic

programming equations take the following form:

$$v(S_i) = \min_{u \in U(S_i)} \Big\{ \overbrace{\sum_{y \in \mathscr{X}} \big(v(y) - d(S_i,u)\big) P(y|S_i,u)}^{\mu} + \kappa \sum_{z \in \mathscr{X}} \Big(v(z) - d(S_i,u) - \mu\Big)_+ P(z|S_i,u) \Big\}, \quad S_i \in \{S_1, \ldots, S_N\} \tag{4.3}$$

$$v(D) = 0 \tag{4.4}$$

$$v(L) = 0. \tag{4.5}$$

In a similar way, when mean-AVaR is used as the one-step conditional measure of risk, using (4.1), (4.2) and (3.7), dynamic programming equations are formulated as follows:

$$v(S_i) = \min_{u \in U(S_i), \eta \in \mathbb{R}} \Big\{ \lambda \sum_{y \in \mathscr{X}} \Big(v(y) - d(S_i,u)\Big) P(y|S_i,u) + (1-\lambda)\Big(\eta + \frac{1}{1-\alpha} \sum_{z \in \mathscr{X}} \Big(v(z) - d(S_i,u) - \eta\Big)_+ P(z|S_i,u)\Big)\Big\},$$

$$S_i \in \{S_1, \ldots, S_N\} \tag{4.6}$$

$$v(D) = 0 \tag{4.7}$$

$$v(L) = 0. \tag{4.8}$$

Here note that $v(L)$ and $v(D)$ are set to zero since these states are absorbing states. No more life days are collected once the absorbing states are reached.

In order to be able to calculate $v(S_i), S_i \in \{S_1, \ldots, S_N\}$, we need to compute $d(S_i, W)$ and $d(S_i, T)$, rewards associated with waiting and transplantation at each period respectively. The waiting rewards, $d(S_i, W)$, $S_i \in \{S_1, \ldots, S_N\}$ are determined beforehand by assigning (if needed) quality of life adjusted lifetime that could be obtained by waiting a day without transplantation, which varies between 0 and 1. The post-transplant rewards $d(S_i, T), S_i \in \{S_1, \ldots, S_N\}$ are calculated using a finite horizon undiscounted MDP model (see [10]). Let the post-transplant horizon have K periods and $-v_k(S_i)$ be the risk-averse lifetime left k days after transplanting at state S_i. It is assumed that the patient lives at most K days after the transplantation. Accordingly, $v_K(S_i) = 0$ and $d(S_i, T) = -v_0(S_i)$. Probability that a patient lives another day given that she has lived k days after the transplantation is given as $P(k+1|k, S_i)$.

Accordingly, the post-transplant risk-averse lifetimes, i.e. $d(S_i, T)$, $S_i \in \{S_1, \ldots, S_N\}$ can be found using a finite horizon MDP as discussed in [10]. In this study we use first-order mean-semi-deviation and mean-AVaR as the one-step coherent measures of risk. When first-order mean-semi-deviation is used as the dynamic risk measure, dynamic programming equations in the finite horizon MDP are stated as follows:

$$\begin{aligned} v_{S_i}(k) = & \overbrace{\Big(-q + v_{S_i}(k+1)\Big)P(k+1|k,S_i) + (-q)\Big(1 - P(k+1|k,S_i)\Big)}^{\mu} \\ & + \kappa\Big(\big(-q + v_{S_i}(k+1) - \mu\big)_+ P(k+1|k,S_i) + (-q-\mu)_+(1 - P(k+1|k,S_i)\Big), \\ & i = 1,2,\ldots,N, \quad k = 0,1,\ldots,K. \end{aligned} \tag{4.9}$$

Here μ corresponds to expected value at stage $k+1$ and the quantity multiplied by κ is the upper-semi-deviation. For mean-AVaR risk measure, dynamic programming equations take the following form:

$$\begin{aligned} v_{S_i}(k) = \min_{\eta \in \mathbb{R}} \Bigg\{ & \lambda\Bigg(\Big(-q + v_{S_i}(k+1)\Big)P(k+1|k,S_i) + (-q)\Big(1 - P(k+1|k,S_i)\Big)\Bigg) \\ & + (1-\lambda)\Bigg(\eta + \frac{1}{1-\alpha}\Big(\big(-q + v_{S_i}(k+1) - \eta\big)_+ P(k+1|k,S_i) \\ & + \big(-q - \eta\big)_+ \big(1 - P(k+1|k,S_i)\big)\Big)\Bigg)\Bigg\} \\ & i = 1,2,\ldots,N, \quad k = 0,1,\ldots,K. \end{aligned} \tag{4.10}$$

In this equation the quantity weighted by λ is the expected value at stage $k+1$ and the quantity weighted by $(1-\lambda)$ corresponds to AVaR. The quantity q corresponds to the reward obtained by patient by living another day. It corresponds to one day if QALY is not used, otherwise it corresponds to a proportion of a day. In the next section, we will discuss the results obtained by implementing this model.

Chapter 5

Solution Methods

The dynamic programming equations (4.1) and (4.2) have a unique solution that gives us the negative of maximum risk-adjusted lifetime for any initial state (see Cavus and Ruszczynski [10]). There exist a number of methods to compute this unique solution (see Cavus and Ruszczynski [9] for details). These are value iteration and policy iteration algorithms. In the policy evaluation step of the policy iteration algorithm, a nonsmooth equation system has to be solved. Cavus and Ruszczynski [9] suggest two different methods to solve this equation system resulted from risk-averse undiscounted transient Markov chains: convex optimization and Newton's nonsmooth method. We adapted these algorithms into the liver transplantation model described throughout Section 4.1. Aside from the risk-averse algorithms, we also implement the risk-neutral algorithms in order to obtain a thorough comparison between risk-neutral and risk-averse results. The quantitative comparison of the algorithms in terms of performance is provided in Section 6.2.1. All the algorithms mentioned in this section rely on the reward parameters $d(S_i, W)$, $d(S_i, T)$, $S_i \in \{S_1, \ldots, S_N\}$ and the health state transition probability matrix $P(y|x, u), x \in \mathscr{X}, y \in \mathscr{X}, u \in U(x)$ as data. In the following sections, we describe the value iteration and policy iteration algorithms.

5.1 Value Iteration

The value iteration method was initially suggested by Bellman [76]. The risk-averse extension of the value iteration algorithm for undiscounted infinite horizon problems can be found in [9]. Value iteration algorithm starts from an arbitrary value function v_0 and it finds a new approximation in each iteration, i.e. v_k for $k = 1, 2, 3, \ldots$. The algorithm returns the optimal value function once $v_{k+1} \equiv v_k$. However, in practice, the algorithm stops once a certain tolerance level is satisfied. We use $v(x), x \in \mathscr{X}$ as the value function whose definition can be seen in Section 4.2. The stopping condition for the value iteration algorithms in our study is defined as $\max_{i=1}^{N} |v_{k+1}(S_i) - v_k(S_i)| \leq \epsilon$ where $\epsilon \in \mathbb{R}$ is an arbitrary number indicating the tolerance. In our value iteration implementations, ϵ is taken to be 10^{-4}. A small epsilon value ensures that the optimal values are well approximated; however, it leads to more iterations until convergence. Bellman's risk-neutral value iteration algorithm can be implemented as follows for our problem under the risk-neutral case:

Algorithm 5.1 Risk-neutral value iteration algorithm.

procedure RISK-NEUTRAL VALUE ITERATION(v_0)
  $k \leftarrow 0$;
  $v_0(x) \leftarrow 0, \quad x \in \mathscr{X}$ ▷ Any initial value function with $v(L) = v(D) = 0$ is acceptable.
  do
    $v_{k+1}(S_i) \leftarrow \min_{u \in U(S_i)} \{ \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) \}, \quad S_i \in \{S_1, \ldots, S_N\}$
    $v_{k+1}(L) \leftarrow 0, \quad v_{k+1}(D) \leftarrow 0$
    $k \leftarrow k + 1$
  while $\max_{i=1}^{N} |v_k(S_i) - v_{k-1}(S_i)| > \epsilon$
  $v^*(x) \leftarrow v_k(x), \quad x \in \mathscr{X}$
  $\pi^*(S_i) \leftarrow \arg \min_{u \in U(S_i)} \{ \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) \}, \quad S_i \in \{S_1, \ldots, S_N\}$
  $\pi^*(L) \leftarrow C, \quad \pi^*(D) \leftarrow C$
  return v^*, π^*
end procedure

Our main goal in our implementations is to obtain the optimal policies under risk-aversion. If first-order mean-semi-deviation is used as the one-step conditional risk measure, the dynamic programming equations required to solve our problem get the form (4.3), (4.4), (4.5). Accordingly, we obtain Algorithm 5.2.

Algorithm 5.2 Value iteration algorithm with first-order mean-semi-deviation as the one-step conditional risk measure.

procedure VALUE ITERATION WITH FIRST-ORDER MEAN-SEMI-DEVIATION(v_0)
  $k \leftarrow 0$
  $v_0(x) \leftarrow 0, \quad x \in \mathscr{X}$ ▷ Any initial value function with $v(L) = v(D) = 0$ is acceptable.
  do

$$v_{k+1}(S_i) \leftarrow \min_{u \in U(S_i)} \Big\{ \underbrace{\sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u)}_{\mu} + \kappa \sum_{y \in X} \Big(v_k(S_i) - d(S_i, u) - \mu\Big)_+ P(y|S_i, u) \Big\}, \quad S_i \in \{S_1, \ldots, S_N\}$$

    $v_{k+1}(L) \leftarrow 0, \quad v_{k+1}(D) \leftarrow 0$
    $k \leftarrow k + 1$
  while $\max_{i=1}^{N} |v_k(S_i) - v_{k-1}(S_i)| > \epsilon$
  $v^*(x) \leftarrow v_k(x), \quad x \in \mathscr{X}$
  $\pi^*(S_i) \leftarrow \arg \min_{u \in U(S_i)} \Big\{ \mu + \kappa \sum_{y \in X} \Big(v_k(y) - d(S_i, u) - \mu\Big)_+ P(y|S_i, u) \Big\}, \quad S_i \in \{S_1, \ldots, S_N\}$
  $\pi^*(L) \leftarrow C, \quad \pi^*(D) \leftarrow C$
  return v^*, π^*
end procedure

The alternative is to use mean-AVaR as the one step conditional risk measure. The dynamic programming equations in this case are formulated as in (4.6), (4.7), (4.8). The value-at-risk, η, is implemented as $\eta_k^{\alpha}(S_i, u)$ for health state $S_i \in \{S_1, \ldots, S_N\}$, iteration k, under action $u \in U(S_i)$ and parameter α. Such a value iteration algorithm is formulated as in Algorithm 5.3.

Algorithm 5.3 Value iteration algorithm with mean-AVaR as the one-step conditional risk measure.

procedure VALUE ITERATION WITH MEAN-AVAR(v_0)
  $k \leftarrow 0$
  $v_0(x) \leftarrow 0, \quad x \in \mathscr{X}$ ▷ Any initial value function with $v(L) = v(D) = 0$ is acceptable.
  do
    $\eta_k^\alpha(S_i, u) \leftarrow \inf_{y \in \mathscr{X}} \Big\{ w \in \mathbb{R} | P(v_k(y) - d(S_i, u) \leq w) \geq \alpha \Big\}, \quad S_i \in \{S_1, \ldots, S_N\}$

$$v_{k+1}(S_i) \leftarrow \min_{u \in U(S_i)} \Bigg\{ \lambda \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) + (1-\lambda)\Big(\eta_k^\alpha(S_i, u) + \frac{1}{1-\alpha}\Big(\sum_{y \in \mathscr{X}} (v_k(y) - d(S_i, u) - \eta_k^\alpha(S_i, u))_+ P(y|S_i, u)\Big)\Big) \Bigg\}, \quad S_i \in \{S_1, \ldots, S_N\}$$

    $v_{k+1}(L) \leftarrow 0, \quad v_{k+1}(D) \leftarrow 0$
    $k \leftarrow k + 1$
  while $\max_{i=1}^N |v_k(S_i) - v_{k-1}(S_i)| > \epsilon$
  $v^*(x) \leftarrow v_k(x), \quad x \in \mathscr{X}$

$$\pi^*(S_i) \leftarrow \arg\min_{u \in U(S_i)} \Bigg\{ \lambda \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) + (1-\lambda)\Big(\eta_{k-1}^\alpha(S_i, u) + \frac{1}{1-\alpha}\Big(\sum_{y \in \mathscr{X}} (v_k(y) - d(S_i, u) - \eta_{k-1}^\alpha(S_i, u))_+ P(y|S_i, u)\Big)\Big) \Bigg\}, \quad S_i \in \{S_1, \ldots, S_N\}$$

  $\pi^*(L) \leftarrow C, \quad \pi^*(D) \leftarrow C$
  return v^*, π^*
end procedure

5.2 Policy Iteration

The basic policy iteration algorithm was developed by Howard [77]. The policy iteration algorithm starts with an initial arbitrary policy π_0. At iteration k it evaluates the policy π_k and finds the value function v_k associated with that policy. Then using the value function v_k, the current policy is improved into another policy π_{k+1}. Once $\pi_{k+1} \equiv \pi_k$, the algorithm stops and returns an optimal policy and its associated value function. The risk-neutral policy iteration algorithm is adapted as in Algorithm 5.4 for our problem. Under risk-neutrality, at the policy evaluation step of each iteration the value function can be found simply by solving a linear system of equations. However, when a risk transition mapping is incorporated instead of the expected value, the problem becomes more complex since a nonsmooth equation system has to be solved. Two methods suggested by [9] are used to solve this nonsmooth equation system. These methods are convex optimization and Newton's method. In the following subsections, the policy

Algorithm 5.4 Risk-neutral policy iteration algorithm.

procedure POLICY ITERATION(π_0)
  $k \leftarrow 0$
  $\pi_0(S_i) \leftarrow T, \quad S_i \in \{S_1, \ldots, S_N\}, \quad \pi_0(L) \leftarrow C, \quad \pi_0(D) \leftarrow C$ ▷ Any initial feasible policy is acceptable.
  do
    Policy Evaluation Step:
    solve

$$\begin{aligned} v(S_i) &= \sum_{y \in \mathscr{X}} \Big(v(y) - d(S_i, \pi_k(S_i))\Big) P(y|S_i, \pi_k(S_i)), \quad S_i \in \{S_1, \ldots, S_N\} \\ v(L) &= 0 \\ v(D) &= 0 \end{aligned}$$

    $v_k(x) \leftarrow v(x), \quad x \in \mathscr{X}$
    Policy Improvement Step:

$$\pi_{k+1}(S_i) \leftarrow \arg \min_{u \in U(S_i)} \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u), \quad S_i \in \{S_1, \ldots, S_N\}$$

    $k \leftarrow k + 1$
  while $\pi_k \not\equiv \pi_{k-1}$
  $v^*(x) \leftarrow v_{k-1}(x), \quad x \in \mathscr{X}$
  $\pi^*(x) \leftarrow \pi_k(x), \quad x \in \mathscr{X}$
  return v^*, π^*
end procedure

iteration algorithm with each of these methods will be demonstrated explicitly.

5.2.1 Policy Iteration Using Convex Optimization

As mentioned before, we use first-order mean-semi-deviation and mean-AVaR as one-step conditional risk measures. Risk-averse policy iteration algorithm, when first-order mean-semi-deviation is used, can be used to compute the unique solution of the dynamic programming equations (4.3), (4.4), (4.5). It can also provide an optimal policy. In this algorithm, at each policy evaluation step of iteration k, the value function v_k corresponding to the current policy π_k is found

as an optimal solution $\bar{v}$ of a convex program which is formulated as follows:

$$\min_{v} \sum_{x \in \mathscr{X}} v(x)$$
s.t.
$$\mu = \sum_{y \in \mathscr{X}} \big(v(y) - d(S_i, \pi_k(S_i))\big) P(y|S_i, \pi_k(S_i))$$
$$v(S_i) \geq \mu + \kappa \sum_{\bar{y} \in \mathscr{X}} \big(v(\bar{y}) - d(S_i, \pi_k(S_i)) - \mu\big)_+ P(\bar{y}|S_i, \pi_k(S_i)),$$
$$S_i \in \{S_1, \ldots, S_N\}$$
$$v(L) = 0$$
$$v(D) = 0.$$

For our convenience, we use a linearized equivalent program. The linearization can be given as follows:

$$\min_{v,z} \sum_{x \in \mathscr{X}} v(x) \tag{5.1}$$
s.t.
$$v(S_i) \geq \sum_{y \in \mathscr{X}} (v(y) - d(S_i, \pi_k(S_i))) P(y|S_i, \pi_k(S_i)) \tag{5.2}$$
$$+\kappa \sum_{y \in \mathscr{X}} z(S_i, y) P(y|S_i, \pi_k(S_i)), \quad S_i \in \{S_1, \ldots, S_N\} \tag{5.3}$$
$$z(S_i, y) \geq v(y) - d(S_i, \pi_t(S_i)) \tag{5.4}$$
$$- \sum_{\bar{y} \in \mathscr{X}} (v(\bar{y}) - d(S_i, \pi_k(S_i))) P(\bar{y}|S_i, \pi_k(S_i)), \tag{5.5}$$
$$S_i \in \{S_1, \ldots, S_N\}; y \in \mathscr{X} \tag{5.6}$$
$$z(S_i, y) \geq 0, \quad S_i \in \{S_1, \ldots, S_N\}; \quad y \in \mathscr{X} \tag{5.7}$$
$$v(L) = 0 \tag{5.8}$$
$$v(D) = 0. \tag{5.9}$$

Accordingly, we construct Algorithm 5.5 using the linear program in (5.1) - (5.9).

Algorithm 5.5 Policy iteration algorithm using convex optimization with first-order mean-semi-deviation as the one-step conditional risk measure.

procedure POLICY ITERATION WITH FIRST-ORDER MEAN-SEMI-DEVIATION (CO)(π_0)
  $k \leftarrow 0$
  $\pi_0(S_i) \leftarrow T, \quad S_i \in \{S_1, \ldots, S_N\}, \quad \pi_0(L) \leftarrow C, \quad \pi_0(D) \leftarrow C$ ▷ Any initial feasible policy is acceptable.
  do
    Policy Evaluation Step:
    solve (5.1) - (5.9) to find an optimal solution $\bar{v}$
    $v_k(x) \leftarrow \bar{v}(x), \quad x \in \mathscr{X}$
    Policy Improvement Step:

$$\pi_{k+1}(S_i) \leftarrow \arg\min_{u \in U(S_i)} \underbrace{\sum_{y \in X} \Big(v_k(y) - d(S_i, u)\Big) P(y|S_i, u)}_{\mu} + \kappa \sum_{y \in X} \Big(v_k(y) - d(S_i, u) - \mu\Big)_+ P(y|S_i, u), \quad S_i \in \{S_1, \ldots, S_N\}$$

    $k \leftarrow k + 1$
  while $\pi_k \not\equiv \pi_{k-1}$
  $v^*(x) \leftarrow v_{k-1}(x), \quad x \in \mathscr{X}$
  $\pi^*(x) \leftarrow \pi_k(x), \quad x \in \mathscr{X}$
  return v^*, π^*
end procedure

Alternatively, mean-AVaR can be used as the one-step conditional risk measure instead of first-order mean-semi-deviation. We can use the risk-averse policy iteration algorithm with mean-AVaR to find the unique solution to the dynamic programming equations (4.6), (4.7), (4.8). At each policy evaluation step of iteration k, the convex program whose optimal value $\bar{v}$ provides the value function v_k corresponding to the current policy π_k is formulated as follows:

$$\min_{v, \eta^\alpha} \sum_{x \in \mathscr{X}} v(x)$$

s.t.

$$v(S_i) \geq \lambda \sum_{y \in \mathscr{X}} (v(y) - d(S_i, \pi_k(S_i))) P(y|S_i, \pi_k(S_i)) + (1 - \lambda)\Big(\eta^\alpha(S_i, \pi_k(S_i)) + \frac{1}{1 - \alpha} \sum_{y \in \mathscr{X}} \big(v(y) - d(S_i, \pi_k(S_i)) - \eta^\alpha(S_i, \pi_k(S_i))\big)_+ P(y|S_i, \pi_k(S_i))\Big), \quad S_i \in \{S_1, \ldots, S_N\}$$

$$v(L) = 0$$

$$v(D) = 0$$

Note that the optimal $\eta^\alpha(S_i, \pi_k(S_i))$, denoted as $\bar{\eta}^\alpha(S_i, \pi_k(S_i))$, $S_i \in$

$\{S_1, \ldots, S_N\}$ provides the value at risk at level α. As in the first-order mean-semi-deviation case mentioned above, we use a linearized equivalent program. The linearized program can be given as follows:

$$\min_{v,z,\eta^\alpha} \sum_{x \in \mathscr{X}} v(x) \tag{5.10}$$

s.t.

$$v(S_i) \geq \lambda \sum_{y \in \mathscr{Z}} (v(y) - d(S_i, \pi_k(S_i)))P(y|S_i, \pi_k(S_i)) + (1-\lambda)\Big(\eta^\alpha(S_i, \pi_k(S_i)) + \frac{1}{1-\alpha} \sum_{y \in \mathscr{Z}} z(S_i, y)P(y|S_i, \pi_k(S_i))\Big),$$
$$S_i \in \{S_1, \ldots, S_N\} \tag{5.11}$$

$$z(S_i, y) \geq v(y) - d(S_i, \pi_k(S_i)) - \eta^\alpha(S_i, \pi_k(S_i)),$$
$$S_i \in \{S_1, \ldots, S_N\}; \quad y \in \mathscr{X} \tag{5.12}$$

$$z(S_i, y) \geq 0, \quad S_i \in \{S_1, \ldots, S_N\}; \quad y \in \mathscr{X} \tag{5.13}$$

$$v(L) = 0 \tag{5.14}$$

$$v(D) = 0 \tag{5.15}$$

Accordingly, we obtain Algorithm 5.6.

5.2.2 Policy Iteration with Newton's Nonsmooth Method

The dynamic programming equations for both first-order mean-semi-deviation (4.3), (4.4), (4.5) and mean-AVaR (4.6), (4.7), (4.8) risk mappings are nonsmooth formulations. To solve these systems of equations, we will use the specialized nonsmooth Newton method suggested in Cavus and Ruszczynski [9] for infinite horizon problems. The risk transition mappings (3.6) and (3.7) admit the dual representation (3.8). This method involves solving linear approximations of the dual problem and iteratively converging to a solution. At each policy evaluation step, the values associated with the current policy are approximated via Newton iterations. The stopping condition is similar to what we have in value iteration algorithm, as soon as the maximum difference between successive Newton iteration value functions is below a threshold ϵ, the Newton iterations stop. However,

Algorithm 5.6 Policy iteration algorithm using convex optimization with mean-AVaR as the one-step conditional risk measure.

procedure POLICY ITERATION WITH MEAN-AVAR (CO)(π_0)
  $k \leftarrow 0$
  $\pi_0(S_i) \leftarrow T, \quad S_i \in \{S_1, \ldots, S_N\}, \quad \pi_0(L) \leftarrow C, \quad \pi_0(D) \leftarrow C$ ▷ Any initial feasible policy is acceptable.
  do
    Policy Evaluation Step:
    solve (5.10) - (5.15) to find an optimal solution $\bar{v}$ and $\bar{\eta}^{\alpha}$
    $\eta_k^{\alpha}(S_i, u) \leftarrow \bar{\eta}^{\alpha}(S_i, u), \quad S_i \in \{S_1, \ldots, S_N\}; \quad u \in U(S_i)$
    $v_k(x) \leftarrow \bar{v}(x), \quad x \in \mathscr{X}$
    Policy Improvement Step:

$$\pi_{k+1}(S_i) \leftarrow \arg\min_{u \in U(S_i)} \Bigg\{ \lambda \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) + (1-\lambda)\Big(\eta_k^{\alpha}(S_i, u) + \frac{1}{1-\alpha}\Big(\sum_{y \in \mathscr{X}} (v_k(y) - d(S_i, u) - \eta_k^{\alpha}(S_i, u))_+ P(y|S_i, u)\Big)\Big)\Bigg\}, \quad S_i \in \{S_1, \ldots, S_N\}$$

    $k \leftarrow k + 1$
  while $\pi_k \not\equiv \pi_{k-1}$
  $v^*(x) \leftarrow v_{k-1}(x), \quad x \in \mathscr{X}$
  $\pi^*(x) \leftarrow \pi_k(x), \quad x \in \mathscr{X}$
  return v^*, π^*
end procedure

unlike in the value iteration algorithm, Newton iterations converge significantly faster. In our implementations we set $\epsilon = 10^{-9}$. Accordingly, using the definitions in (3.9) and (3.10) Algorithms 5.7 and 5.8 for first-order mean-semi-deviation and mean-AVaR are established, respectively. A more detailed description of Newton's nonsmooth method for infinite horizon problems can be found in Cavus and Ruszczynski [9].

Algorithm 5.7 Policy iteration algorithm with Newton's method under first-order mean-semi-deviation risk measure.

procedure POLICY ITERATION UNDER FIRST-ORDER MEAN-SEMI-DEVIATION (NM)(π_0)

$k \leftarrow 0$

$v_0(x) \leftarrow 0, \quad x \in \mathscr{X}$ ▷ Any initial value function is acceptable.

$\pi_0(S_i) \leftarrow T, \quad S_i \in \{S_1, \ldots, S_N\}, \quad \pi_0(L) \leftarrow C, \quad \pi_0(D) \leftarrow C$ ▷ Any initial feasible policy is acceptable.

do

Policy Evaluation Step:

$l \leftarrow 0$

$v^l(x) \leftarrow v_k(x), \quad x \in \mathscr{X}$

do

for $S_i \in \{S_1, \ldots, S_N\}$ **do**

solve the following for μ^*

$$\begin{aligned}
&\min_{\mu,h} -\sum_{y\in\mathscr{X}} (v^l(y) - d(S_i, \pi_k(S_i)))\mu(S_i, y) \\
&\textbf{s.t.} \\
&\mu(S_i, y) = P(y|S_i, \pi(S_i))\big(1 + h(S_i, y) - \sum_{\bar{y}\in\mathscr{X}} h(S_i, \bar{y})P(\bar{y}|S_i, \pi_k(S_i))\big), \quad \forall y \in \mathscr{X} \\
&\sum_{y\in\mathscr{X}} \mu(S_i, y) = 1 \\
&h(S_i, y) \leq \kappa, \quad \forall y \in \mathscr{X} \\
&\mu(S_i, y), h(S_i, y) \geq 0, \quad \forall y \in \mathscr{X}
\end{aligned}$$

end for

$\mu^l(S_i, y) \leftarrow \mu^*(S_i, y), \quad S_i \in \{S_1, \ldots, S_N\}; \quad y \in \mathscr{X}$

solve

$$\begin{aligned}
&v^l(S_i) = \sum_{y\in\mathscr{X}} (v^l(y) - d(S_i, \pi_k(S_i)))\mu^l(S_i, y), \quad S_i \in \{S_1, \ldots, S_N\} \\
&v^l(L) = 0 \\
&v^l(D) = 0
\end{aligned}$$

$l \leftarrow l + 1$

while $\max_{i=1}^N |v^l(S_i) - v^{l-1}(S_i)| > \epsilon$

$v_k(x) \leftarrow v^{l-1}(x), \quad x \in \mathscr{X}$

Policy Improvement Step:

$$\begin{aligned}
\pi_{k+1}(S_i) \leftarrow \arg \min_{u\in U(S_i)} &\underbrace{\sum_{y\in X} \Big(v_k(y) - d(S_i, u)\Big)P(y|S_i, u)}_{\mu} \\
&+ \kappa \sum_{y\in X} \Big(v_k(y) - d(S_i, u) - \mu\Big)_+ P(y|S_i, u), \quad S_i \in \{S_1, \ldots, S_N\}
\end{aligned}$$

$k \leftarrow k + 1$

while $\pi_k \not\equiv \pi_{k-1}$

$v^*(x) \leftarrow v_{k-1}(x), \quad x \in \mathscr{X}$

$\pi^*(x) \leftarrow \pi_k(x), \quad x \in \mathscr{X}$

return v^*, π^*

end procedure

Algorithm 5.8 Policy iteration algorithm with Newton's method under mean-AVaR risk measure.

procedure POLICY ITERATION UNDER MEAN-AVAR (NM)(π_0)

$k \leftarrow 0$

$v_0(x) \leftarrow 0, \quad x \in \mathscr{X}$ ▷ Any initial value function is acceptable.

$\pi_0(S_i) \leftarrow T, \quad S_i \in \{S_1, \ldots, S_N\}, \quad \pi_0(L) \leftarrow C, \quad \pi_0(D) \leftarrow C$ ▷ Any initial feasible policy is acceptable.

do

Policy Evaluation Step:

$l \leftarrow 0$

$v^l(x) \leftarrow v_k(x), \quad x \in \mathscr{X}$

do

for $S_i \in \{S_1, \ldots, S_N\}$ **do**

solve the following for μ^*

$$\begin{aligned}
&\min_{\mu} -\sum_{y \in \mathscr{X}} (v^l(y) - d(S_i, \pi_k(S_i)))\mu(S_i, y) \\
&\textbf{s.t.} \\
&\mu(S_i, y) = \lambda P(y|S_i, \pi_k(S_i)) + (1-\lambda) z(S_i, y), \quad \forall y \in \mathscr{X} \\
&z(S_i, y) \leq \frac{1}{(1-\alpha)} P(y|S_i, \pi_k(S_i)), \quad \forall y \in \mathscr{X} \\
&\sum_{y \in \mathscr{X}} \mu(S_i, y) = 1 \\
&\mu(S_i, y) \geq 0, \quad \forall y \in \mathscr{X}
\end{aligned}$$

end for

$\mu^l(S_i, y) \leftarrow \mu^*(S_i, y), \quad S_i \in \{S_1, \ldots, S_N\}; \quad y \in \mathscr{X}$

solve

$$\begin{aligned}
&v^l(S_i) = \sum_{y \in \mathscr{X}} (v^l(y) - d(S_i, \pi_k(S_i)))\mu^l(S_i, y), \forall S_i \in \{S_1, \ldots, S_N\} \\
&v^l(L) = 0 \\
&v^l(D) = 0
\end{aligned}$$

$l \leftarrow l + 1$

while $\max_{i=1}^{N} |v^l(S_i) - v^{l-1}(S_i)| > \epsilon$

$v_k(x) \leftarrow v^{l-1}(x), \quad x \in \mathscr{X}$

Policy Improvement Step:

$\eta_k^{\alpha}(S_i, u) \leftarrow \inf_{y \in \mathscr{X}} \Big\{ w \in \mathbb{R} | P(v_k(y) - d(S_i, u) \leq w) \geq \alpha \Big\}, \quad S_i \in \{S_1, \ldots, S_N\}$

$$\begin{aligned}
\pi_{k+1}(S_i) \leftarrow \arg \min_{u \in U(S_i)} \Bigg\{ & \lambda \sum_{y \in \mathscr{X}} \big(v_k(y) - d(S_i, u)\big) P(y|S_i, u) + (1-\lambda) \Big(\eta_k^{\alpha}(S_i, u) \\
& + \frac{1}{1-\alpha} \Big(\sum_{y \in \mathscr{X}} (v_k(y) - d(S_i, u) - \eta_k^{\alpha}(S_i, u))_+ P(y|S_i, u) \Big) \Big) \Bigg\}, \quad S_i \in \{S_1, \ldots, S_N\}
\end{aligned}$$

$k \leftarrow k + 1$

while $\pi_k \not\equiv \pi_{k-1}$

$v^*(x) \leftarrow v_{k-1}(x), \quad x \in \mathscr{X}$

$\pi^*(x) \leftarrow \pi_k(x), \quad x \in \mathscr{X}$

return v^*, π^*

end procedure

Chapter 6

Computational Study

In this chapter, we cover the numerical results of our model. As our intention was made clear before, we use clinical data to obtain the numerical results. Using these data, we compare the efficiency of the algorithms used in this study and then we provide the optimal policies under different measures of risk and different risk-aversion levels. These policies in our results, assume a control-limit type, i.e. the optimal policies recommend waiting up to a certain health state and transplantation for any worse health state. There is no exceptional case observed. We find a concrete relationship between the risk-preferences of the decision maker and the optimal policies. We also investigate the effect of other problem parameters on the optimal policies. We provide a sensitivity analysis to show that the optimal policies show differences for different transition probability matrices with different disease progression rates. Finally, we provide the results of a simulation model to quantify the efficiency of the mean-risk measures that we use.

6.1 Data

The model is implemented using clinically induced transition probability matrices and post-transplant daily survival distributions for different types of patients and

organs. As mentioned in Section 4.1, we use MELD score to assess the health state of a patient. Recall that MELD score varies between 6 and 40. In the model that is implemented, there are 18 health states (S_1, S_2, ..., S_{18}) each associated with a certain MELD score range — S_1 corresponds to MELD score 6-7, S_2 to MELD score 8-9, ... ,S_{17} to MELD score 38-39 and S_{18} to MELD score 40.

There are two sets of transition probability matrices associated with two disease groups: cirrhotic diseases (disease group 1) and hepatitis B (disease group 2). By its nature, hepatitis B is a more progressive disease. As previous studies make clear, the transition probabilities heavily depend on the type of the end-stage liver disease (see [5] and the references therein). The data used to estimate transition probability matrices are obtained from the Thomas E. Starzl Transplantation Institute within the University of Pittsburgh Medical Center (UPMC) and are not publicly available.

Organ Num.	Age	Sex	Blood Type	Race	Cytomegalovirus (CMVGR)
1	20	Male	A	White	Negative
2	30	Male	A	White	Negative
3	40	Male	A	White	Negative
4	50	Male	A	White	Negative
5	60	Male	A	White	Negative
6	70	Male	A	White	Negative

Table 6.1: Organs used in test problems.

Patient Num.	Age	Sex	Blood Type	CMVGR	Encephalopathy	Prior Transplant
1	20	Male	A	Positive	Yes	No
2	30	Male	A	Positive	Yes	No
3	40	Male	A	Positive	Yes	No

Table 6.2: Patient types in test problems.

We use post-transplant survival distribution data for three patients and six organ types, all differing by age (see Table 6.1 and Table 6.2). An organ obtained from an aged donor is expected to have lower quality. Likewise, an aged patient is likely to live less than a young patient after transplantation. Given a patient and

organ type, there is an associated post-transplant lifetime distribution for each health state. These distributions are obtained using publicly available UNOS post-transplant survival data and UPMC data. In the next section, we describe the problem parameter estimation procedures using these data.

6.1.1 Parameter Estimation

In this section we describe how the data mentioned in the previous section are utilized to estimate the parameters used in the risk-averse MDP model. These parameters are the health state transition probability matrices, post-transplant rewards, quality of life (QOL) scores for each health state before transplantation and QALY in the post-transplantation stage.

6.1.1.1 Transition Probability Matrices

We use transition probability matrices obtained using UPMC clinical data. The transition probability data has been also previously used in the dissertation by Alagoz [78]. These data are obtained from 3009 patients having end-stage liver disease between the years 1991 and 2000. The data consist of clinical information of the patients alongside demographical specifications. The clinical information consists knowledge such as type of the liver disease, laboratory test results, event of pre-transplant death. The medical condition of these patients have been measured throughout the years and occasional gaps in the data are estimated using cubic spline functions. More information on these cubic spline functions can be found in the study by Alagoz et al. [79]. The UPMC data are categorized with respect to end-stage liver disease etiology, i.e. disease type. As mentioned in the previous section, in our study, we use transition probability matrices for two disease types.

6.1.1.2 Post-Transplant Lifetime

We use the Cox proportional hazards model of Roberts et al. [80] to estimate the expected lifetime when a patient is transplanted with an organ. This model uses UPMC and UNOS data for estimating post-transplant rewards. We estimate the distributions of the survival rates of patients given that they are transplanted with a particular organ. For this purpose, we firstly run the post-transplant simulation model used in Shechter et al. [81] many times for thousands of patients and estimate the expected lifetime of each patient for a given organ. We then assume that the lifetime after transplantation follows exponential distribution with a mean that is equal to this expected lifetime calculated by the model used in Shechter et al. [81]. Exponential distribution is widely accepted for survival models [82], [83], [84]. As we use days as our time unit, we discretize the survival distribution in a daily basis. Then we use the exponential distribution to generate a distribution of the death times for patients after transplant. This post-transplant survival distribution is used in estimating the immediate transplant reward function $d(S_i, T)$, $S_i \in \{S_1, S_2, \ldots, S_{18}\}$.

6.1.1.3 Quality-Adjusted Lifetime

We estimate the quality of life scores of patients with end-stage liver disease (ESLD) associated with MELD score using the studies by Shaw et al. [85] and Younoissi et al. [86] and the UPMC data. Namely, the study by Younoissi et al. [86] reports the utilities and health-related quality of life in patients with chronic liver disease by surveying 353 patients with ESLD. They report the reductions in quality of life due to cholestatic liver disease, hepatitis, and cancer using two different quality of life (QOL) surveys: generic QOL survey called Short Form-36 (scaled between 0 and 100) and a disease-specific health-related quality of life instrument (Chronic Liver Disease Questionnaire (CLDQ)). Among these surveys, we used the reduction in CLDQ since it is specific to ESLD. We found that mean reduction in utility due to ESLD is 23% . Then using UPMC data, we estimated the average MELD score of patients that experienced these 3 diseases and found

that 23% reduction in QOL corresponds to the MELD score of 18. The study by Shaw et al. [85] reports the U.S. valuation of the EQ-5D health states, which is a multiattribute, preference-based health status measure. Using the estimates from Shaw et al. [85] we estimate that the QOL score associated with MELD score of 40 and over is 0.31. Then, assuming that the QOL score for MELD score 6 is equal to 1, we estimate MELD score-specific QOL scores that would generate the 23% reduction in QOL for MELD score of 18. That is, we interpolate the QOL scores for each MELD score between 6 and 40 to match this mean reduction. This calculation provides us the final estimates for QOL by MELD score.

QOL is a significant factor also in post-transplant stage. We synthesize the findings of two study in order to perform QOL adjustments in post-transplant rewards. According to Longworth and Bryan [87], EQ-5D increases from 0.517 to 0.608 after transplantation. This corresponds to a $(0.608 - 0.517)/0.517 = \%18$ increase in QOL. Pinston et al. [88] show that SF-36 gives the results 31 and 44 for pre-transplant patients. After transplantation this value is observed as 38 and 49. This implies an increase of $(38+49-31-44)/(31+44) = \%16$ in QOL. Using these two studies, we assume that the average increase in QOL is 17%. Another study by Ahmad et al. [89] find that the average MELD score at transplantation is 23, which has a QOL score of 0.66 according to our study. If we increase 0.66 by 17%, we obtain 0.77 and thus we use 0.77 for QALY after transplantation.

In the following sections, we discuss risk-neutral and risk-averse results obtained using these data. Risk-averse results are obtained for both first-order mean-semi-deviation and mean-AVaR risk measures.

6.2 Computational Results

In this section we interpret the computational results obtained using clinical data. We start by comparing the efficiency of the solution methods discussed in Chapter 5 in terms of CPU time and number of iterations. We then move onto the optimal policies obtained using different data sets and risk preferences. In all

data and parameter settings, the optimal policies were control-limit type policies and we have found that these control-limits are affected by disease, patient and organ type, usage of QALY; most convincingly of all, the risk preferences of the decision makers, i.e. patients or physicians. We also provide an analysis where we discuss the sensitivity of the control-limits to the transition probability matrix. In addition to these, we provide the results from a simulation of the liver-transplantation problem in order to be able to discuss the effectiveness of risk-averse optimal control-limits.

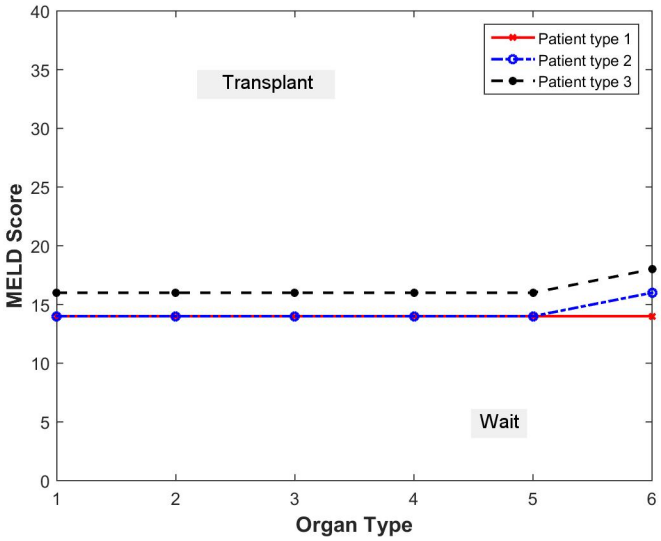
6.2.1 Comparison of Algorithms

One of the aims of our study is to compare, in a problem depending on actual data, the computational methods for risk-averse undiscounted MDPs suggested by Cavus and Ruszczynski [9]. In order to be able to make a quantitative comparison between the methods in consideration, we have obtained results under different risk measures and various risk preference settings. We have assumed that QALY is 1. We have recorded the CPU time and number of iterations for each of these settings for value iteration and policy iteration algorithms. The detailed performance measures are given in Table A.1 for the first-order mean-semi-deviation risk measure, and the comparison under mean-AVaR risk measure can be seen in Table A.2. The tables allow us to make definitive comparisons between the methods. It is immediately clear that the value iteration takes a lot of iterations to converge to a near-optimal solution within the tolerance range $\epsilon = 10^{-4}$, and therefore it is the slowest method except for the risk-neutral case (see A.1 for $\kappa = 0$). For risk-averse formulations, it can be observed that policy iteration method provides optimal results quickly and with very low variance in CPU time. It is also evident that policy iteration with convex optimization method is exceptionally fast for all risk-averse problems.

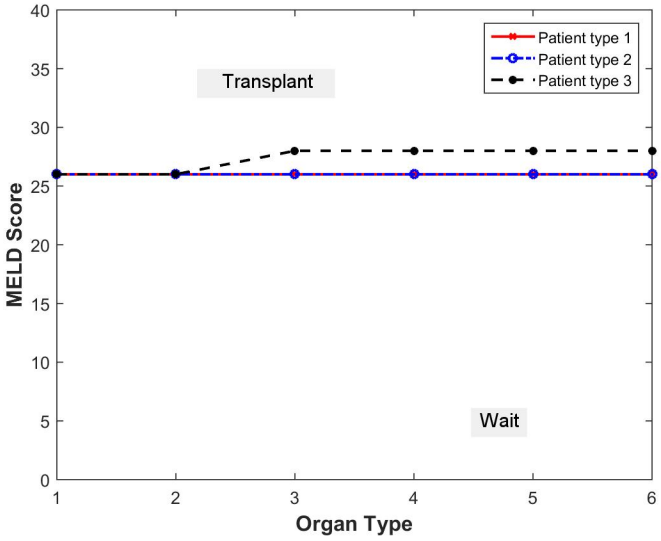
6.2.2 Risk-Neutral Optimal Policies

The results of the risk-neutral problem are given in Figure 6.1. Since end-stage liver diseases are characterized as progressive diseases, what we mean by a later stage of a disease is a sicker health state, i.e. a health state with a higher MELD score. In Figure 6.1, the first noticeable effect is the major shift in the control-limits as we look at the graphs for disease groups 1 and 2. The reason the control-limits for disease group 2 are significantly later than those in disease group 1 is the fact that the post-transplant rewards for disease group 1 are drastically higher than post-transplant rewards of disease group 2. The difference is so great that among all the patient-organ combinations, disease group 1 post-transplant rewards dominate disease group 2 post-transplant rewards without exception. We also observe that the difference is higher for healthier states. For disease group 2, it is possible to state the marginal benefit from transplantation in risk-neutral problem, compared with disease group 1, is overall low and the patient tends to wait until the later stages of the disease. It will be observed in following sections that in risk-averse cases, the effect of the disease is not as pronounced as in the risk-neutral case.

The other factors in the risk-neutral problem to be considered are the patient age and the organ quality. As organ type goes from 1 to 6, the quality of the organ decreases. Similarly, while patient type 1 is the youngest patient, patient 3 is the oldest. It can be seen in Figure 6.1, for a certain patient, as the organ quality decreases, the model favors transplanting at later stages. Another consistent observation in the result is that as the patient age increases, transplantation becomes optimal at later stages of the disease for the same organs. This is most likely due to the fact that the survival benefit obtained from transplantation decreases as the patient or donor age increases. Indeed, the observations on the post-transplant survival distributions of different patient types also confirm that the expected lifetime gained by transplantation decreases as patient age increases. Our risk-neutral results on disease type, patient age and organ quality are consistent with those in the risk neutral study offered by Alagoz et al. [5].



(a) Disease Group 1



(b) Disease Group 2

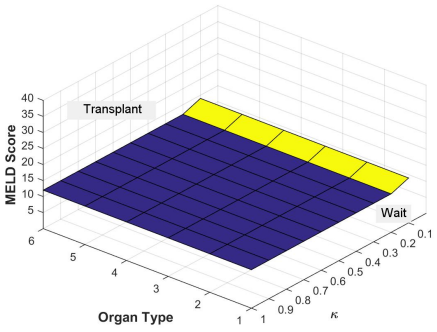
Figure 6.1: Risk-neutral optimal control limits without using QALY.

The risk-neutral results obtained using quality-adjusted lifetime are illustrated in Figure C.1. It can be observed that although the results do not significantly differ from the problem where quality-adjusted lifetime is not used, the transplantation decisions tend to shift to later stages of the disease for all organs, if they do shift at all. Since the control-limits move to later states of the disease, it is possible to state that the survival benefit from the transplantation decreases.

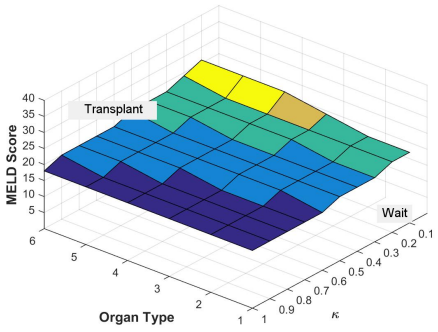
6.2.3 Optimal Control-Limits Obtained under First-Order Mean-Semi-Deviation Risk Measure

The most important finding in our computational results is the effect of risk preferences on transplantation decisions. In the first-order mean-semi-deviation risk measure, the risk-aversion is parametrized by $\kappa \in [0, 1]$ as mentioned in (3.1). Remember that a higher κ value indicates more risk-aversion. This parametrization of risk preferences gives us an opportunity to observe the effect of risk-aversion by changing the parameter κ along arbitrary increments in the interval $[0, 1]$. Figure 6.2 shows the optimal control-limits when first-order mean-semi-deviation is used as the risk measure. Comparing with the results obtained for the risk-neutral problem in Figure 6.1, it can be noticed that the more κ increases, the more the optimal control-limits deviate from the risk-neutral case. As it is observed in Figure 6.2, high κ values result in earlier transplants for each patient type and disease group. This observation mainly originates from the tendency of risk-averse patients to prefer a relatively certain reward, transplantation, over risk of death while waiting. The threshold of this preference shifts across different MELD scores as κ changes.

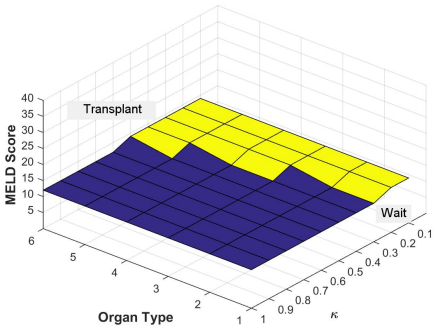
Similar to the risk-neutral results in Figure 6.1, the optimal control-limits under first-order mean-semi-deviation risk measure are also significantly sensitive to organ quality. As the organ quality increases, early transplantation becomes more likely. The post-transplant reward is lower for organs of low quality, and for risk-averse decision makers, this trend is kept. However, it is also clear in Figure 6.2 that an increasing degree of risk-aversion decreases the control-limits for all



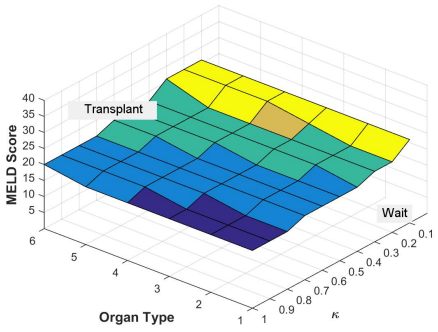
(a) Disease Group 1, Patient type 1



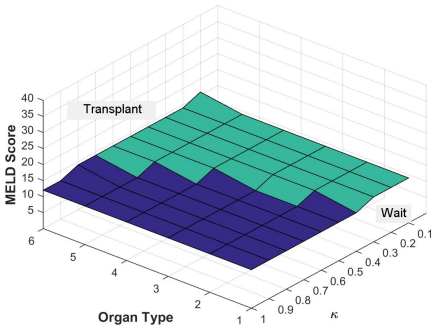
(b) Disease Group 2, Patient type 1



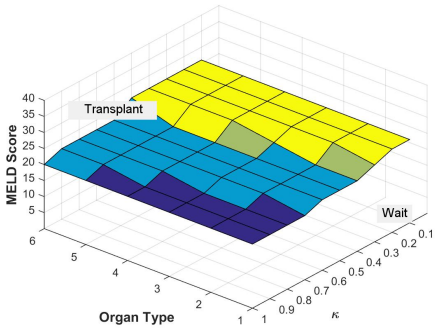
(c) Disease Group 1, Patient type 2



(d) Disease Group 2, Patient type 2



(e) Disease Group 1, Patient type 3



(f) Disease Group 2, Patient type 3

Figure 6.2: Optimal control limits under first-order mean-semi-deviation risk measure without using QALY.

organs while maintaining the result for organ quality.

The effect of disease type is also clear in results obtained under first-order mean-semi-deviation risk measure. From Figure 6.2, we observe that the magnitude of the shift in control-limits mentioned in the previous section becomes gradually less as we increase risk-aversion via κ. This is because risk-aversion leads to proportionally less risk-adjusted post-transplant rewards. It can also be observed that the optimal policies for disease group 2 (hepatitis) are considerably more risk sensitive than disease group 1 (cirrhotic diseases). This outcome can be justified as hepatitis is known to be a more progressive disease than cirrhosis. The effect of patient age on the optimal policies is similar to the findings for risk-neutral case due to the same possible reasons — increasing patient age results in later stage transplantations because survival benefit obtained from transplantation diminishes for aged patients.

The effect of quality-adjustment in lifetime is easily observable if one were to compare the Figures 6.2 and C.2. The usage of quality-adjusted lifetime makes incremental changes, if there are changes at all, in optimal control-limits for disease group 2. On the other hand, since disease group 1 offers significantly more risk-adjusted post-transplant lifetime which would suffer more from quality-adjustment, the control-limits for disease group 1 in Figure C.2 are noticeably higher than the ones in Figure 6.2. Another important observation is that usage of quality-adjusted lifetime renders the optimal control-limits slightly more sensitive to changing risk preferences.

6.2.4 Optimal Control-Limits Obtained under Mean-AVaR Risk Measure

The other risk measure that we use in our study is mean-AVaR. In mean-AVaR risk measure, the risk preferences are parametrized by $\lambda \in [0, 1]$ and $\alpha \in [0, 1)$, as mentioned in (3.2). It is necessary here to remind that as λ decreases and α increases, a decision maker, patient or physician, having a mean-AVaR objective

becomes more risk-averse. For $\lambda = 1$, we observe the risk-neutral case. Figures B.1- B.6 show the optimal control-limits under mean-AVaR risk measure with incrementally changing λ and α parameters for different patient and organ types. Consistently with our finding in the previous section, we observe that, for each patient and organ type couple, increased risk-aversion leads to earlier transplantations in optimal policies.

It is also easily noticeable that the effect of the organ quality is comparable to the findings in risk-neutral problem and first-order mean-semi-deviation cases. As observed throughout the Figures B.1 - B.6, for the same risk preferences and patient types, the transplantation decision occurs at higher MELD scores as the organ quality decreases. Another factor determining the control-limits is the patient age. As clearly seen in Figure B.7, increasing patient age results in later transplantations. This effect can also be observed in Figures B.1- B.6 for all organs. The reason that lower organ quality and higher patient age leads to later transplantations is because in such cases the marginal benefit from transplantation decreases, as explained in the previous sections.

As in risk-neutral problem or first-order mean-semi-deviation case, we observe the effect of disease type on optimal control-limits. Figures B.1 - B.6 demonstrate that optimal control-limits for disease group 2 are significantly higher than those for disease group 1 for all patient-organ couples and under all risk preferences. The difference gradually decreases as the risk-aversion increases. This result is very comparable to what we find under first-order mean-semi-deviation risk measure. Another similar result is that disease group 2 control-limits are more risk-sensitive than disease group 1 results, the reason of which is explained in the previous section.

Throughout the Figures C.3 - C.8, we provide the optimal control-limits obtained under mean-AVaR risk measure with quality-adjusted lifetime. The changes in optimal control-limit policies brought by introduction of quality-adjusted lifetime are consistent with our findings under first-order mean-semi-deviation. We observe that quality-of-life adjustment increases risk-sensitivity of disease group 1 patients for all organ types, and the optimal control-limits for

the disease group 2 remain less affected.

6.2.5 Sensitivity Analysis

The observation in the previous sections regarding the sensitivity of the optimal control-limits towards the disease type, that is, the transition probability matrix, has incurred a need to perform a sensitivity analysis with transition probability matrices whose certain attributes could be incrementally changed in order to see control-limit policy outcomes. Assume we smooth the disease group 1 transition probability matrix in such a manner that, for each health state, the probabilities of getting to a worse health state or getting to a better health state are aggregated into the probability of getting to the next or previous state, respectively. Such a transition probability matrix resembles that of a random-walk process across the health states. Based on this smoothed transition probability matrix, we have obtained four additional transition probability matrices. These matrices have increased or decreased probabilities of getting worse by 10% or 20%. Probabilities of staying on the same state are left unchanged. For a health state, let the getting worse probability be p. Then, the the new getting worse probabilities are set as $p' = 1 - (1 - p)^r$ where r can be 80%, 90%, 110% or 120%. Getting better probabilities are set so that the sum of the transition probabilities from the current state add up to 1. These additional matrices provide us a means to observe how control-limits change when the transition probability matrix is changed monotonically in terms disease aggressiveness. We have conducted the sensitivity analysis on patient type 2, across the six organs. In Figure D.1 we observe that increased probabilities of getting worse lead to earlier transplantations for all organ types. The same effect can also be observed in risk-averse problems, demonstrated in Figure D.2 under first-order mean-semi-deviation risk measure and in Figures D.3 - D.7 under mean-AVaR risk measure. The results indicate that a transition probability matrix with lower probabilities of getting worse, with the available data, increases risk sensitivity. This observation may be the result of the fact that in the pessimistic transition matrices, the exaggeration of the probabilities of getting worse is so vast that it reduces the uncertainty to

obtain distinct policies under different risk preferences.

6.2.6 Simulation

The Markovian nature of the liver-transplantation problem can be simulated effectively. Assume that a patient enters the model described in Section 4.1 at a health state $S_i \in \{S_1, S_2, \ldots, S_{18}\}$. Also assume that we have a predefined control-limit S_τ. The patient moves through the health states of the problem until either death or a health state S_x such that S_x is a sicker health state than S_τ. Transplantation is decided if a state sicker than S_τ is reached before the death of the patient. The transitions occur according to the transition probability matrix of the disease. After the transplantation has occurred, the post-transplantation stage is simulated using the post-transplant lifetime distributions. We implement this simulation model using disease group 2 transition probability matrix and patient 2 - organ 5 post-transplant data. We do not use quality-adjusted lifetime. The initial health state is set as S_1 and we obtain results for $\tau = 1, 2, \ldots, 18$. For each control-limit policy, we run a sample of 10^5 patients. For each control-limit, we record the sample mean and standard deviation of the simulated lifetime for these 10^5 patients. We also record the pre-transplant mortality rate among these 10^5 patients, which is the ratio of the patients who die before a transplantation to the sample size. The sample means and standard deviations of the total simulated lifetimes for each control-limit are as in Table 6.3.

According to the results described in Figure 6.1, under disease group 2 transition probability matrix, the risk-neutral control-limit for patient 2 and organ 5 is a MELD score of 26 which is represented by health state S_{11}. The simulation results in Table 6.3 confirm that the sample mean is indeed the highest for control-limit S_{11}. However, as observed in Table 6.3, control-limit S_{11} does not promise the lowest sample standard deviation. Sample standard deviation initially increases as the control-limit changes to sicker health states, then it becomes stagnant, and then it slightly decreases.

We expect a risk-averse policy to aim a compromise between high expected

Table 6.3: Sample mean and standard deviation of total lifetime across the control-limits for disease group 2, patient 2, organ 5, with pre-translant mortality rates.

S_τ	Sample Mean (μ_{S_τ})	Sample Std. Dev. (σ_{S_τ})	Mortality (θ_{S_τ})
S_1	1684.83	1688.74	0.0000
S_2	1838.99	1642.57	0.0000
S_3	2052.55	1643.40	0.0238
S_4	2246.48	1657.06	0.0524
S_5	2467.74	1702.66	0.0665
S_6	2635.57	1734.79	0.0793
S_7	2821.20	1821.31	0.0943
S_8	2949.78	1893.23	0.1153
S_9	3052.69	1965.96	0.1594
S_{10}	3131.66	2032.78	0.2122
S_{11}	3181.25	2097.16	0.2658
S_{12}	3175.90	2135.60	0.3492
S_{13}	3116.22	2140.81	0.4419
S_{14}	3032.64	2134.95	0.5580
S_{15}	2938.50	2101.81	0.6560
S_{16}	2858.90	2068.63	0.7336
S_{17}	2771.47	2018.21	0.7966
S_{18}	2708.26	1990.16	0.8741

lifetime and low standard deviation, as well as low mortality rate. Under first-order mean-semi-deviation risk measure, we determine optimal control-limits as indicated in Figure 6.2. Table 6.4 shows these control-limits alongside their corresponding sample means, standard deviations and pre-transplant mortality rates obtained from Table 6.3. In order to facilitate the comparison with the risk-neutral results, we also provide the percentage difference between the risk-averse results and risk-neutral results. Let μ_{S_τ} be the sample mean corresponding to the control limit S_τ, $\tau = 1, 2, \ldots, 18$ whose values are given in Table 6.3. For instance, as observed in Table 6.4, the sample mean for $\kappa = 1$ is μ_{S_7}. This sample mean is $\mu_{S_{11}}$ for the risk-neutral optimal control limit. Then the percentage difference of $\mu_{S_{11}}$ and μ_{S_7} equals $\frac{\mu_{S_{11}} - \mu_{S_7}}{\mu_{S_{11}}}(100)\%$. The percentage differences for sample standard deviation σ_{S_τ} and pre-transplant mortality rate θ_{S_τ}, $\tau = 1, 2, \ldots, 18$ are calculated similarly.

Under first-order mean-semi-deviation case, we show the optimal control-limits for each risk preference in Figure 6.2. We tabulate the sample means, standard deviations and pre-transplant mortality rates corresponding to the optimal control-limits of each risk-preference into Table 6.4. We notice in Table 6.4 that as risk-aversion increases, sample mean lifetime decreases up to 11.32% at $\kappa = 1$. However, this decrease in sample mean lifetime is less than the decrease in sample standard deviation for all risk preferences. We observe that for the most risk-averse case, reduction in the sample standard deviation is 13.15%. We also observe that as the risk-aversion increases, pre-transplant mortality rate decreases significantly. Table 6.4 shows that even slight risk-aversion, namely a control-limit of S_9 for $\kappa = 0.3$ is sufficient to decrease the pre-transplant mortality rate by 40.02% whereas only 6.26% of the mean lifetime is ceded. The reduction in pre-transplant mortality rate increases up to 64.54% for control limit S_7 at $\kappa = 1$ and mean lifetime is reduced by only 11.32%.

The optimal control-limits under mean-AVaR risk measure are also obtained and shown in Figure B.5. We build Table 6.5 with the same metrics as under first-order mean-semi-deviation, only the risk is parametrized using λ and α. Table 6.5 demonstrates that, as under first-order mean-semi-deviation, the optimal control-limits vary between S_7 and S_{11} based on risk preferences. As λ decreases

Table 6.4: Simulation results under first-order mean-semi-deviation risk measure for disease group 2, patient 2, organ 5.

κ	Sample Mean	% Diff.	Sample Std. Dev.	% Diff.	Mortality Rate	% Diff.	Opt. Control Limits
0.1	3181.25	0.00%	2097.16	0.00%	0.2658	0.00%	S_{11}
0.2	3181.25	0.00%	2097.16	0.00%	0.2658	0.00%	S_{11}
0.3	3052.69	4.04%	1965.96	6.26%	0.1594	40.02%	S_9
0.4	3052.69	4.04%	1965.96	6.26%	0.1594	40.02%	S_9
0.5	3052.69	4.04%	1965.96	6.26%	0.1594	40.02%	S_9
0.6	3052.69	4.04%	1965.96	6.26%	0.1594	40.02%	S_9
0.7	2949.78	7.28%	1893.23	9.72%	0.1153	56.63%	S_8
0.8	2949.78	7.28%	1893.23	9.72%	0.1153	56.63%	S_8
0.9	2949.78	7.28%	1893.23	9.72%	0.1153	56.63%	S_8
1	2821.2	11.32%	1821.31	13.15%	0.0943	64.54%	S_7

and α increases, the optimal control-limit moves to a healthier state and we see reasonable reductions in mean lifetime, alongside a greater reduction in standard deviation and a far greater reduction in pre-transplant mortality rate.

This simulation has provided us the means to quantitatively show the relationship between risk-aversion and the standard deviation and pre-transplantation mortality rate. The whole range of risk-preferences is captured by incrementally changing the risk parameters for both risk measures. In the Tables 6.4 and 6.5, it is observed that as the degree of risk-aversion increases, the standard deviation and pre-transplant mortality rate decrease. It can also be observed that the reduction in mortality rate as the risk-aversion increases is exceptionally high under both risk measures. It is possible to conclude that risk-averse optimal control-limits can reduce standard deviation and mortality rate while not making substantial cessions of expected lifetime.

Table 6.5: Simulation results under mean-AVaR risk measure for disease group 2, patient 2, organ 5.

Optimal Control-Limits									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7
0.2	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7
0.3	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7
0.4	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7	S_7
0.5	S_8	S_8	S_8	S_7	S_7	S_7	S_7	S_7	S_7
0.6	S_8	S_8	S_8	S_8	S_8	S_7	S_7	S_7	S_7
0.7	S_9	S_9	S_9	S_8	S_8	S_8	S_8	S_7	S_7
0.8	S_{10}	S_9	S_9	S_9	S_9	S_9	S_8	S_8	S_7
0.9	S_{11}	S_{11}	S_{11}	S_{11}	S_{11}	S_{10}	S_9	S_9	S_8

Sample Mean									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20
0.2	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20
0.3	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20
0.4	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20
0.5	2949.78	2949.78	2949.78	2821.20	2821.20	2821.20	2821.20	2821.20	2821.20
0.6	2949.78	2949.78	2949.78	2949.78	2949.78	2821.20	2821.20	2821.20	2821.20
0.7	3052.69	3052.69	3052.69	2949.78	2949.78	2949.78	2949.78	2821.20	2821.20
0.8	3131.66	3052.69	3052.69	3052.69	3052.69	3052.69	2949.78	2949.78	2821.20
0.9	3181.25	3181.25	3181.25	3181.25	3181.25	3131.66	3052.69	3052.69	2949.78

% Difference of the Sample Mean from the Values under Risk-Neutral Optimal Policies									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%
0.2	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.318%	11.32%	11.32%
0.3	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%
0.4	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%
0.5	7.276%	7.276%	7.276%	11.32%	11.32%	11.32%	11.32%	11.32%	11.32%
0.6	7.276%	7.276%	7.276%	7.276%	7.276%	11.32%	11.32%	11.32%	11.32%
0.7	4.041%	4.041%	4.041%	7.276%	7.276%	7.276%	7.276%	11.32%	11.32%
0.8	1.559%	4.041%	4.041%	4.041%	4.041%	4.041%	7.276%	7.276%	11.32%
0.9	0.000%	0.000%	0.000%	0.000%	0.000%	1.559%	4.041%	4.041%	7.276%

Sample Standard Deviation									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31
0.2	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31
0.3	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31
0.4	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31
0.5	1893.23	1893.23	1893.23	1821.31	1821.31	1821.31	1821.31	1821.31	1821.31
0.6	1893.23	1893.23	1893.23	1893.23	1893.23	1821.31	1821.31	1821.31	1821.31
0.7	1965.96	1965.96	1965.96	1893.23	1893.23	1893.23	1893.23	1821.31	1821.31
0.8	2032.78	1965.96	1965.96	1965.96	1965.96	1965.96	1893.23	1893.23	1821.31
0.9	2097.16	2097.16	2097.16	2097.16	2097.16	2032.78	1965.96	1965.96	1893.23

% of Reduction in Standard Deviation from the Values under Risk-Neutral Optimal Policies									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
0.2	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
0.3	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
0.4	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
0.5	9.72%	9.72%	9.72%	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
0.6	9.72%	9.72%	9.72%	9.72%	9.72%	13.15%	13.15%	13.15%	13.15%
0.7	6.26%	6.26%	6.26%	9.72%	9.72%	9.72%	9.72%	13.15%	13.15%
0.8	3.07%	6.26%	6.26%	6.26%	6.26%	6.26%	9.72%	9.72%	13.15%
0.9	0.00%	0.00%	0.00%	0.00%	0.00%	3.07%	6.26%	6.26%	9.72%

Mortality Rate									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943
0.2	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943
0.3	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943
0.4	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943
0.5	0.1153	0.1153	0.1153	0.0943	0.0943	0.0943	0.0943	0.0943	0.0943
0.6	0.1153	0.1153	0.1153	0.1153	0.1153	0.0943	0.0943	0.0943	0.0943
0.7	0.1594	0.1594	0.1594	0.1153	0.1153	0.1153	0.1153	0.0943	0.0943
0.8	0.2122	0.1594	0.1594	0.1594	0.1594	0.1594	0.1153	0.1153	0.0943
0.9	0.2658	0.2658	0.2658	0.2658	0.2658	0.2122	0.1594	0.1594	0.1153

% of Reduction in Mortality Rate from the Values under Risk-Neutral Optimal Policies									
λ \ α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%
0.2	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%
0.3	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%
0.4	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%
0.5	56.63%	56.63%	56.63%	64.54%	64.54%	64.54%	64.54%	64.54%	64.54%
0.6	56.63%	56.63%	56.63%	56.63%	56.63%	64.54%	64.54%	64.54%	64.54%
0.7	40.02%	40.02%	40.02%	56.63%	56.63%	56.63%	56.63%	64.54%	64.54%
0.8	20.18%	40.02%	40.02%	40.02%	40.02%	40.02%	56.63%	56.63%	64.54%
0.9	0.00%	0.00%	0.00%	0.00%	0.00%	20.18%	40.02%	40.02%	56.63%

Chapter 7

Conclusion

In this study, we approach the end-stage living-donor liver transplantation problem from a risk-averse perspective. The aim is to find the optimal transplantation policy such that the risk-adjusted lifetime is maximized. We model the problem as a discrete-time undiscounted infinite horizon transient MDP. We use MELD score information to model the transient states of the MDP model. Since pre-transplant stage of the problem ends in either transplantation or death, we use absorbing states representing post-transplant state and death.

We assume that decision makers, patients or physicians, pursue higher expected lifetime and also avoid risk. Therefore, risk-aversion is incorporated into this MDP model using the conditional mean-risk measures in the dynamic programming equations instead of conditional mean. The two types of mean-risk measures which we use are first-order mean-semi-deviation and mean-AVaR.

In order to be able to compute the optimal policies of this risk-averse MDP, we adapt risk-averse value iteration and policy iteration algorithms suggested by Cavus and Ruszczynski [9]. Within these algorithms, as both first-order mean-semi-deviation and mean-AVaR risk-transition mappings offer non-smooth formulations, for computational tractability, we use either linearized convex programs or Newton's non-smooth method, both proposed by Cavus an Ruszczynski [9].

We implement the risk-averse MDP using clinical data. Part of these data consists of health state transition probability matrices for two disease groups. The rest of the data is the post-transplant survival rates of different types of patients and organs where the patients and donated organs differ by age. We estimate problem parameters using these actual data. We also incorporate the effect of quality of life in pre-transplant and post-transplant stages of the problem.

The results of our implementation indicate that the optimal control-limits show great sensitivity to risk-preferences of the decision makers, as well as to the type of the end-stage liver disease, patient age and organ quality. We show that the optimal control-limits, under risk-neutrality, move to later stages of the disease as organ quality decreases or patient age increases. This result is observed also in risk-averse optimal control-limits. We also find that the optimal control-limits for disease group 2 (hepatitis B) are attained at significantly higher MELD scores than those for disease group 1 (cirrhotic diseases). When risk-aversion is incorporated using first-order mean-semi-deviation or mean-AVaR risk measures, we observe that increasing risk-aversion lead to earlier transplantation decisions. We also observe that risk-sensitivity is higher in disease group 2. When we incorporate quality of life years (QALY) into our computations, we observe that the optimal control-limits tend to shift to later stages of the disease, and this effect is more clearly observed in disease group 1. Our results also indicate that using QALY leads to more risk-sensitive optimal policies.

Our results confirm a significant relationship between the type of the end-stage liver disease and optimal control-limits. Therefore, we also conduct a sensitivity analysis to demonstrate the effect of the disease. We allow ourselves a health state transition probability matrix built using disease group 1 original matrix which resembles to a random-walk process. We adjust the getting worse probabilities in such a way that they are incrementally increased, and we obtain additional matrices which represent diseases with different progression rates. We find that the optimal control-limits show significant sensitivity to disease type and an aggressive disease with high getting worse probabilities leads to earlier transplantation decisions. We also observe that risk-sensitivity is higher in milder diseases with low getting worse probabilities.

In order to further validate our results, we also implemented a simulation model. We observed that, the risk-averse optimal control-limits under first-order mean-semi-deviation and mean-AVaR risk measures, although they offer slightly less expected lifetime, provide significant reductions in standard deviation and exceptionally lower pre-transplant mortality rates.

To our best knowledge, our study offers the first risk-averse MDP model using coherent risk measures in a clinical problem based on actual data. We utilize the parametrization of risk inherent to coherent risk-measures such as first-order mean-semi-deviation and mean-AVaR to conveniently represent the varying risk-preferences of the decision makers in a clinical environment. It is possible to use similar models in order to solve problems with Markovian nature which require sequential decision making while considering different degrees of risk-aversion.

Bibliography

[1] OPTN, "Transplants by donor type." Retrieved from `http://optn.transplant.hrsa.gov`, 2016. Accessed at 23th January, 2016.

[2] S. Cykert, "Risk acceptance and risk aversion: patients' perspectives on lung surgery," *Thoracic Surgery Clinics*, vol. 14, no. 3, pp. 287–293, 2004.

[3] S. Massin, B. Ventelou, A. Nebout, P. Verger, and C. Pulcini, "Cross-sectional survey: Risk-averse french general practitioners are more favorable toward influenza vaccination," *Vaccine*, vol. 33, no. 5, pp. 610–614, 2015.

[4] P. Wakker and D. Deneffe, "Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown," *Management Science*, vol. 42, no. 8, pp. 1131–1150, 1996.

[5] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts, "The optimal timing of living-donor liver transplantation," *Management Science*, vol. 50, no. 10, pp. 1420–1430, 2004.

[6] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts, "Determining the acceptance of cadaveric livers using an implicit model of the waiting list," *Operations Research*, vol. 55, no. 1, pp. 24–36, 2007.

[7] B. Sandikci, L. M. Maillart, A. J. Schaefer, O. Alagoz, and M. S. Roberts, "Estimating the patient's price of privacy in liver transplantation," *Operations Research*, vol. 56, no. 6, pp. 1393–1410, 2008.

[8] A. Ruszczynski, "Risk-averse dynamic programming for markov decision processes," *Mathematical Programming*, vol. 125, no. 2, pp. 235–261, 2010.

[9] O. Cavus and A. Ruszczynski, "Computational methods for risk-averse undiscounted transient Markov models," *Operations Research*, vol. 62, no. 2, pp. 401–417, 2014.

[10] O. Cavus and A. Ruszczynski, "Risk-averse control of undiscounted transient Markov models," *SIAM Journal on Control and Optimization*, vol. 52, no. 6, pp. 3935–3966, 2014.

[11] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[12] R. Wiesner, "Meld and peld: Application of survival models to liver allocation," *Liver Transplantation*, vol. 7, no. 7, pp. 567–580, 2001.

[13] D. E. Schaubel, M. K. Guidinger, S. W. Biggins, J. D. Kalbfleisch, E. A. Pomfret, P. Sharma, and R. M. Merion, "Survival benefit-based deceased-donor liver allocation," *American Journal of Transplantation*, vol. 9, no. 4p2, pp. 970–981, 2009.

[14] M. Akan, O. Alagoz, B. Ata, F. S. Erenay, and A. Said, "A broader view of designing the liver allocation system," *Operations Research*, vol. 60, no. 4, pp. 757–770, 2012.

[15] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts, "Choosing among living-donor and cadaveric livers," *Management Science*, vol. 53, no. 11, pp. 1702–1715, 2007.

[16] R. Righter, "A resource allocation problem in a random environment," *Operations Research*, vol. 37, no. 2, pp. 329–338, 1989.

[17] I. David, "A sequential assignment match process with general renewal arrival times," *Probability in the Engineering and Informational Sciences*, vol. 9, no. 03, p. 475, 1995.

[18] I. David and U. Yechiali, "A time-dependent stopping problem with application to live organ transplants," *Operations Research*, vol. 33, no. 3, pp. 491–504, 1985.

[19] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1947.

[20] G. W. Torrance, "Social preferences for health states: An empirical evaluation of three measurement techniques," *Socio-Economic Planning Sciences*, vol. 10, no. 3, pp. 129–136, 1976.

[21] G. W. Torrance, W. H. Thomas, and D. L. Sackett, "A utility maximization model for evaluation of health care programs," *Health Services Research*, vol. 7, no. 3, pp. 118–133, 1972.

[22] A. M. Stiggelbout, G. M. Kiebert, J. Kievit, J. W. Leer, G. Stoter, and J. C. De Haes, "Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores," *Medical Decision Making*, vol. 14, no. 1, pp. 82–90, 1994.

[23] J. S. Pliskin, D. S. Shepard, and M. C. Weinstein, "Utility functions for life years and health status," *Operations Research*, vol. 28, no. 1, pp. 206–224, 1980.

[24] E. Karni, "A theory of medical decision making under uncertainty," *Journal of Risk and Uncertainty*, vol. 39, no. 1, pp. 1–16, 2009.

[25] E. Karni, M. Leshno, and S. Rapaport, "Helping patients and physicians reach individualized medical decisions: theory and application to prenatal diagnostic testing," *Theory and Decision*, vol. 76, no. 4, pp. 451–467, 2013.

[26] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, p. 263, 1979.

[27] A. Tversky and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.

[28] H. Bleichrodt and J. L. Pinto, "A parameter-free elicitation of the probability weighting function in medical decision analysis," *Management Science*, vol. 46, no. 11, pp. 1485–1496, 2000.

[29] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts, *Operations Research and Health Care: A Handbook of Methods and Applications*, ch. Modeling Medical Treatment Using Markov Decision Processes, pp. 593–612. Springer US, 2004.

[30] C. Lefevre, "Optimal control of a birth and death epidemic process," *Operations Research*, vol. 29, no. 5, pp. 971–982, 1981.

[31] D. M. Faissol, P. M. Griffin, H. E. Kirkizlar, and J. L. Swann, "Timing of testing and treatment of hepatitis c and other diseases." 2006.

[32] R. Alterovitz, M. Branicky, and K. Goldberg, "Motion planning under uncertainty for image-guided medical needle steering," *The International Journal of Robotics Research*, vol. 27, no. 11-12, pp. 1361–1374, 2008.

[33] S. Kumar and N. Ghildayal, "Markov decision process for comparative economic analysis of immunization strategies against hepatitis a virus," *IEEE Engineering Management Review*, vol. 44, no. 1, pp. 33–46, 2016.

[34] M. J. Barry, "Watchful waiting vs immediate transurethral resection for symptomatic prostatism. the importance of patients' preferences," *JAMA: The Journal of the American Medical Association*, vol. 259, no. 20, pp. 3010–3017, 1988.

[35] J.-H. Ahn and J. C. Hornberger, "Involving patients in the cadaveric kidney transplant allocation process: A decision-theoretic perspective," *Management Science*, vol. 42, no. 5, pp. 629–641, 1996.

[36] S. M. Shechter, M. D. Bailey, A. J. Schaefer, and M. S. Roberts, "The optimal time to initiate hiv therapy under ordered health states," *Operations Research*, vol. 56, no. 1, pp. 20–33, 2008.

[37] J. Chhatwal, O. Alagoz, and E. S. Burnside, "Optimal breast biopsy decision-making based on mammographic features and demographic factors," *Operations Research*, vol. 58, no. 6, pp. 1577–1591, 2010.

[38] B. T. Denton, M. Kurt, N. D. Shah, S. C. Bryant, and S. A. Smith, "Optimizing the start time of statin therapy for patients with diabetes," *Medical Decision Making*, vol. 29, no. 3, pp. 351–367, 2009.

[39] M. Kurt, B. T. Denton, A. J. Schaefer, N. D. Shah, and S. A. Smith, "At what lipid ratios should a patient with type 2 diabetes initiate statins?," *IIE Transactions on Healthcare Engineering*, vol. 1, no. 1, pp. 49–65, 2009.

[40] K. J. Van Arendonk, E. K. H. Chow, N. T. James, B. J. Orandi, T. A. Ellison, J. M. Smith, P. M. Colombani, , and D. L. Segev, "Choosing the order of deceased donor and living donor kidney transplantation in pediatric recipients," *Transplantation*, vol. 99, no. 2, pp. 360–366, 2015.

[41] C. Hu, W. S. Lovejoy, and S. L. Shafer, "Comparison of some suboptimal control policies in medical drug therapy," *Operations Research*, vol. 44, no. 5, pp. 696–709, 1996.

[42] M. Hauskrecht and H. Fraser, "Planning treatment of ischemic heart disease with partially observable markov decision processes," *Artificial Intelligence in Medicine*, vol. 18, pp. 221–244, 2000.

[43] L. M. Maillart, J. S. Ivy, S. Ransom, and K. Diehl, "Assessing dynamic breast cancer screening policies," *Operations Research*, vol. 56, no. 6, pp. 1411–1427, 2008.

[44] J. E. Kreke, M. D. Bailey, A. J. Schaefer, D. C. Angus, and M. S. Roberts, "Modeling hospital discharge policies for patients with pneumonia-related sepsis," *IIE Transactions*, vol. 40, no. 9, pp. 853–860, 2008.

[45] T. Ayer, O. Alagoz, and N. K. Stout, "OR forum—a POMDP approach to personalize mammography screening decisions," *Operations Research*, vol. 60, no. 5, pp. 1019–1034, 2012.

[46] C. C. Bennett and K. Hauser, "Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach," *Artificial Intelligence in Medicine*, vol. 57, no. 1, pp. 9–19, 2013.

[47] D. J. Cher, J. Miyamoto, and L. A. Lenert, "Incorporating risk attitude into markov-process decision models: Importance for individual decision making," *Medical Decision Making*, vol. 17, pp. 340–350, 1997.

[48] V. Tilson and D. A. Tilson, "Use of a Markov decision process model for treatment selection in an asymptomatic disease with consideration of risk sensitivity," *Socio-Economic Planning Sciences*, vol. 47, no. 3, pp. 172–182, 2013.

[49] R. A. Howard and J. E. Matheson, "Risk-sensitive markov decision processes," *Management Science*, vol. 18, no. 7, pp. 356–369, 1972.

[50] S. J. Jaquette, "A utility criterion for undiscounted markov decision processes," *Management Science*, vol. 23, no. 1, 1976.

[51] T. Bielecki, D. Hernandez-Hernandez, and R. S. Pliska, "Risk sensitive control of finite state markov chains in discrete time, with applications to portfolio management," *Mathematical Methods of Operations Research*, vol. 50, no. 2, pp. 167–188, 1999.

[52] R. Cavazos-Cadena and E. Fernandez-Gaucherand, "Controlled markov chains with risk-sensitive criteria: Average cost, optimality equations, and optimal solutions," *Mathematical Methods of Operations Research*, vol. 49, no. 2, pp. 299–324, 1999.

[53] S. P. Coraluppi and S. I. Marcus, "Risk-sensitive and minimax control of discrete-time, finite-state markov decision processes," *Automatica*, vol. 35, no. 2, pp. 301–309, 1999.

[54] S. D. Patek, "On terminating markov decision processes with a risk-averse objective function," *Automatica*, vol. 37, no. 9, pp. 1379–1386, 2001.

[55] V. S. Borkar and S. P. Meyn, "Risk-sensitive optimal control for Markov decision processes with monotone cost," *Mathematics of Operations Research*, vol. 27, no. 1, pp. 192–209, 2002.

[56] R. Cavazos-Cadena and R. M. de Oca, "The value iteration algorithm in risk-sensitive average Markov decision chains with finite state space," *Mathematics of Operations Research*, vol. 28, no. 4, pp. 752–776, 2003.

[57] N. Bauerle and U. Rieder, "More risk-sensitive Markov decision processes," *Mathematics of Operations Research*, vol. 39, no. 1, pp. 105–120, 2014.

[58] G. B. Di Masi and L. Stettner, "Infinite horizon risk sensitive control of discrete time Markov processes with small risk," *Systems & Control Letters*, vol. 40, no. 1, pp. 15–20, 2000.

[59] Y. Ohtsubo, "Minimizing risk models in stochastic shortest path problems," *Mathematical Methods of Operations Research (ZOR)*, vol. 57, no. 1, pp. 79–88, 2003.

[60] Y. Ohtsubo, "Optimal threshold probability in undiscounted markov decision processes with a target set," *Applied Mathematics and Computation*, vol. 149, no. 2, pp. 519–532, 2004.

[61] C. Wu and Y. Lin, "Minimizing risk models in markov decision processes with policies depending on target values," *Journal of Mathematical Analysis and Applications*, vol. 231, no. 1, pp. 47–67, 1999.

[62] K. Boda, J. Filar, Y. Lin, and L. Spanjers, "Stochastic target hitting time and the problem of early retirement," *IEEE Transactions on Automatic Control*, vol. 49, no. 3, pp. 409–419, 2004.

[63] J. A. Filar, L. C. M. Kallenberg, and H.-M. Lee, "Variance-penalized markov decision processes," *Mathematics of Operations Research*, vol. 14, no. 1, pp. 147–161, 1989.

[64] N. Bauerle and A. Mundt, "Dynamic mean-risk optimization in a binomial model," *Mathematical Methods of Operations Research*, vol. 70, no. 2, pp. 219–239, 2008.

[65] W. B. Haskell and R. Jain, "A convex analytic approach to risk-aware Markov decision processes," *SIAM Journal on Control and Optimization*, vol. 53, no. 3, pp. 1569–1598, 2015.

[66] N. Bauerle and J. Ott, "Markov decision processes with average-value-at-risk criteria," *Mathematical Methods of Operations Research*, vol. 74, no. 3, pp. 361–379, 2011.

[67] Y. Chow, A. Tamar, S. Mannor, and M. Pavone, "Risk-sensitive and robust decision-making: a cvar optimization approach," in *NIPS*, 2015.

[68] A. Ruszczynski and A. Shapiro, "Conditional risk mappings," *Mathematics of Operations Research*, vol. 31, no. 3, pp. 544–561, 2006.

[69] W. Ogryczak and A. Ruszczynski, "From stochastic dominance to mean-risk models: Semideviations as risk measures," *European Journal of Operational Research*, vol. 116, no. 1, pp. 33–50, 1999.

[70] W. Ogryczak and A. Ruszczynski, "On consistency of stochastic dominance and mean-semideviation models," *Mathematical Programming*, vol. 89, no. 2, pp. 217–232, 2001.

[71] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Multistage Risk Averse Optimization*, ch. 6.7, pp. 308–328. Society for Industrial & Applied Mathematics (SIAM), 2009.

[72] W. Ogryczak and A. Ruszczynski, "Dual stochastic dominance and related mean-risk models," *SIAM J. Optim.*, vol. 13, no. 1, pp. 60–78, 2002.

[73] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[74] A. Ruszczynski and A. Shapiro, "Optimization of convex risk functions," *Mathematics of Operations Research*, vol. 31, no. 3, pp. 433–452, 2006.

[75] P. Kamath, "A model to predict survival in patients with end-stage liver disease," *Hepatology*, vol. 33, no. 2, pp. 464–470, 2001.

[76] R. E. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[77] R. A. Howard, *Dynamic Programming and Markov Processes*. John Wiley & Sons, 1960.

[78] O. Alagoz, *Optimal Policies for the Acceptance of Living- and Cadaveric-Donor Livers*. PhD thesis, University of Pittsburgh, 2004.

[79] O. Alagoz, C. L. Bryce, S. Shechter, A. Schaefer, C. C. Chang, D. C. Angus, and M. S. Roberts, "Incorporating biological natural history in simulation models: Empirical estimates of the progression of end-stage liver disease," *Medical Decision Making*, vol. 25, no. 6, pp. 620–632, 2005.

[80] M. S. Roberts, D. C. Angus, C. L. Bryce, Z. Valenta, and L. Weissfeld, "Survival after liver transplantation in the united states: A disease-specific analysis of the unos database," *Liver Transplantation*, vol. 10, no. 7, pp. 886–897, 2004.

[81] S. M. Shechter, C. L. Bryce, O. Alagoz, J. E. Kreke, J. E. Stahl, A. J. Schaefer, D. Angus, and M. Roberts, "A clinically based discrete-event simulation of end-stage liver disease and the organ allocation process," *Medical Decision Making*, vol. 25, no. 2, pp. 199–209, 2005.

[82] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc., 2nd ed., 2003.

[83] J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag New York, Inc., 2nd ed., 2003.

[84] D. Machin, Y. B. Cheung, and M. K. B. Parmar, *Survival Analysis: A Practical Approach*. John Wiley & Sons Ltd, 2nd ed., 2006.

[85] J. W. Shaw, J. A. Johnson, and S. J. Coons, "Us valuation of the eq-5d health states: development and testing of the d1 valuation model," *Medical Care*, vol. 43, no. 3, pp. 203–220, 2005.

[86] Z. M. Younossi, N. Boparai, L. L. Price, M. L. Kiwi, M. McCormick, and G. Guyatt, "Health-related quality of life in chronic liver disease: the impact of type and severity of disease," *The American Journal of Gastroenterology*, vol. 96, no. 7, pp. 2199–2205, 2001.

[87] L. Longworth and S. Bryan, "An empirical comparison of eq-5d and sf-6d in liver transplant patients," *Health Economics*, vol. 12, no. 12, pp. 1061–1067, 2003.

[88] C. W. Pinston, I. D. Feurer, J. L. Payne, P. E. Wise, S. Shockley, and T. Speroff, "Health-related quality of life after different types of solid organ transplantation," *Annals of Surgery*, vol. 232, no. 4, p. 597, 2000.

[89] J. Ahmad, K. K. Downey, M. Akoad, and T. V. Cacciarelli, "Impact of the meld score on waiting time and disease severity in liver transplantation in united states veterans," *Liver Transplantation*, vol. 13, no. 11, pp. 1564–1569, 2007.

Appendix A

Comparison of Algorithms

	Value iter. ($\epsilon = 10^{-4}$)		Policy iter. w/ Newton's method ($\epsilon = 10^{-9}$)			Policy iter. w/ convex opt.	
κ	# of value iter.	CPU time (sec)	# of policy iter.	# of Newton iter.	CPU time (sec)	# of policy iter.	CPU time (sec)
0	7341	0.469	4	2, 2, 2, 2	0.781	4	0.016
0.1	6276	3.094	4	3, 2, 2, 2	0.187	4	0.109
0.2	5448	2.656	4	4, 2, 2, 2	0.187	4	0.062
0.3	4781	2.234	4	4, 2, 2, 2	0.172	4	0.047
0.4	4206	2.156	4	5, 2, 2, 2	0.375	4	0.047
0.5	3449	1.578	3	5, 2, 2	0.187	3	0.047
0.6	3168	1.453	3	5, 2, 2	0.187	3	0.047
0.7	2928	1.531	3	5, 2, 2	0.187	3	0.031
0.8	2720	1.266	3	5, 2, 2	0.187	3	0.047
0.9	2538	1.391	3	5, 2, 2	0.187	3	0.031
1	2378	1.078	3	5, 2, 2	0.187	3	0.031

Table A.1: Performance comparisons of methods under first-order mean-semi-deviation risk measure.

Table A.2: Performance comparisons of methods for certain parameters under mean-AVaR risk measure.

Parameters		Value Iter. ($\epsilon = 10^{-4}$)		Policy Iter. w/ Newton's method ($\epsilon = 10^{-9}$)			Policy Iter. w/ convex opt. method	
λ	α	# of value iter.	CPU time (sec)	# of policy iter.	# of Newton iter.	CPU time (sec)	# of policy iter.	CPU time (sec)
0.1	0.1	3375	10.328	3	3, 2, 2	0.641	3	0.047
0.1	0.2	2997	9.141	3	3, 2, 2	0.641	3	0.047
0.1	0.3	2621	8.156	3	3, 2, 2	0.641	3	0.047
0.1	0.4	2247	7.188	3	3, 2, 2	0.641	3	0.063
0.1	0.5	1874	5.922	3	3, 2, 2	0.641	3	0.047
0.1	0.6	1501	4.734	3	3, 2, 2	0.719	3	0.063
0.1	0.7	1127	3.422	3	4, 2, 2	0.734	3	0.063
0.1	0.8	751	2.406	3	4, 2, 2	0.734	3	0.047
0.1	0.9	375	1.141	3	5, 2, 2	0.828	3	0.109
0.2	0.1	3648	11.156	3	3, 2, 2	0.656	3	0.047
0.2	0.2	3261	10.266	3	3, 2, 2	0.641	3	0.047
0.2	0.3	2870	8.938	3	3, 2, 2	0.625	3	0.047
0.2	0.4	2474	7.953	3	3, 2, 2	0.625	3	0.063
0.2	0.5	2073	6.406	3	3, 2, 2	0.625	3	0.063
0.2	0.6	1668	5.281	3	3, 2, 2	0.641	3	0.063
0.2	0.7	1258	3.969	3	4, 2, 2	0.734	3	0.063
0.2	0.8	842	2.578	3	4, 2, 2	0.734	3	0.047
0.2	0.9	421	1.328	3	4, 2, 2	0.734	3	0.063
0.3	0.1	3921	12.016	3	3, 2, 2	0.625	3	0.031
0.3	0.2	3536	10.859	3	3, 2, 2	0.656	3	0.063
0.3	0.3	3137	10.25	3	3, 2, 2	0.641	3	0.063
0.3	0.4	2727	8.328	3	3, 2, 2	0.641	3	0.063
0.3	0.5	2304	7.063	3	3, 2, 2	0.641	3	0.063
0.3	0.6	1868	5.813	3	3, 2, 2	0.641	3	0.047
0.3	0.7	1418	4.406	3	4, 2, 2	0.734	3	0.063

Table A.2: Performance comparisons of methods for certain parameters under mean-AVaR risk measure.

Parameters		Value Iter. ($\epsilon = 10^{-4}$)		Policy Iter. w/ Newton's method ($\epsilon = 10^{-9}$)			Policy Iter. w/ convex opt. method	
λ	α	# of value iter.	CPU time (sec)	# of policy iter.	# of Newton iter.	CPU time (sec)	# of policy iter.	CPU time (sec)
0.3	0.8	955	2.922	3	3, 2, 2	0.656	3	0.047
0.3	0.9	481	1.500	3	5, 2, 2	0.828	3	0.063
0.4	0.1	4197	13.297	3	3, 2, 2	0.641	3	0.063
0.4	0.2	3824	12.016	3	3, 2, 2	0.625	3	0.063
0.4	0.3	3429	10.500	3	3, 2, 2	0.641	3	0.047
0.4	0.4	3013	9.219	3	3, 2, 2	0.641	3	0.063
0.4	0.5	2573	7.859	3	3, 2, 2	0.625	3	0.063
0.4	0.6	2108	6.641	3	3, 2, 2	0.656	3	0.063
0.4	0.7	1618	4.938	3	4, 2, 2	0.734	3	0.063
0.4	0.8	1101	3.484	3	3, 2, 2	0.641	3	0.063
0.4	0.9	559	1.703	3	6, 2, 2	0.938	3	0.047
0.5	0.1	4477	14.156	3	3, 2, 2	0.656	3	0.047
0.5	0.2	4128	12.609	3	3, 2, 2	0.703	3	0.047
0.5	0.3	3750	11.906	3	3, 2, 2	0.641	3	0.047
0.5	0.4	3339	10.516	3	3, 2, 2	0.641	3	0.097
0.5	0.5	2891	9.156	3	3, 2, 2	0.625	3	0.063
0.5	0.6	2404	7.578	3	3, 2, 2	0.641	3	0.047
0.5	0.7	1873	6.234	3	4, 2, 2	0.750	3	0.047
0.5	0.8	1294	3.984	3	4, 2, 2	0.734	3	0.063
0.5	0.9	666	2.203	3	4, 2, 2	0.734	3	0.063
0.6	0.1	4765	14.828	3	3, 2, 2	0.656	3	0.063
0.6	0.2	4453	13.688	3	3, 2, 2	0.656	3	0.063
0.6	0.3	4105	12.531	3	3, 2, 2	0.641	3	0.031
0.6	0.4	3714	11.313	3	3, 2, 2	0.641	3	0.047
0.6	0.5	3274	10.078	3	3, 2, 2	0.641	3	0.047

Table A.2: Performance comparisons of methods for certain parameters under mean-AVaR risk measure.

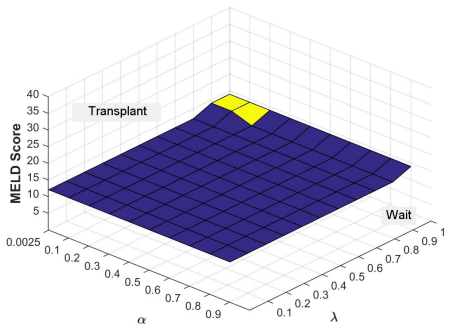
Parameters		Value Iter. ($\epsilon = 10^{-4}$)		Policy Iter. w/ Newton's method ($\epsilon = 10^{-9}$)			Policy Iter. w/ convex opt. method	
λ	α	# of value iter.	CPU time (sec)	# of policy iter.	# of Newton iter.	CPU time (sec)	# of policy iter.	CPU time (sec)
0.6	0.6	2776	8.469	3	3, 2, 2	0.641	3	0.063
0.6	0.7	2208	6.703	3	4, 2, 2	0.781	3	0.047
0.6	0.8	1560	4.750	3	3, 2, 2	0.641	3	0.063
0.6	0.9	821	2.578	3	4, 2, 2	0.734	3	0.063
0.7	0.1	5628	17.922	4	3, 2, 2, 2	0.813	4	0.063
0.7	0.2	5140	16.25	4	3, 2, 2, 2	0.844	4	0.047
0.7	0.3	4500	13.891	3	3, 2, 2	0,656	3	0.047
0.7	0.4	4151	12.656	3	3, 2, 2	0.656	3	0.047
0.7	0.5	3741	11.359	3	3, 2, 2	0.656	3	0.047
0.7	0.6	3253	9.922	3	3, 2, 2	0.672	3	0.063
0.7	0.7	2666	8.453	3	4, 2, 2	0.766	3	0.063
0.7	0.8	1949	5.922	3	3, 2, 2	0.656	3	0.063
0.7	0.9	1065	3.359	3	4, 2, 2	0.781	3	0.063
0.8	0.1	6206	18.969	4	3, 2, 2, 2	0.859	4	0.078
0.8	0.2	5941	18.188	4	3, 2, 2, 2	0.984	4	0.078
0.8	0.3	5622	17.375	4	3, 2, 2, 2	0.813	4	0.078
0.8	0.4	5226	15.875	4	3, 2, 2, 2	0.828	4	0.063
0.8	0.5	4321	13.344	3	3, 2, 2	0,656	3	0.094
0.8	0.6	3886	12.031	3	3, 2, 2	0.641	3	0.063
0.8	0.7	3322	10.156	3	3, 2, 2	0.641	3	0.063
0.8	0.8	2563	7.875	3	3, 2, 2	0.656	3	0.047
0.8	0.9	1500	4.609	3	4, 2, 2	0.75	3	0.047
0.9	0.1	6763	20.719	4	3, 2, 2, 2	0.844	4	0.094
0.9	0.2	6614	20.313	4	3, 2, 2, 2	0.828	4	0.063
0.9	0.3	6430	20.016	4	3, 2, 2, 2	0.813	4	0.078

Table A.2: Performance comparisons of methods for certain parameters under mean-AVaR risk measure.

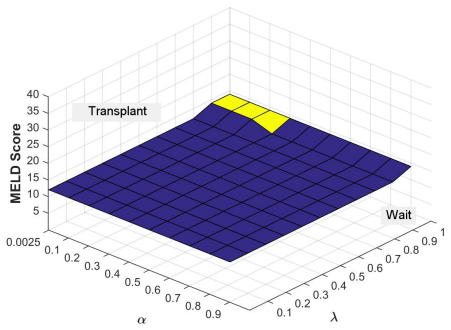
Parameters		Value Iter. ($\epsilon = 10^{-4}$)		Policy Iter. w/ Newton's method ($\epsilon = 10^{-9}$)			Policy Iter. w/ convex opt. method	
λ	α	# of value iter.	CPU time (sec)	# of policy iter.	# of Newton iter.	CPU time (sec)	# of policy iter.	CPU time (sec)
0.9	0.4	6198	19.125	4	3, 2, 2, 2	0.844	4	0.063
0.9	0.5	5895	18.063	4	3, 2, 2, 2	0.828	4	0.078
0.9	0.6	5483	16.781	4	3, 2, 2, 2	0.813	4	0.063
0.9	0.7	4879	15.422	4	4, 2, 2, 2	0.906	4	0.063
0.9	0.8	3653	11.141	3	4, 2, 2	0.766	3	0.063
0.9	0.9	2466	8.141	3	4, 2, 2	0.734	3	0.078

Appendix B

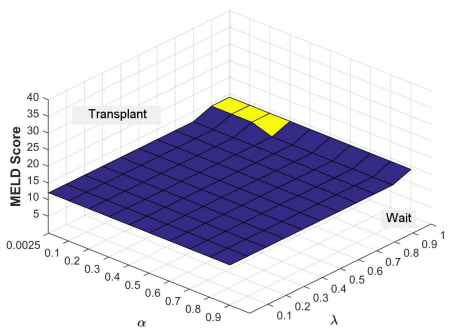
Optimal Control-Limits Obtained under Mean-AVaR Risk Measure without Using QALY



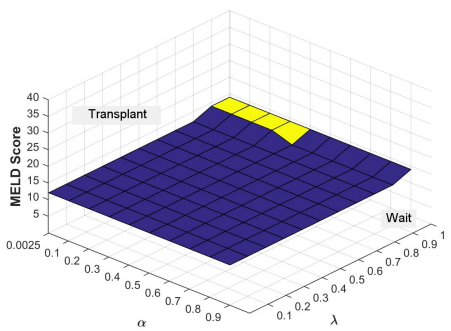
(a) Organ 1



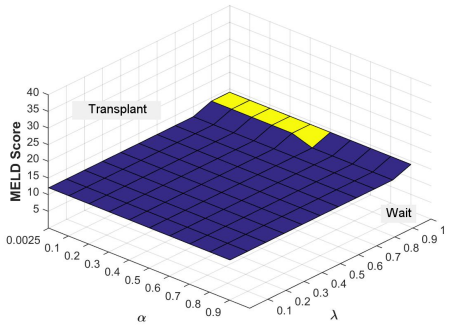
(b) Organ 2



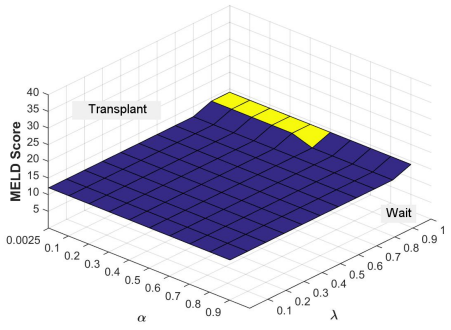
(c) Organ 3



(d) Organ 4

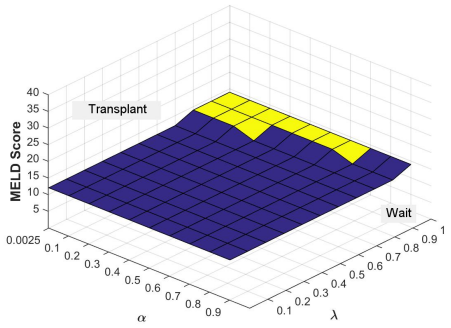


(e) Organ 5

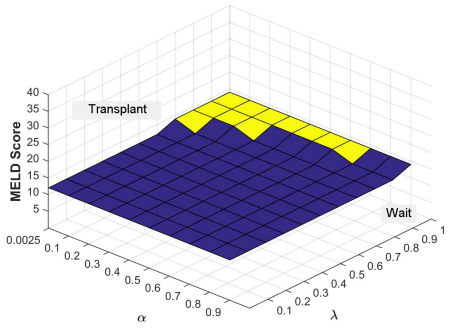


(f) Organ 6

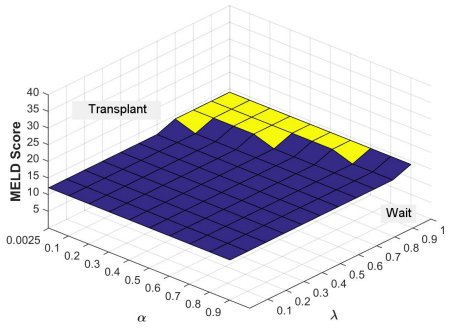
Figure B.1: Optimal control limits for disease group 1, patient type 1 under mean-AVaR risk measure.



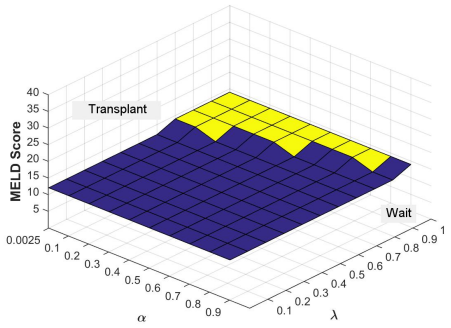
(a) Organ 1



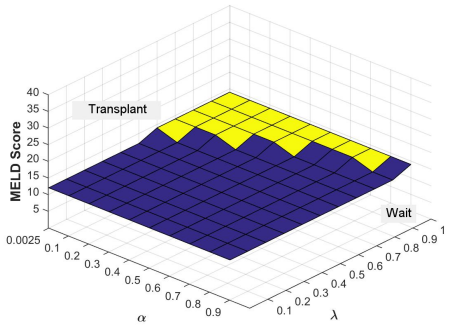
(b) Organ 2



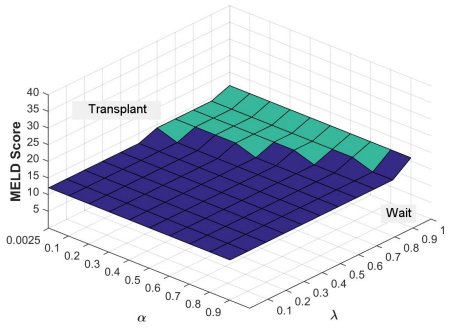
(c) Organ 3



(d) Organ 4

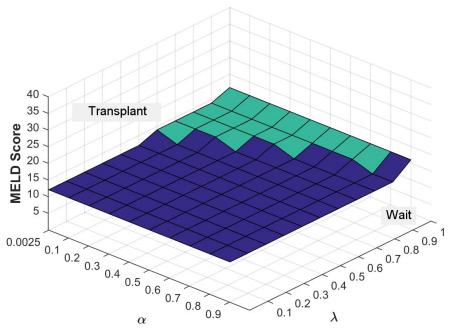


(e) Organ 5

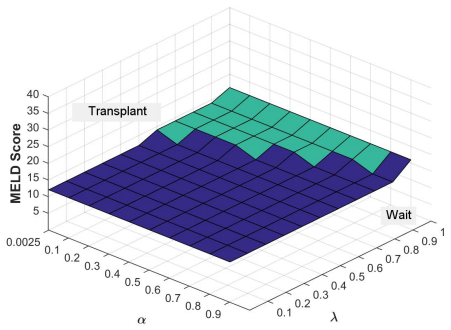


(f) Organ 6

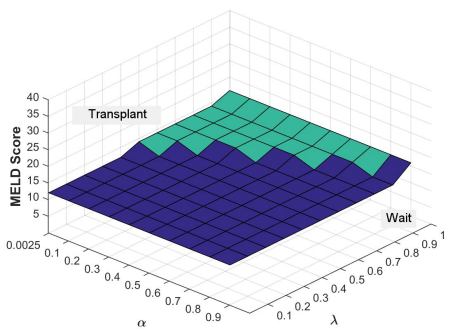
Figure B.2: Optimal control limits for disease group 1, patient type 2 under mean-AVaR risk measure.



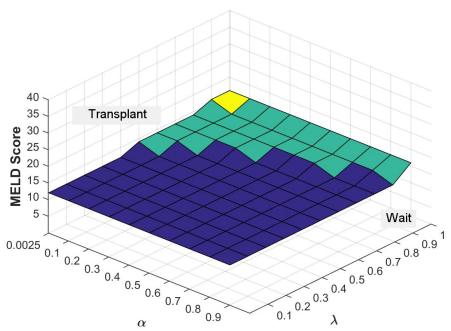
(a) Organ 1



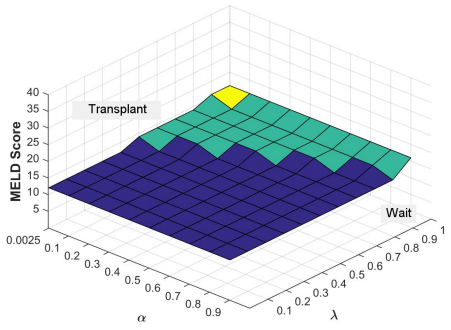
(b) Organ 2



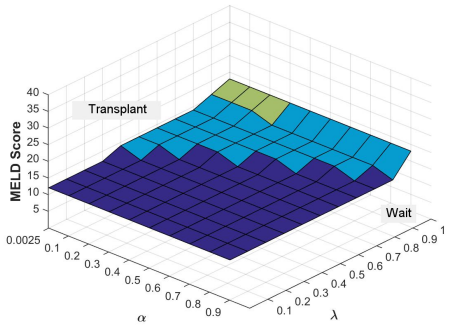
(c) Organ 3



(d) Organ 4

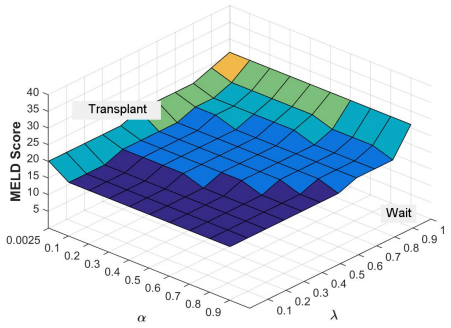


(e) Organ 5

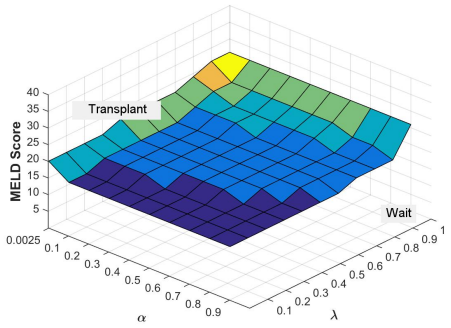


(f) Organ 6

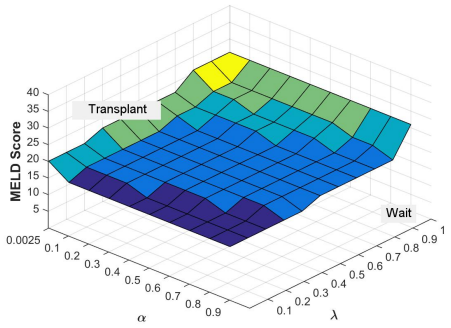
Figure B.3: Optimal control limits for disease group 1, patient type 3 under mean-AVaR risk measure.



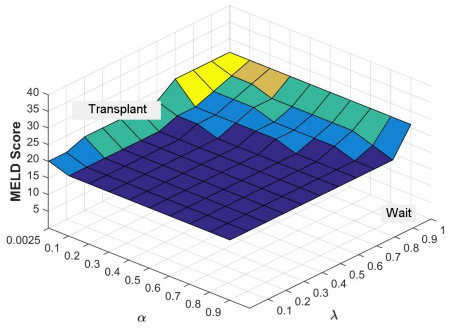
(a) Organ 1



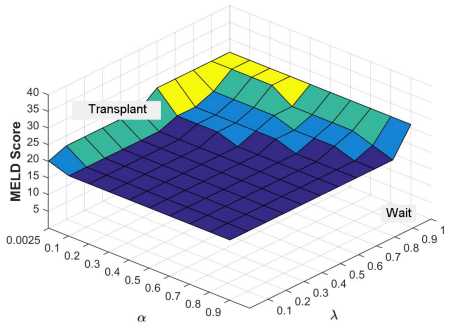
(b) Organ 2



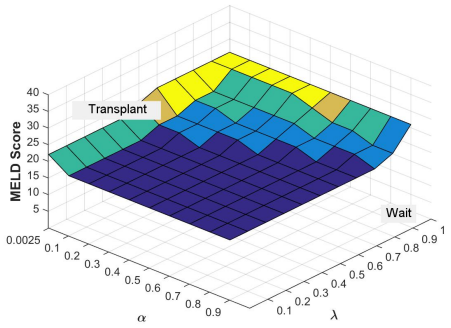
(c) Organ 3



(d) Organ 4

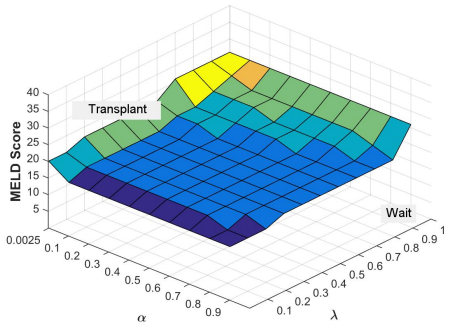


(e) Organ 5

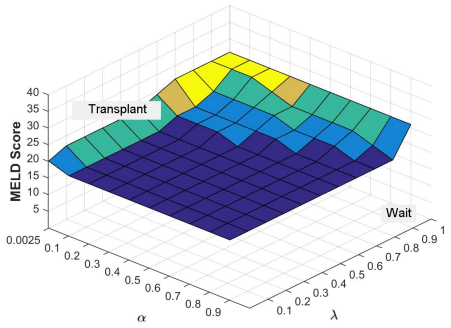


(f) Organ 6

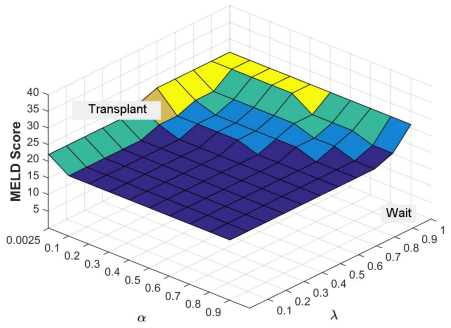
Figure B.4: Optimal control limits for disease group 2, patient type 1 under mean-AVaR risk measure.



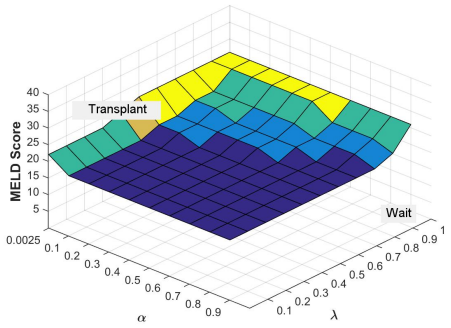
(a) Organ 1



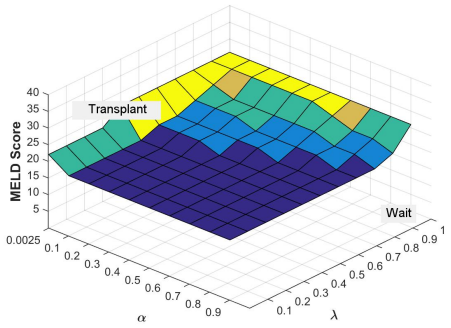
(b) Organ 2



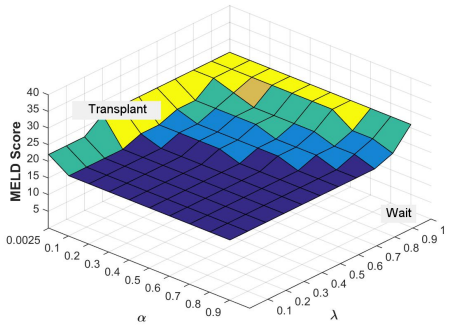
(c) Organ 3



(d) Organ 4

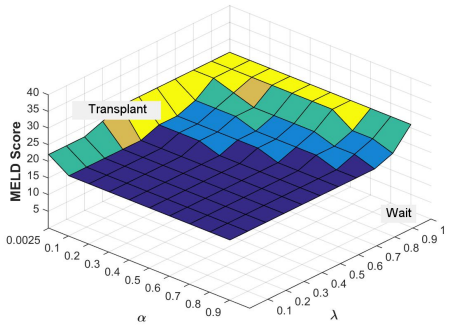


(e) Organ 5

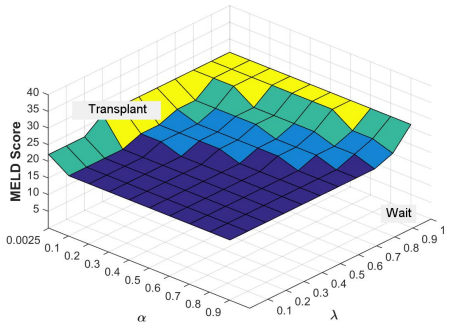


(f) Organ 6

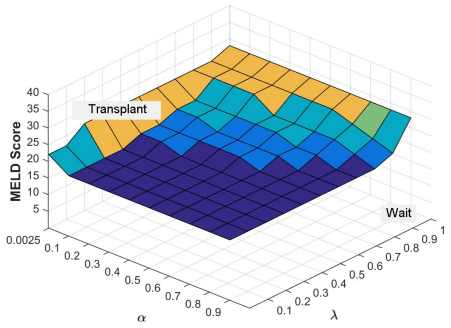
Figure B.5: Optimal control limits for disease group 2, patient type 2 under mean-AVaR risk measure.



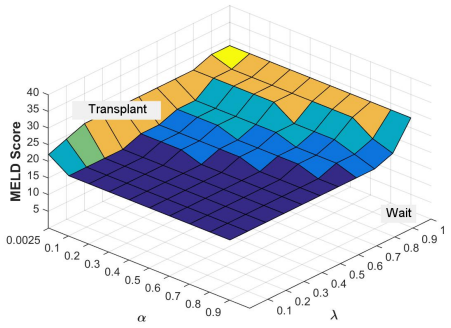
(a) Organ 1



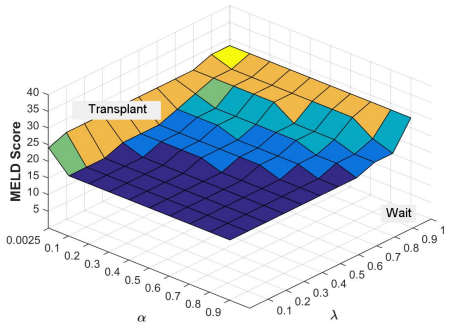
(b) Organ 2



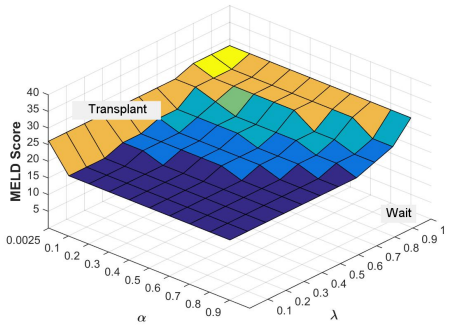
(c) Organ 3



(d) Organ 4

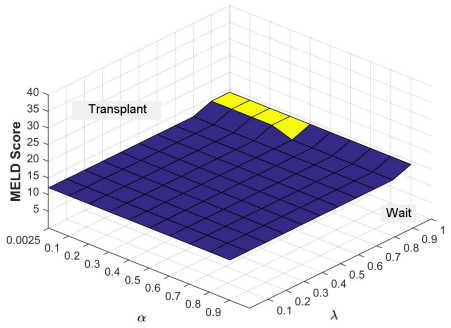


(e) Organ 5

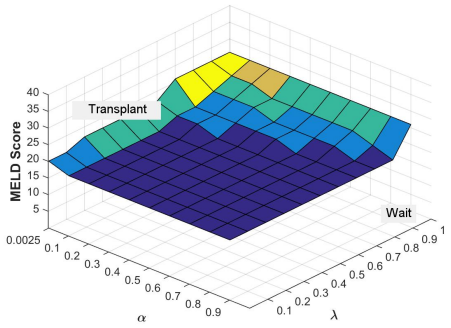


(f) Organ 6

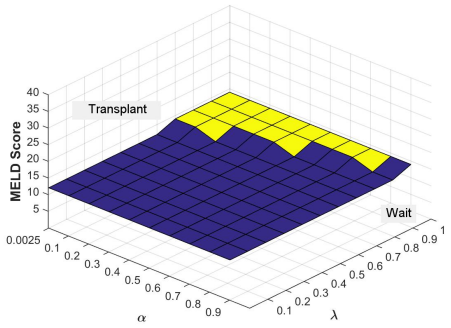
Figure B.6: Optimal control limits for disease group 2, patient type 3 under mean-AVaR risk measure.



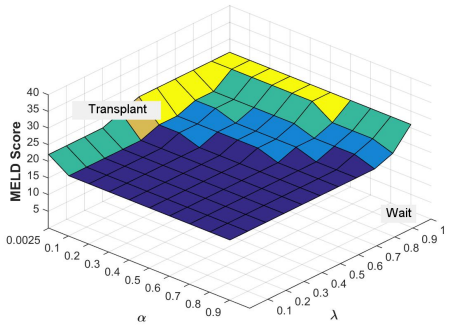
(a) Patient type 1, Disease Group 1



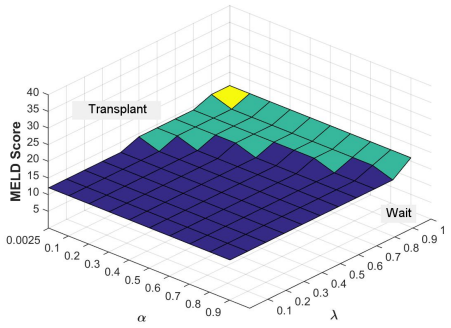
(b) Patient type 1, Disease Group 2



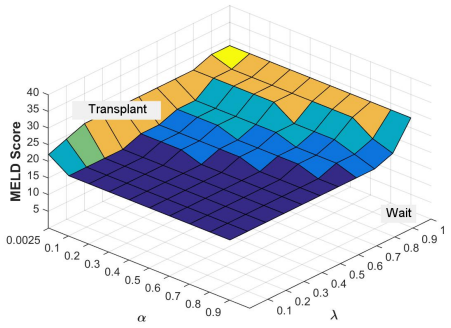
(c) Patient type 2, Disease Group 1



(d) Patient type 2, Disease Group 2



(e) Patient type 3, Disease Group 1

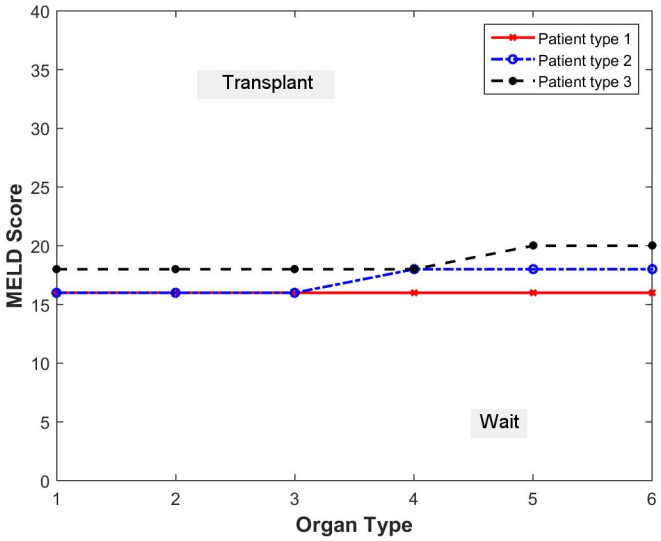


(f) Patient type 3, Disease Group 2

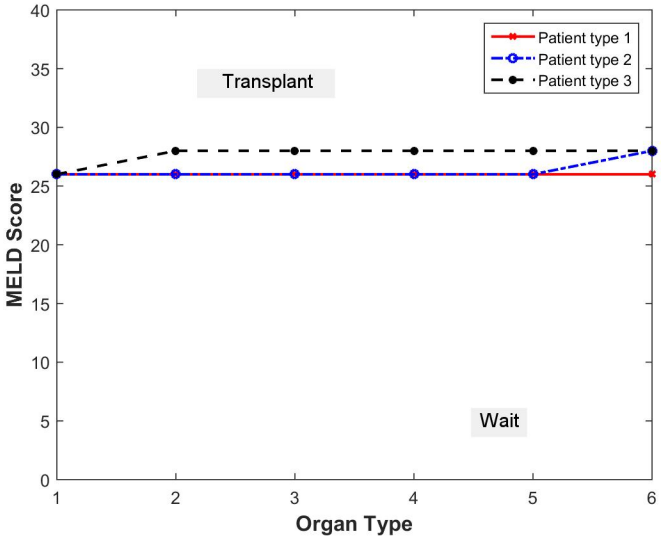
Figure B.7: Optimal control limit comparisons among patients when organ type 4 is used under mean-AVaR risk measure.

Appendix C

Optimal Control-Limits Obtained Using QALY

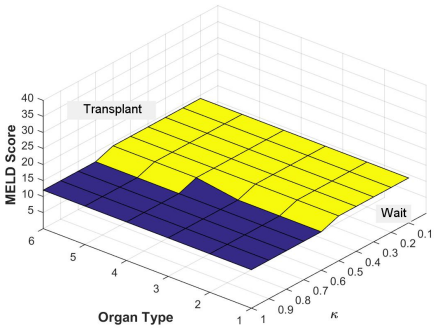


(a) Disease Group 1

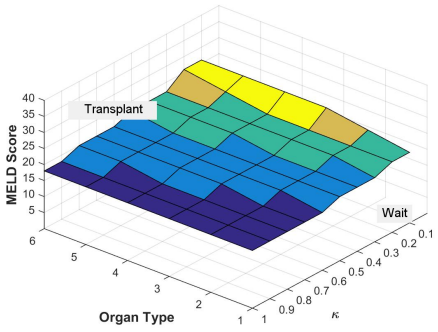


(b) Disease Group 2

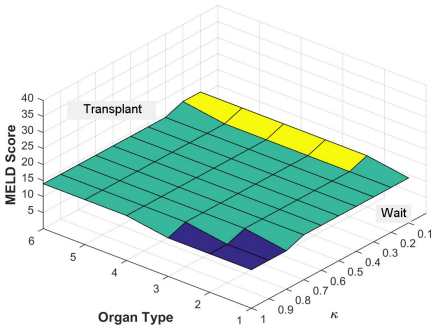
Figure C.1: Risk-neutral optimal control limits.



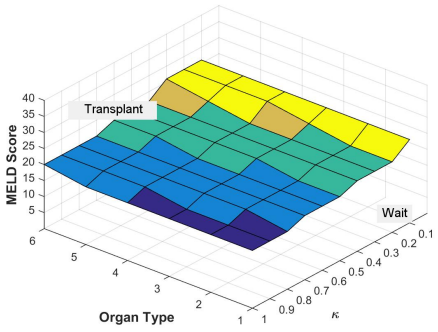
(a) Disease Group 1, Patient type 1



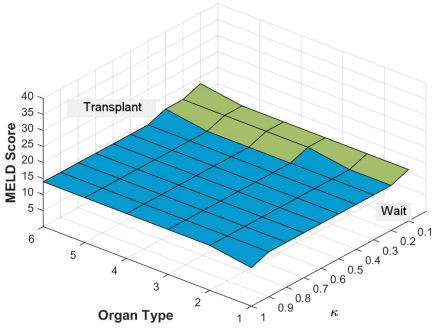
(b) Disease Group 2, Patient type 1



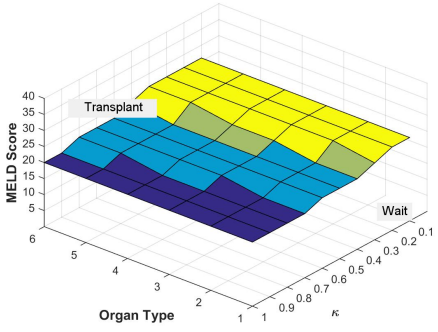
(c) Disease Group 1, Patient type 2



(d) Disease Group 2, Patient type 2

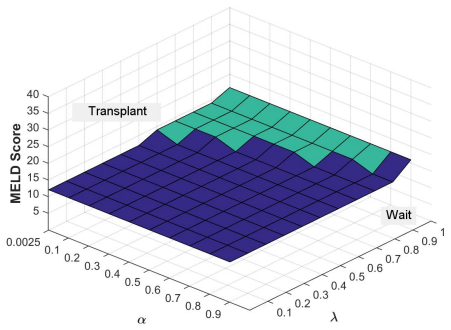


(e) Disease Group 1, Patient type 3

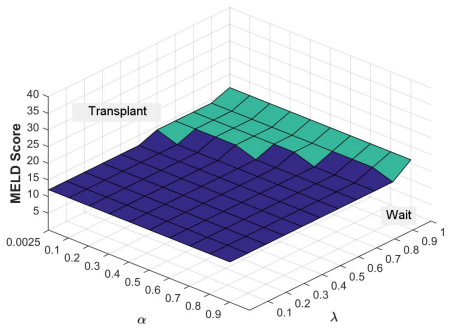


(f) Disease Group 2, Patient type 3

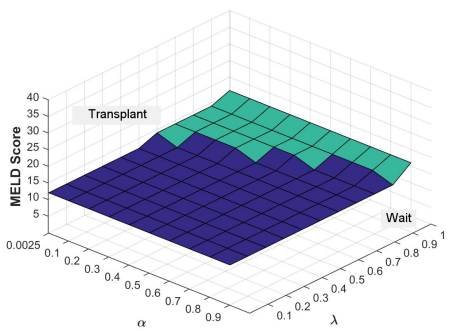
Figure C.2: Optimal control limits under first-order mean-semi-deviation risk measure.



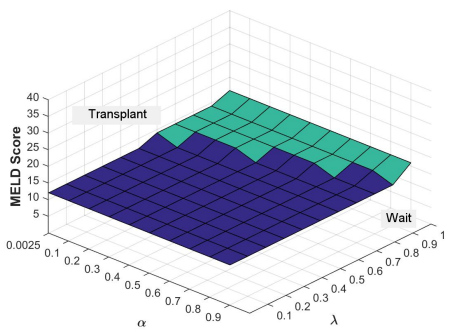
(a) Organ 1



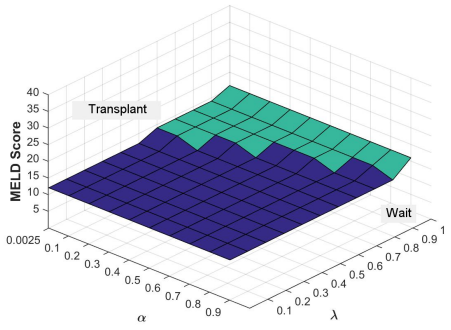
(b) Organ 2



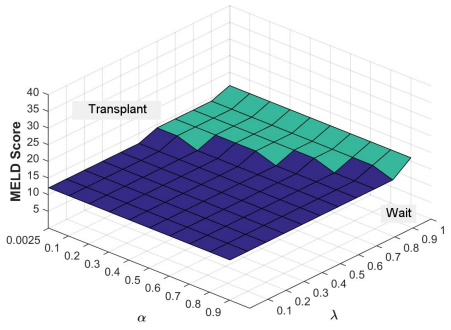
(c) Organ 3



(d) Organ 4

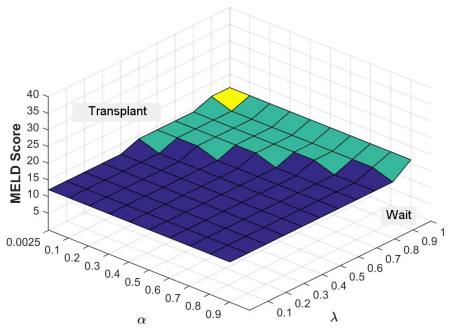


(e) Organ 5

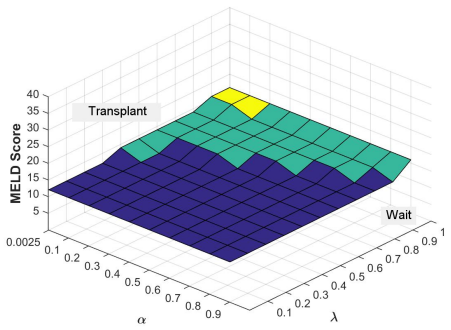


(f) Organ 6

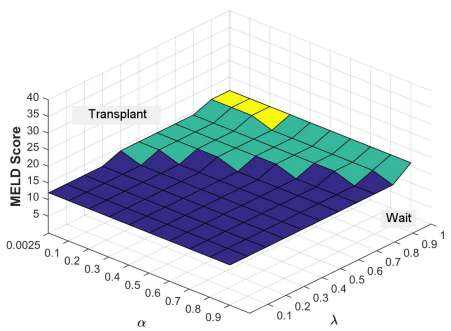
Figure C.3: Optimal control limits for disease group 1, patient type 1 under mean-AVaR risk measure.



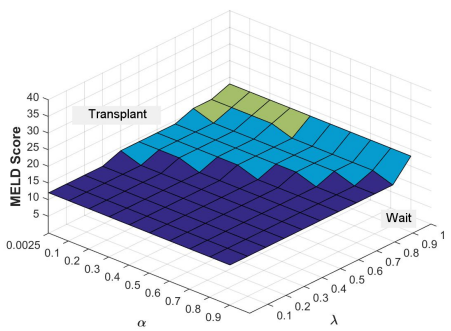
(a) Organ 1



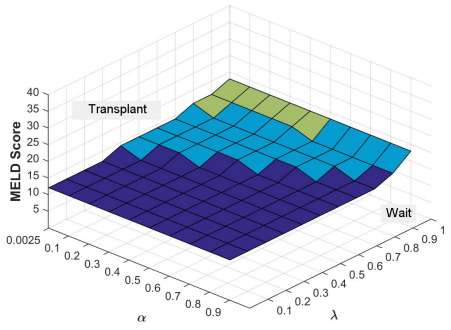
(b) Organ 2



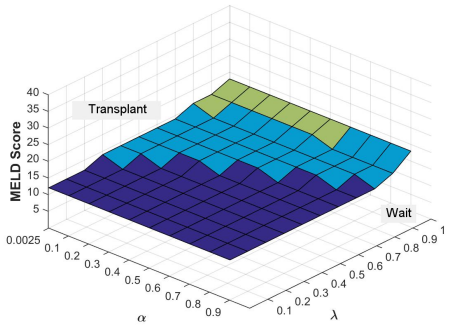
(c) Organ 3



(d) Organ 4

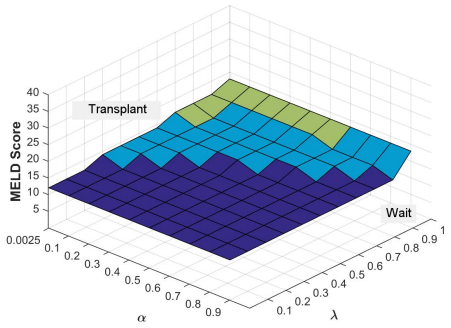


(e) Organ 5

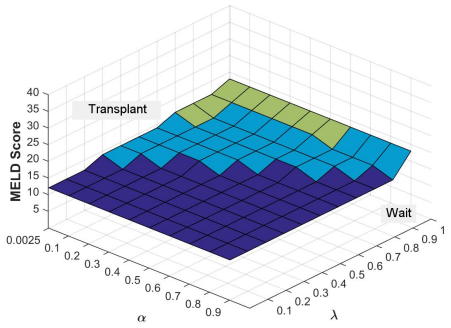


(f) Organ 6

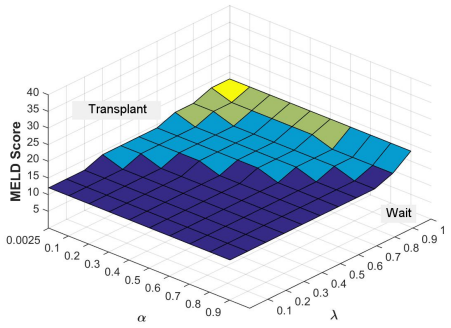
Figure C.4: Optimal control limits for disease group 1, patient type 2 under mean-AVaR risk measure.



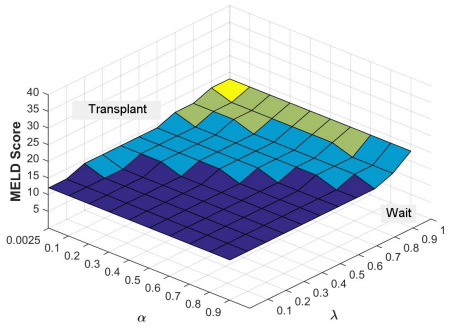
(a) Organ 1



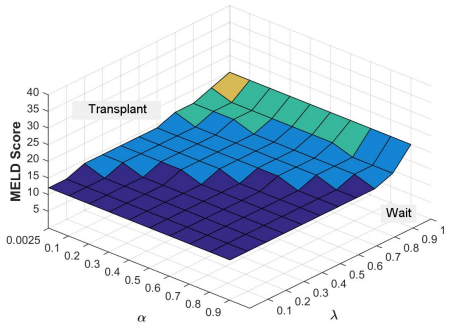
(b) Organ 2



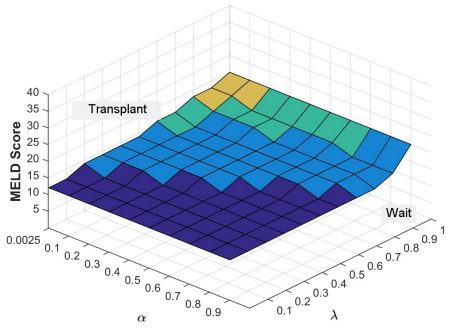
(c) Organ 3



(d) Organ 4

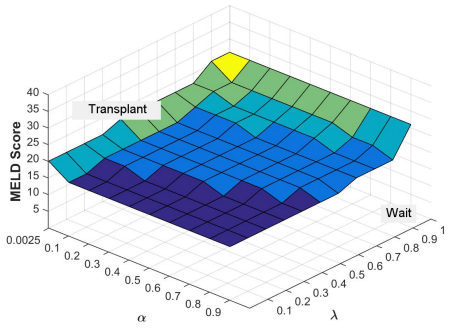


(e) Organ 5

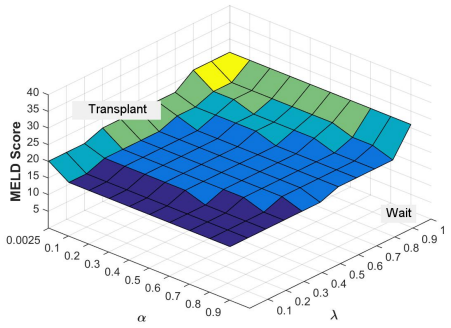


(f) Organ 6

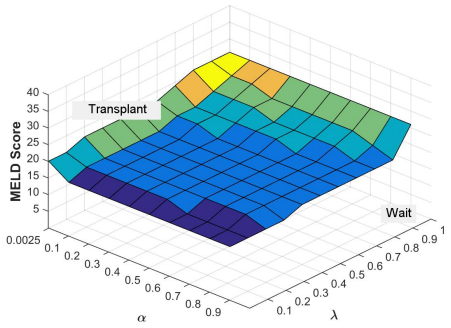
Figure C.5: Optimal control limits for disease group 1, patient type 3 under mean-AVaR risk measure.



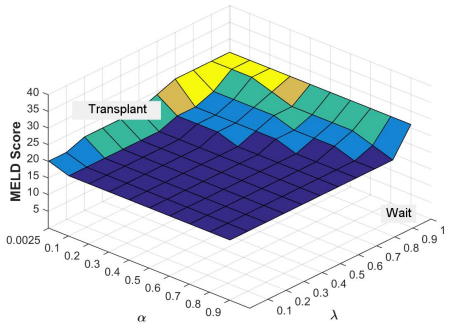
(a) Organ 1



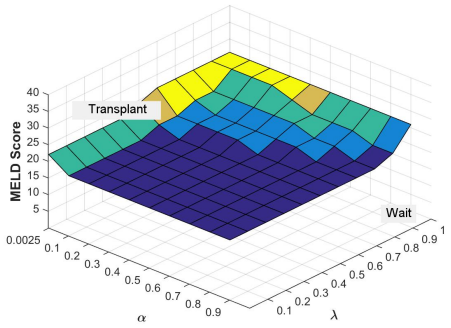
(b) Organ 2



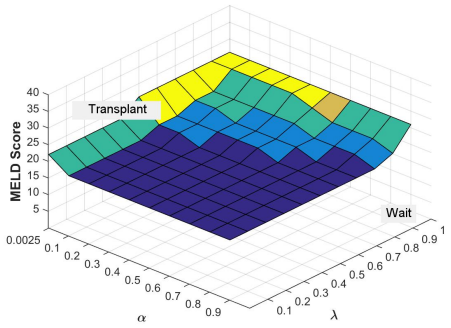
(c) Organ 3



(d) Organ 4

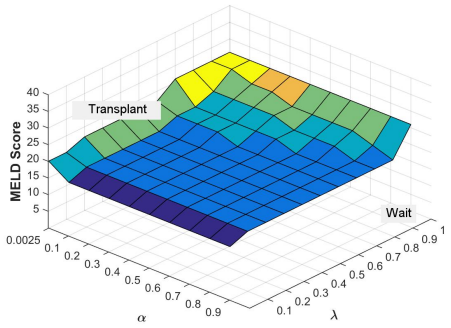


(e) Organ 5

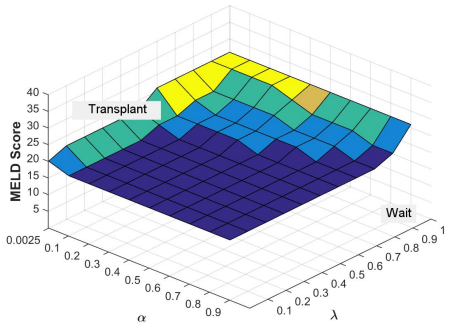


(f) Organ 6

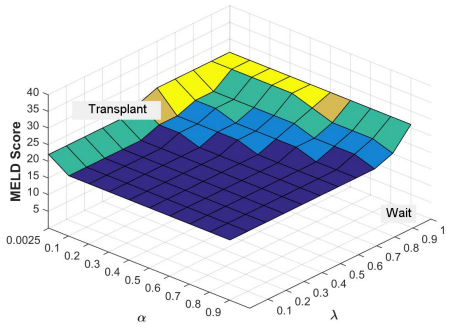
Figure C.6: Optimal control limits for disease group 2, patient type 1 under mean-AVaR risk measure.



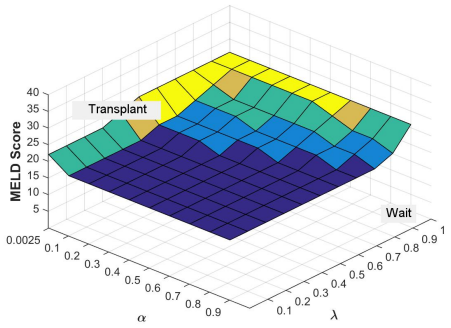
(a) Organ 1



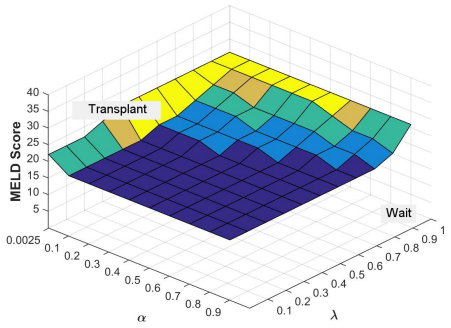
(b) Organ 2



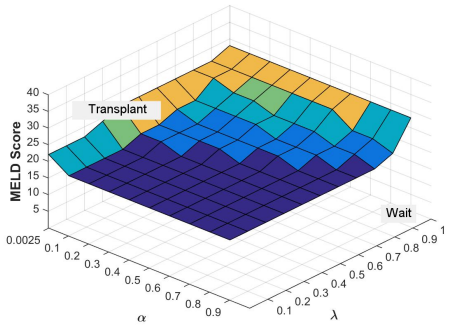
(c) Organ 3



(d) Organ 4

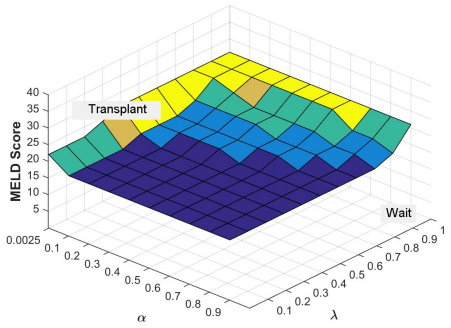


(e) Organ 5

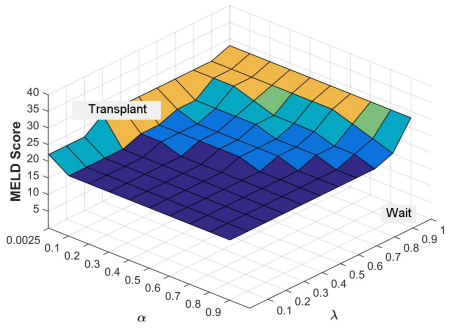


(f) Organ 6

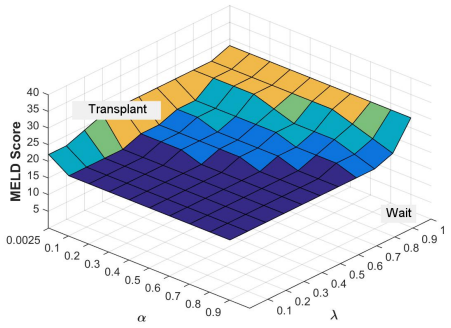
Figure C.7: Optimal control limits for disease group 2, patient type 2 under mean-AVaR risk measure.



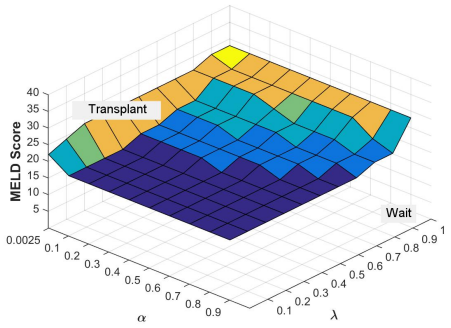
(a) Organ 1



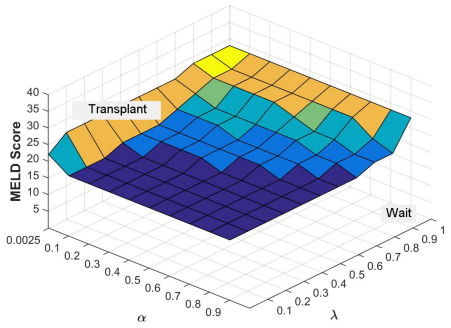
(b) Organ 2



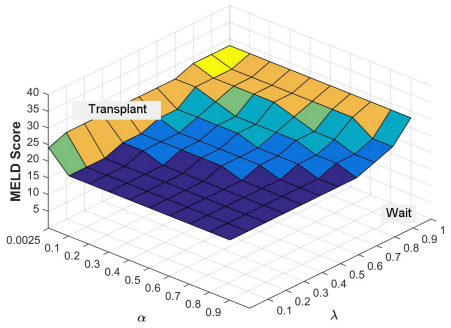
(c) Organ 3



(d) Organ 4



(e) Organ 5



(f) Organ 6

Figure C.8: Optimal control limits for disease group 2, patient type 3 under mean-AVaR risk measure.

Appendix D

Sensitivity Analysis

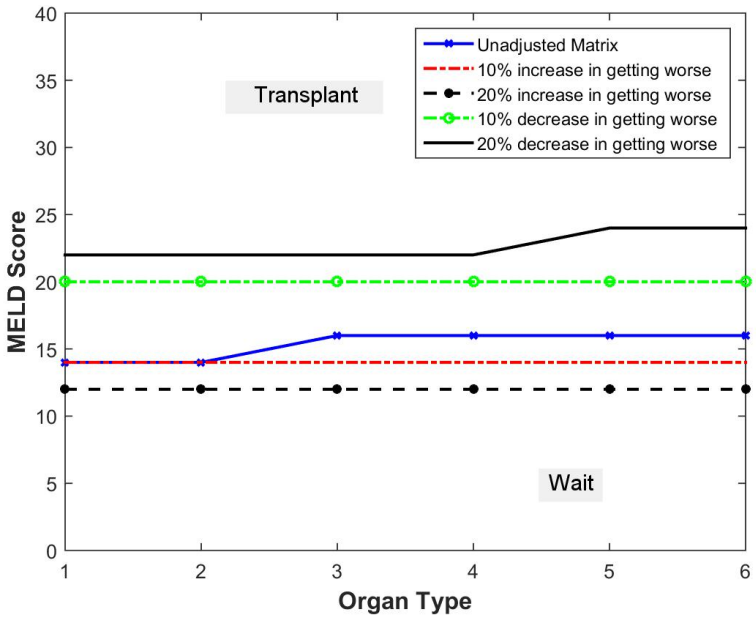
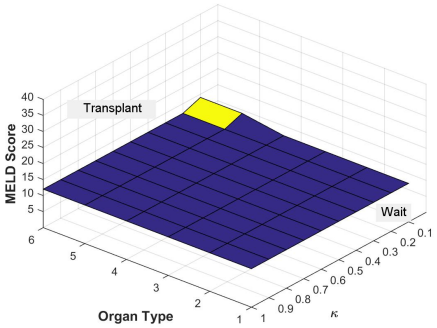
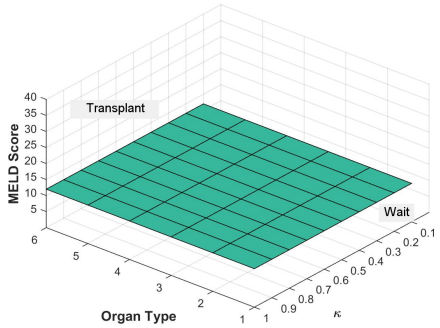


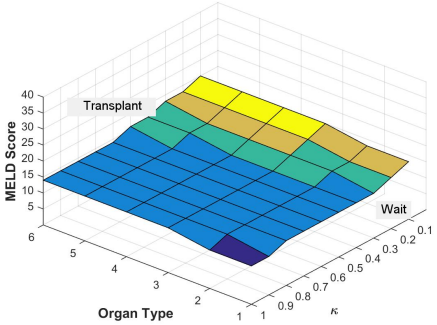
Figure D.1: Sensitivity analysis for the risk-neutral problem using disease group 1, patient 2 data.



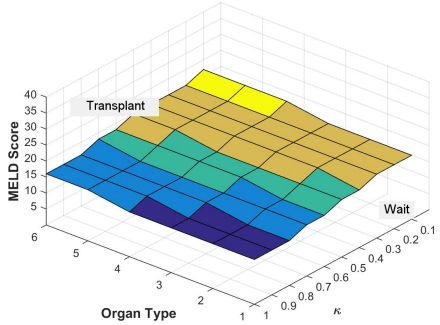
(a) 10 % increase in getting worse probabilities



(b) 20 % increase in getting worse probabilities

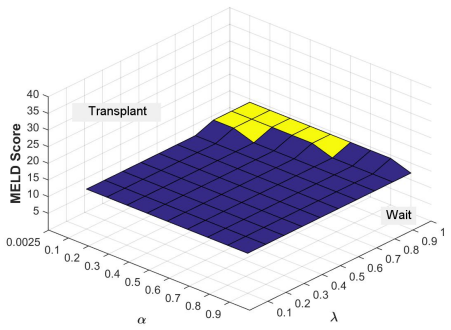


(c) 10 % decrease in getting worse probabilities

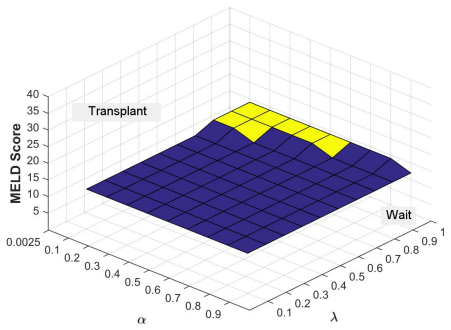


(d) 20 % decrease in getting worse probabilities

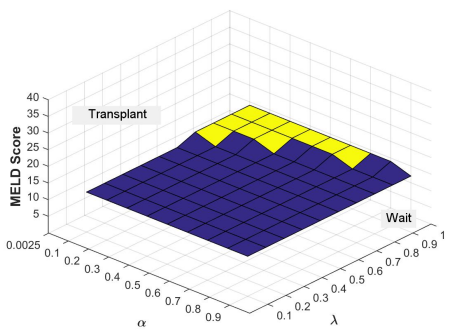
Figure D.2: Sensitivity analysis under first-order mean-semi-deviation risk measure using disease group 1, patient 2 data.



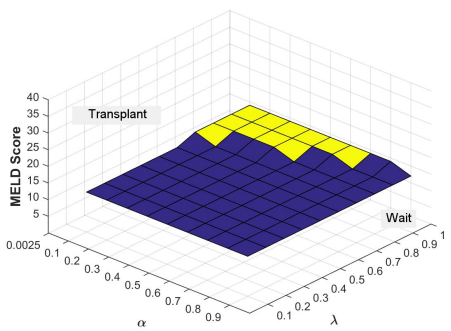
(a) Organ 1



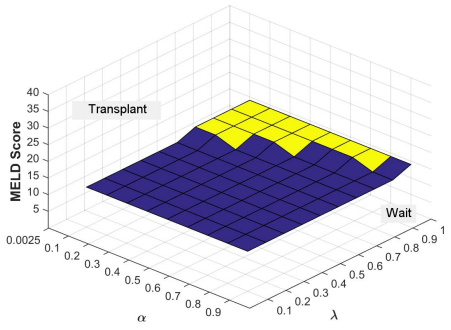
(b) Organ 2



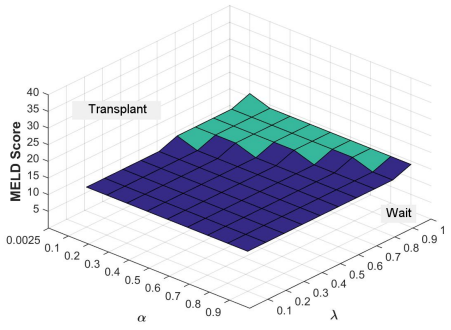
(c) Organ 3



(d) Organ 4

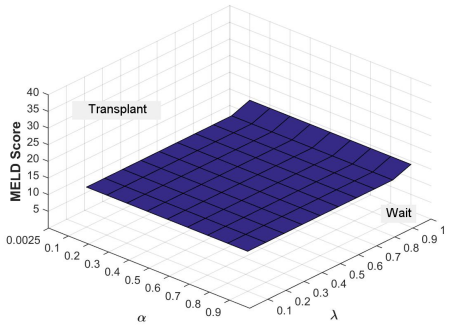


(e) Organ 5

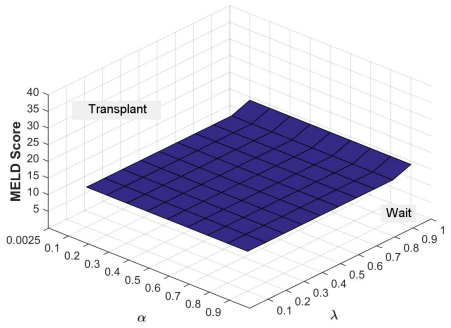


(f) Organ 6

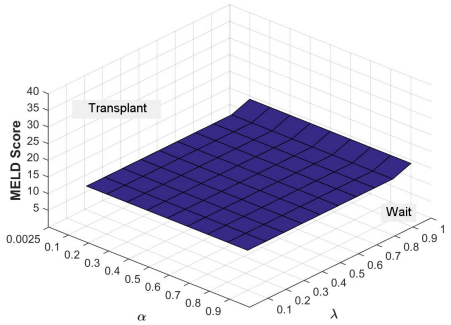
Figure D.3: Sensitivity analysis under mean-AVaR risk measure using patient 2 data and smoothed matrix with unadjusted probabilities for disease group 1.



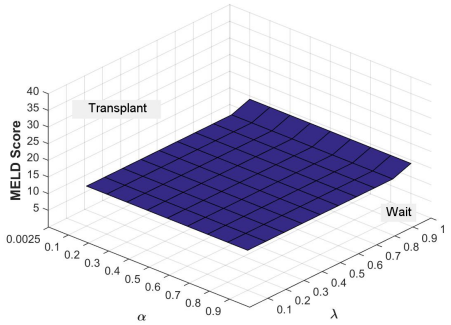
(a) Organ 1



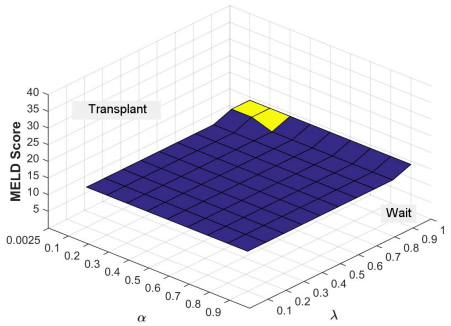
(b) Organ 2



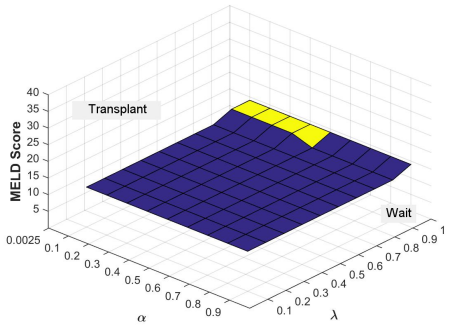
(c) Organ 3



(d) Organ 4

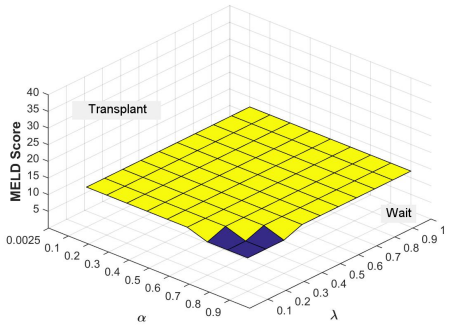


(e) Organ 5

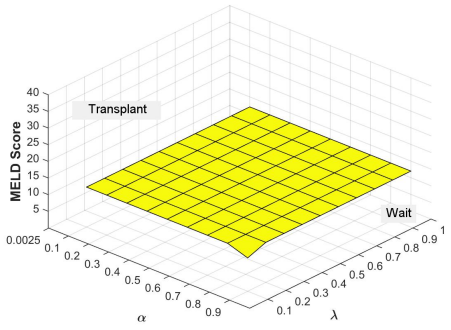


(f) Organ 6

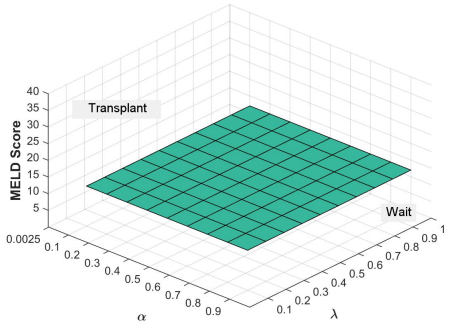
Figure D.4: Sensitivity analysis under mean-AVaR risk measure using patient 2 data and the matrix with 10 % increased probabilities of getting worse for disease group 1.



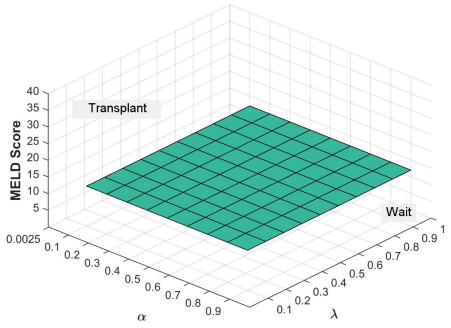
(a) Organ 1



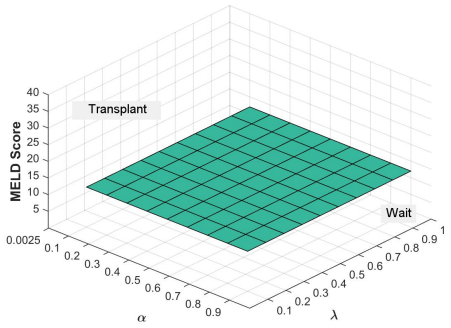
(b) Organ 2



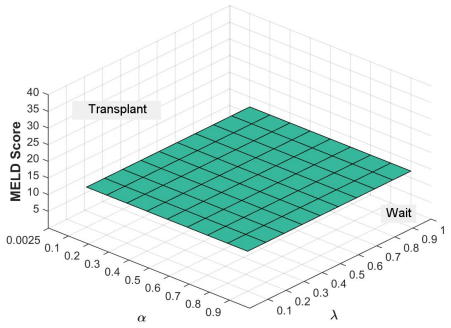
(c) Organ 3



(d) Organ 4

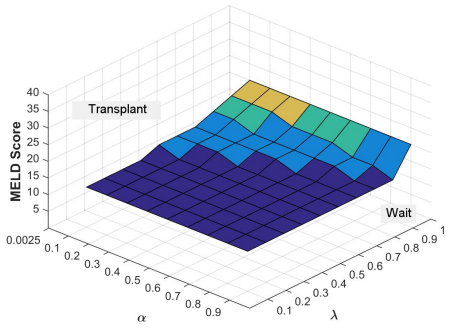


(e) Organ 5

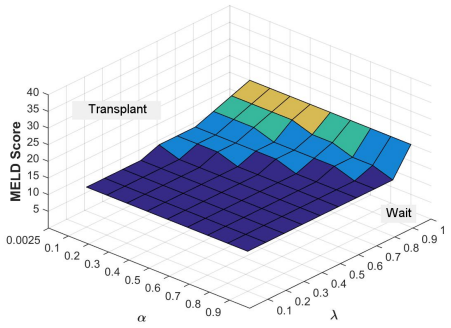


(f) Organ 6

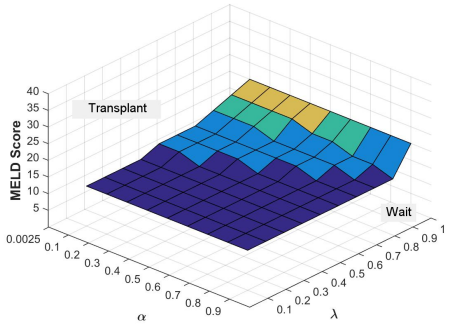
Figure D.5: Sensitivity analysis under mean-AVaR risk measure using patient 2 data and the matrix with 20 % increased probabilities of getting worse for disease group 1.



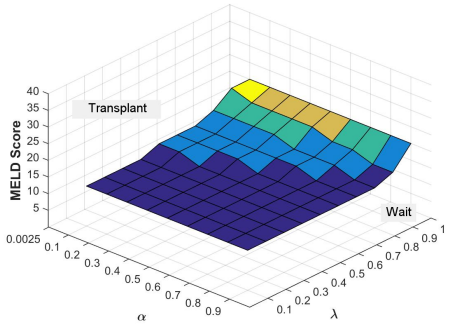
(a) Organ 1



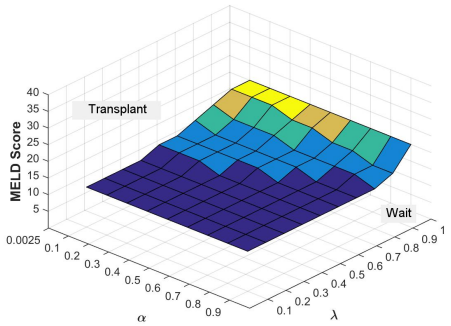
(b) Organ 2



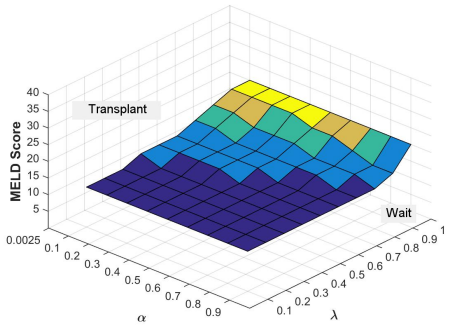
(c) Organ 3



(d) Organ 4

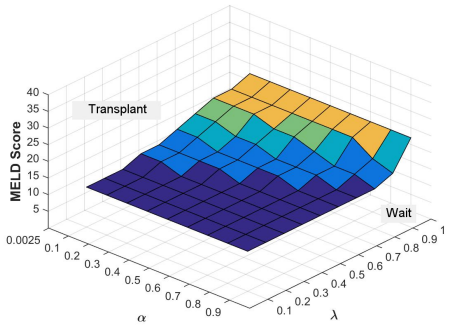


(e) Organ 5

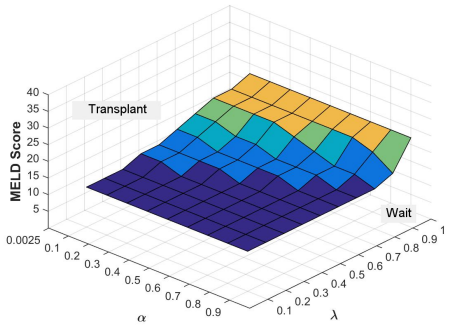


(f) Organ 6

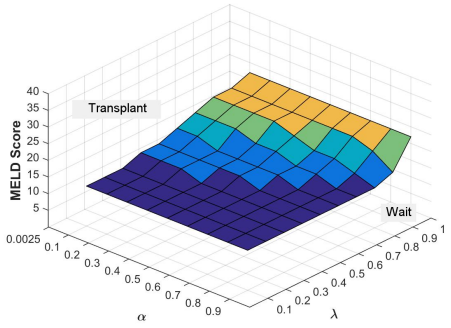
Figure D.6: Sensitivity analysis under mean-AVaR risk measure using patient 2 data and the matrix with 10 % decreased probabilities of getting worse for disease group 1.



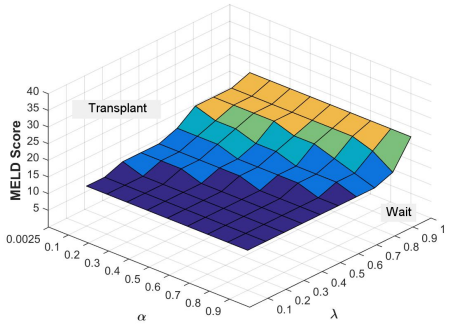
(a) Organ 1



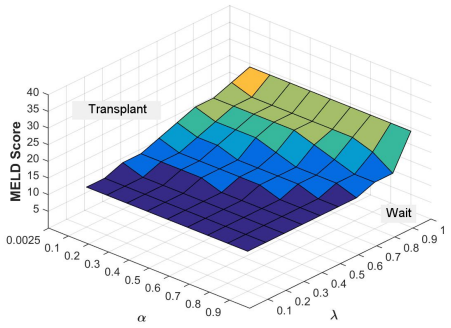
(b) Organ 2



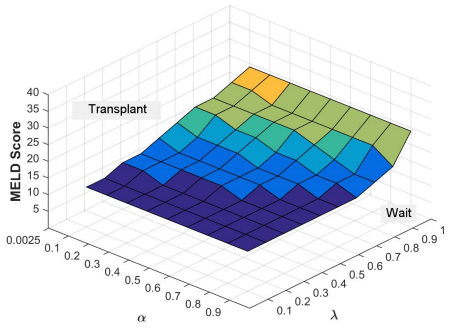
(c) Organ 3



(d) Organ 4



(e) Organ 5



(f) Organ 6

Figure D.7: Sensitivity analysis under mean-AVaR risk measure using patient 2 data and the matrix with 20 % decreased probabilities of getting worse for disease group 1.