WHEN TO INITIATE, WHEN TO SWITCH, AND HOW TO SEQUENCE HIV THERAPIES: A MARKOV DECISION PROCESS APPROACH

by

Steven M. Shechter

BS, Loyola University Chicago, 1997

MS, Georgia Institute of Technology, 1999

Submitted to the Graduate Faculty of

the School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

SCHOOL OF ENGINEERING

This dissertation was presented

by

Steven M. Shechter

It was defended on

July 14, 2006

and approved by

Andrew J. Schaefer, PhD, Assistant Professor, School of Engineering

Matthew D. Bailey, PhD, Assistant Professor, School of Engineering

R. Scott Braithwaite, MD, Assistant Professor, School of Medicine, Yale University

Jonathan P. Caulkins, PhD, Professor, Heinz School of Public Policy, Carnegie Mellon University

Chung-Chou H. Chang, PhD, Assistant Professor, School of Medicine

Mainak Mazumdar, PhD, Professor, School of Engineering

Mark S. Roberts, MD, Associate Professor, School of Medicine

Dissertation Director: Andrew J. Schaefer, PhD, Assistant Professor, School of Engineering

ABSTRACT

WHEN TO INITIATE, WHEN TO SWITCH, AND HOW TO SEQUENCE HIV THERAPIES: A MARKOV DECISION PROCESS APPROACH

Steven M. Shechter, PhD

University of Pittsburgh, 2006

HIV and AIDS are major health care problems throughout the world, with 40 million people living with HIV by the end of 2005. In that year alone, 5 million people acquired HIV, and 3 million people died of AIDS. For many patients, advances in therapies over the past ten years have changed HIV from a fatal disease to a chronic, yet manageable condition. The purpose of this dissertation is to address the challenge of effectively managing HIV therapies, with a goal of maximizing a patient's total expected lifetime or quality-adjusted lifetime.

Perhaps the most important issue in HIV care is when a patient should initiate therapy. Benefits of delaying therapy include avoiding the negative side effects and toxicities associated with the drugs, delaying selective pressures that induce the development of resistant strains of the virus, and preserving a limited number of treatment options. On the other hand, the risks of delayed therapy include the possibility of irreversible damage to the immune system, development of AIDS-related complications, and death. We develop a Markov decision process (MDP) model that examines this question, and we solve it using clinical data. Because of the development of resistance to administered therapies over time, an extension to the initiation question arises: when should a patient switch therapies? Also, inherent in both the initiation and switching questions is the question of which therapy to use each time. We develop MDP models that consider the switching and sequencing problems, and we discuss the challenges involved in solving these models.

Keywords: Markov Decision Processes, Dynamic Programming, Stochastic Optimization, Optimal Therapy Planning, HIV, Health Care Policy, Medical Decision Making.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I came to the University of Pittsburgh in January of 2002, excited to pursue my PhD studies at the intersection of operations research and health care. I got what I wanted and a lot more. The academic experiences I have had since arriving have been incredibly rewarding, and for that I thank my advisor, Andrew Schaefer. He knew exactly why I was leaving work in industry to pursue my PhD, and he gave me both the support and freedom to pursue topics that interested me most so as to achieve my goals. I appreciate his encouragement to write papers and grants, get involved in various projects, and present my work at conferences, all of which have undoubtedly prepared me well for being an assistant professor. I came to Pitt knowing I would be working on health care applications without knowing who I would be working with from health care. That wild card turned out to be a major reason I have been rewarded beyond what I could have imagined. Mark Roberts has been an incredible mentor and colleague, teaching me the ways of medical decision making and bridging my gaps between operations research and health care. I have also thoroughly enjoyed working with Joyce Chang, Cindy Bryce, and Scott Braithwaite on various projects. I especially thank Scott for allowing me the opportunity to work closely with him on an HIV simulation model, which inspired my dissertation topic. Returning to the operations research side of my dissertation, I thank Matt Bailey for all of his diligence in checking the theoretical results of my dissertation and for our discussions on methodological developments. Our off-topic discussions were also valuable for giving me the right amount of down time to return to work re-energized. Since deciding on my dissertation topic, I have had the fortune of having Jon Caulkins serve on my committee, providing fresh, insightful perspectives with respect to both my dissertation research as well as my academic job search. And I thank Mainak Mazumdar for always welcoming me into his office and clearly answering my random statistical questions over the years.

In addition to the great faculty support, I have worked alongside some stellar peers at Pitt. For their academic and moral support, I would like to thank Oguzhan Alagoz, Halil Bayrak, Mehmet Demirci, Nan Kong, Rob Koppenhaver, Jen Kreke, Chen Li, Aytun Ozturk, Claudio Rubial, Gorkem Saka, Gillian Sanders, Burhan Sandikci, Natasa Vidic, and Zhouyan Wang. I also appreciate the excellent undergraduate research assistants I have worked with

over the years: Lindsay Destefano, Vicki Gettinger, and Danielle Williams. Best of luck to all of you, and I look forward to crossing paths again.

About half way through my PhD studies, I was blessed with meeting my wife, Maria Anduanet. You have kept me smiling ever since I met you. Thank you for all your love, support, and laughter, which give me the strength to move afar and the confidence that our time in Vancouver will be successful and enjoyable. I have been surrounded by incredibly loving family and friends throughout my life. To Lori and David; to Larry and Cheryl; to Margot, Josh, and Amy; to Ross, Neil, and Joseph; to Mimi and Papa Florence; to Bubby and Papa Shechter: thank you for the unique part each of you has played in bringing me to this point of my life. And I have added such wonderful family by marrying Anduanet: her mom and dad, Mary and Luis, and her brother José Luis. Thank you for welcoming me with open arms from day one. In closing, I dedicate my dissertation to my parents, Ron and Sandy Shechter. You have always looked out for my best interests. To explain your everlasting love, your constant support, and your encouragement to pursue my interests though they may take me many miles from you, is beyond the scope of this work.

Finally, I would also like to thank the Agency for Healthcare Research and Quality for their financial support through their Grant for Health Services Dissertation Research (1 R36 HS15533-01).

I have exceeded my two minutes, and the orchestra is starting to play. With that, I present my thesis.

1.0 INTRODUCTION

Acquired Immune Deficiency Syndrome (AIDS) and its cause, the Human Immunodeficiency Virus (HIV), are among the most pressing health care problems in the world, with approximately 40 million people living with HIV by the end of 2005. In that year alone, 5 million people became infected with HIV while 3 million people died of AIDS [176]. Unfortunately, the number of people living with HIV continues to grow [176].

HIV's fatal effects arise from its attack of a person's CD4 white blood cells. As these cells become depleted, HIV patients become more vulnerable to certain infections, resulting in AIDS and eventually death [169]. Antiretroviral therapy involves the administration of drugs that are designed to inhibit HIV replication so as to preserve the vital CD4 cells. The first drug for fighting HIV, AZT, was approved for use in 1986 [162]. Although promising at first, the success of using a single therapy was not long lasting due to the rapid development of resistant strains of the virus [162]. Therefore, new classes of drugs were developed in the early 1990s, leading to the widespread use of highly active antiretroviral therapy (HAART) in the mid-to-late 1990s. HAART, the standard of care today, is also called combination therapy or cocktail therapy because it involves patients taking three or more drugs simultaneously from the various drug classes (we will also just refer to it as "therapy"). The increased use of HAART in the past ten years has led to significant reductions in HIV-related morbidity, mortality, and health care utilization [127, 170]. As such, many patients have seen HIV transform from a fatal disease to a chronic, yet manageable condition [153].

This dissertation focuses on the effective management of HAART. Despite great advances in HIV therapies, there is still considerable debate about the best way to use them. For example, the best time for a patient to initiate HAART is an open question [2, 42, 75, 78, 79, 81, 87, 106, 113, 124, 131, 150, 171]. According to Dr. Anthony Fauci, the director of the

National Institute of Allergy and Infectious Diseases, the question of when to initiate therapy is "the most important question in HIV therapy" [79]. Benefits of delaying therapy include avoiding the negative side effects and toxicities associated with the drugs, delaying selective pressures that induce the development of resistance strains of the virus, and preserving a limited number of treatment options [123]. On the other hand, the risks of delayed therapy include the possibility of irreversible damage to the immune system, development of AIDS-related complications, and death [123]. Of course, with the various therapies available, there is also a question of which therapy to use first.

Along with the question of when to initiate therapy comes the question of when to switch therapies. Upon initiating an effective therapy, viral load (the number of HIV RNA copies/mL of blood) typically drops and CD4 count (the number of CD4 cells/mm^3 of blood) rises for some period of time. However, in the face of selective pressures, therapies typically lose effectiveness due to the build-up of resistant mutations, resulting in a rise of viral load (or virologic failure), a loss of CD4 cells (or immunologic failure), or the development of AIDS-related complications (or clinical progression) [123]. These failures of therapies may also happen because patients experience intolerable side effects from their therapies, do not adhere well to their therapies, or have problems with drug absorption [123]. For these reasons, deciding when to switch therapies and which therapy to use next are also major considerations in HIV care.

Decisions involving the management of HAART are currently based on clinical judgment and national guidelines for HIV care, which in turn are based on the outcomes of clinical studies and expert opinion [123]. In Section 2.1 we review some of the studies that may have influenced the current U.S. Department of Health and Human Services (DHHS) guidelines, which are described in Section 2.2. Although clinical studies should be undertaken prior to making any significant changes to treatment policy, such studies are not practical for narrowing the field of potentially beneficial treatment plans. For example, cost, time, and ethical considerations render randomized controlled trials unrealistic as the number of treatment strategies increases. In such situations, mathematical models may prove quite useful to test

many possible treatment options with relatively little cost, in a relatively short amount of time, and with no risk to patients. Therefore, the ultimate goal of our research is to provide analyses and methods that can inform clinical studies.

1.1 PURPOSE AND CONTRIBUTIONS OF PRESENT RESEARCH

The purpose of this dissertation is to address the challenge of effectively managing HIV therapies, with a goal of maximizing a patient's total expected lifetime or quality-adjusted lifetime. We shall do this by developing Markov decision process (MDP) models that capture essential aspects of HIV progression and treatment. Generally, an MDP model is applicable whenever a decision maker observes the state of a system at multiple points in time, and at each time, chooses an action to meet a goal (for example, maximize the expected reward over the entire time horizon of the problem). Based on the current state and action taken, a reward is received between time periods. Furthermore, due to uncertainty in the way the system reacts to the chosen action, the decision maker has only a probabilistic sense of how the system will evolve until the next period [140].

HIV therapy planning fits this framework. A physician sees an HIV patient periodically, observing the patient's state of HIV through laboratory measurements of key prognostic variables such as CD4 count or viral load. Based on these measurements, the physician decides (with the patient) whether to initiate therapy (and which one), switch to a different therapy (and to which one), or continue as is until the next patient visit. Typically, decisions are made with an overall goal of maximizing the patient's expected lifetime or quality-adjusted lifetime. For the former goal, the reward may equal the time between visits, and for the latter, it may equal a quality-adjusted time between visits. Also, at the time of the decision, it is not known how the prognostic variables will change over the next period or whether the patient will even survive until the next visit. However, probabilities of the different outcomes can be estimated based on clinical data.

To our knowledge, we are the first to utilize MDPs to consider optimal HIV therapy planning. For a number of reasons, we believe this provides a natural and clinically valid ap-

proach to the treatment problem. First, MDPs are designed to solve discrete-time stochastic dynamic problems, and although patients may be seen at any time for specific problems, most HIV care occurs at fixed intervals of time. Our MDP will consider actions taken at these discrete time intervals (e.g., every month). Second, CD4 counts are known to vary considerably day to day [169], so it is important to use a framework that considers the stochastic progression of the disease. Third, MDPs include summable rewards and can consider the possibility of patient death between time periods, both of which fit well with our objective of maximizing expected lifetime or quality-adjusted lifetime. Finally, in addition to their natural fit for HIV therapy planning and a variety of other decision problems, MDPs can be used to gain deeper insight into a problem by investigating structural properties; that is, an understanding of how certain structure on the input may guarantee certain structure on the optimal solution. We refer the reader to [140] for extensive coverage of MDPs and references to other literature on the topic. For a review of MDPs applied to health care, see [152].

The specific goals of this research are to:

1. **Develop MDP models of the optimal time to initiate HIV therapy,**

2. **Use clinical data to solve for the optimal time to initiate therapy as a function of a patient's CD4 count,**

3. **Develop MDP models of the optimal sequencing and switching of HIV therapies, and**

4. **Explore structural properties of each model.**

The overall contribution of this research is to:

- **Develop and solve the first HIV optimization models that aim to maximize a patient's lifetime or quality-adjusted lifetime.**

The remainder of this dissertation is organized as follows. In Chapter 2, we describe relevant literature, including clinical studies, mathematical studies, and the current national guidelines for HIV therapy planning. Chapter 3 develops clinically based models of the probabilistic progression of CD4 counts prior to initiating therapy, along with estimates of remaining survival upon initiating therapy. These are used in Chapter 4, where we develop an MDP of the optimal time to initiate therapy as a function of a single prognostic variable (CD4 count). We explore structural properties and solve the model with the data-based components of Chapter 3. Chapter 5 extends the framework to consider a two-dimensional state space of CD4 count and viral load. Chapters 6 and 7 then consider the questions of the optimal switching and sequencing of therapies. Chapter 6 considers these questions assuming knowledge only about the lifetime distributions induced by the therapies, while Chapter 7 takes the perspective of informative, periodic observations of a patient's health. We discuss conclusions, limitations, and future extensions of the dissertation in Chapter 8.

2.0 LITERATURE REVIEW

The literature on HIV and AIDS is extremely broad and deep. Here, we focus on research related to the effective management of HAART. We divide our review into clinical studies and mathematic models, and we give an overview of the U.S. Department of Health and Human Services (DHHS) guidelines for HIV therapy management.

2.1 CLINICAL STUDIES

We review two major types of clinical studies that inform therapeutic decision making for HIV: observational cohort studies and randomized controlled trials.

2.1.1 Observational Cohort Studies

In an observational cohort study, the outcomes of a patient's chosen therapeutic course are analyzed retrospectively, and then artificial treatment categories are created for comparison. For example, outcomes of patients who initiated therapy when their CD4 counts were between 200 and 350 cells/mm^3 of blood may be compared to similar outcomes for patients who initiated therapy with a CD4 count greater than 350. Note that in this type of study, there is no attempt to actually place patients into different treatment groups (which is the case for randomized controlled trials). Relevant studies include [2, 8, 35, 61, 80, 106, 124, 126, 132, 168]. We review just some of them here.

Lepri et al. [106] analyzed a cohort of patients who initiated HAART for the first time between 1997 and 1998. They grouped patients according to three categories of CD4 count

around the time of therapy initiation (referred to as the baseline CD4 count): ≤200, 201-350, and >350. The authors found no significant difference in the risks of virological failure between patients initiating therapy from the two higher CD4 categories. However, they did find a notable difference in the risk between those initiating from the lower two categories.

Phillips et al. [132] aggregated patients from three cohort studies, and considered the CD4 categories <200, 200-349, and ≥350, along with three categories of baseline viral load (<10,000, 10,000-99,999, and ≥100,000 copies/mL of blood). They did not find a significant correlation between lower baseline CD4 count and poorer virological outcomes. However, they did find that patients initiating therapy with a viral load greater than 100,000 had slower rates of obtaining viral suppression.

Egger et al. [61] analyzed 13 cohort studies and found that baseline CD4 category (<50, 50-99, 100-199, 200-349, ≥350) was highly associated with progression to AIDS or death. Baseline viral loads above 100,000 were also associated with clinical progression.

Many cohort studies compare survival rates of patients from the time they initiate therapy from, say, CD4 category 200-350 with the survival rates of patients from the time they initiate therapy from CD4 category <200. Palella et al. [126], however, compared survival rates of patients from the time they initiated therapy from the higher CD4 category with the survival rates of patients from the time they delayed therapy from those same CD4 categories and initiated therapy from lower categories. In other words, the authors compared patients from the same starting point to estimate the effect of earlier versus delayed therapy. This consideration of the survival of patients prior to initiating therapy (also referred to as lead time) in conjunction with their survival after initiating therapy is essential for informing the question of when to initiate therapy. The authors compared mortality rates for patients who initiated versus delayed therapy from CD4 categories <200, 201-350, 351-500, and >500. They found a clear survival benefit for patients who initiated as opposed to delayed therapy in the CD4 category 201-350, and found that initiating versus delaying therapy when the CD4 count is between 351 and 500 may also confer a benefit (results for the latter were not statistically significant, with a p-value of .17). With respect to achieving an undetectable viral load during three to four years of follow-up, the authors did find strong support for patients initiating versus delaying therapy in the 351-500 range.

Sterling et al. [168] also considered lead-time effects in comparing patients who initiated therapy when their CD4 counts were between 350 and 499 with those who delayed therapy. Kaplan-Meier curves of the time until a new AIDS-defining illness or death for patients in the former group appeared better, though there was not a statistically significant difference ($p = .21$). Based on these results, the authors support the guidelines policy of waiting until the CD4 count falls below 350 before considering initiating HAART. For other studies that consider lead-time effects, see [2, 8].

2.1.2 Randomized Controlled Trials

In a randomized controlled trial (RCT), patients are randomly assigned to different treatment groups, which reduces the selection bias inherent in observational cohort studies. The National Institute of Allergies and Infectious Diseases recently conducted an RCT, called Strategies for Management of Antiretroviral Therapy (SMART), to compare two approaches to HAART management: taking therapy immediately or waiting until the CD4 count falls below 250 [44]. Patients in the latter group who initiate therapy also discontinue therapy if their CD4 rises above 350. This trial was stopped in January 2006 because investigators found that patients delaying and interrupting therapy experienced twice the risk of developing AIDS or dying compared to patients taking therapy immediately and continuously [62, 120].

Robbins et al. [144] conducted an RCT that considered which three-drug regimen to use as initial therapy. Within the same trial, Shafer et al. [155] examined the effectiveness of initiating therapy with a four-drug regimen as compared to two sequential three-drug regimens. Results supported the use of a particular three-drug regimen over others and did not find support for initiating therapy with a four-drug regimen.

Martinez et al. [112] led an RCT to test the effect of a proactive switching policy that switches between two drug regimens every three months, even if the viral load is still suppressed at these times. They found that the switching policy yielded better virologic

outcomes than the standard of care policy that switches therapy only when virologic failure is detected. This trial was inspired by the mathematical modeling work of D'Amato et al. [51], described more in Section 2.3.1.

An RCT conducted by Stebbing et al. [165] found that the recycling of certain drugs in heavily pre-treated patients led to significant drops in viral load over a twelve week study period.

Relationship to Present Research

In the introductory chapter, we noted that clinical studies should be carried out before making any significant change to a treatment strategy. Therefore, we do not suggest that the results of mathematical models, such as those developed in the following chapters, be used directly to effect immediate change in treatment policy. Rather, we believe that well constructed models provide efficient means for determining a feasible set of treatment options to examine in a clinical study. Therefore, it is our hope to present our models, receive comments on them, and refine them to the point that they are trusted tools to motivate clinical studies, which in turn may justify changes to treatment policy.

2.2 NATIONAL GUIDELINES POLICIES

The following summarizes the most recent DHHS guidelines on initiating, switching, and sequencing HIV therapies [123]. These guidelines are often revised as the results of new studies, such as those discussed above, become available. Before proceeding, it will be helpful to describe a little more about HIV therapies. There are three major classes of drugs commonly used to combat the virus: Nucleoside Reverse Transcriptase Inhibitors (NRTIs), Non-nucleoside Reverse Transcriptase Inhibitor (NNRTIs), and Protease Inhibitors (PIs). Currently, there are approximately 13 individual drugs within the NRTI class, 3 in the NNRTI class, and 9 in the PI class. For more information on the different drug classes, refer to [121].

2.2.1 Initiating Therapy

If a patient has had an AIDS-defining illness or experiences severe HIV symptoms, then the guidelines indicate a patient should initiate therapy. Otherwise, the recommendations are primarily driven by the patient's CD4 count. When it is below 200, therapy should be initiated, and when it is greater than 350, therapy should be delayed (although if the viral load is greater than 100,000 copies/mL of blood, some clinicians recommend initiating therapy). For CD4 counts between 201 and 350, the situation highly depends on the physician and patient. For example, a physician may recommend a delay in therapy if the patient may have trouble adhering to it.

Based on results of clinical studies, the DHHS recommends specific regimens to start with and regimens that should be avoided. Recommended combination therapies contain a "nucleoside backbone," which consists of two drugs within the NRTI class. For initial therapy, the third drug should come from either the PI or NNRTI class (refer to the DHHS guidelines for names of specific drugs recommended).

2.2.2 Switching Therapy

The DHHS guidelines for when to switch antiretroviral therapy are much more complex than their guidelines on when to initiate. For example, there are different ways in which therapy may be thought of as failing: virological, immunological, and clinical. With treatment-naive patients (patients never before on treatment), virologic failure occurs when there are repeated viral load measurements of greater than 400 after 24 weeks or greater than 50 after 48 weeks. Virologic failure also occurs if the viral load is repeatedly detected after being suppressed below detectable limits. Immunologic failure is defined as the failure of the CD4 count to increase by 25-50 cells/mm^3 of blood above the baseline count over the first year of therapy or a decrease of the CD4 count below the baseline at any time. Finally, clinical progression is the occurrence of an AIDS-related illness or death at least three months after being on a regimen. The three types of failure have distinct time courses and may occur independently of one another. Typically, though, virologic failure is followed by immunologic failure and then clinical progression.

If some type of treatment failure is observed, the guidelines recommend that physicians investigate the likely reason for the failure. If the problem seems to be with the patient's adherence to therapy, the physician may consider simplifying the regimen by decreasing the quantity of pills or the frequency of taking them. For problems with side effects, the physician may consider treating the symptom directly, changing from one drug to another within the same class, or changing drug classes. For drug absorption problems, the physician should check that the patient is taking the drugs in accordance with meal restrictions and that there are no interactions between the patient's HIV therapies and other medications or dietary supplements. If the above do not present difficulties, then the likely reason for failure is the development of drug resistance, and physicians should obtain resistance test results while the patient is taking the failing regimen.

The guidelines state, "there is no consensus on the optimal time to change therapy for virologic failure." An aggressive approach would be to change therapy immediately for any repeated detectable viral load levels, while some approaches allow detectable viral load up to some level (e.g., between 1,000 and 5,000 copies/mL of blood). If drug resistance is present, then the guidelines suggest switching therapies earlier rather than later, though this is vague. For patients who have been on many therapies, the guidelines indicate that keeping them on the same regimen may be reasonable if few treatment options remain. However, if the CD4 count is less than 100, it is recommended to switch therapies if any effective ones remain. Even with no treatment options remaining, it may benefit patients to remain on a failing therapy, as some evidence suggests that HIV is less fit as a mutated strain compared to the original "wild-type" strain [117, 137, 141]. If the viral load is suppressed but there is apparent immunologic failure, perhaps just one drug may need to be changed. In the case of viral load suppression with clinical progression, an immune reconstitution syndrome may be present (an inflammatory response sometimes triggered when a patient's immune system improves [36]), which may respond to anti-inflammatory treatment instead of a change of therapy.

Because drug resistance tends to be cumulative for a patient, upon deciding to switch therapy, all prior treatment history and resistance test results should be considered to ensure proper selection of effective drugs [123]. The guidelines recommend trying to find at least two drugs that should be active against the current strain of virus. If that does not work, then adding a PI with the drug ritonavir may be considered.

The DHHS guidelines emphasize careful consideration before switching to a new therapy, as there are a limited number of effective regimens available to a patient. Therefore, although maximal suppression of viral load is one stated goal, one needs to balance an inclination to switch out a failing therapy against the preservation of a limited supply.

2.3 MATHEMATICAL MODELS

We next describe relevant literature in mathematical modeling. We begin by covering models directly related to HIV therapy planning. Then we discuss well-studied areas of operations research that, to our knowledge, have never been considered in the context of drug therapy management despite having some natural connections. These areas are machine maintenance and inventory depletion management. We also describe a class of problems that do not have natural connections to therapy planning but which includes our models for the optimal time to initiate HIV therapy. These are known as optimal stopping problems.

2.3.1 HIV Models

Several continuous-time control-theoretic models address optimal HIV therapy planning [7, 20, 21, 29, 33, 66, 87, 89, 97, 98, 181, 182]. Each of these models attempts to reduce viral load to low levels or maintain CD4 counts within desired ranges. For example, Jeffrey et al. [87] took a control-theoretic approach for evaluating the effect of initiating HIV therapy during the acute infection stage, the asymptomatic stage, or the later stages of the disease. They found that it is easiest to control viral load when it is at high levels during the acute infection stage.

Wein et al. [182] took a control-theoretic approach for deciding when to switch combination therapy and which therapy to use next. Using ordinary differential equations to model the progression of uninfected CD4 cells, viral strains that may emerge, and CD4 cells infected by each viral strain, their objective was to minimize the total viral load over some time horizon, T, where at each time point, t, one of a finite set of therapies may be used. Two key assumptions were 1) that therapy can be changed at any instant of time, and 2) that a clinician can perfectly observe a patient's levels of CD4 cells and viral strains. Due to the difficulty of finding a closed-form solution to their control-theoretic formulation, the authors used approximation methods and dynamic programming to derive switching policies. Their proposed strategy was a dynamic index policy that at each point in time finds the drug regimen that yields the largest index, where the index value for a certain regimen represents that therapy's effectiveness in preventing viral load increases. Using a simple example of two viral strains and two drug therapies, they used a Monte Carlo simulation to evaluate various policies including a static policy (of using just one therapy for the entire time horizon), a simplified dynamic index policy, and the original dynamic index policy. The two dynamic index policies yielded similar results and performed significantly better than the static policy.

Brandt and Chen [29] modeled HIV, CD4 cell, and CD8 cell (another type of white blood cell) dynamics using nonlinear continuous differential equations, and they used time-delay feedback control to suppress viral load below detectable levels. Noting that actual patient measurements are typically taken one to three months apart, they compared a more realistic discrete-time version of their control problem with the continuous version and found that the former performed quite well. The authors mentioned that feedback control, which lowers drug doses as viral load declines, has the benefit of reducing side-effects from the drugs and possibly increasing adherence.

Berman and Dubin [21] considered both a continuous-time deterministic model and a discrete-time stochastic model for choosing the CD4 count at which to initiate therapy so as to maximize the duration of time the CD4 count remains above a certain level (such as 200). The deterministic model assumed that the CD4 count declines linearly with some slope m_1 until therapy is introduced, at which point it levels out for an amount of time

that may depend on the CD4 count. This plateau is supposed to represent the duration of the therapy's effectiveness. When this time expires, they assumed the CD4 level continues to decline linearly with another slope m_2. They found relationships between m_1, m_2, and the duration function that determine whether initiating therapy earlier or later is better. Their stochastic model considered random errors around the CD4 measurements at equally spaced point in time and produced similar results to their deterministic model. Berman [20] expanded on this model by forming a continuous stochastic model that employs diffusion processes to model the change in CD4 count. Because these models were developed in the pre-HAART era, they considered only one therapy (AZT).

Kamina [91] considered a patient's immune response to HIV in deriving growth models of the virus. Recent studies demonstrate the importance of the immune system response in fighting HIV (see [91] for a list of such studies). Most of Kamina's research was devoted to developing a set of differential equations that describes the dynamics of the uninfected CD4 cells, the various viral strains, the actively and latently infected CD4 cells, and the immune cells that attack the various viral strains. Kamina did, however, use Monte Carlo simulation to test three different treatment strategies: 1) using a single drug, 2) alternating two equally effective drugs at predetermined switching times, and 3) using those two drugs and switching only when the resistant mutations exceeded a certain level. The second and third strategy outperformed the first one, while the resistance threshold affected which of the second and third strategy was better.

D'Amato et al. [51] developed a Monte Carlo simulation to evaluate the risks associated with switching therapy too early or too late, based on various switching policies. They assumed that viral load dynamics followed deterministic equations and that the viral load measurements were subject to random error. Then they compared the different policies by evaluating two measures of performance: the probability of switching therapy prior to the viral load nadir, and the mean time until changing therapy for those patients who switch after the nadir. Since the desire to minimize both these measures presents a conflict (making one of these small tends to make the other one larger), they searched the parameter space in discrete jumps to find the parameters that minimize the mean time until post-nadir switching subject to a maximum probability of pre-nadir switching.

D'Amato et al. [52], working within the above framework, developed approximate, closed-form expressions for the two performance measures discussed above. They then tried to obtain the policy that minimizes the expected time between viral rebound and regimen switch, subject to a maximum limit on the probability of pre-nadir switching. Their optimization problem was too difficult to solve exactly, so they used the simulation to evaluate the various policies.

Freedberg et al. [67], using data from major clinical trials, developed a Monte Carlo simulation model of the progression of viral load and CD4 count before and after patients initiate HAART. They performed various cost effectiveness analyses such as deciding whether or not three-drug therapy is cost effective compared to no therapy at all and evaluating the cost effectiveness of initiating three-drug therapy at various CD4 counts. Their model predicted that three-drug therapy is cost effective and that initiating therapy at a CD4 count of 500 was more cost effective than initiating at a CD4 count of 350 or 200.

Richter et al. [143] built a Monte Carlo simulation of HIV patients undergoing up to three sequential drug regimens. Their model considered progression to AIDS and mortality as functions of CD4 count. They incorporated quality-of-life and cost measures into their model and used it to test a hypothetical example comparing patients who do not take therapy with those who do. A drawback of their model is that they only considered an immediate one-time gain of CD4 with successful therapy (as opposed to an upward stochastic progression over multiple time periods).

Braithwaite et al. [26] also developed a Monte Carlo simulation of HIV progression, which differs from other models by explicitly considering patient adherence to therapy and the development of resistant mutations of the virus. These are important features, as adherence and resistance are major factors in the success or failure of HIV therapy [11, 18, 34, 37, 38, 63, 71, 115, 122, 123, 129, 147, 163]. The model has been used to estimate the impact of alcohol consumption on HIV patient survival [27], to estimate the proportion of HIV patients who die of comorbid diseases [25], and to explore the relationship between adherence and mutation accumulation [28].

Relationship to Present Research

Although ours is not the first mathematical modeling approach to HIV therapy planning, we explore optimization models that capture clinical realities not covered by other models. For instance, the control-theoretic models did not consider the possibility of patient death nor model quality of life. Yet one of the studies assumed an infinite time horizon [182]. Also, with the exception of [20, 21], the control-theoretic models assumed a set of deterministic differential equation models and did not consider stochastic progression of patient health. Furthermore, the models implicitly assumed continuous patient monitoring and changes of therapy, which is not realistic in the HIV patient setting (such a framework is more suitable, for example, for surgery patients receiving anesthesia infusions [84]).

The papers by D'Amato et al. [51, 52] on when to switch therapy did not consider the issue in the context of optimizing patient outcomes. Rather, based on a proposed switching strategy, their models considered tradeoffs between two undesirable events: switching therapy earlier than necessary vs. later than one should. However, their models did not consider the possibility of patient death or factor in long-run outcomes associated with switching therapies. Furthermore, they made strong assumptions of deterministic viral load decay and rebound and did not consider fluctuations in these values for reasons other than measurement error.

Here we describe in more detail the advantages and disadvantages of Monte Carlo simulation modeling versus MDP modeling and how they may complement each other. An MDP solves for an optimal decision policy with respect to a single outcome of interest and returns the exact value associated with the optimal policy. Moreover, the solution of an MDP rarely requires an exhaustive search over all the possible policies. On the other hand, a simulation takes as input a single decision policy and approximates various outcomes of interest by averaging the results of many simulation replications. Although simulation models are excellent tools for testing a variety of "what if" scenarios, to seek an optimal decision policy through simulation, one would have to exhaustively test all possible decisions across multiple time periods and compare the results of multiple replications. Furthermore, because the simulation results are estimates, choosing the policy that yields the best outcome is subject to a probability of being the incorrect choice (see [103] for further discussion of this topic).

A clear advantage to simulations, however, is that they can accommodate considerably more detail than MDPs and can easily generate a multitude of outcomes. As a result of the advantages and limitations of both methodologies, Monte Carlo simulations and MDPs may be used together to inform decision making for complex problems. For example, a highly detailed simulation model may suggest what outcome is most important and sensitive to the decision making process. Then we can solve a simpler MDP model for the optimal policy with respect to that outcome. In turn, the resulting policy may be used in the more detailed simulation for further evaluation of a variety of outcomes not optimized for by the MDP.

2.3.2 Machine Maintenance Models

There are well over 1,000 papers discussing optimal policies for maintaining a system subject to stochastic deterioration over time. Most of these are discussed or listed in various surveys and bibliographies of the literature [86, 114, 135, 161, 177, 180]. Surveys and discussions of applying these techniques to real problems can be found in [55, 149]. Also, classic books discussing this subject include [12, 88]. Unless otherwise noted, the descriptions below are based upon aggregating concepts from these sources.

In nearly every machine maintenance optimization model, a system undergoes stochastic deterioration, and actions are available to alter the state of the system to either avoid system failure or attend to failures. The actions are chosen to meet an underlying objective such as minimizing the long-run average cost of the system or the long-run fraction of the time the system is inoperable. Beyond this basic commonality, there are various ways one can dichotomize the research on machine maintenance optimization. For example, one such classification is whether the model is a "preventative" or "preparedness" maintenance problem. In preventative models, it is assumed that one always knows the state of the system and therefore knows when a failure occurs. The decision problem arises because performing a repair or replacement after a failure is costlier than doing so beforehand. The problem then becomes one of choosing the best time to replace or repair a machine. Two types of models often considered are the "age-replacement" and "block-replacement" models. In the age-replacement model, a machine is replaced whenever it reaches an age of T or whenever it

fails. A typical objective is to determine the replacement time, T, that minimizes the long-run average cost of system. With the block-replacement model, replacements are scheduled at calendar times $(T, 2T, \ldots)$ and upon failures. These types of policies are generally considered for systems with multiple parts in which advantages from economies of scale can be obtained by ordering more than one part at a time. In preparedness models, the state of the system is unknown until an inspection or replacement action is performed. Upon inspection, if the system is in a failed state, a replacement is made. Otherwise, the decision maker can choose to perform maintenance on the machine or choose the next time to inspect the machine. There is a tradeoff between frequent and cheaper inspections and infrequent inspections with a costly penalty for undetected system failures.

Another dichotomy is between systems with a single stochastically failing part and systems with multiple stochastically failing parts. For single-part systems, it is assumed that the system depends entirely on the functioning of one vital part; when that part fails, the system fails. For multi-part systems, the system may fail when some number of parts fail, and policies indicate inspection and/or maintenance actions to take for each of the parts. The actions to take for some parts may depend on the state of others, and savings may be realized by replacing several parts at once (as with the block replacement policy mentioned above).

The literature can also be separated into models for which an endless supply of replacements (and infinite time horizon) is assumed versus those with a finite supply (and a finite or random time horizon). Most models consider the infinite replacement model and assume that replacements behave stochastically identical and independent of one another (referred to as a renewal process) which is fairly easily solved for an optimal stationary policy. However, McCall [114] discussed finite time models for which the renewal assumptions do not hold and stationary policies are no longer optimal. Rather, upon each maintenance action, a new optimal policy is obtained that depends on the remaining time left. Derman et al. [58] considered a finite number of replacements available, and we discuss this paper in detail in Chapter 6. We discuss other models that consider a finite stock of replacements in Section 2.3.3.

Another important dichotomy in the machine maintenance literature is between models with two states (functioning or failed) and models with multiple states. In the former case, one knows only whether or not a machine is operating. In the latter case, models typically contain a failed state, and various degrees of functioning, from a new machine to increasing levels of system deterioration. The literature shows two distinct approaches to modeling the two-state vs. multi-state models. The two-state models generally take a lifetime distribution approach in which the cumulative distribution functions describing machine lifetimes are assumed known and the decision maker then chooses the best time along those distributions to perform a maintenance action or inspection. The framework for these models generally implies that the decision maker sets a policy a priori that will not change over time. Solution approaches for the two-state modeling approach may involve a combination of dynamic programming and calculus-based techniques for finding minima or maxima of functions. We discuss scheduling HIV therapies in this context in Chapter 6.

The two-state modeling approach just described is in contrast to the MDP modeling approach generally taken in the multi-state models [56, 83]. These models assume that any information required to predict future system outcomes is contained in the description of the current state of the system and is independent of past states. Whereas optimal policies for the two-state models described above indicate a definite action to take at some predetermined time t, the optimal policies for the multi-state models typically are of the form: "if you observe that the system in state s_i, take action a_i." In other words, the MDP allows for using updates of the system state to make decisions adaptively instead of a priori. Authors typically seek conditions that guarantee optimal structural policies, such as control-limit policies that say a machine should be replaced if it is in some state or worse and to do nothing until the next period otherwise. We discuss scheduling HIV therapies in this context in Chapter 7.

Klein [100] gives two justifications of the Markovian assumption in these models. The first considers the observation of approximate exponential lifetime distributions for various electronic components. From the memoryless property of the exponential, the Markov

property follows. The other justification relates to the general practice in classical dynamics of describing a physical system sufficiently well in order to predict outcomes based on the present state, regardless of the path taken to get there.

Machine maintenance models are also separated by whether model components are assumed to be known completely or partially. For example, in the two-state models, the lifetime distribution of the system may not be known with certainty. In such cases, minimax policies may be sought, or Bayesian updating may be employed [88]. For the Markov models, there may be uncertainty as to the true state of the system, in which case one can use solution techniques derived from MDPs, known as partially observable Markov decision processes (POMDPs) (for a survey of POMDPs, see [118])

Shock models represent one modeling approach that does not fall neatly into any of the dichotomies mentioned above (for a review, see [177]). These models describe a system subject to shocks occurring at random times, causing a random amount of damage to the system. The time between shocks and the amount of damage may depend on the damage accumulated until some time, t. System damage is modeled through failure probabilities and costs. Again, replacements made before system failure cost less than replacements made after system failure, and all replacements act like new with identical stochastic deterioration. Most authors seek optimal control-limit policies in which a replacement is made when the accumulated damage reaches some threshold level or worse (unless a failure occurs before that time, which leads to system replacement).

Relationship to Present Research

Although there is significant potential to apply ideas from machine maintenance theory to health care, there are few examples in the literature. In his 1965 survey, McCall [114] points out that medical applications of the analyses presented in his paper are virtually nonexistent. Little has changed since then. Christer and Scarf [40] considered maintenance optimization theory in the context of deciding when to replace medical ventilator equipment. Although the burden to patients was considered using a penalty cost for inaccessible equipment, the problem was primarily a cost minimization problem rather than a patient health optimization problem. Some authors have applied inspection policies to the problem of scheduling medical exams. For example, Lincoln and Weiss [109] discussed the scheduling of cervical cancer

examinations and the mean time until detection of a tumor under a given scheduling policy. They sought the maximum inter-examination times that satisfied various objectives regarding the time until tumor detection (e.g., the policy needed to ensure that the mean time until detection of a tumor be less than some time T). Kirch and Klein [96], Zelen [189], and Biswas and Dewanji [23] took a slightly different approach by framing the detection delay optimization problems in the context of a constraint on the number of examinations.

In their 1976 survey, Pierskalla and Voelker [135] discuss how the theory of machine maintenance can be applied to human maintenance, thereby cutting costs and prolonging patient lifetimes. Just as a systems manager has options for improving the "health" of a system, physicians generally have options which can "repair" the health of a patient in the short term. To our knowledge, Alagoz et al. [5, 6] are the only ones to have made an explicit connection between human health optimization and machine maintenance planning (with regard to liver transplantation). In subsequent chapters, we draw connections between machine maintenance problems and HIV therapy planning.

2.3.3 Inventory Depletion Management

Another area of research that has connections to HIV therapy planning is inventory depletion management (IDM), also referred to as optimal issuing policies (see, for example, [30, 31, 57, 64, 108, 134, 133, 188]). The basic framework of these problems is that there are a finite number of similar items, with various ages, waiting in a stockpile. Items are selected from the stockpile one at a time to be used in the field, at which time they generate a field life that depends on the item's age at issuance. An item is issued immediately from the stockpile upon the termination of the previous item's lifetime in the field. The objective is then to maximize the field life (or expected field life) obtained by the sequence of selected items from the stockpile. Under varying assumptions about the field life as a function of the item's age at issuance, authors give conditions under which LIFO (last in, first out) or FIFO (first in, first out) policies are optimal. By LIFO, they mean to always select the youngest item and by FIFO, to always select the oldest.

For example, a paper by Brown and Solomon [31] considered n independent items, with possibly different lifetime distributions that are multiplied by a factor $d(t)$ when the item is issued at time t (with $d(0) = 1$). That is to say, at time $t = 0$, item i has a field lifetime distributed as X_i and at time t it has a field lifetime distributed as $X_i d(t)$ (presumably $d(t)$ decreases over time, although that is not required). An optimal policy is one that yields a total field lifetime distribution which is stochastically greater than that of any other policy. They considered distributions at time $t = 0$ that can be ordered in a certain stochastic sense. For example, if X_i and X_j are continuous random variables with common support $[a, b]$, then $X_i \prec X_j$ if for all $a < r < s < b$, $f_{X_j}(s)/f_{X_i}(s) \geq f_{X_j}(r)/f_{X_i}(r)$. Their main result is that if $d(t)$ is positive and strictly concave (convex) then LIFO (FIFO) is the unique optimal policy.

Relationship to Present Research

We may be inclined to adapt the IDM framework to HIV therapy planning by selecting n different HIV therapies from a stockpile in a way that maximizes the expected patient lifetime from the sequencing. Suppose the n therapies have different levels of effectiveness and that we can order them in the sense given above for the patient lifetime distributions they induce at time $t = 0$. Also, suppose we have a function $d(t)$, with a range between zero and one, to indicate a decreasing therapy effectiveness over time (resistance to certain therapies taken can confer resistance to other therapies not yet taken (known as "cross-resistance"), thus decreasing the unused therapies' effectiveness). Then if $d(t)$ is strictly convex, we may say that the optimal policy is to always select the most effective therapy currently available. However, there is a key difference between the traditional IDM problem and a therapeutic optimization problem: the IDM problem assumes that an item can be replaced after its lifetime ends, whereas if a patient's lifetime ends, leftover therapies in the stockpile become worthless. The essence of the therapeutic optimization problem is to delay the end of a patient's lifetime by strategically replacing therapies. If the IDM models had the following two extensions, then they would be more applicable to therapy planning. One is to consider that while item i is activated in the field, there is a probability that it causes system-wide failure in which remaining items in the stockpile become useless. The other

extension is then to consider optimal issuing *and removal* policies that decide which item to issue at what time *and when to remove the item* so as to strategically avoid system failure and maximize the expected field life of the issuing and removal policy.

2.3.4 Optimal Stopping Problems

Optimal stopping problems describe situations in which a decision maker observes a system periodically, accrues rewards and/or costs between periods, and at any period, can stop the process and receive a terminal reward or cost. The decision maker's goal is to decide when to stop the process so as to maximize the total expected reward or minimize the total expected cost [39, 65]. MDPs are often used to solve these problems because of their natural fit. Examples of optimal stopping problems include the secretary problem [140] (when should someone stop interviewing candidates for a job and hire the current candidate?) and the parking problem [138] (when should someone stop searching for a better parking spot and take the next open space?). A recent application to therapy planning is the optimal time for a patient to accept a living-donor [5] or a cadaveric liver transplantation [6].

Relationship to Present Research

Our models for the optimal time to initiate HIV therapy, presented in Chapters 4 and 5, fall into the class of optimal stopping problems. The reward a patient receives when waiting from one month to the next can be defined as that one month of time or one quality-adjusted month. Upon initiating therapy, patients in our model receive a terminal reward equal to the expected remaining lifetime or quality-adjusted lifetime, based on their health state at the time of initiation. When waiting, patients may die and lose the opportunity to receive the larger terminal reward; however, initiating therapy when in a relatively healthy state may be wasteful. The goal, therefore, is to decide the best time to stop the waiting process and initiate therapy.

3.0 STATISTICAL MODELING OF THE NATURAL HISTORY OF CD4 COUNT AND SURVIVAL UPON INITIATING THERAPY

3.1 INTRODUCTION

This chapter discusses the use of clinical data to build a discrete-time Markov model of the natural history of CD4 progression (that is, progression without treatment) as well as a survival model to estimate expected remaining lifetime upon initiating HIV therapy from different CD4 strata [159, 160]. Both components will be used to solve the MDP model of when to initiate therapy discussed in Chapter 4.

Our data are provided by the Veterans Aging Cohort Study (VACS), a prospective, observational cohort study of HIV positive and HIV negative patients from Veterans Health Administration (VA) hospitals across the U.S. [179]. The VA is the largest provider of HIV care in the nation. Our cohort contains 25,550 HIV+ patients with a history of laboratory measurements and 66,840 HIV- patients to draw from for controls (used in the survival model). Because 98% of the HIV+ patients in this cohort are men, we focus our analyses solely on male patients (25,041 HIV+ patients). We discuss the implications of this in Sections 3.4 and 4.6.

Unless otherwise noted, all statistical work is done in the open-source statistical package **R**, available at http://www.r-project.org/ (for references on using **R**, see [50, 178]).

3.2 A DISCRETE-TIME MARKOV MODEL OF THE NATURAL HISTORY OF CD4 COUNT

In addressing the question of the optimal time to initiate therapy in Chapter 4, we shall consider antiretroviral-naive patients; the answer to the question may differ significantly for patients who have prior antiretroviral experience and thus may have developed resistant strains of the virus. We track the progression of patients' CD4 counts because this is the primary variable used to guide when patients should initiate therapy [123, 166]. We focus our analysis on male patients between 40 and 50 years old upon their first CD4 measurement, as this ten year age bracket has the largest number of patients in our cohort. Also, we categorize the CD4 count into four distinct categories: $0-49$, $50-199$, $200-349$, and $\geq$350 cells/mm^3 of blood. Similar groupings can be found in other clinical studies, and these strata allow for results to be interpreted in the context of the guidelines categories of <200, 200-350, and >350 [123] (recall the recommendation that patients initiate therapy when CD4 falls below 200, wait when it is above 350, with debate on what to do between 200 and 350).

Our optimization model considers patients who visit their physician every month to determine whether or not to initiate therapy based on their CD4 count. Although CD4 counts prior to the initiation of therapy generally decline, the actual measurements may vary considerably from month to month. This is seen in the plots of actual CD4 measurements taken over time for two VACS patients, labeled A and B (Figures 1 and 2). Therefore, instead of considering just the mean decrease in CD4 counts over time [116], we want to build a discrete-time Markov model that describes the monthly transition probabilities among the various CD4 categories and death, prior to initiating HIV therapy. Discrete-time Markov modeling is a common technique in medical decision making to predict and track patient progression from one time period to the next, as a function of a patient's current state [16, 145, 164]. The technique is also natural to use when clinical decisions may be considered at discrete time periods. There may be several challenges in constructing transition probability matrices, such as irregular observation times, incomplete data, and censored observations [47, 48, 70, 85, 184]. Details of other discrete-time Markov models of CD4 progression can be found in [1, 32, 68, 110, 111, 113, 154]. Freedberg et al. [68] also considered a Markov

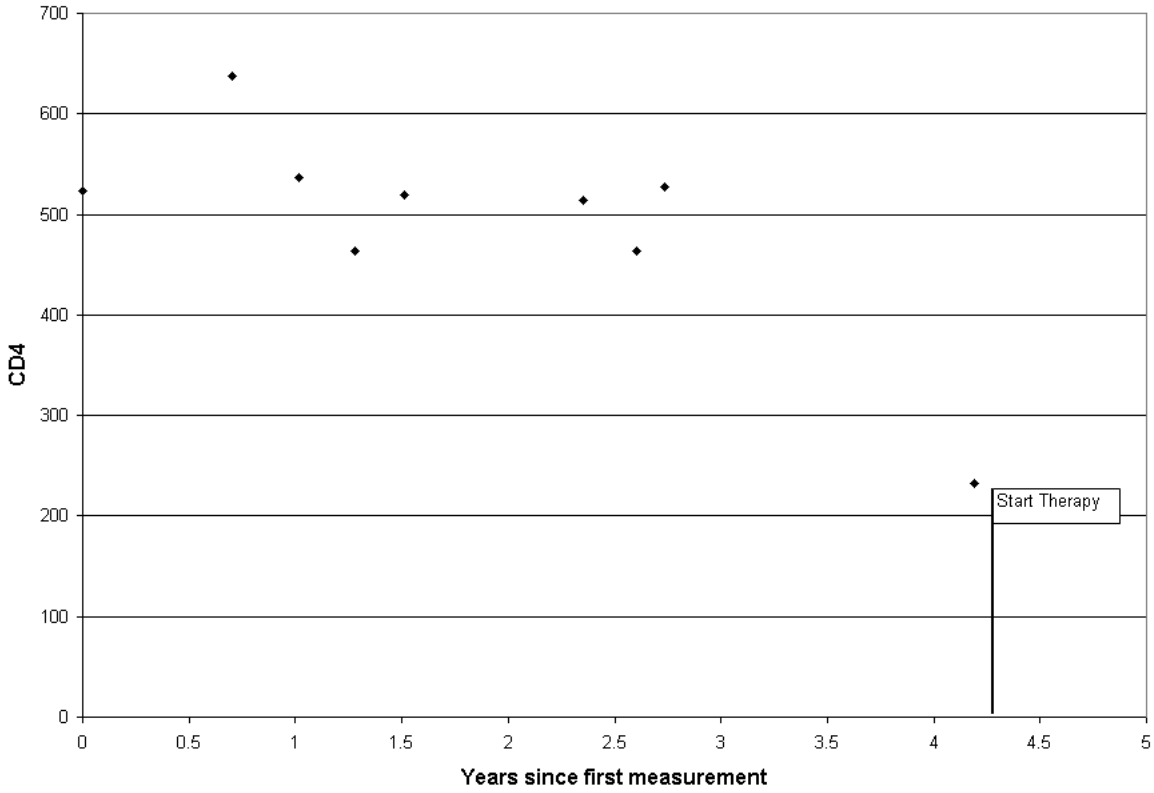

Figure 1: Actual CD4 measurements over time for Patient A

cycle length of one month to reflect the setting for HIV care. Because our MDP model also requires estimates of expected remaining survival upon initiating therapy (discussed in the next section), this chapter develops both parts of the model from a single data source (VACS).

In addition to the variability in the CD4 counts, Figures 1 and 2 also demonstrate irregularity in the time between consecutive CD4 measurements. In fact, many months may elapse between measurements. Therefore, we use a method called smoothing splines to fit a continuous curve to the data in a way that captures the variability of the data [73]. Then we sample CD4 counts from this curve at monthly intervals, categorize them, and build an empirical estimate of transition probabilities among CD4 categories and death from one

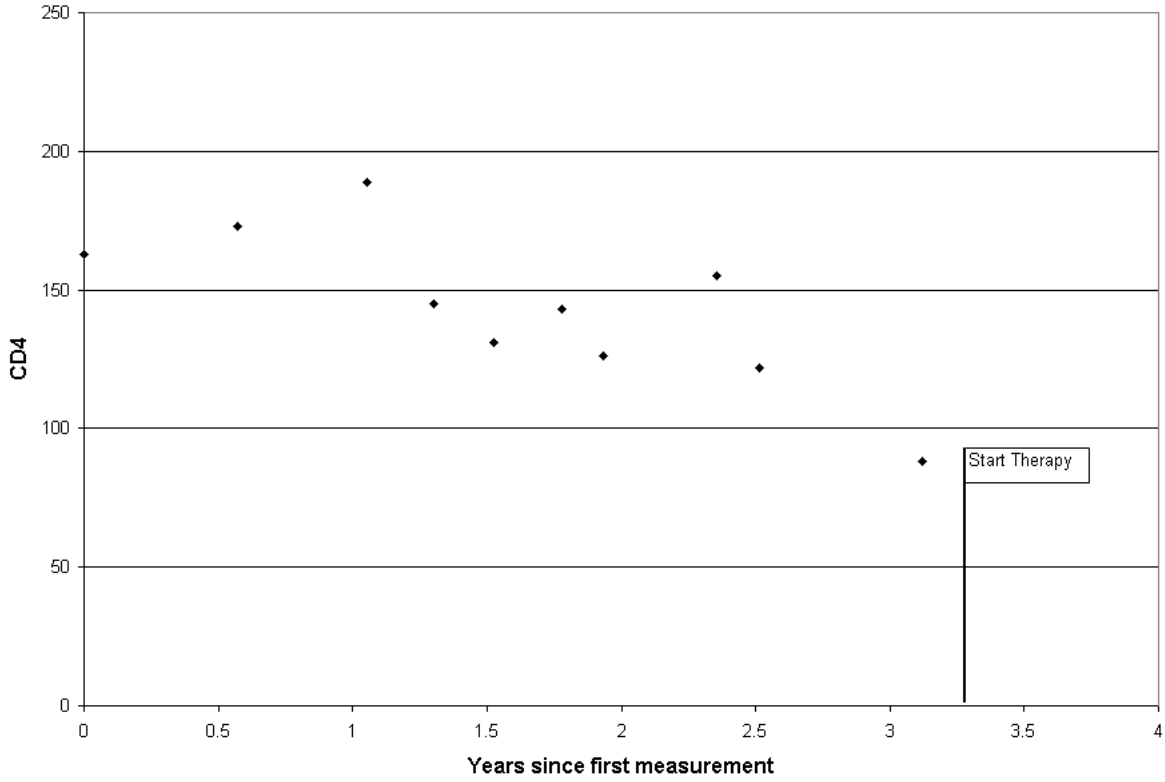

Figure 2: Actual CD4 measurements over time for Patient B

month to the next. A similar approach was used by Alagoz et al. [4] to model the natural history of end-stage liver disease.

A smoothing spline is a curve-fitting technique that allows one to make an explicit tradeoff between how close the curve comes to the actual data (thus reducing the sum of squared residuals) and how smooth the curve is. For example, a linear regression through a set of data is the smoothest of all cubic splines, but with the greatest sum of squared residuals. On the other hand, one can have a cubic spline go through each data point (i.e., interpolate the data), which reduces the sum of squared residuals to zero, but it is considered the roughest of all the splines. The tradeoff can also be made somewhere in between these two extremes. We briefly describe the method more formally, based on [73]. Let $g(\cdot)$ be a twice-differentiable

function and consider a set of n data points $(t_i, Y_i), i = 1, \ldots, n$, along with the following function, called the "penalized sum of squares":

$$S(g) = \sum_{i=1}^{n} [Y_i - g(t_i)]^2 + \alpha \int_{t_1}^{t_n} [g''(x)]^2 dx \tag{3.1}$$

The first term of the sum gives the sum of squared residuals. The component $\int_{t_1}^{t_2} [g''(x)]^2 dx$ measures the roughness of the function g and when multiplied by α, the term is referred to as the "roughness penalty." The parameter α, called the "smoothing parameter", is a user-defined input that represents the tradeoff between how smooth or variable one wants the curve. Under a choice of α, the goal is to find a twice-differentiable function, $\hat{g}$, that minimizes $S(g)$. As $\alpha \to \infty$, $\hat{g}$ approaches the linear regression fit (hereon referred to as the "smoothest spline"), and as $\alpha \to 0$, $\hat{g}$ approaches the interpolating curve (hereon referred to as the "roughest spline"). It can be proven that the function $\hat{g}$ is a natural cubic spline over the interval $[t_1, t_n]$, that is, a piecewise cubic polynomial connected at the points $(t_2, \ldots, t_{n-1})$ (also referred to as "knots"), such that the first and second derivatives of f are continuous at each t_i, and the second derivative equals 0 at t_1 and t_n.

Figure 3 shows the actual CD4 measurements along with three splines for Patient A. Rather than adjust α, one may also adjust a parameter called the "degrees of freedom" (df), which has a direct mathematical relationship with the smoothing parameter α. When df equals 2, we obtain the smoothest spline and when df equals the number of data points, we obtain the roughest spline (labeled as the "0 sum of squared residuals" in the figures). The spline labeled "average", is obtained by setting df to be $(2 + \#datapoints)/2$ (df can take on fractional values). Figure 4 shows the same for Patient B. Because of the significant variability in a patient's CD4 measurements over time, we use the roughest splines. In Section 4.5, we examine the sensitivity of results to the choice of df.

To capture the period of time for which patients are antiretroviral-naive, we proceed as follows. Our data include the date on which the patient first started therapy in VACS, so we consider patients and their lab values prior to this date; if no such date is recorded, then we include those patients and all of their lab values as well. From this set of patients, we exclude anyone whose viral load at any time was below 5,000 copies/mL of blood, as it is

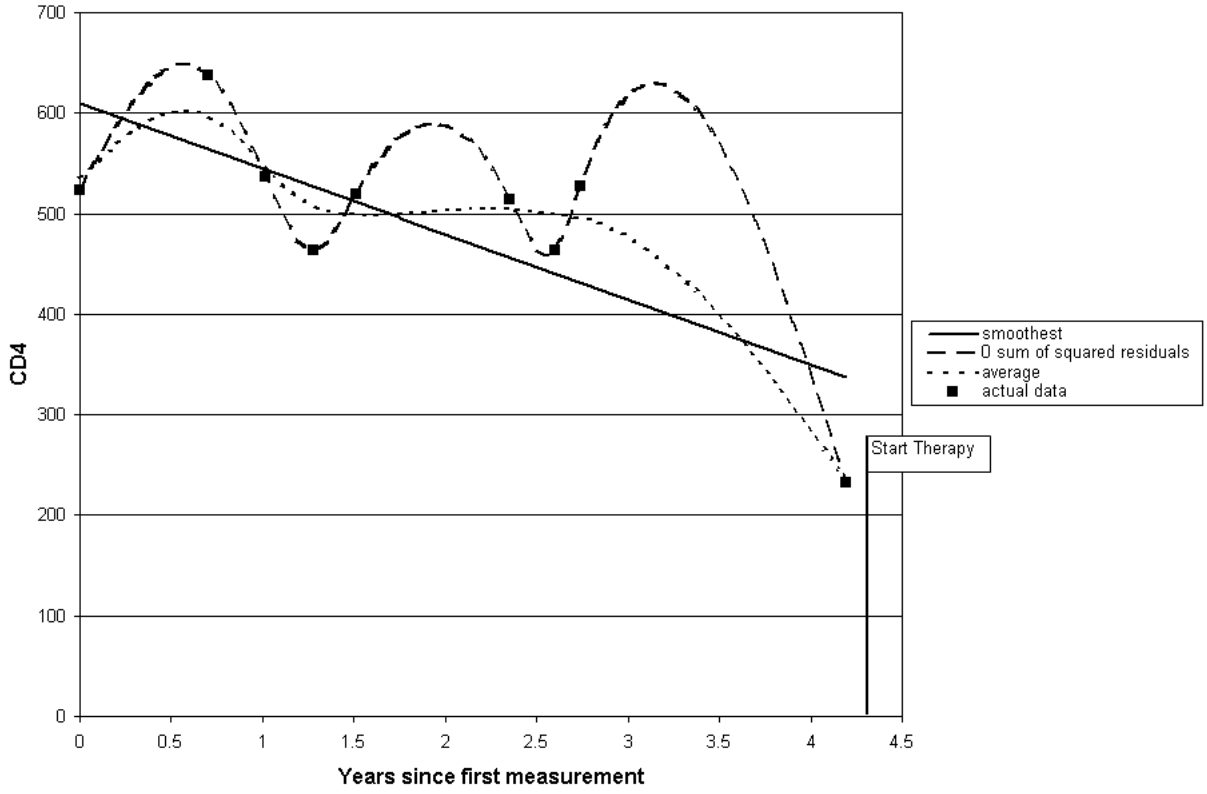

Figure 3: Actual CD4 measurements with splines for Patient A

unlikely that an antiretroviral-naive patient would attain such a low level of viral load [90]. This reduces our patient set from 25,041 to 20,586.

We then build our discrete-time Markov model of the natural history of CD4 progression as follows. Because patients may receive viral load tests at some times and CD4 measurements at others, we keep only those patients for whom CD4 measurements exist prior to the first therapy start date. Over 4,500 patients had viral load measurements but no CD4 measurements prior to initiating therapy, and thus our set reduces to 15,924 patients after this step. Finally, we select patients between 40 and 50 years old upon their first CD4 measurement (recall that this ten year age bracket has the largest number of patients in our cohort), leaving us with 6,749 patients from which to build our natural history model. The

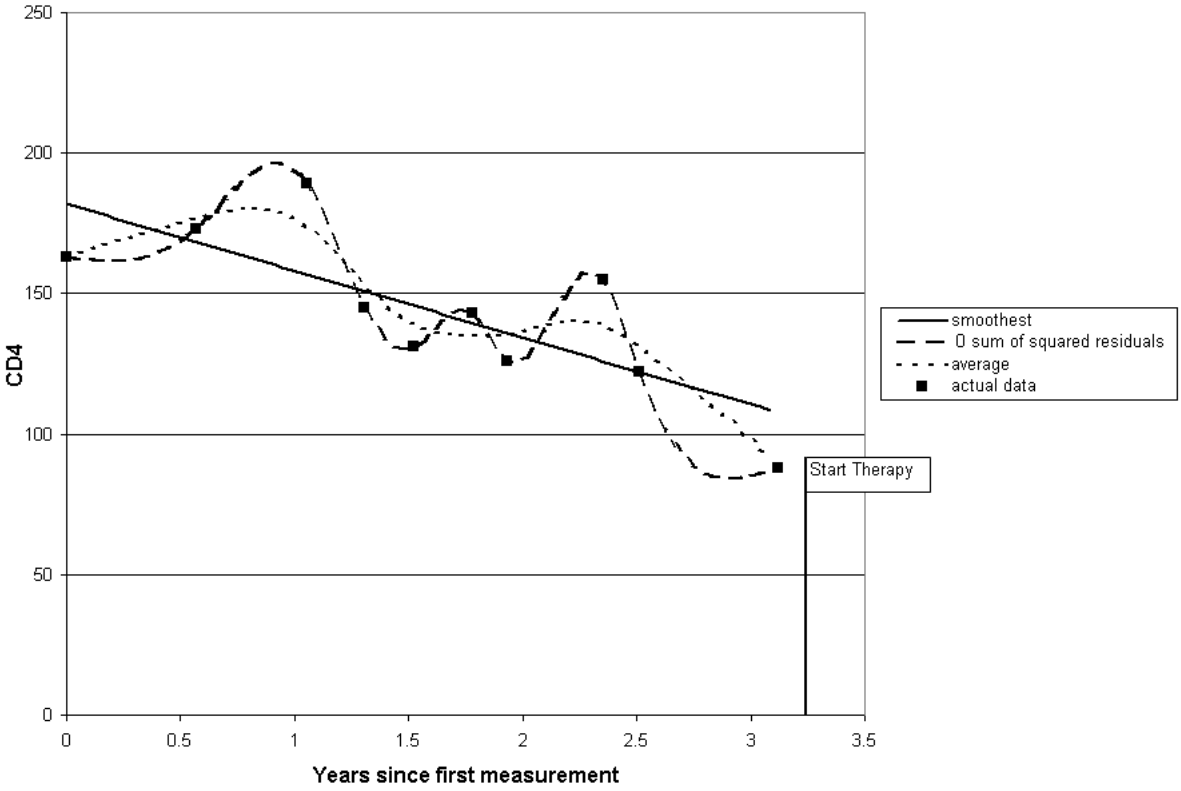

Figure 4: Actual CD4 measurements with splines for Patient B

earliest CD4 measurement among these patients was taken on October 1, 1991 and the most recent measurement was taken February 13, 2003.

For patients with four or more CD4 measurements ($n = 1,351$), we fit a cubic spline (four is the minimum number of data points required to construct a cubic spline); for patients with two or three measurements ($n = 1,999$), we fit a linear regression; and we eliminate patients with only one CD4 measurement ($n = 3,399$). Also, if the time between the patient's final antiretroviral-naive CD4 measurement and the first measurement is less than 25 days, we eliminate the patient from our estimation of monthly transition probabilities. This happens

for only one patient with four or more observations, and happens for 311 of the patients with two or three observations. Therefore, we build our transition probability matrix using 1,350 cubic splines and 1,688 linear regressions.

CD4 counts are non-negative, and CD4 counts greater than 2,000 cells/mm^3 of blood are clinically unrealistic [24, 102]. When building our cubic splines that maximally capture the CD4 fluctuations, the curves may very well extend beyond these limits (this happens for 163 of the splines). When this occurs, we decrease *df* by one to construct a smoother spline and determine if the curve lies within the limits. If not, we repeat this process until we obtain a spline that stays completely within the CD4 range of 0 to 2000. If we reach $df = 2$ (the linear regression) and still find that the the limits are crossed, then we simply cap the estimates at these limits.

After obtaining our curves, we begin tabulating the transitions from the CD4 category at the beginning of one month to the CD4 category at beginning of the next month (or death) that occur as we proceed from the start of one patient's curve to the end of that curve. Our first step in recording the transitions to death is as follows. If the next event beyond the last CD4 measurement is the patient's death, and if this is within about a month of the last CD4 measurement (we chose six weeks, or about 1.5 months, as our cutoff), then we tabulate a transition from that patient's last curve-estimated CD4 category to death ($n = 46$). If the death date is beyond this time ($n = 111$), then we take the average monthly change in CD4 count from the patient's first CD4 measurement until the last, ensure that this is no more than zero (i.e., we do not allow for a positive slope in CD4 counts prior to initiating therapy), and extend the monthly CD4 estimates beyond the final curve-derived estimate by this amount, until arriving within 30 days of the patient's death date. Then we tabulate a transition from the final estimated CD4 category to death.

Figure 5 demonstrates the procedure of tabulating transitions for a hypothetical patient who dies without ever taking therapy. The generated spline for the patient is shown, along with vertical lines representing 30 day intervals along the spline. The table within the figure shows the aggregate monthly transitions from day 0 until the patient dies after day 570. The categories in the first column represent possible ranges of the patient's CD4 count at the beginning of a month, and the categories in the first row show the same categories along with

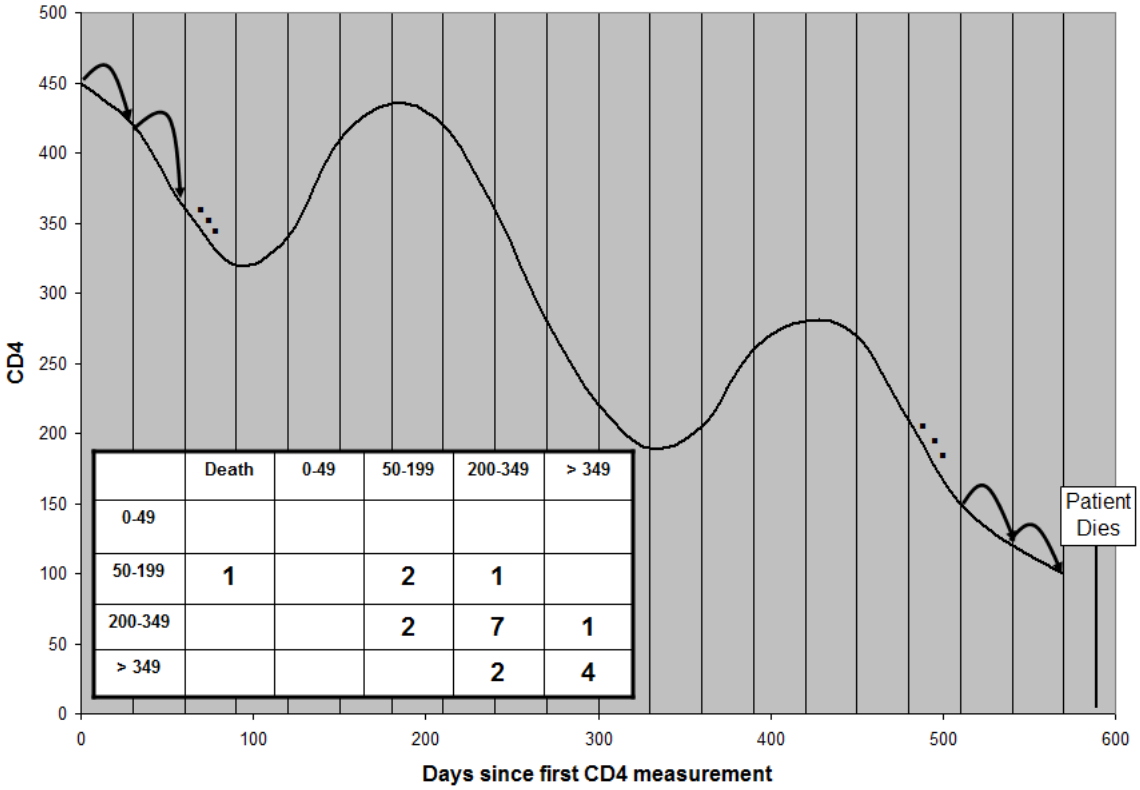

Figure 5: Example of building a transition matrix for a hypothetical patient

the possibility of death by the beginning of the next month. The final transition is from a CD4 count between 50 and 199 to death, and therefore we mark a 1 in the appropriate table cell showing a transition to death. We proceed in this fashion for each patient, aggregate the tables across all patients, and divide each cell count by the row sums to estimate the natural history transition probabilities.

3.2.1 The Problem of Censored Observations

Note that we have only recorded 157 deaths from the 3,038 patients we are examining (5%). The reason for this is the censoring of patients from the natural history progression. Censoring, a concept from survival analysis, describes the situation in which the end-point

Table 1: Percent of monthly transitions from each category that are censored due to the initiation of therapy

	CD4 Categories			
	0-49	50-199	200-349	≥ 350
Censors from Initiating Therapy	7.6%	4.9%	4.4%	2.1%

of interest is not observed for a patient [43]. For example, when tracking the survival of patients prior to initiating therapy, there are three ways their pre-therapy survival may be censored. One is that they have not started therapy and are still alive at the time the data is collected. Another is that for some reason, they leave the study (e.g. moved away). This is known as "loss to follow-up". The third reason is that they initiate therapy, which is a common type of censoring in HIV care because as patients' CD4 counts decline, guidelines suggest they initiate therapy at some point. In other words, few HIV patients go entirely untreated; our data yield 2,690 patients (89%) whose natural history CD4 measurements are censored by the initiation of therapy, and just 191 (6%) who are censored for other reasons (the other 5% are the recorded patient deaths). To tabulate transitions from the various CD4 categories to censors attributable to the initiation of therapy, we consider whether or not the patient initiates therapy within six weeks of the final CD4 measurement and follow a similar procedure as described above (in the case of patient death) for extending the CD4 progression up until the last month before the initiation of therapy. Table 1 shows the percentage of transitions, from each CD4 category, that result in a censored observation attributable to the patient initiating therapy by the beginning of the next month. This percentage is as high as 7.6% for patients with CD4 counts < 50 and decreases monotonically in the CD4 categories to 2.1% for patients with CD4 counts ≥ 350.

Unfortunately, when censors are caused for reasons such as the initiation of therapy at lower levels of CD4 count (referred to as "informative censors"), they invalidate standard methods of survival analysis [43]. Moreover, survival analysis is concerned with estimating

times until events, which is not our primary interest for this part of the model (this will be our main interest in the next section when we want to derive estimates of expected remaining lifetime after initiating therapy). Instead, our objective here is to estimate monthly transition probabilities among the CD4 categories and death. Our challenge, therefore, lies in deciding how to handle these censored observations. It does not appear that other Markov models of HIV have considered this problem (though some of them used cohorts of patients from a time when few or no therapeutic options were available; hence, informative censoring may have been negligible). Alagoz et al. [5] constructed natural history transition probability matrices for patients with end-stage liver disease; however, they did not consider censored observations. Unlike with HIV, in which it is rare for a patient to die before ever receiving therapy, many end-stage liver disease patients do die before ever receiving a transplant, their only viable therapy [5]. Therefore, in that context as well, informative censoring of the natural history progression may not present a significant issue. One study on diabetic retinopathy did address the issue of censors attributable to therapeutic intervention by considering how these censored observations may have been distributed had therapy not been initiated [53]. Also, Craig et al. [47] considered a Bayesian approach for handling informative censors.

First, we demonstrate that it is important to address censored observations. If we ignore the censored observations and estimate remaining lifetimes associated with a policy of never initiating therapy, then for patients whose CD4 counts are in categories $0-49$, $50-199$, $200-349$, and ≥ 350 we estimate expected remaining lifetimes of 14.4, 19.1, 21.7, and 23.6 years, respectively. These estimates are too high for HIV patients who never take therapy. It is difficult to find other estimates of expected lifetime without ever taking therapy; however, one paper reports median survival estimates in the pre-HAART era for CD4 categories ≤ 50, ≤ 200, and >200 of .99, 3.2, and 9.1 years, respectively [95].

Our method for handling the informative censors is similar to that just described in [53]. For each CD4 category, we hypothesize how the censored observations may have been distributed among the various CD4 categories and death, had therapy not been initiated at that time period. We then derive empirical estimates of the natural history transition probabilities. Table 2 presents four different scenarios for redistributing the informative censors. For example, under Scenario 1, which we refer to as our baseline scenario, we

assume that 90% of the transitions from CD4 category 0-49 at the beginning of one month to the initiation of therapy by the beginning of the next month would have resulted in the patient dying had therapy not been initiated, and 10% would have remained in the 0-49 CD4 range. The percentages for the other rows are interpreted similarly. Essentially, the baseline model hypothesizes that for the two lower CD4 categories, a large percentage of the patients would have died that month had they not initiated therapy (though less for the 50-199 category than for the 0-49 category), and for the two higher CD4 categories, a small percentage of censored observations would have resulted in a patient death. For the two higher categories, we hypothesize a higher weighting towards the transitions remaining in the same CD4 category or moving down one (for example, we hypothesize that 40% of the informative censors from CD4 category 200-349 would have remained in that CD4 category, while 60% of the informative censors from the highest CD4 category would have remained in that category). Scenario 2 presents a heavier weighting towards death, Scenario 3 presents a lighter weighting towards death (relative to the baseline), and Scenario 4 presents an even lighter weighting towards death. Note that in each scenario, we assume that a censored observation would not have resulted in the patient's CD4 count improving had therapy not been initiated (shown by the upper triangle of 0s). Table 3 shows the monthly probability of dying from each CD4 category under each censoring redistribution scenario (including the case of ignoring censors discussed above), and Table 4 presents the expected remaining lifetime implied by each scenario for patients never taking therapy, as a function of their CD4 category. We chose Scenario 1 as the baseline because the resulting natural history lifetimes appear clinically plausible, based on the estimates given above, found in [95]. We discuss the optimal initiation policy implications of these scenarios in Section 4.5.

Table 2: Censoring redistributions (Scenario 1: baseline; Scenario 2: heavier weighting towards death; Scenario 3: lighter weighting towards death; Scenario 4: even lighter weighting towards death)

Scenario 1	Death	0-49	50-199	200-349	≥ 350
0-49	90%	10%	0%	0%	0%
50-199	40%	55%	5%	0%	0%
200-349	5%	5%	50%	40%	0%
> 349	2%	8%	10%	20%	60%
Scenario 2	Death	0-49	50-199	200-349	≥ 350
0-49	90%	10%	0%	0%	0%
50-199	70%	20%	10%	0%	0%
200-349	50%	10%	30%	10%	0%
> 349	30%	10%	20%	30%	10%
Scenario 3	Death	0-49	50-199	200-349	≥ 350
0-49	30%	70%	0%	0%	0%
50-199	10%	20%	70%	0%	0%
200-349	2.5%	7.5%	20%	70%	0%
> 349	.9%	1%	5%	13.1%	80%
Scenario 4	Death	0-49	50-199	200-349	≥ 350
0-49	10%	90%	0%	0%	0%
50-199	2.5%	7.5%	90%	0%	0%
200-349	.9%	1%	8.1%	90%	0%
> 349	.1%	1.5%	3%	5.4%	90%

Table 3: Probability of patient death within a month, according to censoring redistribution scenario and CD4 category

Redistribution Scenario	0-49	50-199	200-349	≥ 350
0 (ignoring censors)	.0153	.0027	.0013	.0007
1 (baseline)	.0829	.0221	.0035	.0012
2 (heavier towards death)	.0829	.0366	.0233	.0070
3 (lighter towards death)	.0371	.0075	.0024	.0009
4 (even lighter towards death)	.0218	.0038	.0017	.0008

Table 4: Expected natural history remaining life years, according to censoring redistribution scenario and CD4 category

Redistribution Scenario	0-49	50-199	200-349	≥ 350
0 (ignoring censors)	14.43	19.11	21.67	23.57
1 (baseline)	1.51	2.96	4.68	6.30
2 (heavier towards death)	1.30	2.17	2.95	4.01
3 (lighter towards death)	4.35	7.19	9.21	11.07
4 (even lighter towards death)	8.89	12.93	15.34	17.21

3.3 ESTIMATING SURVIVAL AFTER INITIATING HIV THERAPY

As mentioned, a key component of the MDP model of Chapter 4 is the expected remaining lifetime for a patient initiating HAART from the various CD4 categories. Several authors have compared survival rates across different CD4 categories after patients initiated HAART [8, 61, 80, 126, 139, 167]. However, these studies focused on survival differences in the 3-5 years for which HAART was in use by the time the studies concluded. Because of the great success of HAART, there is a large proportion of patients who are still alive on therapy since starting it in the late 1990s. Figures 6 through 9 show Kaplan-Meier survival curves [93] (along with 95% confidence bands) for 40-50 year old male VACS patients initiating therapy from the various CD4 categories. Because we are interested in lifetime survival estimates, it is necessary to extrapolate survival beyond the 6-7 year limits of our observations. However, any parametric fit to the data will significantly overestimate the expected remaining survival because the data do not reflect the increased age-related (non-HIV) mortality risks these 40-50 year old patients will face as they turn 60, 70, 80, and so on.

Some authors have taken model-based approaches to estimate remaining lifetimes for patients initiating therapy from the different CD4 categories [26, 67, 95, 113]. For example, King et al. [95] report median survival times of 5.5, 8.5, and 15.4 years for patients initiating HAART from CD4 categories $\leq$50, $\leq$200, and $>$ 200, respectively. We describe a statistical approach.

As we did with the natural history development, we start with our cohort of 25,041 male HIV patients and exclude patients that appear to have been on some therapy prior to their recorded date of first therapy (reducing the set to 20,586 patients). To be reasonably sure that patients were initiating HAART (as opposed to monotherapy or double-drug therapy), we select patients whose first therapy start date was on January 1, 1998 or later. Also, because we estimate patient survival without therapy to be less than 10 years, we select patients in the same 40-50 year old age bracket (upon initiating therapy) as with our natural history model. These two steps reduce our set to 2,535. To associate therapy initiation with a CD4 category, we keep only those patients who have a CD4 measurement within 24 days of the first therapy start date, leaving us with 762 patients.

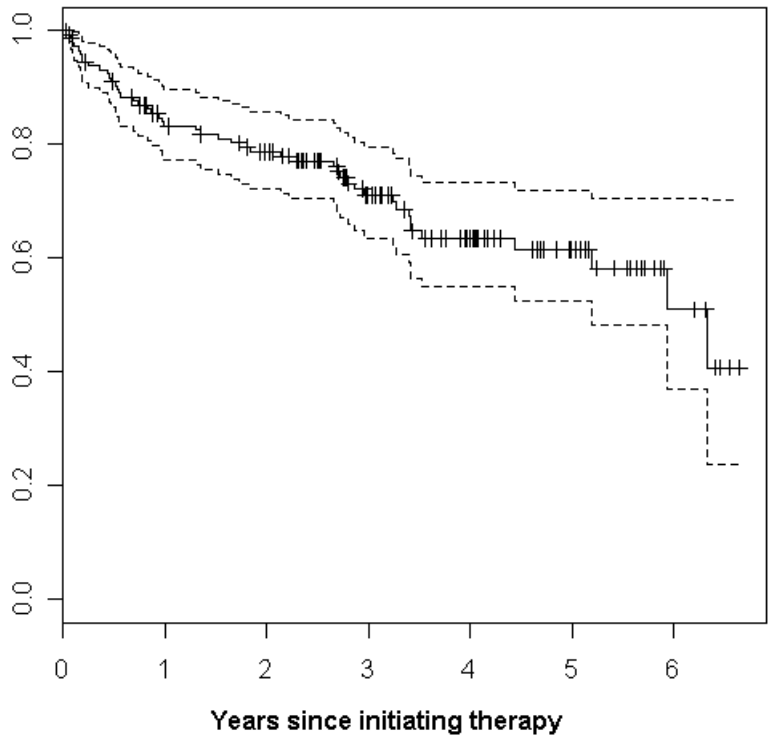

Figure 6: Kaplan-Meier survival curve and its 95% confidence interval for patients initiating therapy from CD4 category 0-49

For each of these patients with a date of death indicated, we record the survival time since the start of therapy. When no date of death is indicated, we record the time between the first therapy start date and the last observation date, and mark these observations as censored. We group each survival or censor time according to the CD4 category near the initiation of therapy to create the Kaplan-Meier survival curves shown in Figures 6 through 9. Table 5 gives a breakdown of the number of patients in each strata, the number of survival times ending in death, and the number of censored times.

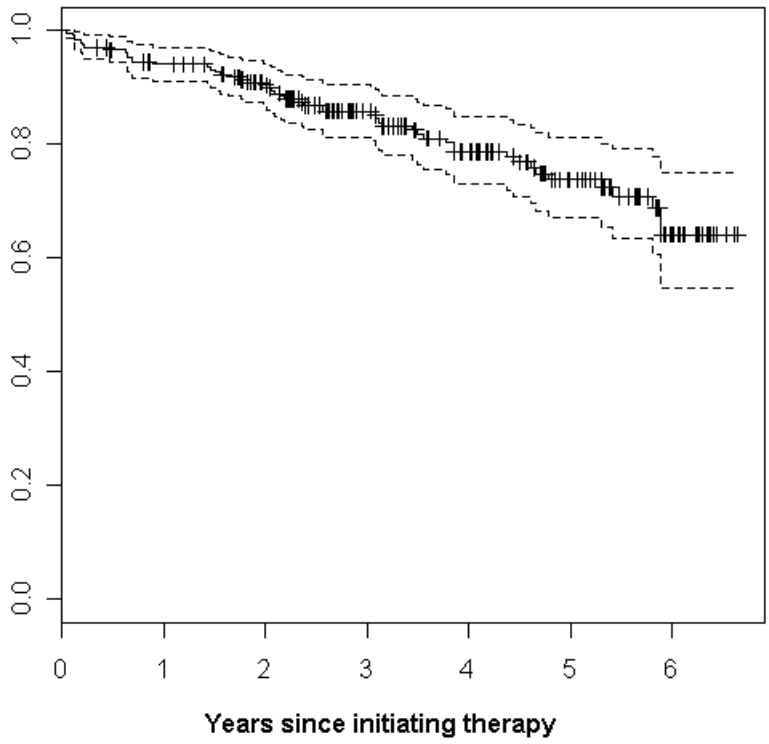

Figure 7: Kaplan-Meier survival curve and its 95% confidence interval for patients initiating therapy from CD4 category 50-199

To address the problem of the 6-7 year limits of our survival curves in deriving estimates of expected lifetimes, we take the following steps. First, we use Cox proportional hazards models [46] to calculate hazard ratios between male VACS HIV+ patients and HIV- controls (further controlling for age and race) across the various CD4 categories (these calculations were performed with the statistical package STATA). A hazard ratio compares the hazard rate (in this case, the hazard rate of death) between two different groups, such as between a "standard treatment" group and a "new treatment" group (in this case, between two different types of VACS populations: HIV+ and HIV- males) [43]. We then apply these hazard ratios

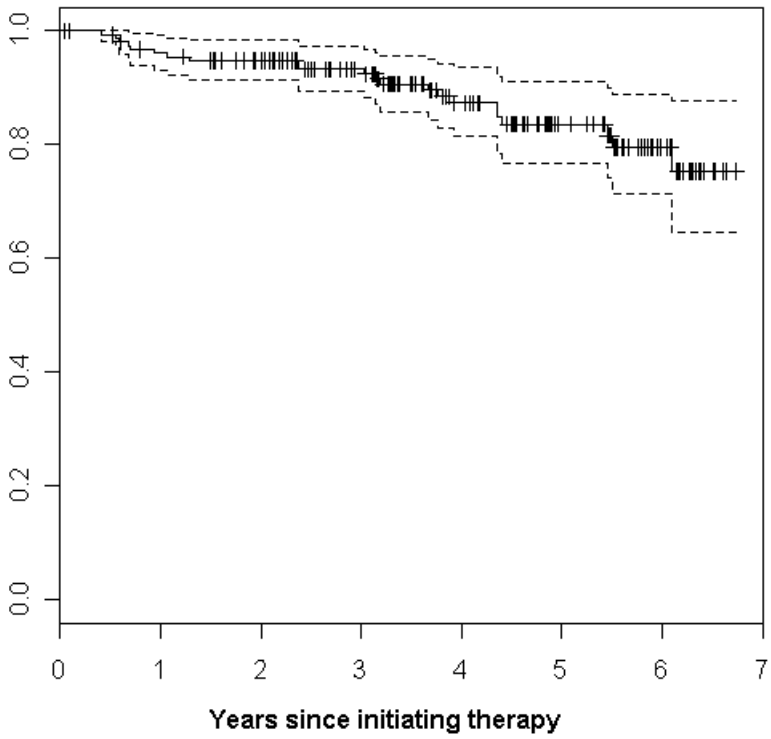

Figure 8: Kaplan-Meier survival curve and its 95% confidence interval for patients initiating therapy from CD4 category 200-349

to standard life table data and use a life table construction method [9] to estimate the expected remaining lifetimes for 45-year-old HIV+ men from VACS who initiate therapy from the various CD4 categories. Table 6 shows the results, with an expected remaining survival time of 6.34 years for patients initiating therapy when their CD4 is below 50, to 24.79 years for those initiating when their CD4 count is at least 350.

By working with life table estimates of remaining lifetimes, we factor in the aging of the cohort and fix the problem mentioned earlier about the survival curves of Figures 6 through 9 not extending long enough to consider an increasing age-based mortality. Note that we

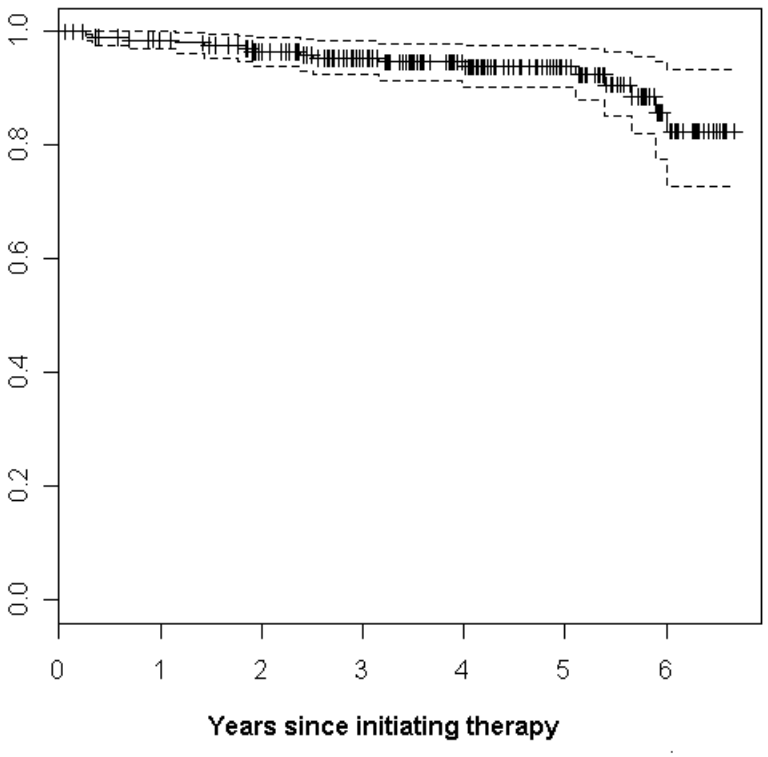

Figure 9: Kaplan-Meier survival curve and its 95% confidence interval for patients initiating therapy from CD4 category $\geq$350

Table 5: Statistics on deaths versus censors when initiating therapy from each CD4 category

	CD4 Categories				
	0-49	50-199	200-349	$\geq$ 350	Totals
Deaths	48	51	22	16	137
Censors	104	190	137	194	625
Patients	152	241	159	210	762

Table 6: Expected remaining life years after initiating HAART from various CD4 categories

	CD4 Categories			
	0-49	50-199	200-349	≥ 350
Mean Remaining Lifetime	6.34	11.61	17.25	24.79

have directly applied the hazard rates for being HIV+ compared to HIV- in VACS to the life table estimates of remaining survival because these estimates are mostly based on the survival of an HIV- population. This makes the implicit assumption that HIV- patients in VACS are similar to HIV- patients in general, which may not be the case. It may be feasible to develop a hazard ratio of being HIV- in VACS compared to being HIV- in the rest of the population, and apply this to the life table estimates before applying the hazard ratios comparing HIV+ and HIV- VACS patients. We leave this for future work.

3.4 CONCLUSIONS

This chapter has focused on using clinical data to develop two key components of the MDP model we present in Chapter 4: a discrete-time Markov model that estimates monthly transition probabilities among different CD4 categories and death, and estimates of expected remaining lifetime after initiating HAART, as a function of those same CD4 categories upon initiation. We have discussed some of the literature and challenges unique to each component. One should keep in mind that these analyses were based on male patients only; it is not clear how outputs would vary when generated from female patients. In Section 4.5, we will solve the MDP model based on the results of this chapter, as well as explore the MDP's sensitivity to variations in our statistical outcomes.

4.0 THE OPTIMAL TIME TO INITIATE HIV THERAPY: ONE-DIMENSIONAL STATE SPACE (CD4 COUNT)

4.1 INTRODUCTION

As discussed in Chapters 1 and 2, the question of when to initiate HIV therapy is central to effective HIV care, and is therefore subject to much research and debate. The treatment strategy has changed over the years, from a paradigm of "hit hard, hit early" [78] in the late 1990s, to a more cautious strategy of "hit hard, but only when necessary" [75]. The latter approach is reflected in current DHHS guidelines that suggest delaying the initiation of HAART until the CD4 count drops to 200-350 cells/mm^3 of blood [123]. However, as discussed below, there are reasons to reconsider the earlier paradigm of treating HIV earlier in its course. Clearly, the best time to initiate HAART is unresolved and evolving.

Because CD4 count is arguably the most important factor in considering when to initiate therapy [123, 166], this chapter develops an MDP of the optimal time to initiate therapy as a function of this variable alone. However, the model is general enough to handle any important prognostic variable. We consider objectives of maximizing a patient's total expected lifetime or quality-adjusted lifetime. In Chapter 5, we consider a two dimensional state space of CD4 count and viral load, as the latter is also an important prognostic indicator of HIV outcomes [116]. The current chapter is the basis of a paper by Shechter et al. [158].

4.2 MODELING FRAMEWORK

We consider a patient in a chronic stage of HIV (the period starting about three weeks after being infected with the virus [10]) who must decide (with his or her physician) when to initiate therapy. We assume the patient visits a physician periodically, and at each visit, the patient's CD4 count is measured. The frequency of these visits may depend on the physician as well as the stage of HIV, but we shall assume they occur monthly (as discussed in Section 3.2, Freedberg et al. [68] also considered a model with monthly cycles to reflect the setting for HIV care). Based on the measurement, a decision is made either to initiate therapy or to wait and reevaluate the situation at the next visit (if the patient is alive at that time). If the patient initiates therapy, a terminal reward is received and the process terminates. We consider terminal rewards such as expected lifetime or expected quality-adjusted lifetime from the time of initiating therapy from various categories of CD4 count. If the patient waits, a reward is accrued between physician visits, and the patient transitions to another CD4 category or death at the next visit with some known probability distribution. The one-period reward may be the time or quality-adjusted time between visits, which may also depend on the patient's CD4 category. Our objective is to maximize the total expected lifetime reward for the patient.

Formally, the components of the MDP are described as follows:

$T = \{1, 2, ..., \infty\}$: the monthly decision epochs.

s: the health state of the patient, represented by a range of CD4 count. We let $s = 0$ indicate an absorbing state for which no further rewards accrue (i.e., death) and let $s \in \{1, \ldots, N\}$ represent different CD4 categories (with higher states representing higher ranges of CD4 count). Let S represent the set of all states.

$a(s)$: the decision taken when the patient is in CD4 category s prior to initiating therapy. $a \in \{W, I\}$ where W indicates to continue waiting and I indicates to initiate therapy.

$r(s)$: the reward the patient receives when waiting in state s. We assign $r(0) = 0$.

$R(s)$: the expected total remaining reward, received when the patient initiates therapy from state s. We assign $R(0) = 0$. As discussed more next through the transition probabilities, $R(s)$ represents a terminal reward for the process.

$p(j|s)$: the probability that the patient's CD4 category goes from s at time t to j at time $t+1$ when waiting another period. Patient death is represented by $p(0|s)$, and we assume the probability structure ensures that state 0 is reachable from every state. Also, state 0 is absorbing, so that $p(0|0) = 1$. If $a(s) = I$, the patient receives reward $R(s)$ and moves with certainty to state 0, an absorbing state of 0 reward. Let P represent the matrix of transitions probabilities when waiting.

$v^*(s)$: the value vector that gives the optimal expected remaining reward when the patient is in state s and has not yet initiated therapy. By our construction, $v^*(0) = 0$.

Our setup fits the framework of a stochastic longest path problem [22], for which it is known that $v^*(s)$ is the unique optimal solution to the following set of recursive optimality equations, also known as Bellman's equations [17, 140]:

$$v(s) \quad = \quad \max \left\{ r(s) + \sum_{j \in S} p(j|s) v(j), \quad R(s) \right\} \quad \text{for all } s \in S. \tag{4.1}$$

For example, if the objective is to maximize a patient's total expected lifetime, then we let $r(s) = 1$ month, we let $R(s) \equiv L(s) =$ the expected remaining life months upon initiating therapy from state s, and we obtain the optimal expected remaining lifetime from each CD4 category as the solution of:

$$v(s) \quad = \quad \max \left\{ 1 + \sum_{j \in S} p(j|s) v(j), \quad L(s) \right\} \quad \text{for all } s \in S. \tag{4.2}$$

We note that the values of $L(s)$ will depend on the patient's age. In Section 3.3, we estimated the expected remaining lifetimes upon initiating therapy for 45-year old male patients. In Section 4.6, we discuss the assumption that $L(s)$ does not decrease from period to period.

An issue that often arises in the medical decision making literature is that of discounting future health outcomes [59, 72]. Although it is intuitive to include a discounting factor for future monetary outcomes, it is less clear that one should do so for future health outcomes. However, it is standard practice in cost-effectiveness analyses to discount both (and at the same rate) [94, 183]. We note that incorporating a discount factor presents no conceptual or computational challenges for our model. It is well known that if we include a discount factor $\lambda < 1$, any solution to Bellman's equations is the unique optimal solution [140].

Furthermore, under our finite state and action sets, any decision rule that satisfies Bellman's equations (called a conserving decision rule) forms an optimal stationary policy [140]. With no discounting ($\lambda = 1$), then in general one cannot say that a solution to Bellman's equations yields an optimal solution. Moreover, an optimal policy need not even exist [146]. However, in the case that there is an absorbing state with a reward of 0 that is eventually reached with probability 1, a solution to Bellman's equations is in fact optimal and a conserving decision rule forms an optimal stationary policy [140, 146]. This is the situation we present here; we indicated above that the death state is absorbing, yields a reward of 0, and is eventually reached from every other state. Also, for both discounting and no discounting, the value and policy iteration algorithms converge to the optimal solution [140]. Therefore, for the sake of clarity, we do not include a discount factor in our equations (i.e., we maximize total expected undiscounted rewards).

4.3 STRUCTURAL PROPERTIES

Before solving the MDP using the clinically based components developed in Chapter 3, we consider how certain structure on the model input may guarantee certain structure on the model output (i.e. the optimal values and policies). In addition to providing deeper insight into the overall problem, discovering such structural properties can make implementation easier and accelerate solution time.

We first state the following definition [12]:

Definition 4.1. *An $N \times N$ transition probability matrix P is said to be* IFR *(increasing failure rate) if its rows are in increasing stochastic order. That is, P is* IFR *if*

$$y(i) = \sum_{j=s}^{N} p(j|i),$$

is nondecreasing in i for all $s \in \{1, \ldots, N\}$.

The IFR property is often used in the machine maintenance literature. In the context of our framework, the IFR property implies that patients in better health states have a higher probability of moving to any particular health state or better. Conversely, patients in worse health states have a higher probability of going to any particular health state or worse (including death).

We will also make use of the following lemma, whose proof is similar to Lemma 4.7.2 in [140]. We show the proof because it will be instructive for the proof of Lemma 7.1 in Chapter 7.

Lemma 4.1. *Suppose* $\{v_i\}$ *(i=0,...,N) is a sequence of numbers, and* $\{p_i\}$ *and* $\{q_i\}$ *(i=0,...,N) are two discrete probability distributions such that*

$$\sum_{i=k}^{N} q_i \geq \sum_{i=k}^{N} p_i \quad \textit{for all } k \in \{0, \ldots, N\}. \tag{4.3}$$

Then, if $\{v_i\}$ *is nondecreasing (nonincreasing) in* i*,*

$$\sum_{i=0}^{N} q_i v_i \geq (\leq) \sum_{i=0}^{N} p_i v_i$$

.

Proof. We prove the lemma for the case that $\{v_i\}$ is nondecreasing in i. Define $v_{-1} = 0$. Then:

$$\begin{aligned}
\sum_{i=0}^{N} q_i v_i &= \sum_{i=0}^{N} q_i \sum_{j=0}^{i} (v_j - v_{j-1}) && (4.4)\\
&= \sum_{i=0}^{N} (v_i - v_{i-1}) \sum_{j=i}^{N} q_j && (4.5)\\
&= \sum_{i=1}^{N} (v_i - v_{i-1}) \sum_{j=i}^{N} q_j + v_0 \sum_{j=0}^{N} q_j && (4.6)\\
&\geq \sum_{i=1}^{N} (v_i - v_{i-1}) \sum_{j=i}^{N} p_j + v_0 \sum_{j=0}^{N} p_j && (4.7)\\
&= \sum_{i=0}^{N} p_i v_i, && (4.8)
\end{aligned}$$

where (4.4), (4.5), and (4.6) follow by a rearrangement of terms, and (4.7) follows by the assumption that $\{v_i\}$ is nondecreasing in i, the condition of (4.3), the fact that probabilities are nonnegative, and the fact that $\sum_{j=0}^{N} q_j = \sum_{j=0}^{N} p_j$. Finally, (4.8) follows by similar steps to go from (4.4) to (4.6). □

Now consider the following assumptions:

($As4.1$) $r(s)$ and $R(s)$ are both nonnegative and nondecreasing in s.

($As4.2$) P is IFR.

$As4.1$ states that as the CD4 category increases, the reward over one period and the terminal reward do not decrease, and $As4.2$ indicates that patients waiting in better health states have better transitions probabilities than those waiting in worse health states.

In the following intuitively appealing theorem, we show that lower health states are "worse," and higher states are "better". In the proof that follows and other places in the dissertation, we refer to MDP solution algorithms such as value iteration and policy iteration. We refer the reader to [140] for details on these algorithms.

Theorem 4.1. *Under assumptions $As4.1$ and $As4.2$, $v^*(s)$ is nonnegative and nondecreasing in s.*

Proof. We prove this by induction. Let $i \in \{0, 1, \ldots\}$ represent iteration i of the value iteration algorithm, and let v^i be the resulting value vector of that iteration. We suppose that $v^i(s)$ is nonnegative and nondecreasing in s (note that $v^0(s) = 0$ for all s satisfies this property for $i = 0$). By this assumption and assumption $As4.2$ (i.e., P is IFR), we apply Lemma 1 to obtain:

$$\sum_{j \in S} p(j|s+1)v^i(j) \geq \sum_{j \in S} p(j|s)v^i(j) \geq 0 \quad \text{for all } s \in S.$$

Combining this with assumption $As4.1$ (i.e., $r(s)$ is nonnegative and nondecreasing in s for all s), we have:

$$r(s+1) + \sum_{j \in S} p(j|s+1)v^i(j) \geq r(s) + \sum_{j \in S} p(j|s)v^i(j) \geq 0 \quad \text{for all } s \in S. \tag{4.9}$$

We also know by assumption $As4.1$ that:

$$R(s+1) \geq R(s) \geq 0 \quad \text{for all } s \in S. \tag{4.10}$$

Combining (4.9) and (4.10) yields:

$$\max\left\{r(s+1) + \sum_j p(j|s+1)v^i(j), \quad R(s+1)\right\} \tag{4.11}$$

$$\geq \quad \max\left\{r(s) + \sum_j p(j|s)v^i(j), \quad R(s)\right\} \geq 0 \quad \text{for all } s \in S. \tag{4.12}$$

According to the value iteration algorithm,

$$v^{i+1}(k) = \max\left\{r(k) + \sum_j p(j|k)v^i(j), \quad R(k)\right\} \quad \text{for all } k \in S. \tag{4.13}$$

Therefore, (4.12) and (4.13) imply:

$$v^{i+1}(s+1) \geq v^{i+1}(s) \geq 0 \quad \text{for all } s \in S.$$

Thus, $v^{i+1}(s)$ is nonnegative and nondecreasing in s. Then, because $\lim_{n\to\infty} v^n(s) = v^*(s)$, it follows that $v^*(s)$ is nonnegative and nondecreasing in s. □

In addition to establishing structure on the optimal value vector, it is common to find conditions on the input parameters that lead to structured optimal policies. The following establishes both a necessary and sufficient condition for initiating therapy in every state to be an optimal policy. The condition in the theorem states that the value of initiating therapy in each state is at least as great as the value of waiting one period in that state and initiating therapy the next period. As discussed in Section 4.5, our data satisfy this condition.

Theorem 4.2. *$a^*(s) = I$ for all s if and only if:*

$$R(s) \geq r(s) + \sum_j p(j|s)R(j) \quad \textit{for all } s \in S.$$

Proof. If $a^*(s) = I$ for all s, then $v^*(s) = R(s)$ for all s. Furthermore, v^* must satisfy Bellman's equations:

$$v(s) \quad = \quad \max\left\{ r(s) + \sum_j p(j|s)v(j), \quad R(s) \right\} \quad \text{for all } s \in S,$$

which implies:

$$R(s) \geq r(s) + \sum_j p(j|s)R(j) \quad \text{for all } s \in S.$$

Now suppose

$$R(s) \geq r(s) + \sum_j p(j|s)R(j) \quad \text{for all } s \in S.$$

Letting $v(s) = R(s)$, we see that v satisfies Bellman's equations given in (4.1). Therefore, $v(s) = R(s)$ is an optimal value function which is achieved by the policy $a^*(s) = I$ for all s. □

Similarly, the following gives a sufficient condition for waiting to be uniquely optimal for a particular health state. In other words, it is not also optimal to initiate therapy from that state.

Corollary 4.1. *If* $R(s') < r(s') + \sum_j p(j|s')R(j)$ *for some* $s' \in S$*, then* $a^*(s') = W$*, uniquely.*

Proof. Bellman's equations indicate that $v^*(s) \geq R(s)$ for all $s \in S$. Therefore, if $R(s') < r(s') + \sum_j p(j|s')R(j)$ it follows that $R(s') < r(s') + \sum_j p(j|s')v^*(j)$ which implies that $a^*(s') = W$, uniquely. □

Another structured policy that may arise is that of an optimal *control-limit policy*, in which states are given a meaningful ordering, there are two actions available, and for states below a threshold state it is optimal to choose one action, whereas for states at or above the threshold it is optimal to choose the other action. We call this threshold state the *control limit*. In the case of HIV therapy, such a policy may imply that below a certain CD4 category, a patient should initiate therapy; otherwise, it is best to wait. As discussed in Section 2.2, the current DHHS guidelines appear to suggest a control-limit policy: patients with lower CD4 counts should initiate therapy, whereas patients with higher CD4 counts should wait.

Other authors have explored sufficient conditions for the existence of optimal monotone policies (of which a control-limit policy is a special case) in general, infinite horizon MDPs with possibly many actions [83, 140]. For example, some of the conditions involve both the reward and transition probability structure having subadditivity properties or both having superadditivity properties (see [140] for an explanation of these properties). However, by the fact that $R(s)$ represents a lifetime reward from initiating therapy and $r(s)$ represents a single period reward, it can be shown that the reward structure in our problem satisfies the superadditivity property. On the other hand, one can also check that the transition probability structure satisfies the subadditivity property. Alagoz et al. [5] presented a set of sufficient conditions to guarantee an optimal control-limit policy for a model with similar structure to ours; however, our data-driven model is far from satisfying those conditions. Here we explore a different situation that guarantees an optimal control-limit policy.

Consider the clinically plausible case of a patient's health never improving prior to initiating therapy, i.e. consider the following assumption:

$(As4.3)$ $p(j|s) = 0$ whenever $j > s$.

Some authors have explicitly made this assumption in their models [32, 68, 111, 113]. Our data indicate a low probability of upward movement (approximately .04).

Additionally, consider the following assumption stating that the difference in value, between successive states, of a policy that waits in the current state and initiates therapy in all lower states is at least as great as the difference in value, between those states, of a policy that initiates therapy immediately:

$(As4.4)$ $\frac{r(s+1)+\sum_{j<s+1} p(j|s+1)R(j)}{1-p(s+1|s+1)} - \frac{r(s)+\sum_{j<s} p(j|s)R(j)}{1-p(s|s)} \geq R(s+1) - R(s)$ for all $s \in S$.

$As4.4$ also has intuitive appeal as it essentially implies that with a better health state, delaying therapy is relatively more appealing than initiating therapy (this is a type of subadditivity condition).

We can now prove the existence of an optimal control-limit policy.

Theorem 4.3. *Suppose assumptions As*4.3 *and As*4.4 *hold. Then an optimal control-limit policy exists.*

Proof. Either $a^*(s) = I$ for all s, or there exists a smallest CD4 category, s', for which $a^*(s') = W$, uniquely. The former is vacuously a control-limit policy. Considering the latter, if we prove that $a^*(s') = W$ implies $a^*(s'+1) = W$, then the result follows by induction. By Bellman's equations, we know that $a^*(s') = W$ and *As*4.3 imply:

$$\frac{r(s') + \sum_{j<s'} p(j|s')R(j)}{1 - p(s'|s')} \geq R(s'). \tag{4.14}$$

If we add

$$\frac{r(s'+1) + \sum_{j<s'+1} p(j|s'+1)R(j)}{1 - p(s'+1|s'+1)} - \frac{r(s') + \sum_{j<s'} p(j|s')R(j)}{1 - p(s'|s')}$$

to the left-hand side of (4.14) and add $R(s'+1) - R(s')$ to the right-hand side of (4.14), then by assumption *As*4.4 and (4.14), we obtain:

$$\frac{r(s'+1) + \sum_{j<s'+1} p(j|s'+1)R(j)}{1 - p(s'+1|s'+1)} \geq R(s'+1).$$

By induction, we can see that for all $s \geq s'$,

$$\frac{r(s) + \sum_{j<s} p(j|s)R(j)}{1 - p(s|s)} \geq R(s).$$

Because $v^*(j) \geq R(j)$ for all j, it follows that

$$\frac{r(s) + \sum_{j<s} p(j|s)v^*(j)}{1 - p(s|s)} \geq R(s) \quad \text{for all } s \geq s',$$

which implies that $a^*(s) = W$ for all $s \geq s'$. □

In Theorem 4.8 of Section 4.4 below, we show how knowing the existence of a control-limit policy accelerates solution time.

4.3.1 Patient-Specific Considerations

While national guidelines policies are based on studies from entire cohorts of patients, there is a push to make HIV therapy more patient-focused [77]. Therefore, we incorporate two important patient factors for deciding when to initiate HIV therapy: quality of life and adherence.

Quality of Life

One reason patients may want to delay initiating therapy is to avoid negative side effects of the drugs such as nausea, fatigue, or lipodystrophy syndrome [37, 60, 105]. Of course, even without therapy, advanced stages of any disease will reduce patients' quality of life. Patients with advanced stages of HIV, for example, may experience "opportunistic infections" such as mycobacterium avium complex, tuberculosis, Kaposi's sarcoma, pneumocystis carinii pneumonia, or cytomegalovirus [3]. Tengs and Lin [172] provide a meta-analysis of studies eliciting patient utilities for different stages of HIV. Estimates by CD4 strata can be found in [68, 151]. However, none of the papers we reviewed [14, 15, 41, 45, 68, 69, 82, 105, 119, 125, 128, 142, 148, 151, 174, 175, 185, 186] presents patient-based utility estimates according to both stage of HIV as well as whether or not the patient is on therapy. For example, a patient with a high CD4 count taking HAART would presumably elicit a lower quality of life than a patient with the same CD4 count who is not taking therapy. This is a key distinction to make in evaluating the optimal time to initiate therapy. Lenert et al. [105], have come the closest to uncovering this distinction by estimating a utility decrement of .20, beyond the reduced utility for having HIV, for patients with lipodystrophy complications from therapy. It is not clear, however, to what extent the baseline estimates included patients on HAART and thus experiencing the burden of taking therapy or experiencing side effects other than lipodystrophy syndrome. In Section 4.5, we solve our MDP by performing sensitivity analyses around CD4-based and therapy-based utility estimates.

For now, we let $0 \leq u_w(s) \leq 1$ be the utility associated with waiting to initiate therapy when in state s, and we let $0 \leq u_i(s) \leq 1$ be the average utility for the remainder of the patient's life, when initiating therapy from state s. We note that this average remaining utility may depend on the patient's age. Applying these utilities to Bellman's equations that

maximize expected remaining lifetime (4.2) yields Bellman's equations for the problem of maximizing quality-adjusted lifetime:

$$y(s) \quad = \quad \max \left\{ u_w(s) + \sum_j p(j|s) y(j), \quad u_i(s) L(s) \right\} \quad \text{for all } s \in S,$$

where $y(s)$ represents the expected remaining quality-adjusted lifetime from state s, and $L(s)$ is the expected remaining (unadjusted) lifetime after initiating therapy from state s. We shall use these equations in Section 4.5 to obtain the solution of the MDP under a quality-adjusted lifetime framework.

Adherence

Another important issue that may delay the initiation of therapy is the degree to which a physician believes a patient may adhere to the prescribed therapy [123]. Various studies have demonstrated associations between lower adherence and worse outcomes such as higher viral load [76], lower CD4 count [129], higher incidence of AIDS [11], more days in the hospital [129], and higher mortality rates [34]. Though by most conventions taking at least 80% of prescribed medication implies compliance with the therapy, a study by Paterson et al. [129] exposed significant differences in outcomes even for adherence levels that differ in the 80-100% range. However, adherence rates are often lower than this as various studies have reported 40-50% of patients taking less than 80% of their medication [18, 63, 71].

Our model implicitly represents the effects of partial adherence to the extent this is represented in the data-based estimates of the expected remaining lifetime upon initiating therapy. In other words, the estimates reflect the partial adherence exhibited by VACS patients. We now *explicitly* consider how a particular patient's tendency to adhere may affect optimal policies. Specifically, we consider a patient-dependent multiplier, m, that applies to $L(s)$ for each s. If the patient tends to adhere better than the average adherence level of the cohort, then $m > 1$ and the estimates of $L(s)$ are increased proportionally to reflect better outcomes for this patient relative to the average patient. If the patient adheres poorly, then $m < 1$ and the estimates of $L(s)$ are reduced to reflect worse outcomes. We explore structural properties of this model and consider building more complex models of

patient adherence in future research. For example, the model explicitly considers adherence under the framework of maximizing expected lifetime, and we discuss the inclusion of adherence in a quality-adjusted lifetime framework in Section 4.6.

Note that by our model formulation, the adherence factor only affects the estimated reward for initiating therapy, $R(s)$; it has no affect on the pre-therapy components P and $r(s)$. Explicit consideration of patient adherence under the objective of maximizing expected lifetime leads to the following Bellman's equations:

$$z(s|m) \;=\; \max\left\{1+\sum_j p(j|s)z(j|m),\;\; mL(s)\right\} \quad \text{for all } s\in S$$

where $z(s|m)$ represents the expected remaining lifetime from state s for a patient whose adherence parameter is m.

Suppose the optimal policy for the problem of maximizing expected lifetime without explicit consideration of patient adherence is to initiate therapy for all states (i.e., the condition of Theorem 4.2 is satisfied). Then the following result allows us to identify conditions on m that will ensure such a policy is no longer optimal when explicitly considering adherence. First we define the following:

$$\begin{aligned}
\Delta_s &\equiv L(s) - [1+\sum_j p(j|s)L(j)] \quad \text{for all } s\in S,\\
\Delta_{min} &\equiv \min_{s\in S}\Delta_s, \text{ and}\\
\Delta_{max} &\equiv \max_{s\in S}\Delta_s.
\end{aligned}$$

In words, Δ_s is the difference (for the expected lifetime problem without explicit consideration of adherence) between the terminal reward of initiating therapy from state s and the value associated with waiting one period in state s and initiating therapy the next period.

The following results should be interpreted in the context of the problem of maximizing total expected lifetime under an explicit consideration of patient adherence.

Theorem 4.4. *For every state s such that $1/m > 1+\Delta_s$, it is uniquely optimal to wait.*

Proof. As a consequence of Corollary 4.1, it is uniquely optimal to wait in state s if

$$mL(s) < 1 + \sum_j p(j|s)mL(j).$$

This is equivalent to

$$L(s) < 1/m + \sum_j p(j|s)L(j),$$

which is equivalent to

$$1/m > L(s) - \sum_j p(j|s)L(j).$$

By the definition of Δ_s, the above is equivalent to:

$$1/m > 1 + \Delta_s.$$

□

It follows that if $1/m > 1 + \Delta_{max}$, it is optimal to wait in every state. The following corollary provides the least restrictive condition to guarantee an optimal policy for which it is optimal to wait in some state. We let s_{min} be a state that minimizes Δ_s.

Corollary 4.2. *$a^*(s_{min}) = W$, uniquely, if and only if $1/m > 1 + \Delta_{min}$.*

Proof. The sufficiency part of the proof follows directly from Theorem 4.4.

To prove necessity, suppose $a^*(s_{min}) = W$, uniquely. Also, suppose (towards a contradiction) that $1/m \leq 1 + \Delta_{min}$. Then it follows that $1/m \leq 1 + \Delta_s$ for all s. By the definition of Δ_s, we have:

$$1/m \leq L(s) - \sum_j p(j|s)L(j) \quad \text{for all } s \in S.$$

Multiplying through by m and rearranging terms yields:

$$mL(s) \geq 1 + \sum_j p(j|s)mL(j) \quad \text{for all } s \in S.$$

By Theorem 4.2, this implies that $a^*(s) = I$ for all s, which contradicts our first assumption that $a^*(s_{min}) = W$, uniquely. Therefore, $1/m > 1 + \Delta_{min}$. □

Note that by Theorem 4.2, if the optimal policy in the expected lifetime problem without explicit consideration of patient adherence is to initiate therapy in every state, then $\Delta_{min} > 0$. In that case, it follows from Corollary 4.2 that a necessary condition for waiting to be uniquely optimal in state s_{min} with explicit consideration of patient adherence is that $m < 1$. In other words, the optimal policy of initiating therapy from each state will change only if the patient adheres poorly.

We prove the following additional results. The first one says that patients with greater levels of adherence have greater expected remaining lifetimes, as expected.

Theorem 4.5. *$z(s|m)$ is nondecreasing in m for each s.*

Proof. Let $m_1 \leq m_2$. We prove this by performing parallel iterations of the value iteration algorithm to solve for $z(s|m_1)$ and $z(s|m_2)$ for all s. Suppose for some iteration, i, of the algorithm $z^i(s|m_1) \leq z^i(s|m_2)$ for all s (note that starting each problem with zero vectors satisfies this). Then for each s,

$$z^{i+1}(s|m_1) = \max\left\{1 + \sum_j p(j|s)z^i(j|m_1),\ \ m_1 L(s)\right\}, \text{ and} \tag{4.15}$$

$$z^{i+1}(s|m_2) = \max\left\{1 + \sum_j p(j|s)z^i(j|m_2),\ \ m_2 L(s)\right\}. \tag{4.16}$$

By the inductive assumption, $1 + \sum_j p(j|s)z^i(j|m_1) \leq 1 + \sum_j p(j|s)z^i(j|m_2)$ and by the assumption that $m_1 \leq m_2$, $m_1 L(s) \leq m_2 L(s)$. Therefore, $z^{i+1}(s|m_1) \leq z^{i+1}(s|m_2)$. Taking the limit of the value iterates proves the result. □

Our final result for explicit considerations of adherence concerns control limit policies. It says that if optimal control-limit policies hold for two patients with adherence levels m_1 and m_2, respectively, then assuming patient health does not improve while waiting ($As4.3$), the patient with the greater adherence level will initiate therapy whenever the other patient does, and perhaps earlier.

Theorem 4.6. *Suppose the optimal policies for patients with adherence levels m_1 and m_2 ($m_1 \leq m_2$) have control limits given by c_1 and c_2, respectively. Then if As4.3 holds, $c_1 \leq c_2$.*

Proof. First, let us assume that $a^*(1|m_1) = I$ (otherwise, the result holds vacuously). Then by $As4.3$,

$$m_1 L(1) \geq \frac{1}{1 - p(1|1)}, \tag{4.17}$$

and by the assumption that $m_1 \leq m_2$,

$$m_2 L(1) \geq \frac{1}{1 - p(1|1)}. \tag{4.18}$$

Therefore, $a^*(1|m_2) = I$ as well. Now let $s' > 1$, arbitrarily, and assume that for $s < s'$, $a^*(s|m_2) = I$. We complete the proof by showing that if $a^*(s'|m_1) = I$, then $a^*(s'|m_2) = I$ as well. By $As4.3$, $a^*(s'|m_1) = I$ implies

$$m_1 L(s') \geq \frac{1 + \sum_{j<s'} p(j|s') m_1 L(j)}{1 - p(s'|s')}, \tag{4.19}$$

which is equivalent to

$$L(s') \geq \frac{1/m_1 + \sum_{j<s'} p(j|s') L(j)}{1 - p(s'|s')}. \tag{4.20}$$

Because $m_1 \leq m_2$ by assumption, it follows that

$$L(s') \geq \frac{1/m_2 + \sum_{j<s'} p(j|s') L(j)}{1 - p(s'|s')}, \tag{4.21}$$

which, is equivalent to

$$m_2 L(s') \geq \frac{1 + \sum_{j<s'} p(j|s') m_2 L(j)}{1 - p(s'|s')}. \tag{4.22}$$

By the inductive assumption that $a^*(s|m_2) = I$ for $s < s'$, it follows that $a^*(s'|m_2) = I$. $\square$

4.4 COMPUTATIONAL CONSIDERATIONS

In this section, we make some observations about solving optimal stopping problems represented by Bellman's equations in (4.1). The model we solve in the next section is small enough that we do not encounter any computational difficulties in solving it. However, the results of this section apply to optimal stopping time problems for which large state spaces may impose computational burdens.

It is known that if v^i is the value vector generated by the i^{th} iteration of the policy iteration algorithm, and if v^{i+1} is the value vector generated by the $(i+1)^{st}$ iteration, then $v^{i+1} \geq v^i$, componentwise [140]. Because the optimal stopping problem contains 2^N policies, in theory it takes at most 2^N iterations to guarantee an optimal policy. However, we can show that it actually takes at most N iterations to obtain an optimal policy.

Theorem 4.7. *The optimal stopping problem represented by (4.1) can be solved in at most N iterations of the policy iteration algorithm.*

Proof. Let d^0 be the decision rule that accepts the terminal reward, $R(s)$, for every state s. Then the policy evaluation step gives $v^0(s) = R(s)$ for every s. Now consider the policy improvement step. If there is no policy improvement, then the algorithm terminates. Otherwise, there exists a state s' such that $r(s') + \sum_j p(j|s')R(j) > R(s')$, and we set the action for that state to W. Because we have the property that each successive iteration of the algorithm yields a value vector no smaller than the previous one, it must be that the optimal action in state s' is to wait. In other words, once an iteration is reached in which the action switches from I to W for some state, all subsequent iterations retain the W decision for that state. Hence, there can be at most N switches from the initial decision rule. Because the algorithm terminates when two successive iterations yield the same policy, the algorithm takes at most N steps. □

If the lower triangular property given in $As4.3$ holds, then the optimal solution is obtained in just one iteration of the Gauss-Seidel variant of the value iteration algorithm (in which the value vector is updated after each state evaluation and used in the evaluation of the next state). The reason for this is that $As4.3$ implies that the value of a given state does not

depend on the value of any states above it, so we can start with $s = 1$, solve for $v^*(1)$, use this to obtain $v^*(2)$ and continue until solving the entire problem after reaching state N and solving for $v^*(N)$ using the values of $v^*(s)$ obtained along the way. A similar result is found in [22]. If $As4.4$ also holds, then finding an optimal policy requires even less work.

Theorem 4.8. *If both $As4.3$ and $As4.4$ hold, then an optimal policy can be determined in at most $\lfloor log_2(N) + 1 \rfloor$ steps.*

Proof. From Theorem 3, we know that an optimal control-limit policy exists. Then, by the lower triangular property of $As4.3$ and the fact that $v^*(k) \geq R(k)$ for all $k \in S$, we know that $a^*(s) = I$ if and only if $\frac{1+\sum_{j<s} p(j|s)R(j)}{1-p(s|s)} \leq R(s)$. It takes a binary search to find the smallest s' such that this latter condition does not hold, implying that $a^*(s) = W$ for $s \geq s'$ and $a^*(s) = I$ for $s < s'$. This search takes at most $\lfloor log_2(N) + 1 \rfloor$ steps. □

4.5 DATA-DRIVEN IMPLEMENTATION

In Chapter 3, we discussed the use of clinical data to build two key components of an MDP of the optimal time to initiate HIV therapy: the CD4 natural history transition probability matrix and the expected remaining lifetimes upon initiating therapy from the different CD4 categories. We constructed both components from 40-50 year old male patients in the cohort, and therefore, we present results of optimal policies and expected remaining rewards for those patients.

Results

Recall from Chapter 3 that we examined various scenarios for handling the censored observations in the natural history Markov model. In the present analysis, we shall use three of those scenarios: the baseline weighting of redistributing censored observations towards

death, a heavier weighting, and a lighter weighting (not the lightest weighting that we presented). Table 6 of that chapter also presented what we refer to as the baseline results for the survival model. Here we also consider estimates at 90% and 110% of those values.

Table 7 gives the optimal policy along with the optimal value function (in life years) for each combination of the natural history and survival models, under an objective of maximizing expected lifetime. For the optimal policy vector, "I" indicates to initiate therapy from that CD4 category, while "W" indicates to continue waiting. The table does not show the case of the heavier censoring redistribution towards death, because this yields the same optimal policies and values as for the case of the baseline natural history model, namely to initiate therapy immediately. Under our baseline natural history model and by letting $r(s) = 1$ month for all s, assumptions $As4.1$ and $As4.2$ hold. Therefore, as proven by Theorem 1, our solution produces an optimal value vector that is nonnegative and nondecreasing in the CD4 category. Furthermore, our MDP components satisfy the sufficient condition of Theorem 4.2, and hence, we obtain an optimal policy of initiating therapy from each of the CD4 categories. In fact, it is not until reducing the survival estimates to .43 of our baseline values that we first obtain an optimal policy other than initiating therapy from ever state. Note that the lighter redistribution of censors to death along with the baseline and 90% survival models yield counter-intuitive optimal policies that exhibit a control-limit structure in the opposite direction than expected. However, neither of the value functions for these cases differs substantially from the value function of the policy that initiates therapy from each state. We comment on this more in the next section.

As indicated in Section 3.2, our natural history Markov model depends on the choice of degrees of freedom (df) for the spline, and we chose the parameter yielding the roughest curve: i.e., the one that maximally captures the data fluctuations. Here we consider using the average spline instead (the spline with df half way between the value yielding the smoothest and roughest splines). Table 8 shows the expected remaining lifetime implied by each censoring scenario for patients never taking therapy, as a function of their CD4 category. The values in parentheses correspond to the values from Table 4 of Section 3.2, which were based on the rougher spline. These values exceed the values from the smoother spline in every case but one. Table 9 presents the optimal policy and value function using the

Table 7: Optimal policy and value vector under various combinations of the natural history and survival models

NH model	Survival Model	Output	0-49	50-199	200-349	$\geq$ 350
baseline	baseline	Optimal Policy	I	I	I	I
		Optimal Values	6.34	11.61	17.25	24.79
baseline	110%	Optimal Policy	I	I	I	I
		Optimal Values	6.97	12.77	18.98	27.27
baseline	90%	Optimal Policy	I	I	I	I
		Optimal Values	5.71	10.45	15.53	22.31
lighter	baseline	Optimal Policy	W	W	I	I
		Optimal Values	6.38	11.70	17.25	24.79
lighter	110%	Optimal Policy	I	I	I	I
		Optimal Values	6.97	12.77	18.98	27.27
lighter	90%	Optimal Policy	W	W	W	I
		Optimal Values	5.97	10.76	15.56	22.31

smoother spline, and now every combination of natural history and survival model indicates a patient should initiate therapy from every state.

Next we consider the objective of maximizing a patient's expected quality-adjusted lifetime by incorporating utility weights associated with the different CD4 categories and whether or not the patient is on therapy. As discussed in Section 4.3.1, it is not clear from the literature how utility estimates for various stages of HIV may be distinguished between patients on versus off therapy. For the present analysis, we use patient-derived estimates found in [151] and [68] to estimate off-HAART utilities according to our CD4 categories (Table 10). We incorporate these into the MDP by multiplying the 1-month reward associated with waiting by the CD4-based utility weights presented. We then perform sensitivity analyses around the on-therapy utilities by considering various weighings of the off-therapy

Table 8: Expected natural history remaining life years, according to censoring redistribution scenario and CD4 category, using a smoother spline (with the results of the rougher spline in parentheses)

Redistribution Scenario	0-49	50-199	200-349	≥ 350
0 (ignoring censors)	8.53 (14.43)	12.73 (19.11)	15.80 (21.67)	18.81 (23.57)
1 (baseline)	1.10 (1.51)	2.24 (2.96)	3.99 (4.68)	6.22 (6.30)
2 (heavier towards death)	1.05 (1.30)	1.83 (2.17)	2.63 (2.95)	4.03 (4.01)
3 (lighter towards death)	2.84 (4.35)	5.19 (7.19)	7.34 (9.21)	10.02 (11.07)
4 (even lighter towards death)	5.45 (8.89)	8.95 (12.93)	11.69 (15.34)	14.48 (17.21)

utilities. For example, in one analysis, we multiply each CD4-based off-therapy utility by .9 to represent CD4-based on-therapy utilities. We then take these utilities and multiply them by our estimates of expected remaining lifetime to generate terminal rewards in terms of expected remaining quality-adjusted lifetime. Table 11 shows the on-HAART utilities associated with three different multiplicative factors of the off-therapy utilities (.9, .7, and .5).

Table 12 shows the results of the MDP solution for the different ratios of on-HAART to off-HAART utility, under our baseline natural history and survival models. Note that the value vector is now in terms of quality-adjusted life years instead of life years. We see that even if we estimate on-HAART utility to be half as much as off-HAART utility, the MDP still returns an optimal policy of initiating therapy from each CD4 category, using both utility references. In fact, further testing reveals that it is not until reducing the ratio to .34 (for the set of utilities in Schackman et al. [151]) or .29 (for the set of utilities in Freedberg et al. [68]) that we obtain a solution for which it is optimal to wait in some CD4 category. Using the Schackman et al. [151] utilities for being off HAART and applying the .34 multiplier leads to on-therapy utility estimates of .30, .31, .33, and .33 for the CD4

Table 9: Optimal policy and value vector under our baseline run and various combinations of the natural history and survival models, using a smoother spline

NH model	Survival Model	Output	0-49	50-199	200-349	≥ 350
baseline	baseline	Optimal Policy	I	I	I	I
		Optimal Values	6.34	11.61	17.25	24.79
baseline	110%	Optimal Policy	I	I	I	I
		Optimal Values	6.97	12.77	18.98	27.27
baseline	90%	Optimal Policy	I	I	I	I
		Optimal Values	5.71	10.45	15.53	22.31
lighter	baseline	Optimal Policy	I	I	I	I
		Optimal Values	6.34	11.61	17.25	24.79
lighter	110%	Optimal Policy	I	I	I	I
		Optimal Values	6.97	12.77	18.98	27.27
lighter	90%	Optimal Policy	I	I	I	I
		Optimal Values	5.71	10.45	15.53	22.31

Table 10: Estimates of off-HAART utilities

	CD4 Categories			
Reference	0-49	50-199	200-349	≥ 350
Schackman et al. [151]	.88	.91	.97	.97
Freedberg et al. [68]	.79	.84	.94	.94

Table 11: Estimates of on-HAART utilities

		CD4 Categories			
Reference	Utility Scenario	0-49	50-199	200-349	≥ 350
Schackman et al. [151]	Scenario 1 (.9)	.79	.82	.87	.87
	Scenario 2 (.7)	.62	.64	.68	.68
	Scenario 3 (.5)	.44	.46	.49	.49
Freedberg et al. [68]	Scenario 1 (.9)	.71	.76	.85	.85
	Scenario 2 (.7)	.55	.59	.66	.66
	Scenario 3 (.5)	.40	.42	.47	.47

categories $0-49$, $50-199$, $200-349$, and $\geq$350, respectively. These appear too low based on existing literature of utility estimates for a variety of health conditions. Of 82 quality-of-life estimates for patients with HIV or AIDS, only three were at or below the .30-.33 range of estimates [173]. To compare with another disease and its drug treatment, of 44 utility estimates for breast cancer with chemotherapy, only five were at or below the .30-.33 range [173]. Also, as discussed in Section 4.3.1, Lenert et al. [105] estimated that therapy-related complications may reduce the utility for HIV patients by .20.

In another analysis, we incorporate utility estimates from VACS patients based on different CD4 categories along with whether or not the patients experienced side effects from HAART (unpublished data not shown). We use the utilities of side effects versus no side effects as proxies for utilities associated with being on therapy versus off therapy [24]. In doing so, we still obtain an optimal policy of initiating therapy from each CD4 category, and this policy does not change until the ratio of on-HAART to off-HAART utility falls to approximately .43.

Table 12: Optimal policy and value vector under different utilities

Reference	On-HAART Utility	Output	0-49	50-199	200-349	$\geq$ 350
Schackman et al. [151]	Scenario 1 (.9)	Optimal Policy	I	I	I	I
		Optimal Values	5.01	9.52	15.01	21.57
	Scenario 2 (.7)	Optimal Policy	I	I	I	I
		Optimal Values	3.93	7.43	11.73	16.86
	Scenario 3 (.5)	Optimal Policy	I	I	I	I
		Optimal Values	2.79	5.34	8.45	12.15
Freedberg et al. [68]	Scenario 1 (.9)	Optimal Policy	I	I	I	I
		Optimal Values	4.50	8.82	14.66	21.07
	Scenario 2 (.7)	Optimal Policy	I	I	I	I
		Optimal Values	3.49	6.85	11.39	16.36
	Scenario 3 (.5)	Optimal Policy	I	I	I	I
		Optimal Values	2.54	4.88	8.11	11.65

4.6 CONCLUSIONS

Because therapeutic options and the understanding of HIV have increased markedly over the last twenty years, the prognosis for HIV patients has changed from a fatal disease to a serious yet manageable, chronic condition [153]. As such, the proper administration of these therapies has become extremely complex and open to debate. We have proposed the first application of MDPs for examining the contentious issue of the optimal time to initiate HIV therapy.

Several of our results support the former strategy of treating earlier in the course of disease as opposed to the more recent approach of treating later. With an objective of maximizing expected remaining lifetime, our baseline run of the natural history model yielded

optimal policies of initiating therapy from all states, under each survival model (see Table 7). Under the lighter redistribution of censors to death and the baseline survival model, we obtained a counter-intuitive optimal policy of waiting when the CD4 count is less than 200, and initiating when it is at or above 200. However, the values associated with CD4 categories <50 and $50-199$, 6.38 and 11.70 life years, do not differ much from the value associated with initiating therapy from those states (6.34 and 11.61 life years). In other words, although technically the solution yields a counter-intuitive policy, it is not substantially different from the policy of initiating therapy from all states. Therefore, because of the sensitivity of results to particular data estimates, for the sake of clarity and ease of implementation, one may argue that it is best to use the "suboptimal" policy of initiating therapy from all CD4 categories over the policy generated from the MDP solution. A similar argument applies to the results of the lighter natural history model and the 90% survival model. On the other hand, consider a policy of waiting when the CD4 count is > 350 and initiating therapy whenever it is in a lower category (this is one interpretation of the DHHS guidelines [123]). In that case, the value associated with CD4 category >350 (under the baseline natural history and survival models) is 18.0 years, as opposed to the 24.8 years under the optimal policy. This large difference strengthens the support for initiating therapy from the highest CD4 categories under the objective of maximizing expected lifetime.

Even under the quality-adjusted lifetime framework, our results strongly favor a policy of initiating therapy immediately. As shown in Table 12 and the discussion that surrounded it, under the baseline natural history and survival models, it takes an unrealistically low ratio of on-HAART to off-HAART utility to obtain an optimal policy *other than* initiating therapy from all CD4 categories. Interestingly, a simulation model by Schackman et al. [150] also found that it took a 70% decrease in on-therapy quality-of-life before initiating treatment at a CD4 count of 350 appeared worse than initiating treatment at a CD4 count of 200.

Yet further support for our "hit hard, hit early" results comes from the fact that our model biased the results toward waiting in some states. We did not decrement the terminal reward over time, which means that waiting in some states would be even less appealing if we were to reduce the rewards associated with initiating therapy from other states at later times.

In addition to the study by Schackman et al. [150], other recent work suggests that earlier treatment may be better than delayed treatment. Holmberg et al. [81] argued against recent trends toward delayed treatment by citing various studies demonstrating survival benefits, immunologic benefits, and reduced toxicities associated with earlier treatment. At a recent conference, Lichtenstein et al. [107] reported that concerns about toxicities associated with earlier treatment may be unfounded and suggested that earlier initiation of HAART may be better. Also, as mentioned in Section 2.1.2, the SMART trial was recently terminated because it was found that continuous use of HAART was better than a strategy of having patients stay off therapy until their CD4 counts fell below 250 and having them go off therapy again when their CD4 counts rose above 350. We wonder to what degree the worse outcomes of this latter strategy had to do with the initial delay of therapy.

There are areas for refinement with our model. Our explicit consideration of patient adherence assumed that a single multiplier applied equally to all state-based terminal rewards. Future work may consider state-based adherence multipliers instead. Also, in the above analyses, we considered utilities and adherence separately. Future research will consider a model that simultaneously incorporates utilities and adherence under an objective of maximizing expected quality-adjusted lifetime. Such a complex model will require understanding how patient utilities affect adherence levels which in turn affect the length and quality of life. Some of the limitations of our model also lie within the statistical modeling of the natural history health progression and the survival estimates after initiating therapy. We addressed these limitations in the previous chapter and note that testing suggests that our results are sensitive to assumptions on the redistribution of censored observations in the natural history model. We also note that the estimates of remaining survival in our present model implicitly assume that after initiating therapy, patients progress according to the treatment decisions presently made in practice with regard to choice of initial therapy, time to switch to subsequent therapies, and choice of those therapies. In other words, we took a modeling perspective of determining the optimal time to initiate therapy as it is administered today, as opposed to the optimal time to initiate therapy under a completely optimized treatment process. Chapters 6 and 7 consider the latter framework.

Our model assumed that the transition probability matrix was Markovian with respect to the CD4-based health state, though we did not verify this statistically. One study found that the history of CD4 counts did not inform the short term probability of death differently than using just the current level of CD4 count [54]; however, the history may influence the transitions to other living health states [49]. Our state space is small enough that there would be no computational difficulty to include some of the CD4 history, but we decided to first explore structural properties and solve the problem as a function of a patient's current CD4 count before expanding the state space to include recent CD4 measurements as well. We also note that current DHHS guidelines do not consider a patient's recent CD4 count progression in their recommendations; they just focus on the current CD4 count [123].

We made two time-homogeneity assumptions in our model. We assumed that P, our transition probability matrix for the natural history of CD4, is time homogenous. Although patients surely have an increased risk of death with time, we assumed that the time prior to initiating therapy was not so long as to significantly increase a patient's risk of death. Indeed, the Markov chain induced by P (under our baseline model) yields expected remaining lifetimes of 1.51, 2.96, 4.68, and 6.30 years for patients in CD4 categories $0-49$, $50-199$, $200-349$, and ≥ 350, respectively. We did capture the increased risk of death with age through our estimates of survival after initiating therapy (the terminal rewards), as described in the previous chapter. Our other time-homogeneity assumption was that patients receive these terminal rewards regardless of when therapy is initiated. In order to consider a time-based decrement to the terminal reward, the model would have to become significantly more complex and require much more data. As noted above, however, our assumption biased the model towards waiting in some states.

Recall that our MDP solutions were based on components generated from an all-male cohort of patients. Because we do not know how the natural history of CD4 counts and survival upon initiating HAART differ in the female VACS population, care should be used in applying our results to such patients. Similarly, because VACS patients may tend to have different health problems than other patient populations, one should also be careful in applying our results to non-VACS populations. We note, however, that we can easily generate the MDP components and solutions based on data from other cohorts.

In summary, an MDP is both a natural and useful approach for modeling the optimal time to initiate HIV therapy: upon HIV diagnosis, patients see their physicians periodically until the decision is made to initiate therapy, and there is uncertainty in how the patient's health will progress until that time. Previous analytical approaches have not modeled this problem as a sequential, dynamic and stochastic decision problem with an objective of maximizing a patient's expected lifetime or quality-adjusted lifetime. The MDP framework allows us to gain deeper insight into how changes in model inputs affect the outputs. Of particular interest to policy makers is the strength of our results in suggesting that we move away from recent trends of initiating therapy later in the course of disease and to "hit early" instead.

5.0 THE OPTIMAL TIME TO INITIATE HIV THERAPY: TWO-DIMENSIONAL STATE SPACE (CD4 COUNT AND VIRAL LOAD)

5.1 INTRODUCTION

Although CD4 count is the single most important variable in the decision of when to initiate therapy, another significant prognostic variable is the patient's viral load [116]. As discussed in Section 2.2.1, when the viral load exceeds 100,000 copies/mL of blood, some clinicians may recommend their patients initiate therapy regardless of their CD4 count [123]. Therefore, this chapter considers a state space that includes both the patient's CD4 count as well as the viral load.

With the two-dimensional aspect of the state space, a natural ordering does not exist. For example, with the single dimension, it was intuitive that a higher CD4 count is better than a lower CD4 count. However, now we have to consider comparing states such as higher CD4 count and higher viral load with lower CD4 count and lower viral load. The order is not clear and may depend on the differences between each dimension. We present a plausible framework that leads to certain structural properties.

5.2 MODELING FRAMEWORK

As with the CD4-based model of the previous chapter, we consider a patient with chronic HIV who must decide when to initiate therapy. Now we assume that upon each visit to the

physician, measurements are taken of both the CD4 count and the viral load. We consider rewards and probability transitions that are functions of both of these variables, and seek to maximize the total expected lifetime of a patient.

The components of the MDP are described as follows.

- $T = \{1, 2, ..., \infty\}$: the monthly decision epochs.
- (c, l): the state of the patient, where c indicates the CD4 category and l indicates the viral load category. We let $c \in \{0, 1, \ldots, C\}$ (where 0 indicates an absorbing state of no further reward, such as death), and we let $l \in \{1, \ldots, L+1\}$ (where $L+1$ is also an absorbing state of no further reward). Higher CD4 categories are associated with better health states and higher viral load categories are associated with worse health states. Also by convention, we say that $c = 0$ if and only if $l = L + 1$.
- $a(c, l)$: the decision taken when patient is in state (c, l). $a \in \{W, I\}$ where W indicates to continue waiting, and I indicates to initiate therapy.
- $r(c, l)$: the reward the patient receives when waiting in state (c, l). We assume $r(0, L + 1) = 0$.
- $R(c, l)$: the expected total remaining reward, received when the patient initiates therapy from state (c, l). We assume $R(0, L + 1) = 0$.
- $\alpha[(c', l')|(c, l)]$: the probability of transitioning to state (c', l') at time $t+1$ given that the patient waits in state (c, l) at time t. Death is indicated by the absorbing state $(0, L+1)$, and we assume the probability structure ensures that the death state is reachable from every state. Note that when $a(c, l) = I$, the patient receives $R(c, l)$ and moves to the absorbing state $(0, L + 1)$ with certainty.
- $v^*(c, l)$: the value vector that gives the optimal expected remaining reward when the patient is in state (c, l) and has not yet initiated therapy.

With this setup, $v(0, L + 1) = 0$, and for all $(c, l) \neq (0, L + 1)$, Bellman's equations are as follows:

$$v(c, l) = \max \begin{cases} r(c, l) + \sum_{c'} \sum_{l'} \alpha[(c', l')|(c, l)]v(c', l'), \\ R(c, l). \end{cases} \tag{5.1}$$

At times it will be convenient to think of the state transitions while waiting as a two-step process: first there is a transition at the next time period to one of the two variables and then there is a transition to the other variable conditional on the first movement. We explain this more by creating the following components.

- $\psi[c'|(c,l)]$: the probability of transitioning to a state with CD4 category c' at time $t+1$ given that at time t the patient is in state (c,l) and chooses to wait.
- Ψ_l: the transition probability matrix representing the above transitions over CD4 categories when the viral load category at time t is l.
- $p[c'|l',(c,l)]$: the conditional probability of transitioning to a state with CD4 category c' at time $t+1$ given that at time $t+1$ the patient's viral load category is l' and at time t the patient is in state (c,l) and chooses to wait.
- $P_{l,l'}$: the transition probability matrix representing the above transitions over CD4 categories when the viral load category at time t is l and the viral load category at time $t+1$ is l'.
- $q[l'|(c,l)]$: the probability of transitioning to a state with viral load category l' at time $t+1$ given that at time t the patient is in state (c,l) and chooses to wait.
- Q_c: the transition probability matrix representing the above transitions over viral load categories when the CD4 category at time t is c.
- $\gamma[l'|c',(c,l)]$: the conditional probability of transitioning to a state with viral load category l' at time $t+1$ given that at time $t+1$ the patient's CD4 category is c' and at time t the patient is in state (c,l) and chooses to wait.
- $\Gamma_{c,c'}$: the transition probability matrix representing the above transitions over viral load categories when the CD4 category at time t is c and the CD4 category at time $t+1$ is c'.

Therefore, we can express $\alpha[(c',l')|(c,l)]$ as either $q[l'|(c,l)]p[c'|l',(c,l)]$ (by conditioning the CD4 transition on the viral load category the patient moves to) or $\psi[c'|(c,l)]\gamma[l'|c',(c,l)]$ (by conditioning the viral load transition on the CD4 category the patient moves to). Letting $v(0,L+1)=0$, then for all $(c,l)\neq(0,L+1)$, we can rewrite Bellman's equations as

$$v(c,l)=\max\begin{cases} r(c,l)+\sum_{l'} q[l'|(c,l)]\sum_{c'} p[c'|l',(c,l)]v(c',l'), \\ R(c,l), \end{cases} \tag{5.2}$$

or equivalently,

$$v(c,l) = \max \begin{cases} r(c,l) + \sum_{c'} \psi[c'|(c,l)] \sum_{l'} \gamma[l'|c',(c,l)] v(c',l'), \\ R(c,l). \end{cases} \tag{5.3}$$

5.3 STRUCTURAL PROPERTIES

In this section, we develop conditions that lead to structured optimal policies. We first state a definition used in what follows [101].

Definition 5.1. *Let P and P' be two N x N transition probability matrices. We say P' dominates P, $P' \succeq P$, if $\sum_{j=k}^{N} p'[j|i] \geq \sum_{j=k}^{N} p[j|i]$ for $1 \leq i,k \leq N$.*

In other words, if higher states are associated with better outcomes, then one would prefer to transition according to the probabilities in P' than P.

We now consider the following assumptions:

($As5.1$) $r(c,l)$ is nondecreasing in c for each l and nonincreasing in l for each c.

($As5.2$) $R(c,l)$ is nondecreasing in c for each l and nonincreasing in l for each c.

$As5.1$ and $As5.2$ say that as CD4 categories improve (i.e., move to higher levels) or viral load categories improve (i.e., move to lower levels), the one-period reward ($As5.1$) and the terminal reward ($As5.2$) do not decrease.

($As5.3$) $\Psi_1 \succeq \Psi_2 \ldots \succeq \Psi_L$.

$As5.3$ says that a patient would prefer to be in a state with lower levels of viral load than higher levels with respect to transitions on CD4 categories.

($As5.4$) $P_{l,l'}$ is IFR.

$As5.4$ says that given the viral load category is l at time t and is l' at time $t+1$, a patient has better transition probabilities over CD4 categories when coming from a higher CD4 category at time t.

($As5.5$) $P_{l,1} \succeq P_{l,2} \ldots \succeq P_{l,L}$.

$As5.5$ says that a patient with a viral load category l at time t would prefer to transition to a lower viral load category at time $t+1$ with respect to transitions on CD4 categories.

The next three assumptions are analogous to $As5.3$-$As5.5$, respectively. They differ in the variable (CD4 category or viral load category) that is considered in the transition.

($As5.6$) $Q_1 \succeq Q_2 \ldots \succeq Q_C$.

($As5.7$) $\Gamma_{c,c'}$ is IFR.

($As5.8$) $\Gamma_{c,1} \succeq \Gamma_{c,2} \ldots \succeq \Gamma_{c,C}$.

We can now prove the following:

Theorem 5.1. *Suppose assumptions* $As5.1$-$As5.8$ *hold. Then:*

(a) $v^*(c,l)$ *is nondecreasing in* c *for each* l.

(b) $v^*(c,l)$ *is nonincreasing in* l *for each* c.

Proof. We prove this by induction. Let $i \in 0, 1, \ldots$ and suppose $v^i(c,l)$ is nondecreasing in c for each l and nonincreasing in l for each c. Note that $v^0(c,l) = 0$ for all c, l satisfies this property. Fix $l' \in \{1, \ldots, L\}$, and for all c, define:

$$v^{i+1}(c,l') = \max \begin{cases} r(c,l') + \sum_l q[l|(c,l')] \sum_{c'} p[c'|l,(c,l')] v^i(c',l), \\ R(c,l'), \end{cases}$$

and

$$v^{i+1}(c+1,l') = \max \begin{cases} r(c+1,l') + \sum_l q[l|(c+1,l')] \sum_{c'} p[c'|l,(c+1,l')] v^i(c',l), \\ R(c+1,l'). \end{cases}$$

Note the similarity between these equations and the Bellman equations of the type given in (5.2).

By the assumption that $v^i(c,l)$ is nondecreasing in c for each l and assumption $As5.4$, we apply Lemma 4.1 for each l to obtain:

$$\sum_{c'} p[c'|l,(c+1,l')] v^i(c',l) \geq \sum_{c'} p[c'|l,(c,l')] v^i(c',l). \tag{5.4}$$

Now let $l_1 \leq l_2$. Then we have:

$$\begin{aligned} \sum_{c'} p[c'|l_1,(c+1,l')] v^i(c',l_1) &\geq \sum_{c'} p[c'|l_2,(c+1,l')] v^i(c',l_1) && (5.5) \\ &\geq \sum_{c'} p[c'|l_2,(c+1,l')] v^i(c',l_2), && (5.6) \end{aligned}$$

where (5.5) holds by the assumption that $v^i(c, l_1)$ is nondecreasing in c, $As5.5$, and Lemma 4.1. Then (5.6) follows by the assumption that $v^i(h', l)$ is nonincreasing in l. Therefore:

$$\sum_{c'} p[c'|l, (c+1, l')]v^i(c', l) \text{ is nonincreasing in } l. \tag{5.7}$$

Furthermore,

$$\begin{aligned} & \sum_{l} q[l|(c+1, l')] \sum_{c'} p[c'|l, (c+1, l')]v^i(c', l) && (5.8)\\ \geq & \sum_{l} q[l|(c, l')] \sum_{c'} p[c'|l, (c+1, l')]v^i(c', l) && (5.9)\\ \geq & \sum_{l} q[l|(c, l')] \sum_{c'} p[c'|l, (c, l')]v^i(c', l), && (5.10) \end{aligned}$$

where (5.9) follows from (5.8) by $As5.6$, (5.7) and Lemma 4.1, and (5.10) follows from (5.9) by (5.4).

By assumption $As5.1$ and the relationships in (5.8)-(5.10), we have:

$$\begin{aligned} & r(c+1, l') + \sum_{l} q[l|(c+1, l')] \sum_{c'} p[c'|l, (c+1, l')]v^i(c', l) \\ \geq & \; r(c, l') + \sum_{l} q[l|(c, l')] \sum_{c'} p[c'|l, (c, l')]v^i(c', l). && (5.11) \end{aligned}$$

Next, by assumption $As5.2$,

$$R(c+1, l') \geq R(c, l'). \tag{5.12}$$

Combining the relationships expressed in (5.11) and (5.12) we get:

$$\begin{aligned} & \max\left\{ r(c+1, l') + \sum_{l} q[l|(c+1, l')] \sum_{c'} p[c'|l, (c+1, l')]v^i(c', l), \quad R(c+1, l') \right\} \\ \geq & \; \max\left\{ r(c, l') + \sum_{l} q[l|(c, l')] \sum_{c'} p[c'|l, (c, l')]v^i(c', l), \quad R(c, l') \right\}, \end{aligned}$$

which in turn implies:

$$v^{i+1}(c+1, l') \geq v^{i+1}(c, l').$$

Thus, $v^{i+1}(c, l)$ is nondecreasing in c for an arbitrary l. In almost identical fashion, we can also show that $v^{i+1}(c, l)$ is nonincreasing in l for an arbitrary c by switching the concept of

CD4 categories with viral load categories, using the form of Bellman's equation expressed in (5.3), and proceeding using assumptions $As5.7$, $As5.8$, and $As5.3$, in place of $As5.4$, $As5.5$, and $As5.6$, respectively. Then, taking the limit of the iterates proves that $v^*(c,l)$ is nondecreasing in c for each l and nonincreasing in l for each c. □

The following theorem and corollary are analogous to Theorem 4.2 and Corollary 4.1. The theorem gives necessary and sufficient conditions for an optimal policy of initiating therapy from all states, and the corollary gives sufficient conditions for waiting to be uniquely optimal for a particular health state. They are stated without proof.

Theorem 5.2. *$a^*(c,l) = I$ for all (c,l) if and only if:*

$$R(c,l) \geq r(c,l) + \sum_{c'} \sum_{l'} \alpha[(c',l')|(c,l)]R(c',l') \quad \textit{for all } (c,l).$$

Corollary 5.1. *If $R(c',l') < r(c',l') + \sum_c \sum_l \alpha[(c,l)|(c',l')]R(c,l)$ for some (c',l'), then $a^*(c',l') = W$, uniquely.*

Next we explore conditions under which an optimal control-limit policy exists. Analogous to Assumption $As4.3$, we consider the following two assumptions.

($As5.9$) $p[c'|(c,l)] = 0$ for $c' > c$.

($As5.10$) $q[l'|(c,l)] = 0$ for $l' < l$.

$As5.9$ and $As5.10$ imply that the CD4 category and viral load category, respectively, cannot improve while waiting to initiate HIV therapy, which has intuitive appeal.

We define the following:

$$f(c,l) \equiv \frac{r(c,l) + \sum_{c'<c} \sum_{l' \geq l} \alpha[(c',l')|(c,l)]R(c',l') + \sum_{l'>l} \alpha[(c,l')|(c,l)]R(c,l')}{1 - \alpha[(c,l)|(c,l)]}.$$

In words, $f(c,l)$ is the value associated with a patient being in state (c,l), under a policy that waits whenever in that state, but initiates therapy when a transition is made to a worse state (this is under the assumptions $As5.9$ and $As5.10$). We present two additional assumptions.

($As5.11$) $f(c+1,l) - f(c,l) \geq R(c+1,l) - R(c,l)$ for all (c,l).

($As5.12$) $f(c,l-1) - f(c,l) \geq R(c,l-1) - R(c,l)$ for all (c,l).

$As5.11$ can be thought of as saying that as the CD4 category improves, there is more to be gained by waiting relative to initiating therapy. Conversely, it says that as the CD4

category worsens, there is more to be gained by initiating compared to waiting. *As*5.12 makes the same statement with respect to changes in viral load.

We are now ready to prove sufficient conditions for the existence of a two-dimensional optimal control-limit policy, along with structure on the control limits themselves.

Theorem 5.3. *Under assumptions As*5.9-*As*5.12
(a) an optimal CD4-based control-limit policy exists for each level of viral load. That is to say, for each l there exists $CL_C(l)$, called the "CD4-based control limit," such that whenever $c < CL_C(l)$, $a^*(c,l) = I$, and whenever $c \geq CL_C(l)$, $a^*(c,l) = W$.
(b) an optimal viral-load-based control-limit policy exists for each level of CD4. That is to say, for each c there exists $CL_L(c)$, called the "viral load-based control limit," such that whenever $l \leq CL_L(c)$, $a^*(c,l) = W$, and whenever $l > CL_L(c)$, $a^*(c,l) = I$.
(c) $CL_L(c)$ is nondecreasing in c and $CL_C(l)$ is nondecreasing in l.

Proof. Consider $l = L$, and let c' be the smallest c such that $a^*(c,L) = W$ (if no such c' exists, then a trivial CD4-based control-limit policy of $a^*(c,L) = I$ for all c exists). By assumptions *As*5.9 and *As*5.10, and Bellman's optimality equations, this implies that

$$f(c',L) \geq R(c',L).$$

By assumptions *As*5.11 and *As*5.12, it follows that

$$f(c,l) \geq R(c,l) \qquad \text{for all } (c,l) \text{ such that } c \geq c' \text{ and } l \leq L. \tag{5.13}$$

Because $f(c,l) \geq R(c,l)$ for all (c,l) with $c \geq c'$ and $l \leq L$, and because $f(c,l)$ represents the value of waiting in state (c,l) and initiating therapy for all other states that may follow, the value of waiting in state (c,l) and taking the *optimal* action for all other states that may follow can be no worse. Therefore, $v^*(c,l) \geq R(c,l)$ and $a^*(c,l) = W$ for all (c,l) with $c \geq c'$ and $l \leq L$. By definition, $\mathrm{CL}_C(L) = c'$ and $\mathrm{CL}_L(c) = L$ for all $c \geq c'$. Next, consider $l = L-1$. Let c'' be the smallest value of c such that $a^*(c, L-1) = W$. By the previous step, we know that $c'' \leq \mathrm{CL}_C(L)$. If $c'' < \mathrm{CL}_C(L)$, then it follows that $f(c'', L-1) \geq R(c'', L-1)$, and by the same reasoning as above, we can conclude that $a^*(c,l) = W$ for all (c,l) with $c \geq c''$ and $l \leq L-1$. Iterating in this fashion, from $l = L$ down to $l = 1$, proves the theorem. □

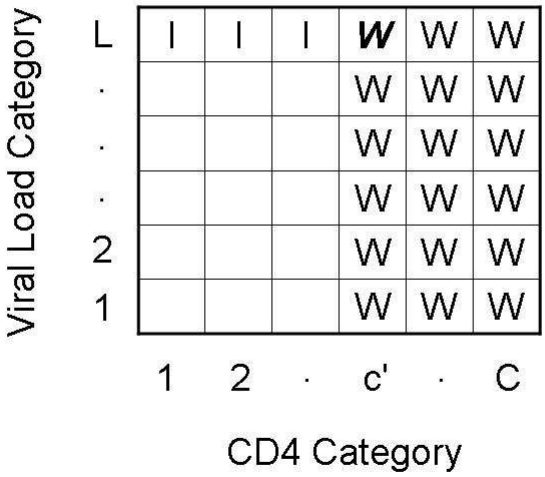

Figure 10: Partial optimal policy after first inductive step

Figure 10 shows how the optimal policy matrix may look after completing the first step of the proof (where $l = L$), and Figure 11 shows how it may look after the final step of the proof.

Comment on Control-Limit Policies in Two Dimensions

Although establishing control-limit policies arises in many problems, there is little in the literature regarding existence of such policies in multidimensional state spaces. Benyamini and Yechiali [19] considered a nonstationary machine maintenance problem with a system variable indicating the state of the machine and another variable indicating the age. They established sufficient conditions for an optimal control-limit policy (with respect to actions of wait or replace) in the following sense: for a fixed age t, there is a control-limit policy with respect to the machine states, and for a fixed machine state i, there is a control-limit policy with respect to the machine's age. There are also some references to two-dimensional

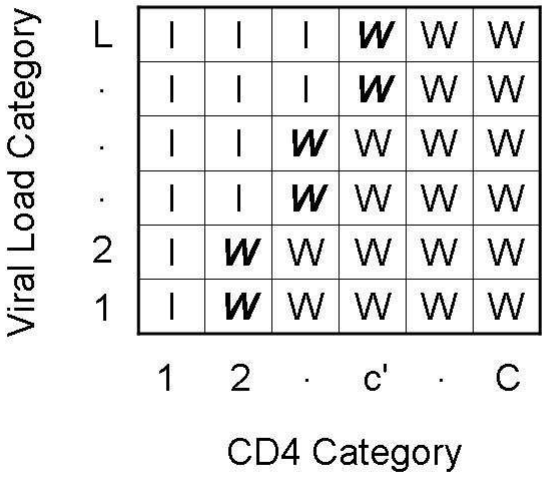

Figure 11: Example of an optimal policy under Theorem 5.3

control-limit policies, or "switching curves", in the queueing control literature. Hajek [74] considered a two-station queueing network, which is controlled as follows: a new customer has some probability of going to each queue, an extra worker spends some portion of time servicing each queue, and customers needing rework are to be sent from one queue to the other according to some station-dependent probability. Under assumptions on the cost and probability transition structures of the network, Hajek proved the existence of a switching curve policy that says: for a given level of customers at station 1, if the number of customers at station 2 is below a certain value, send new arrivals to station 2; otherwise send them to station 1. Furthermore, this limit increases with the number waiting at station 1. Wu et al. [187] considered a two-stage tandem queueing system in which configurable resources can be assigned to work on either queue. They also proved that the optimal policy has a switching curve structure in terms of which queue the configurable resource should service. Unlike

in [74], they were unable to prove monotonic structure of the curve; however, over 30,000 simulations supported this property. Finally, Alagoz et al. [6] developed a model of patient acceptance of cadaveric liver donations with a two-dimensional state space of patient health and liver quality. The authors established conditions guaranteeing an optimal control-limit policy with respect to liver qualities for a given health level, and vice-versa.

5.4 CONCLUSIONS

This chapter expanded on the previous chapter by including an extra variable in the state space. We presented clinically plausible assumptions that led to structural properties analogous to those obtained in the one-dimensional case of the previous chapter. We have not yet constructed the MDP model with actual data; however, we plan to do that in the near future. Our contribution with this chapter was to prove structural properties for a problem with a multidimensional state space. Despite many real problems requiring a multidimensional state space, little theoretical work has been done in this area.

6.0 LIFETIME DISTRIBUTION APPROACH TO THE OPTIMAL SEQUENCING AND SWITCHING OF THERAPIES

6.1 INTRODUCTION

Whereas the previous two chapters focused on the optimal time to initiate therapy, this chapter and the next explore methodological frameworks for considering the optimal switching and sequencing of HIV therapies. This problem is significantly more challenging than the question of when to initiate therapy, both from a computational standpoint as well as a data standpoint. Instead of considering an optimal time to go from being off of therapy to initiating therapy, here we have the problem of timing the start of the first therapy, the start of the second therapy, and so on. Moreover, we have all the permutations of therapy sequences to contend with. Perhaps even more challenging is building a data-driven model. For example, we are limited to available data for actual sequences of therapies that have been offered in clinical settings; it is difficult or even infeasible to estimate the effects of certain sequences of therapies that have never been tried before. The rest of this chapter forms the basis of a paper by Shechter et al. [157].

In thinking of the body as a system to be maintained and therapies as components to be replaced over time, we are able to glean insights from machine maintenance problems and extend them to the context of therapeutic optimization. Despite some natural connections, very little work has been done in this area. We begin by describing one paper in particular from the machine maintenance literature and then extend it to the scheduling of therapies.

Most machine maintenance models assume an infinite supply of replacements and assume that each one behaves stochastically identical and independent of the others. As a result, optimal periodic policies can be obtained by modeling the system as a renewal process.

Derman et al. [58], however, considered a system with one vital component and a finite number of spares. Moreover, unlike most machine models that allow replacements after a failure, they assumed that a failure ends the system's lifetime. Under the usual assumption that the lifetime of a spare is independent and identically distributed (iid) to the lifetimes of the other spares, the authors sought a replacement schedule that maximizes the expected lifetime of the system. In other words, they considered the following question: how much of a current part's lifetime should be used up before replacing it with a new part?

We can think of this problem in the context of maintaining a patient's life: the patient is the system, available therapies are the spare parts, and our objective is to maximize the expected lifetime of the patient (system). Because therapies may have different levels of effectiveness, we extend the work by Derman et al. by considering replacing components whose lifetimes may not be identically distributed. Also, because the use of one therapy can alter the effectiveness of another therapy not yet used (due to cross-resistance), ideally one should consider these dependencies when evaluating a particular sequence. The problem becomes not just one of timing replacements, as in Derman et al., but also one of determining an optimal sequence for them. Clearly, if a patient dies while on a given therapy, there is nothing more that can be done; therefore, the problem lies in trying to prolong the time until this happens by strategically scheduling replacements over time. We begin by keeping the assumption of independent lifetime distributions, while relaxing the assumption that they are all identical. In Section 6.3.3, we consider relaxing the independence assumption as well. We note that implicit in both frameworks is that there are no state observations other than whether the patient (system) is alive (functioning) or dead (failed). All we know in this framework are the lifetime distributions. Chapter 7 considers the possibility of periodic informative state observations.

6.2 BACKGROUND

We describe in detail some of the work by Derman et al. [58], which considers a system with one vital component and n identical spare components. The lifetime distributions of

the components are iid with common cumulative distribution $F(t)$, continuous density $f(t)$, support $(0, T)$ $(0 < T \leq \infty)$, and finite mean μ. Once a component is replaced with a spare component, it cannot be used again, and if the component in service fails, the system lifetime ends. The objective is then to determine when to replace functioning components so as to maximize the expected total lifetime of the system. They let v_i represent the maximum expected system lifetime when there are i spare components available. Thus v_0 represents the expected system lifetime when no spare components are available and the system uses only the initially installed component. Hence $v_0 = \mu$. For $i \geq 1$, v_i satisfies the optimality equation given by

$$v_i = \max_{0 \leq t \leq T} \phi_i(t),$$

where

$$\begin{aligned} \phi_i(t) &= \int_0^t \tau f(\tau)\, d\tau + \bar{F}(t)(t + v_{i-1}) && (6.1) \\ &= \int_0^t \bar{F}(\tau)\, d\tau + \bar{F}(t) v_{i-1}, && (6.2) \end{aligned}$$

where $\bar{F}(t) \equiv 1 - F(t)$, and (6.2) is obtained from (6.1) by integration by parts. Let t_i be a switching time that yields v_i.

In some instances it may never be optimal to install a new component. To address this, the authors consider a distribution "new worse than used in expectation" (NWUE) [13], which is when F has a finite mean and

$$\mu \leq \frac{\int_t^T \bar{F}(\tau)\, d\tau}{\bar{F}(t)} \qquad \text{for all } 0 \leq t \leq T. \tag{6.3}$$

In other words, F is NWUE if at every point in time, the conditional expected remaining mean (the right hand side of (6.3)) is at least as great as the mean of a new component (the left hand side of (6.3)). The following is equivalent to (6.3),

$$\mu \geq \int_0^t \bar{F}(\tau)\, d\tau + \bar{F}(t)\mu \qquad \text{for all } 0 \leq t \leq T,$$

which in the context of the above problem is equivalent to being optimal to never switch components. Therefore, the authors assume that F is not NWUE and concentrate on policies for which it is best to eventually replace a component. This implies that $v_1 > v_0$ and leads to the following results:

(i) v_i *is strictly increasing in* i. That is, there is an increase in expected system lifetime with an increase in spares.

(ii) $v_i - v_{i-1} < v_{i-1} - v_{i-2}, \quad n \geq 2$. That is, there are diminishing marginal returns with more spares.

(iii) $0 < t_n < t_{n-1} ... < t_1 < T$. That is, the optimal time to wait before replacing a working component with one of the spares increases as the remaining spares decreases.

6.3 EXTENSION TO NON-IDENTICAL THERAPIES

We now extend the work by Derman et al. to consider the optimal scheduling of n possibly different therapies, while maintaining the assumption of independence between them. We assume a patient is currently not on any therapy and has n therapies available. We label each therapy by i $(i = 1, ..., n)$, assume the cumulative distribution function of patient survival under therapy i is known and given by $F_i(t)$, with continuous density given by $f_i(t)$, support given by $0 \leq T \leq \infty$, and finite mean μ_i. In other words, if the patient starts taking therapy i and remains on it until death, the patient's remaining lifetime distribution is characterized by F_i, and the patient's expected lifetime is μ_i. We allow the possibility that a patient does not initiate therapy immediately and let F_0 be the distribution function representing the patient's remaining lifetime when never taking therapy. By a schedule we mean a sequence of the therapies $\{a_1, a_2, \ldots, a_n\}$ and an associated sequence of durations $\{t_0, t_1, t_2, \ldots, t_{n-1}\}$, such that the patient starts taking therapy a_1 after t_0 units of time pass, remains on therapy a_1 for t_1 units of time, switches to therapy a_2 and remains on that for t_2 units of time, and so on, until the patient is on the final therapy, a_n, and remains on that therapy until death. We assume that once patients initiate therapy, they remain on therapy for the remainder of their lives. DHHS guidelines for therapy recommend against going off therapy later in the course of treatment as rapid progression to AIDS and death has been observed in patients who do so [123]. We also assume that once a therapy is used, it cannot be used again (we discuss this more in Section 7.5).

6.3.1 Special Case: Two Therapies Available

Before considering the optimal scheduling of many therapies, we first consider how to do this for two available therapies. Moreover, we explore structural properties in this framework. Starting with this simplified setting provides insight to our approach for the more complicated situation of more than two therapies available, discussed in Section 6.3.2.

Consider two therapies, A and B, and suppose the patient initiates therapy immediately. We want to know if it is better for the patient to start with therapy A and switch to B at some time, or vice-versa. If the patient begins with therapy A followed by B, then the optimal value associated with this sequence, v_{AB}, is found via:

$$v_{AB} = \sup_{0 \le t \le T} \phi_{AB}(t), \tag{6.4}$$

where

$$\begin{aligned} \phi_{AB}(t) &= \int_0^t \tau f_A(\tau)\, d\tau + \bar{F}_A(t)(t + \mu_B) \\ &= \int_0^t \bar{F}_A(\tau)\, d\tau + \bar{F}_A(t)\mu_B. \end{aligned} \tag{6.5}$$

In words, $\phi_{AB}(t)$ is the patient's expected lifetime when starting with therapy A and switching to therapy B after t units of time, and v_{AB} finds the supremum of this function over the possible values of t. We assume $\phi_{AB}(t)$ (and $\phi_{BA}(t)$) is continuous and bounded; therefore, if T (the upper limit of the distributions) is finite, we can find a value of t, say t_A, that attains the supremum in $\phi_{AB}(t)$. However, if T is infinite, then the supremum might only be attainable in the limit. We provide an example of this in the next section.

As above, if we begin with therapy B, we have:

$$v_{BA} = \sup_{0 \le t \le T} \phi_{BA}(t),$$

where

$$\phi_{BA}(t) = \int_0^t \bar{F}_B(\tau)\, d\tau + \bar{F}_B(t)\mu_A.$$

We then define

$$v^* = \max\{v_{AB}, v_{BA}\}.$$

To gain insight into this framework, we consider three separate conditions for the lifetime distributions of the therapies: 1) independent exponential lifetime distributions with possibly different rates, 2) independent uniform distributions over possibly different ranges, and 3) independent triangular distributions with possibly different modes and ranges.

Exponential Lifetime Distributions

Suppose therapies A and B have independent exponential lifetime distributions with rates λ_A and λ_B, respectively. The memoryless property of the exponential distribution suggests that there is no merit to the notion of "using up" some lifetime under one exponential distribution and switching to another exponential distribution. Therefore, the optimal policy should be to use only the therapy with the smallest exponential rate. Formally, we have

$$\phi_{AB}(t) \;=\; \int_0^t e^{-\lambda_A \tau}\, d\tau + \frac{e^{-\lambda_A t}}{\lambda_B}.$$

Evaluating the integral, and referring to (6.4), results in

$$v_{AB} = \sup_{0 \le t \le \infty} \left[\frac{(\lambda_A - \lambda_B) e^{-\lambda_A t} + \lambda_B}{\lambda_A \lambda_B} \right].$$

Clearly, if $\lambda_A > \lambda_B$ (and hence $\mu_A < \mu_B$), v_{AB} is obtained by letting $t = 0$, which means it is optimal to immediately replace therapy A with therapy B. If $\lambda_A < \lambda_B$, then it is optimal (in a limiting sense) to let t be as large as possible, which in this case means to never replace therapy A with B. As mentioned earlier, this is a case of v_{AB} being attained only in the limit as $t \to \infty$. Note that if $\lambda_A = \lambda_B$, then it does not matter if or when therapy A is replaced with B. Generalizing these results to n therapies with independent and exponentially distributed lifetimes with different rates, it is optimal to use only the therapy with the smallest rate.

Uniform Lifetime Distributions

Suppose therapies A and B have independent uniform lifetime distributions with supports $[0, a]$ and $[0, b]$, with means $\mu_A = a/2$ and $\mu_B = b/2$, respectively. Then,

$$\phi_{AB}(t) \;=\; \int_0^t (a - \tau)/a \, d\tau + \frac{\mu_B (a - t)}{a}.$$

This results in

$$v_{AB} = \sup_{0 \le t \le a} \left[t - \frac{\mu_B t}{a} - \frac{t^2}{4\mu_A} + \mu_B \right],$$

which is maximized when $t = a - \mu_B$. Therefore, if A must precede B, and if $a < \mu_B$, then it is optimal to discard A immediately and only use therapy B. Otherwise, it is optimal to use A for $a - b/2$ time units and then switch to B. By symmetry, if B must precede A and $b < \mu_A$, then it is optimal to discard B immediately and use only A. Otherwise, it is optimal to use B for $b - a/2$ time units and then switch to A. Interestingly, if the patient has to use therapy A before therapy B and if $a < \mu_B$, then the patient should not spend any time on therapy A. However, if the patient must use B before A, then the patient should use therapy A at some point (specifically, after $b - \mu_A$ time units).

In considering the optimal sequencing of A and B, we consider three cases: 1) $b < a/2$, 2) $b > 2a$, and 3) $a/2 \le b \le 2a$ (or, equivalently, $b/2 \le a \le 2b$). By the above discussion, we can see that if case 1 holds, then it is optimal to start with therapy A, wait $a - \mu_B$ units of time, and then switch to therapy B. If case 2 holds, then it is optimal to start with therapy B, wait $b - \mu_A$ units of time, and then switch to therapy A. For case 3, we show that it is optimal to start with therapy A if $a > b$ or to start with therapy B if $b > a$. We have the following:

$$\begin{aligned} v_{AB} &= a - b/2 - \frac{b(a - b/2)}{2a} - \frac{(a - b/2)^2}{2a} + b/2, \quad \text{and} \\ v_{BA} &= b - a/2 - \frac{a(b - a/2)}{2b} - \frac{(b - a/2)^2}{2b} + a/2. \end{aligned}$$

After simplification, we find that $v_{AB} - v_{BA} > 0$ if $a > b$ and $(a + b)^2 < 5ab$. Suppose $a > b$ and consider two functions, $g_1(t) \equiv (b + t)^2$, and $g_2(t) \equiv 5bt$, where $b < t \le 2b$. Then we see that g_2 is the line connecting $(b, 5b^2)$ with $(2b, 10b^2)$, and g_1 is the parabola connecting $(b, 4b^2)$ with $(2b, 9b^2)$. Because the line lies above the parabola for the range of t, we see that if $b < a \le 2b$, then $(a + b)^2 < 5ab$ and thus $v_{AB} > v_{BA}$. By symmetry, if $a < b \le 2a$, it is optimal to sequence B followed by A. Note that if $a = b$, then the sequencing of the therapies does not matter. The optimal policy is achieved by starting with either one and switching after $\mu_A = \mu_B$ units of time.

In summary, if $a \geq b$, it is optimal to start with therapy A and switch to B after $a - \mu_B$ units of time and if $b \geq a$, it is optimal to start with therapy B and switch to A after $b - \mu_A$ units of time. This also demonstrates that it is optimal to sequence the therapies in decreasing order of the means. In some sense, it seems intuitive that if one therapy has a lifetime distribution with a greater mean than another, then it is optimal to get as much life out of that first therapy before switching to the other one at some later time. As we show next, however, this does not always hold.

Triangular Lifetime Distributions

Suppose therapies A and B have independent triangular lifetime distributions L_A and L_B described by $[0, \alpha_A, \beta_A]$ and $[0, \alpha_B, \beta_B]$, respectively, where α_A represents the mode of the lifetime distribution for therapy A, and 0 and β_A represent the lower and upper limits of the distribution (similarly defined for therapy B). The mean lifetime on therapy A is given by $\mu_A = (0 + \alpha_A + \beta_A)/3$, and

$$\bar{F}_A(t) = \begin{cases} 1 - \frac{t^2}{\alpha_A \beta_A} & : \quad 0 \leq t \leq \alpha_A, \\ \frac{(\beta_A - t)^2}{\beta_A(\beta_A - \alpha_A)} & : \quad \alpha_A < t \leq \beta_A. \end{cases}$$

In evaluating $\phi_{AB}(t)$, we have to consider whether $t \leq \alpha_A$ or $t > \alpha_A$. We have,

$$\phi_{AB}(t) = \begin{cases} \int_0^t (1 - \frac{\tau^2}{\alpha_A \beta_A})\, d\tau + \mu_B (1 - \frac{t^2}{\alpha_A \beta_A}) & : \quad 0 \leq t \leq \alpha_A \\ \int_0^{\alpha_A} (1 - \frac{\tau^2}{\alpha_A \beta_A})\, d\tau + \int_{\alpha_A}^t \frac{(\beta_A - \tau)^2}{\beta_A(\beta_A - \alpha_A)}\, d\tau + \mu_B \frac{(\beta_A - t)^2}{\beta_A(\beta_A - \alpha_A)} & : \quad \alpha_A < t \leq \beta_A. \end{cases} \tag{6.6}$$

Consider L_A defined by $[0, 50, 55]$ ($\mu_A = 35$) and L_B defined by $[0, 10, 80]$ ($\mu_B = 30$). Figure 12 shows how $\phi_{AB}(t)$ and $\phi_{BA}(t)$ vary over the switching times, and we can see that $\phi_{AB}(t)$ reaches a maximum of around 47, whereas $\phi_{BA}(t)$ reaches a maximum of around 40. Therefore, it is optimal to sequence A before B. Now consider L_A defined by $[0, 5, 100]$ ($\mu_A = 35$) and L_B defined by $[0, 30, 60]$ ($\mu_B = 30$). In Figure 13, we see that $\phi_{BA}(t)$ reaches a maximum value of around 46, whereas $\phi_{AB}(t)$ reaches a maximum of around 39. As a result, it is optimal to sequence B before A. Therefore, we have found an example of where it is optimal to use the therapy with the smaller mean *before* using the therapy with the larger mean. Next, we consider stronger assumptions on the lifetime distributions of the therapies and whether such assumptions allow us to make general statements about the

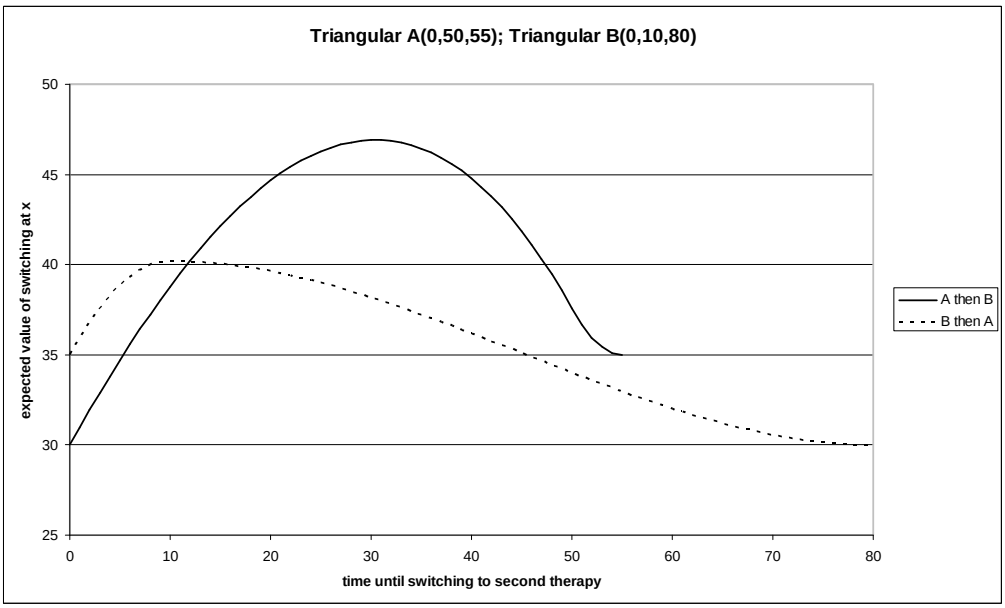

Figure 12: Plots of $\phi_{AB}(t)$ and $\phi_{BA}(t)$ when $A \sim Triang(0, 50, 55)$ and $B \sim Triang(0, 10, 80)$

optimal sequencing of two therapies. Surprisingly, we shall show that even these stronger conditions do not provide sufficient conditions to know a priori the optimal sequencing.

Sequencing Two Therapies under General Stochastic Ordering Relationships

In this section, we consider more restrictive orderings based on information concerning the entire distribution instead of a single parameter such as the mean. We might posit that knowing a stochastic ordering of the therapies can allow us to say something about the optimal sequencing of the therapies. Knowing the optimal sequencing a priori would decrease the overall computational burden of the problem. The following definitions are found in [156].

Definition 6.1. Y is larger than X under the usual stochastic order *($Y \succeq_{st} X$) if for all t,* $\bar{F}_Y(t) \geq \bar{F}_X(t)$.

Definition 6.2. Y is larger than X under the hazard rate order *($Y \succeq_{hr} X$) if* $\bar{F}_Y(t)/\bar{F}_X(t)$ *is nondecreasing in t.*

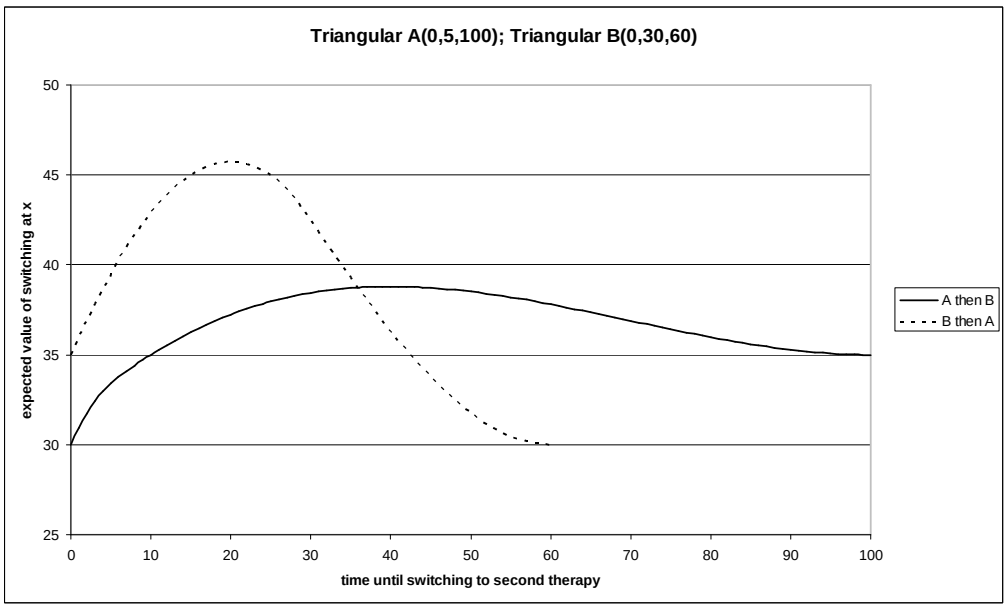

Figure 13: Plots of $\phi_{AB}(t)$ and $\phi_{BA}(t)$ when $A \sim Triang(0, 5, 100)$ and $B \sim Triang(0, 30, 60)$

Definition 6.3. *Y is larger than X under the likelihood ratio order $(Y \succeq_{lr} X)$ if $f_Y(t)/f_X(t)$ is nondecreasing in t over the union of the supports of X and Y.*

In Definitions 6.2 and 6.3, $a/0 \equiv \infty$ when $a > 0$. Now let $m_X(t)$ represent the expected remaining lifetime of X, given it has survived until time t. That is, $m_X(t) \equiv E[X - t|X > t]$. Similarly, let $m_Y(t) \equiv E[Y - t|Y > t]$. We then have the following definition [156].

Definition 6.4. *Y is larger than X under the mean residual life order $(Y \succeq_{mrl} X)$ if $m_Y(t) \geq m_X(t)$ for all t.*

With these definitions, the following relations hold [156]:

$$Y \succeq_{lr} X \Rightarrow Y \succeq_{hr} X \Rightarrow Y \succeq_{st} X \Rightarrow \mu_Y \geq \mu_X,$$

and

$$Y \succeq_{hr} X \Rightarrow Y \succeq_{mrl} X.$$

Returning to the two therapies A and B, intuition may suggest that if $L_B \succeq_{st} L_A$, then it is optimal to use therapy B for some time before switching to therapy A. And if this is true, then the strongest of ordering conditions, ($L_B \succeq_{lr} L_A$), would also guarantee the optimality of starting with therapy B. Unfortunately, as we show next, even the strongest of these stochastic ordering conditions does not guarantee a priori knowledge of the optimal sequencing of two therapies.

Example: Let $L_A \sim \mathit{Triang}(0, 5, 10)$ and $L_B \sim \mathit{Triang}(0, 6, 12)$. We begin by showing that $L_B \succeq_{lr} L_A$. To demonstrate this, we show that $f_B(t)/f_A(t)$ is nondecreasing in t over $(0, 12)$. Based on these distributions, the probability density functions for L_A and L_B are

$$f_A(t) = \begin{cases} t/25 & : \quad 0 \leq t \leq 5 \\ (10-t)/25 & : \quad 5 < t \leq 10 \end{cases}$$

and

$$f_B(t) = \begin{cases} t/36 & : \quad 0 \leq t \leq 6 \\ (12-t)/36 & : \quad 6 < t \leq 12. \end{cases}$$

In the following, we let $r(t) \equiv f_B(t)/f_A(t)$.

Case I: $t \leq 5$

$r(t) = 25/36$ which is clearly nondecreasing.

Case II: $5 < t \leq 6$

$r(t) = \frac{25t}{36(10-t)}$. Note that $\lim_{t \to 5+} r(t) = 25/36 = r(5)$. Furthermore, $r'(t) = \frac{25}{36(10-t)} + \frac{25t}{36(10-t)^2}$, which is positive on $(5, 6]$, and therefore, $r(t)$ is increasing on that interval.

Case III: $6 < t < 10$

$r(t) = \frac{25(12-t)}{36(10-t)}$. Note that $\lim_{t \to 6+} r(t) = 25/24 = r(6)$. Furthermore, $r'(t) = \frac{-25}{36(10-t)} + \frac{25(12-t)}{36(10-t)^2}$, which is positive on $(6, 10)$, and therefore, $r(t)$ is increasing on that interval as well.

Case IV: $10 \leq t < 12$

As discussed above, $r(t) = \infty$.

Hence, $f_B(t)/f_A(t)$ is nondecreasing in t over $(0, 12)$, from which it follows that $L_B \succeq_{lr} L_A$. We now show the counterintuitive result that $v_{BA} < v_{AB}$. First let us derive the

value of v_{AB}. In doing so, we will make use of the expression for $\phi_{AB}(t)$ given in (6.6). In seeking the maximum of $\phi_{AB}(t)$, we consider $\phi'_{AB}(t)$ for $t \leq 5$ and $t > 5$. For $t \leq 5$ we have $\phi'_{AB}(t) = 1 - t^2/50 - 6t/25$, which has a local maximum at $t = \frac{12-\sqrt{344}}{-2} \approx 3.274$. For $t > 5$, $\phi'_{AB}(t) = (10-t)^2/50 - 12(10-t)/50$, so that $\phi_{AB}(t)$ is decreasing on $(\frac{12-\sqrt{344}}{-2}, 10)$. Therefore, if one is to sequence therapy A followed by B, it is optimal to switch from A to B at $t_A = \frac{12-\sqrt{344}}{-2}$. The value associated with doing so is given by:

$$\phi_{AB}(t_A) = \int_0^{t_A} (1 - \tau^2/50)\, d\tau + 6(1 - t_A^2/50),$$

which evaluates to ≈ 7.754. A similar argument shows that $\phi_{BA}(t_B) \approx 7.688$, and therefore $v_{BA} < v_{AB}$.

Intuitively, it would not seem that $L_B \succeq_{lr} L_A$ would imply an optimal policy of sequencing A followed by B. We show that this implication does not hold, concluding that the likelihood ratio ordering does not imply a priori knowledge of the optimal sequencing of therapies. Let $L_A \sim \mathit{Unif}[0, a]$ and $L_B \sim \mathit{Unif}[0, b]$, where $b > a$. Then clearly, $L_B \succeq_{lr} L_A$ since

$$f_B(t)/f_A(t) = \begin{cases} a/b & : \quad 0 \leq t < a, \\ \infty & : \quad a \leq t < b. \end{cases}$$

However, as presented previously with therapies having uniform lifetime distributions, it is optimal to sequence B followed by A and hence $v_{BA} \geq v_{AB}$.

We now consider the more general question of establishing sufficient conditions for a priori sequencing of two therapies A and B based solely on distribution information. One set of sufficient conditions could require that $\phi_{BA}(t) \geq \phi_{AB}(t)$ for all t, from which it would immediately follow that $\phi_{BA}(t_B) \geq \phi_{AB}(t_A)$, implying that it is optimal to start with therapy B before switching to A at some point. However, if $\phi_{BA}(t) \geq \phi_{AB}(t)$ for all t then $\phi_{BA}(0) = \mu_A \geq \phi_{AB}(0) = \mu_B$ and $\phi_{BA}(\infty) = \mu_B \geq \phi_{AB}(\infty) = \mu_A$. These conditions imply that μ_A must equal μ_B. In general, we will be scheduling therapies which could have different lifetime means.

In continuing to explore sufficient conditions for determining the optimal sequencing, let us suppose that $\mu_B \geq \mu_A$, and consider the two functions $\phi_{AB}(t)$ and $\phi_{BA}(t)$. If we switch from A to B immediately, then we have $\phi_{AB}(0) = \mu_B$. Conversely, if we start with B and never switch to A, we have $\phi_{BA}(\infty) = \mu_B$. Therefore, a sufficient condition for $\phi_{BA}(t_B) \geq \phi_{AB}(t_A)$ is that $\mu_B \geq \phi_{AB}(t)$ for all t. In other words, the condition states that it is better to start with therapy B and stay on it indefinitely than to start with therapy A and switch to B at some time. This condition appears to be too strong because it implies that therapy A induces such poor survival that using it for any amount of time is too risky. The existence of other conditions that guarantee a priori knowledge of the optimal sequencing of therapies remains an open question.

The Value of Information

We consider the following situation: Suppose an optimal policy has a patient initiate therapy A and switch to therapy B at time t_A. Does knowing that the patient is alive at some time $0 < t' < t_A$ change the optimal time to switch to therapy B? We show that this information has no effect on the optimal policy, and thus the patient waits another $t_A - t'$ time units before switching therapies.

Theorem 6.1. *Suppose the solution to v_{AB} is obtained at time t_A. Also, suppose that the patient is alive on therapy A at time $t' < t_A$. Then it is still optimal to switch therapy at time t_A.*

Proof. If t_A yields the value v_{AB}, then by (6.4) and (6.5), t_A maximizes the following:

$$\begin{aligned} \phi_{AB}(t) &= \int_0^t \tau f_A(\tau)\, d\tau + \bar{F}_A(t)(t + \mu_B) \\ &= \int_0^t \bar{F}_A(\tau)\, d\tau + \bar{F}_A(t)\mu_B. \end{aligned}$$

By assumption, we observe that the patient is alive on therapy A at time $t' < t_A$. In reevaluating the optimal time to switch therapies, we search for the value of t which maximizes the following:

$$\phi_{AB}(t|t \geq t') = \int_{t'}^t \tau \frac{f_A(\tau)}{\bar{F}_A(t')}\, d\tau + \frac{\bar{F}_A(t)}{\bar{F}_A(t')}(t + \mu_B)$$

which can be rewritten as:

$$\phi_{AB}(t|t \geq t') = \frac{1}{\bar{F}_A(t')}\left[\int_{t'}^{t} \tau f_A(\tau)\, d\tau + \bar{F}_A(t)(t + \mu_B)\right]$$

After one more step, we have the equivalent expression:

$$\phi_{AB}(t|t \geq t') = \frac{1}{\bar{F}_A(t')}\left[\int_{0}^{t} \tau f_A(\tau)\, d\tau + \bar{F}_A(t)(t + \mu_B) - \int_{0}^{t'} \tau f_A(\tau)\, d\tau\right]. \tag{6.7}$$

The last integral in (6.7) is just a constant whose derivative is 0. Furthermore, the outside multiplier of $\frac{1}{\bar{F}_A(t')}$ has no effect on finding where the derivative of $\phi_{AB}(t|t \geq t')$ is equal to 0. In other words, the roots of $\phi'_{AB}(t|t \geq t')$ are the same as those for $\phi'_{AB}(t)$. Moreover, because the outside multiplier is positive, the sign of $\phi''_{AB}(t|t \geq t')$ matches the sign of $\phi''_{AB}(t)$ at every t. Therefore, the value of t that maximizes $\phi_{AB}(t)$ also maximizes $\phi_{AB}(t|t \geq t')$. □

Although knowing the patient is alive does not change the optimal policy, the optimal value function increases.

Theorem 6.2. $\phi_{AB}(t_A|t \geq t') > \phi_{AB}(t_A)$.

Proof. We have:

$$\begin{aligned}
\phi_{AB}(t_A|t \geq t') &> \phi_{AB}(t_A) \text{ iff} \\
\frac{1}{\bar{F}_A(t')}\left[\phi_{AB}(t_A) - \int_{0}^{t'} \tau f_A(\tau)\, d\tau\right] &> \phi_{AB}(t_A) \text{ iff} && (6.8)\\
\phi_{AB}(t_A) - \int_{0}^{t'} \tau f_A(\tau)\, dx &> \phi_{AB}(t_A)\bar{F}_A(t') \text{ iff} && (6.9)\\
\phi_{AB}(t_A) &> \frac{\int_{0}^{t'} \tau f_A(\tau)\, d\tau}{1 - \bar{F}_A(t')} \text{ iff} && (6.10)\\
\phi_{AB}(t_A) &> \frac{\int_{0}^{t'} \tau f_A(\tau)\, d\tau}{F_A(t')}. && (6.11)
\end{aligned}$$

where (6.8) follows by definitions, (6.9) follows by the positivity of $\bar{F}_A(t')$, (6.10) follows by a rearrangement of terms and the positivity of $1 - \bar{F}_A(t')$, and (6.11) follows by the definition of F in terms of $\bar{F}$. Therefore, if we show that (6.11) holds, then we are done. The right hand side of (6.11) is the expectation of $X \sim f_A(x)$, conditional on $X \leq t'$. This is less than the unconditional expectation of $X \sim f_A(x)$, which is the expected value of starting on

therapy A at time 0 and never switching to therapy B. This is no greater than the best one can do by having the option of switching to B at some point of time, which equals $\phi_{AB}(t_A)$. Therefore, (6.11) holds. □

6.3.2 General Case: Multiple Therapies Remaining

Having seen the difficulty in establishing sufficient conditions for determining a priori the optimal sequence of therapies, we return to the general problem described at the beginning of Section 6.3 to see how we may solve it and what other structural properties may exist. Recall the situation we described: a patient not currently on therapy has n therapies available and wishes to determine the optimal sequencing of therapies and times to switch from one therapy to the next one on the list. A naive way to do this is to evaluate all $n!$ sequences of the therapies and solve for the optimal switching times. We present a more efficient dynamic programming approach. Recall that for each therapy i, μ_i is the mean of its lifetime distribution. We let v_S^* represent the optimal value associated with initiating therapy immediately when the set S of therapies are available. Then $v_{\{i\}}^* = \mu_i$. Next, we consider all $\binom{n}{2}$ pairs of therapies and use the methods of Section 6.3.1 to determine the optimal schedule of each pair. For example, for therapies i and j, we determine v_{ij} and v_{ji}. Then $v_{\{ij\}}^* = max(v_{ij}, v_{ji})$. In the next step of the algorithm, we consider all $\binom{n}{3}$ triplets of therapies and perform three evaluations necessary to determine the best optimal scheduling of the three available therapies. For example, consider three available therapies: i, j, and k. We first evaluate the optimal time to switch from therapy i to the best schedule of the remaining therapies j and k. To do so, we consider an evaluation analogous to that given in (6.5). In other words, we determine the best time to stop taking therapy i and follow it up with the best schedule of j and k and its corresponding value $v_{\{j,k\}}^*$. Then we do the same evaluation when starting with j and following up with the best schedule of i and k, and finally we consider switching from k to the best schedule of i and j. Taking the max of these yields $v_{\{i,j,k\}}^*$. We proceed along these lines until obtaining $v_{\{1,2,\ldots,n\}}^*$ which gives the optimal expected remaining lifetime upon initiating therapy. For our final evaluation, we consider

the possibility of the patient waiting some time before initiating therapy and evaluate (6.5) using $\bar{F}_0$ (the lifetime survival function for a patient never initiating therapy) and $v^*_{\{1,2,\ldots,n\}}$. Solving this yields the value of t_0, the time to wait before initiating therapy.

We now provide some results under this extended framework. Theorem 6.3 states that expanding upon a set of available therapies cannot make a patient worse off.

Theorem 6.3. *Let S_1 and S_2 be two sets of therapies such that $S_2 \supset S_1$. Then $v^*_{S_2} \geq v^*_{S_1}$.*

Proof. Consider any sequencing of therapies which lists the therapies in $S_2 \backslash S_1$ before using the optimal sequencing of therapies that yields $v^*_{S_1}$. Then by scheduling the switching times of the therapies in $S_2 \backslash S_1$ to occur at $t = 0$, we obtain a schedule whose value is $v^*_{S_1}$. Therefore, the optimal schedule of the therapies in S_2 yields a value at least as great as $v^*_{S_1}$. □

We now consider a similar condition to the assumption that F is not NWUE made by Derman et al. For every pair of therapies i and j, we assume the following:

$$\frac{\int_t^T \bar{F}_i(\tau)\, d\tau}{\bar{F}_i(t)} < \mu_j \quad \text{for some } 0 \leq t < T, \tag{6.12}$$

which is equivalent to

$$\mu_i < \int_0^t \bar{F}_i(\tau)\, d\tau + \bar{F}_i(t)\mu_j \quad \text{for some } 0 \leq t < T. \tag{6.13}$$

In the context of our problem, (6.13) says that if the patient starts on therapy i, there is a time for which it is optimal to switch from i to j. The basis of this assumption is that most therapies lose effectiveness over time, after which a patient can benefit from using an unused therapy.

Note that although $\phi_{ij}(t)$ is maximized at a finite value of time that satisfies (6.12), it is not necessarily true that a time satisfying (6.12) also maximizes $\phi_{ij}(t)$. Indeed, by continuity, if t_i is a maximizer of $\phi_{ij}(t)$, then there exists a neighborhood $N(t_i)$ of t_i which contains some $t \in N(t_i)$ such that $\phi_{ij}(t) < \phi_{ij}(t_i)$ and such that

$$\frac{\int_t^T \bar{F}_i(\tau)\, d\tau}{\bar{F}_i(t)} < \mu_j. \tag{6.14}$$

Consequently, although we may be observing a patient on therapy i and come to a value of t such that (6.14) holds, it may be optimal to wait some more time before switching to therapy j. For example, consider two therapies both having a $Unif[0, 10]$ distribution. After two units of time, the conditional expected remaining mean of the first distribution is equal to 4, which is less than 5 (the mean of the $Unif[0, 10]$ distribution). However, it is optimal to wait until time 5 and then switch to the second therapy.

With the assumption given in (6.12), there is a strict increase in value with an increase in therapies.

Theorem 6.4. *Under the condition of (6.12), if* $S_2 \supset S_1$, *then* $v^*_{S_2} > v^*_{S_1}$.

Proof. Consider a sequencing of therapies which lists the therapies in S_1 in the order that yields $v^*_{S_1}$ followed by an optimal sequencing of the therapies in $S_2 \backslash S_1$. Without loss of generality, let i be the final therapy in the first sequence used prior to using the first therapy in the optimal sequencing of $S_2 \backslash S_1$. The assumption in (6.12) implies that there is a time for which it is optimal to switch from i to the optimal sequencing of $S_2 \backslash S_1$ (for the assumption directly implies that it is better to switch from i to any single therapy k in $S_2 \backslash S_1$ at some time, and $v^*_{S_2 \backslash S_1} \geq v^*_k$ by Theorem 6.3). Therefore, the value associated with the optimal scheduling of S_1 ($v^*_{S_1}$) increases when we allow the possibility of switching from i at some point of time to the optimal sequencing of $S_2 \backslash S_1$. As this sequencing of the therapies in S_2 was just one consideration, it must follow that $v^*_{S_2} > v^*_{S_1}$. □

The above discussion has taken the perspective of comparing a given set of therapies with a set that includes those same therapies with the addition of more options. Now we take a slightly different perspective. We let S^i be the set of all subsets of S with cardinality equal to i, and we define $y(i)$ to be the maximum of all the values associated with the optimal scheduling of each of the $\binom{n}{i}$ elements in S^i. Formally, we have the following definitions:

$$\begin{aligned} S^i &\equiv \{S' : S' \subseteq S \text{ and } |S'| = i\}, \\ y(i) &\equiv \max_{S' \in S^i} \{v^*_{S'}\}, \\ S^*(i) &\in \operatorname{argmax}_{S' \in S^i} \{v^*_{S'}\}. \end{aligned}$$

We now prove the following:

Theorem 6.5. *If $i < j$, then $y(i) < y(j)$.*

Proof. Let $i < j$. Consider the set $S^*(i)$ and any superset S' of $S^*(i)$ with cardinality j. By Theorem 6.4 we know that $v^*_{S'} > v^*_{S^*(i)} = y(i)$. Because S' is in S^j, it follows that $y(j) \geq v^*_{S'} > v^*_{S^*(i)} = y(i)$. □

One of the interesting results in Derman et al. was that if it is optimal to wait t_x units of time to switch from the current component to a spare component, then the optimal time to replace the new spare component with another spare component is some time $t_y > t_x$ [58]. In other words, fewer spares means longer times to use the currently functioning component. We explore this idea in the current framework. Consider four therapies, A, B, C and D, with lifetime distributions given as $L_A \sim Triang[0, 5, 10]$, $L_B \sim Triang[0, 6, 12]$, $L_C \sim Triang[0, 2, 17]$, and $L_D \sim Triang[0, 2, 18]$. Table 13 contains the different contenders for the optimal sequencing of one, two, three, and four therapies from the possible set of four. It also contains the approximate value associated with the sequence and the approximate time to switch from the first therapy in the sequence to the second one. For example, if the patient can only use one of the therapies, then it is best to use therapy D and expect to live 6.67 more units of time. If the patient can only use two of the four therapies, it is best to use B followed by D, with an expected lifetime of 8.89. Also, the table indicates it is best to switch from B to D after 4.12 units of time. If three therapies may be used, it is best to take B, followed by A, and finish with D. The expected lifetime under this scenario is 10.18 and the patient should stay on therapy B for 3.57 units of time before switching to therapy A. Finally, with all four therapies the optimal sequencing is $BACD$. The expected remaining lifetime under this sequence is 11.03, and the patient should switch from B to A after 3.29 units of time. After switching to A, we can use the table to see that the best time to switch from A to C is after 2.71 units of time, and the best time to switch from C to D is after 3.67 units of time. We see, therefore, that within an optimal schedule of M out of N available therapies, the optimal times to use each subsequent therapy in the list may not be monotonically increasing. One can check that this is also demonstrated when considering the optimal schedule of three out of the four therapies.

6.3.3 Dependence Considerations

Because of the independence assumption, until now we have not needed to concern ourselves with the history of therapies taken when considering the lifetime distribution of a newly administered therapy. In thinking of a general framework for considering dependencies, it is conceivable that we may want to know the following information: the sequence of drugs taken so far, how long ago each one was started, and how long the patient was on each therapy. The reason for wanting such information is that the time spent on a therapy gives information about the *likelihood* of development of resistance against that therapy and others, and the time that has elapsed since that therapy was taken gives information about the current *strength* of resistance that may have developed. Now, when considering a lifetime distribution of a therapy j, instead of dealing with $F_j(t)$ as before, we should think of the form as: $F_j(t|\{a_1, t_1, a_2, t_2..., a_{j-1}, t_{j-1}\})$, where the patient was on therapy a_1 for time t_1, then switched to therapy a_2 and stayed on that for time t_2, until being on therapy a_{j-1} for time t_{j-1} before starting the current therapy.

Solving this kind of problem is extremely difficult. First, the data requirements to support such considerations greatly exceed the data availability in the near term. Second, even with the conditional distribution information immediately available, there is a large computational burden of solving such a problem (this problem may be thought of as a generalization of stochastic machine scheduling [136]). However, there may be considerable simplifications we can make to the general dependency framework we just described. For example, perhaps only the history of therapies until t_x time units ago or the last k therapies need to be kept. We leave for future work these and other simplifications that may render the dependency considerations both clinically plausible and computationally tractable.

6.4 CONCLUSIONS

This chapter has considered switching and sequencing available therapies so as to maximize a patient's expected total lifetime. Our research was motivated by a paper in the machine

Table 13: Sequencing data for four therapies with $L_A \sim Triang[0, 5, 10]$, $L_B \sim Triang[0, 6, 12]$, $L_C \sim Triang[0, 2, 17]$, and $L_D \sim Triang[0, 2, 18]$

# of Therapies Available	Sequence	Optimal Value	First Therapy Usage Time
1	A	5	-
	B	6	-
	C	6.333	-
	D	**6.667**	-
2	AB	7.754	3.274
	BA	7.688	4.849
	AC	8.018	3.159
	CA	6.987	7
	AD	8.287	3.052
	DA	7.245	8
	BC	8.639	4.255
	CB	7.463	5
	BD	**8.891**	**4.124**
	DB	7 .667	6
	CD	7.883	3.667
	DC	7.843	5.333
3	ABC	9.955	2.525
	ABD	10.176	2.469
	ACD	9.302	2.707
	BAC	9.959	3.656
	BAD	**10.180**	**3.574**
	BCD	9.849	3.699
	CAB	8.764	1.948
	CAD	9.241	1.846
	CBD	9.788	1.741
	DAB	8.825	2.493
	DAC	9.053	1.996
	DBC	9.609	1.879
4	$ABCD$	11.026	2.275
	BACD	**11.029**	**3.289**
	$CBAD$	10.974	1.552
	$DBAC$	10.815	1.668

maintenance literature that sought to maximize the expected lifetime of a system with one vital component and n spare parts [58]. In that paper, it was assumed that the lifetime distributions of the parts were independent and identically distributed. We spent most of the chapter exploring the implications of relaxing the assumption of replacements having identical distributions, because different therapies will typically induce different lifetime distributions. In particular, we considered conditions on the distributions that would ensure that we could know the optimal sequencing of therapies in advance of determining the optimal timing of the switches. However, even very strong stochastic ordering conditions fail to give sufficient conditions for a priori optimal sequencing. We considered the more general situation where multiple therapies are available, proved some structural properties regarding the optimal value function, and provided a dynamic programming algorithm based on earlier ideas.

There are some limitations of the model we presented. One is that we did not consider patient aging. We assumed that the lifetime distributions associated with being on a therapy remain the same regardless of when the therapy is taken; in reality, a 60 year old patient initiating a therapy will have a smaller expected remaining lifetime compared to a 20 year old patient initiating the same therapy. One way to handle this may be to include a discount factor in the rewards. We discuss this more in the next chapter. Also, we did not consider the goal of maximizing the total expected *quality-adjusted* life years of a patient. We could consider utility weights associated with being on each therapy without adding any difficulty to the solution methodology. However, implementing utility weights is better suited for a model that considers more patient states besides "alive" and "dead". For example, a model that considers varying degrees of health could better discriminate quality of life valuations. We consider this as well in the next chapter.

7.0 MDP APPROACH TO THE OPTIMAL SEQUENCING AND SWITCHING OF THERAPIES

7.1 INTRODUCTION

Chapter 6 described a methodological approach for the optimal scheduling of HIV therapies when the only information available is the lifetime distribution associated with being on each therapy. The only observations considered were whether the patient was alive or dead. We described a schedule of therapies as a sequence of the therapies $\{a_1, a_2, \ldots, a_n\}$ and an associated sequence of durations $\{t_0, t_1, t_2, \ldots, t_{n-1}\}$, where the patient starts taking therapy a_1 after t_0 units of time pass, remains on therapy a_1 for t_1 units of time, and so on, until the patient is on the final therapy, a_n, and remains on that therapy until death. Furthermore, we showed in Theorem 6.1 that knowing a patient has survived until some time t does nothing to change the optimal time to switch to the next therapy.

This chapter considers a framework in which physicians make periodic patient observations, which allow them to update their prior beliefs of the patient's survival and prognosis and thus make more informed treatment decisions. Rather than an a priori sequence of therapies and usage times, this state-based approach indicates the optimal action to take at each patient visit as a function of the patient's state. This leads to a dynamic solution in which certain states may indicate to continue taking the current therapy, others may indicate to switch to therapy A, others to switch to therapy B, and so on.

As discussed in Chapter 1, MDPs are useful modeling techniques for solving sequential stochastic decision problems. This is the situation with the management of therapies just as it is for the question of when to initiate therapy. Whereas a stationary MDP framework (one in which the components do not change over time) was justifiable for the latter (see

comments in Section 4.6), there are some issues we need to address in the current context to consider a feasible stationary MDP model. Stationary MDPs have significant conceptual, computational, and data-requirement advantages over non-stationary MDPs.

As with the previous chapter, there are connections between this problem and machine maintenance problems, where a machine is periodically inspected and based on its state, a decision is made whether to continue using it or replace it with a new one. Many machine maintenance models assume that a replacement takes the system back to a new state from which the same stochastic deterioration as the previous machine proceeds. There are some exceptions, such as Klein [100], which considers maintenance actions that may take the system to one of various states with certainty, or Hopp and Wu [83], which considers probabilistic movement to various states after a maintenance action. A review by Pham and Wang [130] provides an overview of other machine maintenance literature that address the possibility of imperfect repair.

The previous research, however, has not addressed a situation analogous to that found in HIV therapy planning: after the administration of a new therapy, patients typically see a general stochastic improvement in their health for some time, followed by a general stochastic deterioration after the virus builds resistance to the therapy (if we think about this in terms of survival analysis [99, 104], the hazard rate function for a patient's lifetime on therapy should exhibit a sort of "bathtub" shape). An analogous machine maintenance problem would be if a machine had an initial warmup period prior to peak performance, followed by a general stochastic deterioration. Recall from Section 2.3.2, that Klein used the exponential lifetime distributions observed in machines as justification for considering an MDP approach to optimal machine maintenance problems [100]. However, an exponential distribution implies a constant hazard function. Note that in the models of Chapters 4 and 5, patients are in a general stochastically deteriorating state of health until they begin therapy, at which time they receive the expected remaining lifetime reward associated with being on therapy. Therefore, it was not necessary to explicitly consider the two-phase aspect of the transitions as we do now. An even more sophisticated approach would involve breaking the progression into multiple phases according to degrees of resistance accumulation. For this chapter, we shall focus on a two-phase perspective.

Another obstacle to modeling the problem as a stationary MDP is the notion of aging. Patients, as well as machines, age over time, which should have a dynamic effect on the probability of death beyond the state-based probability of death. With the optimal stopping structure of Chapters 4 and 5, the time horizon before initiating therapy was not significantly long, and we captured aging effects through our estimates of expected lifetime after initiating therapy. In the present case, in which we explicitly model the switching of therapies over time, we need to reconsider how to capture the aging effect.

Kao recommends a semi-Markov process model as a way to overcome the no-aging assumption implied by standard Markov models for machine maintenance [92]. As with most Markov models of machine maintenance, his model considers a system that may occupy a number of states between new and failed; however, the difference is that he allows the system to occupy these states for a random amount of time instead of for a fixed interval. Costs are based on the time spent in a state as well as the cost of replacing a machine from a state, and the objective is to minimize the long-run average cost per unit time. He considers three types of policies: state-dependent, state-age-dependent, and age-dependent. As noted in [19], however, Kao's model only considers aging between state transitions and does not consider a cumulative aging of the system that has a deleterious effect.

Benyamini and Yechiali explicitly address the aging problem in the machine maintenance context by including both state and cumulative age in the MDP framework (we introduced their work in Section 5.3) [19]. However, they make no mention of perhaps the biggest obstacle to implementing their model: constructing data-based estimates of the reward and transition probabilities for every state and age combination. Including the system age in the model may render data too sparse to derive reliable age-based estimates of these components. The large state space may present computational difficulties as well.

Yet another issue arises that we did not face in the MDP models of earlier chapters: the fact that patients have a finite number of therapies (or analogously, that systems have a finite number of replacements). The previous chapter discussed the paper by Derman et al. [58], which is the only machine maintenance paper we are aware of that explicitly considers

a finite supply of replacements. Again, Chapters 4 and 5 did not consider this issue because we did not model the physiological progression of HIV after a patient initiated therapy; the patients received a terminal reward and the problem stopped.

7.2 MODELING FRAMEWORK

We address the above issues in a manner that allows us to utilize a stationary MDP model. With respect to aging the patient, we will approximate the aging effect with a discount factor instead of modeling age explicitly. For example, suppose we want to maximize the expected remaining lifetime of a patient who enters the model at age 20. First, we shall compute a monthly discount factor for an "average" person who is 20 years old as follows: from a standard life table, suppose we find that the expected remaining lifetime for a 20 year old is 60 years. Then, we compute the monthly discount factor as $\alpha = 1 - 1/(60 \cdot 12)$. In other words, α is the discount factor such that the net present value of an infinite stream of discounted one-month rewards equals the expected remaining months of life ($60 \cdot 12 = 720$). Therefore, alpha solves $\sum_{i=0}^{\infty} \alpha^i = 720$. We let this represent an HIV-independent increased mortality that the general population faces as they age.

As for the two-phase nature of the disease progression after switching to a new therapy, we do the following. In addition to the usual state variable indicating the patient's current health, we include a state variable that indicates if the patient's health is stochastically improving, or stochastically deteriorating. The former will hold if the patient is on an effective therapy, and the latter will hold if the patient is on an ineffective therapy. For now, we assume therapies are identical and consider therapies with different levels of effectiveness in Section 7.4. Upon switching to a new therapy, we assume it is effective and the patient progresses according to a transition probability matrix representing stochastic improvement. However, at each period, there is some state-dependent probability that the patient's therapy becomes ineffective (e.g., due to resistance buildup) and moves into a state of stochastically deteriorating health. We assume that the only way to get back to a state of stochastically improving health is to initiate a new therapy.

Finally, with respect to the issue of a finite number of therapies available to a patient, we further expand the state space to include a parameter indicating the number of effective therapies remaining. Upon switching to a new therapy, this parameter decreases by one. When this value reaches zero, i.e., there are no more therapies from which to choose, we assume the patient progresses according to that last therapy's two-phase transition probabilities until death.

We summarize the MDP model as follows. We consider monthly decision epochs, where a patient may continue to wait ($a = W$) or decide to switch to a new therapy ($a = S$) (if there are any remaining). We describe the state by (h, d, n), where h represents the health category (between 0 and H, with 0 representing death), d represents the direction status (1 or 0 according to whether the patient is in a state of stochastically improving or deteriorating health, respectively) and n represents the number of therapies remaining (between 0 and N). For now we assume therapies have independent and identical effects on HIV progression. In Section 7.4, we relax this assumption. We will adopt the goal of maximizing total expected lifetime and assign the reward for each action and state (with $h > 0$) to equal 1 month. The resulting Bellman's equations are:

$$v(h,d,n) = \max \begin{cases} 1 + \sum_{d'=0}^{d'=1} \sum_{h'=0}^{h'=H} p[(h',d',n)|(h,d,n),W]v(h',d',n), \\ 1 + \sum_{d'=0}^{d'=1} \sum_{h'=0}^{h'=H} p[(h',d',n-1)|(h,d,n),S]v(h',d',n-1), \end{cases} \tag{7.1}$$

where $p(\cdot|\cdot,a)$ represents the transition probabilities according to whether the patient waits (top line) or switches to a new therapy (bottom line).

7.3 STRUCTURAL PROPERTIES

We begin with some assumptions. Figure 14 represents some of these assumptions through a partial view of the transition probability matrix for a model with $H = 5$. It shows the parts of the matrix relevant to the case of one and zero therapies remaining. P_{II} represents the submatrix of probabilities that patients waiting in a state of improving health transition to

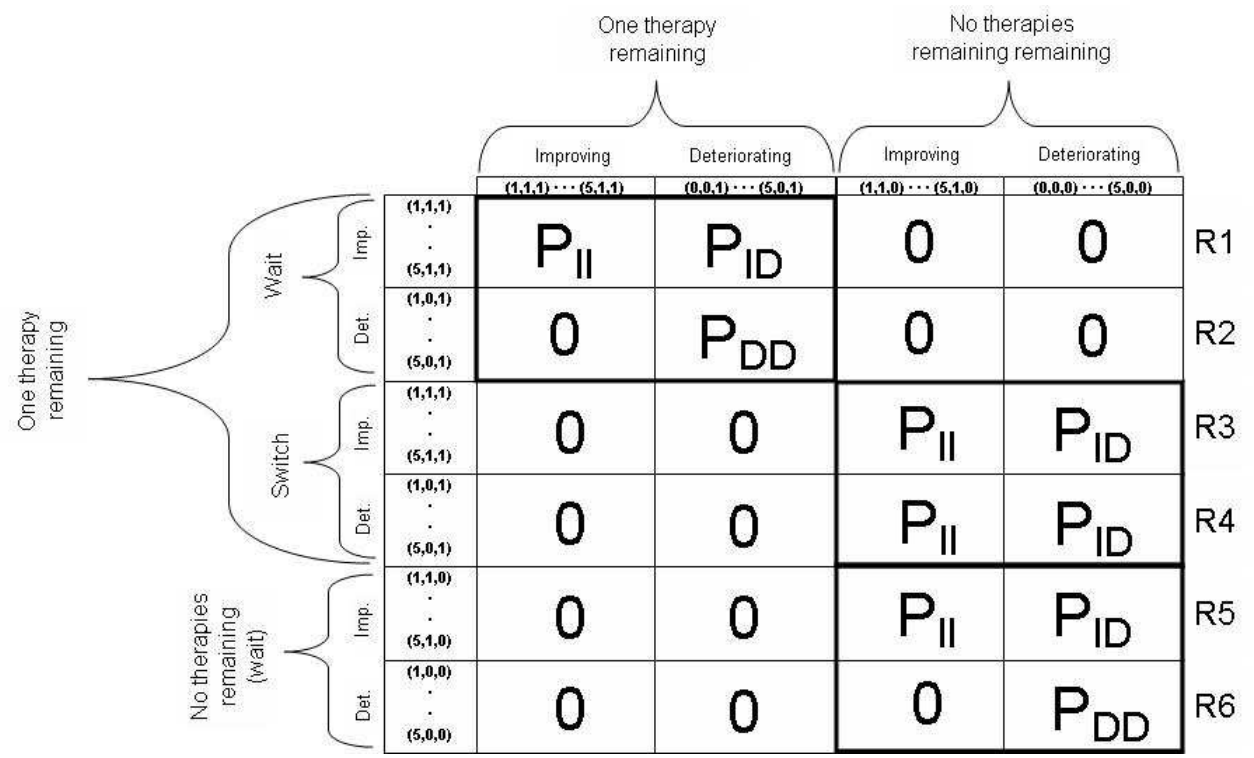

Figure 14: Partial view of the transition probability matrix

the various health states and remain in a state of improving health. For example, P_{II} represents probabilities such as $p[(h', 1, n)|(h, 1, n), W]$. Similarly, P_{ID} represents the submatrix of probabilities that patients waiting in a state of improving health transition to the various health states and move to a state of deteriorating health. For example, P_{ID} represents probabilities such as $p[(h', 0, n)|(h, 1, n), W]$. P_{DD} represents the submatrix of probabilities that patients waiting in a state of deteriorating health transition to the various health states and remain in a state of deteriorating health. For example, P_{DD} represents probabilities such as $p[(h', 0, n)|(h, 0, n), W]$.

($As7.1$) $p[(h', 1, n)|(h, 0, n), W] = 0$ for all h, h', n.

In words, $As7.1$ says that a patient in a state of deteriorating health cannot move to a state of improving health while waiting.

($As7.2$) $p[(h', d', i)|(h, d, i), W] = p[(h', d', j)|(h, d, j), W]$ and $p[(h', d', i-1)|(h, d, i), S] = p[(h', d', j-1)|(h, d, j), S]$ for all i, j, h, d, h', d'.

$As7.2$ says that regardless of the number of therapies remaining, each action induces the same transition probabilities. Essentially, this is the assumption that the effects of therapies are iid. This assumption is represented in the figure by rows $R1$ and $R2$ matching rows $R5$ and $R6$, with respect to feasible moves. For example, a patient can only move to a state of fewer therapies remaining by switching to a new therapy; the action of waiting does not consider the possibility of a loss of available therapies, hence the submatrix of zeroes to the right of $\begin{pmatrix} P_{II} & P_{ID} \\ O & P_{DD} \end{pmatrix}$ in rows $R1$ and $R2$. Similarly, waiting cannot lead to an increase in available therapies, leading to the submatrix of zeroes to the left of $\begin{pmatrix} P_{II} & P_{ID} \\ O & P_{DD} \end{pmatrix}$ in rows $R5$ and $R6$. The second statement in $As7.2$ would be seen by rows $R3$ and $R4$ matching other rows representing switching to a new therapy when more than one therapy remains (not shown in the figure).

(As7.3) $p[(h', d', n-1)|(h, 0, n), S] = p[(h', d', n-1)|(h, 1, n), S] = p[(h', d', n-1)|(h, 1, n-1), W]$ for all h, n, h', d'.

The first equality of $As7.3$ says that switching to a new therapy has the same effect regardless if it is initiated from a state of improving or deteriorating health. This is represented in the figure by rows $R3$ and $R4$ matching each other. The second equality says that the effect of switching to a new therapy when n therapies remain is the same as for a patient waiting in a state of improving health when $n-1$ therapies remain. This is represented by rows $R3$ and $R4$ matching row $R5$ (again, with respect to feasible moves).

Before proceeding with further assumptions, let $p_1(i)$ represent the probability that a patient waiting in health category i with improving health remains in a state of improving health, i.e., $p_1(i) = \sum_{j=0}^{j=H} p[(j, 1, n)|(i, 1, n), W]$. Similarly, let $p_0(i)$ represent the probability that a patient in health category i with improving health transitions to a state of deteriorating health. We shall assume that the probability of dying from a state (h, d, n) is represented by the transition to $(0, 0, n)$ (if waiting) or to $(0, 0, n-1)$ (if switching to

a new therapy). In other words, death occurs in the deteriorating health region only. Then $p_0(i) = 1 - p_1(i)$. We let $\bar{P}_{II}$ represent the transition probabilities among health categories, conditional on a patient in a state of improving health remaining in a state of improving health, and we let $\bar{P}_{ID}$ represent the transition probabilities among health categories, conditional on a patient in a state of improving health moving to a state of deteriorating health. In other words, $\bar{P}_{II}$ is composed of individual probabilities such as $q[(j,1,n)|(i,1,n),W] \equiv p[(j,1,n)|(i,1,n),W]/p_1(i)$, and $\bar{P}_{ID}$ is composed of individual probabilities such as $q[(j,0,n)|(i,1,n),W] \equiv p[(j,0,n)|(i,1,n),W]/p_0(i)$. Continuing with our assumptions, we have:

($As7.4$) P_{DD} is IFR.

($As7.5$) $\bar{P}_{ID} = P_{DD}$.

($As7.6$) $\bar{P}_{II}$ is IFR.

($As7.7$) $\bar{P}_{II} \succeq_{st} P_{DD}$.

($As7.8$) $p_1(i)$ is nondecreasing in i.

$As7.4$ implies that patients waiting in a state of deteriorating health are still better off being in a higher health category than a lower one. $As7.5$ says that given a patient in a state of improving health transitions to a state of deteriorating health, the probabilistic transition among health categories is identical to a patient who waits from a state of deteriorating health. $As7.6$ can be interpreted similarly to $As7.4$, for patients remaining in a state of improving health. $As7.7$ says that for patients remaining in a state of improving health, the transitions among health categories is stochastically larger than the transitions among the health categories for a patient waiting in a state of deteriorating health. The stochastic ordering in this assumption is what conveys our notion of stochastically improving versus deteriorating health. Therefore, our use of the phrase "stochastically improving" need not imply a notion of drift towards higher health categories. Finally, $As7.8$ says that improving patients waiting in higher health categories have a greater chance of remaining in a state of improving health compared to patients waiting in lower health categories. $As7.8$ is equivalent to $p_0(i)$ is nonincreasing in i.

Table 14: Assumptions required by each result

Result	Assumptions used
Theorem 7.1	As7.1 - As7.8
Theorem 7.2	As7.2, As7.3
Theorem 7.3	As7.2, As7.3
Theorem 7.4	As7.1 - As7.7, As7.9 - As7.11
Theorem 7.5	As7.1, As7.3 - As7.8
Lemma 7.1	As7.2 - As7.7, As7.9
Lemma 7.2	As7.1 - As7.7, As7.9 - As7.11

For the remainder of this section we assume we are modeling a system for which *As*7.1-*As*7.8 all hold. For this reason, and for the sake of clarity, we assume they hold for each result. The interested reader can refer to Table 14 to see the minimal set of assumptions required by each result.

We begin with the following intuitive result.

Theorem 7.1. *$v^*(h, d, n)$ is nondecreasing in h, nondecreasing in d, and nondecreasing in n.*

Proof. For some i, let $v^i(h, d, n)$ be nondecreasing in h, nondecreasing in d, and nondecreasing in n. Note that $v^0(h, d, n) = 0$ for all h, d, n satisfies this condition. We fix n' arbitrarily and consider how $v^{i+1}(h, 1, n')$ varies with h (note that we have fixed $d = 1$; we will discuss the case of $d = 0$ below). Let $h_1 < h_2$. For ease of notation and space considerations, we define

the following functions:

$$f(h,d,n,W) \equiv \sum_{j=0}^{M} q[(j,1,n)|(h,d,n),W]v^i(j,1,n), \tag{7.2}$$

$$f(h,d,n,S) \equiv \sum_{j=0}^{M} q[(j,1,n-1)|(h,d,n),S]v^i(j,1,n-1), \tag{7.3}$$

$$g(h,d,n,W) \equiv \sum_{j=0}^{M} q[(j,0,n)|(h,d,n),W]v^i(j,0,n), \tag{7.4}$$

$$g(h,d,n,S) \equiv \sum_{j=0}^{M} q[(j,0,n-1)|(h,d,n),S]v^i(j,0,n-1). \tag{7.5}$$

$f(h,d,n,W)$ represents the expected "cost-to-go" from the next time period on, conditional on the patient moving to a state of stochastically *improving* health, when the patient is currently in state (h,d,n) and takes the action "wait". $g(h,d,n,W)$ is interpreted similarly, but conditional on the patient moving to a state of stochastically *deteriorating* health. $f(h,d,n,S)$ and $g(h,d,n,S)$ are also interpreted similarly, but with the action being "switch".

Then $v^{i+1}(h_2,1,n')$ is the maximum of the following two expressions:

$$1 + p_1(h_2)f(h_2,1,n',W) + p_0(h_2)g(h_2,1,n',W), \tag{7.6}$$

and

$$1 + p_1(h_2)f(h_2,1,n',S) + p_0(h_2)g(h_2,1,n',S), \tag{7.7}$$

where expression (7.6) represents the value update associated with waiting on the current therapy, and (7.7) represents the value update associated with initiating a new therapy (without loss of generality, we assume $n' \geq 1$). Note that (7.7) makes use of assumption $As7.3$ (that patients initiating a new therapy observe the same stochastic transitions over health categories as patients waiting on a therapy and in a state of improving health).

In similar fashion, we can write $v^{i+1}(h_1,1,n')$ as the maximum of the following two expressions:

$$1 + p_1(h_1)f(h_1,1,n',W) + p_0(h_1)g(h_1,1,n',W), \tag{7.8}$$

and

$$1 + p_1(h_1)f(h_1, 1, n', S) + p_0(h_1)g(h_1, 1, n', S). \tag{7.9}$$

We compare expressions (7.6) and (7.8). By *As*7.6 ($\bar{P}_{II}$ is IFR) and the assumption that $v^i(h, d, n)$ is nondecreasing in h,

$$f(h_2, 1, n', W) \geq f(h_1, 1, n, W). \tag{7.10}$$

By *As*7.4, *As*7.5 (which establish that $\bar{P}_{ID}$ is IFR), and the assumption that $v^i(h, d, n)$ is nondecreasing in h,

$$g(h_2, 1, n', W) \geq g(h_1, 1, n, W). \tag{7.11}$$

Also, by *As*7.7 and *As*7.5 (showing that $\bar{P}_{II} \succeq_{st} \bar{P}_{ID}$) and the assumption that $v^i(h, d, n)$ is nondecreasing in h and d,

$$f(h_2, 1, n', W) \geq g(h_2, 1, n', W), \tag{7.12}$$

$$f(h_1, 1, n, W) \geq g(h_1, 1, n, W). \tag{7.13}$$

Finally, combining (7.10)-(7.13) with *As*7.8 establishes that:

$$\begin{aligned} & 1 + p_1(h_2)f(h_2, 1, n', W) + p_0(h_2)g(h_2, 1, n', W) \\ \geq\ & 1 + p_1(h_1)f(h_1, 1, n', W) + p_0(h_1)g(h_1, 1, n', W). \end{aligned} \tag{7.14}$$

A similar exercise shows

$$\begin{aligned} & 1 + p_1(h_2)f(h_2, 1, n', S) + p_0(h_2)g(h_2, 1, n', S) \\ \geq\ & 1 + p_1(h_1)f(h_1, 1, n', S) + p_0(h_1)g(h_1, 1, n', S). \end{aligned} \tag{7.15}$$

Relating these results back to the expressions given in (7.6)- (7.9), shows that $v^{i+1}(h_2, 1, n') \geq v^{i+1}(h_1, 1, n')$. Therefore, $v^{i+1}(h, 1, n')$ is nondecreasing in h. We also need to show that $v^{i+1}(h, 0, n')$ is also nondecreasing in h. If a patient in a state of deteriorating health decides to wait, then it is better to do so from a health category $h_2 > h_1$ by *As*7.1, *As*7.4, and the inductive assumption that $v^i(h, 0, n')$ is nondecreasing in h. If the patient chooses to switch

to a new therapy, then it is still better to be in health category h_2 by the same reasoning that established (7.15). It follows, then, that $v^{i+1}(h, 0, n')$ is nondecreasing in h.

We have shown that $v^{i+1}(h, d, n)$ is nondecreasing in h. Next we show that it is also nondecreasing in d. In other words, we show that for fixed h and n, $v^{i+1}(h, 1, n) \geq v^{i+1}(h, 0, n)$. We arbitrarily fix h' and n' (again, without loss of generality we assume $n' \geq 1$). $v^{i+1}(h', 1, n')$ is the maximum of the following two expressions:

$$1 + p_1(h')f(h', 1, n', W) + p_0(h')g(h', 1, n', W), \tag{7.16}$$

and

$$1 + p_1(h')f(h', 1, n', S) + p_0(h')g(h', 1, n', S). \tag{7.17}$$

Also, $v^{i+1}(h', 0, n')$ is the maximum of these two expressions:

$$1 + \sum_{j=0}^{M} p[(j, 0, n')|(h', 0, n'), W]v^i(j, 0, n'), \tag{7.18}$$

and

$$1 + p_1(h')f(h', 0, n', S) + p_0(h')g(h', 0, n', S). \tag{7.19}$$

Note that in (7.18), there is no consideration of movement to a state of improving health (as assumed in *As*7.1). By *As*7.3 (patients switching to a new therapy have the same stochastic transitions regardless of whether they switch from a state of improving or deteriorating health),

$$\begin{aligned} & 1 + p_1(h')f(h', 1, n', S) + p_0(h')g(h', 1, n', S) \\ = \; & 1 + p_1(h')f(h', 0, n', S) + p_0(h')g(h', 0, n', S). \end{aligned} \tag{7.20}$$

By *As*7.7 (that $\bar{P}_{II} \succeq_{st} P_{DD}$) and the assumption that $v^i(h, d, n)$ is nondecreasing in both h and d,

$$f(h', 1, n', W) \geq \sum_{j=0}^{M} p[(j, 0, n')|(h', 0, n'), W]v^i(j, 0, n').$$

By $As7.5$ (that $\bar{P}_{ID} = P_{DD}$),

$$g(h', 1, n', W) = \sum_{j=0}^{M} p[(j, 0, n')|(h', 0, n'), W]v^i(j, 0, n').$$

Therefore, it follows that

$$\begin{aligned} & 1 + p_1(h')f(h', 1, n', W) + p_0(h')g(h', 1, n', W) \\ \geq \; & 1 + \sum_{j=0}^{M} p[(j, 0, n')|(h', 0, n'), W]v^i(j, 0, n'). \end{aligned} \tag{7.21}$$

By the expressions in (7.16) - (7.19) and the relationships expressed in (7.20) and (7.21), it follows that $v^{i+1}(h, d, n)$ is nondecreasing in d.

It remains to be shown that $v^{i+1}(h, d, n)$ is nondecreasing in n. We begin by fixing h', setting $d' = 1$ (we consider $d' = 0$ later) and consider $1 \leq n_1 < n_2 \leq n$ (we compare the case of one and zero therapies remaining later). $v^{i+1}(h', 1, n_2)$ is the maximum of the following two expressions

$$1 + p_1(h')f(h', 1, n_2, W) + p_0(h')g(h', 1, n_2, W), \tag{7.22}$$

and

$$1 + p_1(h')f(h', 1, n_2, S) + p_0(h')g(h', 1, n_2, S). \tag{7.23}$$

Similarly, $v^{i+1}(h', 1, n_1)$ is the maximum of the following two expressions:

$$1 + p_1(h')f(h', 1, n_1, W) + p_0(h')g(h', 1, n_1, W), \tag{7.24}$$

and

$$1 + p_1(h')f(h', 1, n_1, S) + p_0(h')g(h', 1, n_1, S). \tag{7.25}$$

By $As7.2$ (transitions probabilities among health categories while waiting are iid and similarly for transition probabilities when switching to a new therapy), and the assumption that $v^i(h, d, n)$ is nondecreasing in n, we get that

$$\begin{aligned} & 1 + p_1(h')f(h', 1, n_2, W) + p_0(h')g(h', 1, n_2, W) \\ \geq \; & 1 + p_1(h')f(h', 1, n_1, W) + p_0(h')g(h', 1, n_1, W). \end{aligned} \tag{7.26}$$

For the same reasons, we also get that

$$\begin{aligned} & 1 + p_1(h')f(h', 1, n_2, S) + p_0(h')g(h', 1, n_2, S) \\ \geq \; & 1 + p_1(h')f(h', 1, n_1, S) + p_0(h')g(h', 1, n_1, S). \end{aligned}$$

By these results and expressions (7.22)-(7.25), it follows that $v^{i+1}(h, 1, n)$ is nondecreasing in n for $n \geq 1$. Returning to the comparison of one and zero therapies remaining, for the reasons we established (7.26), we know that the value update associated with waiting in a state with more therapies remaining is no less than the value of waiting in a state with zero therapies remaining. Clearly, having a choice to wait or switch to a new therapy can be no worse than choosing to wait, and therefore, $v^{i+1}(h, 1, 1) \geq v^{i+1}(h, 1, 0)$. We have now proven that $v^{i+1}(h, 1, n)$ is nondecreasing in n for all n. Similar exercises show that $v^{i+1}(h, 0, n)$ is also nondecreasing in n for all n.

In summary, we have shown that $v^{i+1}(h, d, n)$ is nondecreasing in h, nondecreasing in d, and nondecreasing in n. Taking the limit as the value iterates approach infinity, we obtain the desired result that $v^*(h, d, n)$ is nondecreasing in h, nondecreasing in d, and nondecreasing in n. □

The following shows that it is optimal to wait when in a state of improving health.

Theorem 7.2. *For all* $(h, 1, n)$, $a^*(h, 1, n) = W$.

Proof. We assume $n \geq 1$ since there is no option other than to wait when $n = 0$. Then $v^*(h, 1, n)$ is the maximum of the following two expressions:

$$1 + p_1(h)f(h, 1, n, W) + p_0(h)g(h, 1, n, W), \tag{7.27}$$

and

$$1 + p_1(h)f(h, 1, n, S) + p_0(h)g(h, 1, n, S). \tag{7.28}$$

By *As*7.3, the probability components in (7.27) and (7.28) (including the conditional probabilities in the f and g functions) are identical. By the previous result that $v^*(j, 1, n) \geq v^*(j, 1, n-1)$ and $v^*(j, 0, n) \geq v^*(j, 0, n-1)$, the result follows. □

In the following, we show that the value of being in a state of deteriorating health with n therapies remaining is at least as good as being in a state of improving health with $n-1$ therapies remaining.

Theorem 7.3. *For all* $(h,0,n)$, $v^*(h,0,n) \geq v^*(h,1,n-1)$

Proof. Let $n \geq 1$. We have:

$$\begin{aligned} & v^*(h,0,n) \\ \geq\ & 1 + p_1(h)f(h,0,n,S) + p_0(h)g(h,0,n,S) && (7.29) \\ =\ & 1 + p_1(h)f(h,0,n-1,W) + p_0(h)g(h,0,n-1,W) && (7.30) \\ =\ & v^*(h,1,n-1), && (7.31) \end{aligned}$$

where (7.29) follows by the definition of $v^*(h,0,n)$, (7.30) follows from $As7.3$, and (7.31) follows by the result of Theorem 7.2. □

We now develop conditions to guarantee that when there is at least one therapy remaining and the patient is in a state of deteriorating health, there is at least one state for which it is optimal to switch to one of the remaining therapies. In other words, there is a benefit to having the therapies that remain available. We begin with some supporting assumptions and results. $As7.1$-$As7.8$ continue to hold and each additional assumption that follows will hold for the remainder of the section.

The following assumption says that the probability of dying while waiting in a state of improving health is strictly decreasing in the health state.

(As7.9) $p[(0,0,n)|(h,d,n),W]$ is strictly decreasing in h for each d.

Now we can prove that for a patient in a state of stochastically improving health, the optimal value function is strictly increasing in the health category.

Lemma 7.1. $v^*(h,1,n)$ *is strictly increasing in* h.

Proof. Note that by Theorem 7.2, for each h, $v^*(h,1,n)$ is obtained by the action of waiting. Let us consider $v^*(1,1,n)$. Because the immediate reward equals 1 in each state for which $h > 0$, it follows that $v^*(1,1,n) > 0 = v^*(0,1,n)$ (note that by our construction, there really is no chance of a patient even getting to a state $(0,1,n)$ since death is represented only by

state $(0,0,n)$ for each level of n. Therefore we just define the value of that state to be 0.) Now let $1 \leq h' < H$, arbitrarily, and let us compare $v^*(h',1,n)$ with $v^*(h'+1,1,n)$. Referring back to (7.2)-(7.5) (with v^* in place of v^i), we have:

$$v^*(h'+1,1,n) = 1 + p_1(h'+1)f(h'+1,1,n,W) + p_0(h'+1)g(h'+1,1,n,W),$$

and

$$v^*(h',1,n) = 1 + p_1(h')f(h',1,n,W) + p_0(h')g(h',1,n,W).$$

Let us focus on $g(h'+1,1,n,W)$ and $g(h',1,n,W)$. Referring back to lines (4.6) and (4.7) of Lemma 4.1 from Chapter 4 and noting that $v^*(0,0,n) = 0$, we have

$$g(h'+1,1,n,W) = \sum_{i=1}^{H}[v^*(i,0,n) - v^*(i-1,0,n)] \sum_{j=i}^{H} q[(j,0,n)|(h'+1,1,n),W], \quad (7.32)$$

and

$$g(h',1,n,W) = \sum_{i=1}^{H}[v^*(i,0,n) - v^*(i-1,0,n)] \sum_{j=i}^{H} q[(j,0,n)|(h',1,n),W]. \quad (7.33)$$

By the first part of the proof, we know that $v^*(1,0,n) - v^*(0,0,n) > 0$. Also, by $As7.4$, $As7.5$, and $As7.9$, we know that $\sum_{j=1}^{H} q[(j,0,n)|(h'+1,1,n),W] > \sum_{j=1}^{H} q[(j,0,n)|(h',1,n),W]$. Therefore, $g(h'+1,d,n,W) > g(h',d,n,W)$, which along with the steps leading up to (7.14) as well as $As7.9$ (which establishes that $p_1(h) < 1$ for all $1 \leq h < H$), implies:

$$\begin{aligned} & 1 + p_1(h'+1)f(h'+1,1,n,W) + p_0(h'+1)g(h'+1,1,n,W) \\ > \;& 1 + p_1(h')f(h',1,n,W) + p_0(h')g(h',1,n,W), \end{aligned}$$

and hence $v^*(h'+1,1,n) > v^*(h',1,n)$. □

We next provide conditions that allow us to prove that when no therapies remain, the value of being in a state of improving health is strictly greater than being in a state of deteriorating health for each health category greater than 0.

(As7.10) $p_1(h) > 0$ for all $h > 0$.

*As*7.10 says that for every state in which a patient's health is stochastically improving, there is a positive probability that it will remain improving at the next time period.

(As7.11) For each $h > 0$, there exists k such that

$$\sum_{h'=k}^{H} q[(h', 1, 0)|(h, 1, 0), W] > \sum_{h'=k}^{H} q[(h', 1, 0)|(h, 0, 0), W].$$

*As*7.11 says that at least one of the inequalities that go into the condition of *As*7.7 for each $h > 0$ is strict.

We can now prove the following.

Lemma 7.2. *$v^*(h, 1, 0) > v^*(h, 0, 0)$ for all $h > 0$.*

Proof. Let $h > 0$, arbitrarily. When zero therapies remain, the only action is to wait, and hence we have:

$$v^*(h, 1, 0) = 1 + p_1(h) f(h, 1, 0, W) + p_0(h) f(h, 0, 0, W), \tag{7.34}$$

$$v^*(h, 0, 0) = 1 + f(h, 0, 0, W). \tag{7.35}$$

Similar to (7.32), we can write

$$f(h, 1, 0, W) = \sum_{i=1}^{H} [v^*(i, 1, 0) - v^*(i - 1, 0, 0)] \sum_{j=i}^{H} q[(j, 1, 0)|(h, 1, 0), W].$$

By the result of Lemma 7.1 and *As*7.11, it follows that

$$\sum_{h'=0}^{H} q[(h', 1, 0)|(h, 1, 0), W] v^*(h, 1, 0) \tag{7.36}$$

$$> \sum_{h'=0}^{H} q[(h', 0, 0)|(h, 1, 0), W] v^*(h, 1, 0). \tag{7.37}$$

Also, by Theorem 7.1,

$$\sum_{h'=0}^{H} q[(h',0,0)|(h,1,0),W]v^*(h,1,0) \tag{7.38}$$

$$\geq \sum_{h'=0}^{H} q[(h',0,0)|(h,1,0),W]v^*(h,0,0). \tag{7.39}$$

Finally, combining (7.36)-(7.39) with *As*7.10 establishes that the right-hand side of (7.34) is greater than the right-hand side of (7.35) and hence $v^*(h,1,0) > v^*(h,0,0)$. □

We are now ready to prove the result that there is a benefit of having the therapies that remain available.

Theorem 7.4. *For each $n \geq 1$, there exists $h \geq 1$ such that $a^*(h,0,n) = S$, uniquely.*

Proof. Suppose, towards a contradiction, that there exists $n \geq 1$, such that $a^*(h,0,n) = W$ for all $h \geq 1$. Then by *As*7.2, $v^*(h,0,n) = v^*(h,0,0)$ for all h. But Lemma 7.2 states that $v^*(h,1,0) > v^*(h,0,0)$, and by Theorems 7.1 and 7.3, it follows that $v^*(h,0,n) > v^*(h,0,0)$. □

In the following, we suppose that upon switching to a new therapy, the system transitions to one of the various health categories and health directions independently of the state from which the system moves. This generalizes a common assumption in machine replacement problems, where upon replacement, the system moves to the best state with probability 1, regardless of the state from which the replacement is made [22]. While the generalization below is not clinically realistic in the HIV setting, we include it for additional structural analysis.

Theorem 7.5. *Suppose that instead of As*7.3*, we assume that upon switching to a new therapy from a state (h_1,d_1,n_1), patients transition to various states of the form (h_2,d_2,n_1-1) with probabilities independent of (h_1,d_1,n_1). Then for each level of $n \geq 1$, an optimal control-limit policy exists.*

Proof. Let $n \geq 1$ and suppose that there exists h, such that $a^*(h, 0, n) = W$, uniquely, and $a^*(h+1, 0, n) = S$. By the assumption of the current theorem, the value associated with switching to a new therapy from state $(h, 0, n)$ equals the value associated with switching from state $(h+1, 0, n)$. This, along with our assumption that $a^*(h, 0, n) = W$, uniquely, implies that $v^*(h, 0, n) > v^*(h+1, 0, n)$, contradicting the result of Theorem 7.1. Therefore, $a^*(h, 0, n) = W$, uniquely, implies $a^*(h+1, 0, n) = W$, and the result follows. □

7.4 EXTENDING THE FRAMEWORK

Just as we did in Chapter 6, we want to consider the more clinically realistic setting of therapies that are not identical and independent of each other's use. We begin by relaxing the assumption of identically distributed therapies and consider therapies that may have different levels of effectiveness. We develop an algorithm that embeds MDPs within an outer dynamic program, analogous to the algorithm developed in Section 6.3.2. For example, for each therapy, we consider a patient who is on that therapy with no more therapies remaining, and we obtain the value associated with each state (the only action is to continue waiting on that therapy). Next, we consider all $\binom{n}{2}$ pairs of therapies and for each state, evaluate the optimal value associated with being on one therapy of the pair and having the other therapy available. For example, for therapies i and j, we first consider that the patient is on therapy i with j remaining, and for each state we determine whether it is optimal for the patient to remain on i or switch to j. We then do the same thing, assuming the patient is on j with i remaining. For ease of exposition, suppose a patient has never initiated therapy and only has therapies i and j available. Then for each state, the patient may continue waiting, initiate therapy i, or initiate therapy j. From the previous step of the algorithm, we obtained the values of being in each state, on therapy i with j remaining, or on therapy j with i remaining, so we do not need to repeat that work at this level of the algorithm.

Recall that in Chapters 4 and 5 we assumed that once patients initiate therapy, they receive the expected remaining lifetime associated with the sequencing and switching of therapies as done in current practice. In Section 4.6, we mentioned that after the presentation

of the current and previous chapters, we may reconsider the problem of the optimal time to initiate therapy under a framework that optimizes the entire therapy planning process. The algorithm we described does just that: the final step of the algorithm considers a patient never before on therapy and for each state, determines whether the patient should initiate therapy and if so, which one.

Finally, we note that the same reasons we gave for the difficulty of relaxing the independence assumption in Section 6.3.3 apply to this context as well. However, similar simplifications may make the problem feasible to solve. We leave that for future work.

7.5 CONCLUSIONS

There are various phenomena in HIV therapy planning (and therapeutic optimization problems in general) that challenge the often used framework of a stationary MDP. One is the fact that patient aging must be considered when modeling a long time horizon. Another is the reality that therapies are often effective for some period of time but then lose their effectiveness. Finally, patients face a finite number of therapy replacements.

We addressed each of these issues in developing a stationary MDP for the problem of switching and sequencing HIV therapies. Additionally, we demonstrated structural properties of this framework under an assumption of equally effective therapies. We also described an MDP algorithm for the setting of therapies with possibly different levels of effectiveness.

The data requirements for a clinically based implementation of this model are immense and we leave that for future work. Also, some of our modeling assumptions may be reconsidered in future work. For example, we assumed that therapies cannot be reused; however, it may be useful to think about the possibility of reusing therapies at a reduced level of effectiveness. Similarly, we assumed that waiting cannot cause the loss of available therapies. In HIV care, it is possible that the development of resistance to an active drug confers resistance to a drug never before taken, so we may consider the possibility of a stock of unused therapies losing effectiveness over time (this has connections to the review of inventory depletion management problems in Section 2.3.3). Finally, we also assumed that new

therapies are not introduced over time; in reality, new drugs are often under research and development. Therefore, we may want to consider this source of uncertainty as well.

8.0 CONCLUSIONS

This dissertation has focused on applying mathematical techniques to address fundamental and unresolved questions in HIV therapy planning: when to initiate, when to switch, and how to sequence therapies. In Chapter 2, we reviewed the literature addressing these topics, which includes various clinical as well as mathematical studies. For a process as complex as treatment planning, mathematical studies have practical advantages over clinical studies, though the latter are necessary to perform before making significant changes to treatment guidelines. Two types of mathematical approaches have been prominent in HIV therapy planning: Monte Carlo simulation models and control-theoretic models. As discussed, simulation models are excellent tools for testing "what if" scenarios but are not efficient means for seeking optimal policies. On the other hand, while control-theoretic models optimize some objective, they fail to consider essential issues in HIV care: risk of death and quality of life. Because HIV therapy planning is a sequential stochastic decision process which should consider long-term survival and quality of life, we modeled the above questions with a Markov decision process (MDP). MDPs are designed precisely to optimize periodic decisions under an uncertain environment, and they are well suited for handling objectives such as maximizing expected lifetime or quality-adjusted lifetime.

Chapters 4 and 5 examined the question of the optimal time to initiate HIV therapy, with the former considering a single prognostic variable (CD4 count) and the latter consider two variables (CD4 count and viral load). After developing the modeling frameworks for both, we examined structural properties such as increasing value functions, optimal policies of initiating therapy from all states, and optimal control-limit policies. We also considered how to adapt the single-variable model to include patient-specific factors such as quality of

life and adherence. In Chapter 3, we used clinical data to build the key components of the MDP we solved in Chapter 4: a natural history transition probability matrix along with estimates of patient survival after initiating therapy.

The results of the data-driven MDP demonstrated strong support for a strategy of initiating therapy immediately. Under a variety of sensitivity analyses with respect to both objectives of maximizing lifetime and quality-adjusted lifetime, the solutions pointed to a policy of initiating therapy from each CD4 strata. Moreover, we found other examples in the literature to suggest that treatment guidelines reexamine recent trends toward treating HIV later in its course.

Chapters 6 and 7 developed models for addressing the questions of when to switch therapies and how to sequence them. In Chapter 6, we considered a continuous-time approach to switching and sequencing therapies, given that one has information only about the lifetime distributions associated with the therapies. Chapter 7 considered periodic observations and state variables that indicate a patient's state of health, whether or not the patient is on an effective therapy, and the number of effective therapies remaining.

Limitations of the work lie both with the data as well as some of our modeling assumptions. For instance, we built our models based on male patients in the VA. Therefore, our results may not extend readily to women or HIV patients in other settings. A primary data challenge was how to handle the significant number of censored observations attributable to initiating therapy in the natural history model. Future work may consider a more rigorous approach to dealing with this issue. Also, although we shall soon build and solve a data-driven model of the optimal time to initiate therapy as a function of both CD4 count and viral load, building a clinically valid, data-driven model of the switching and sequencing questions will not appear soon due to great data requirements. One modeling assumption we made was of perfect observations of a patient's state. Depending on the amount of uncertainty in the CD4 and viral load measurements, we may want to model the problem as a partially-observable MDP that considers this uncertainty. Moreover, measurements of patient adherence are notoriously inaccurate and hence we should take this into account when explicitly considering patient adherence. Furthermore, we would like to combine an explicit consideration of patient adherence with utilities in a quality-of-life framework. Finally, we

assumed periodic inspections of patient health (e.g., every month) and leave for future research a cost-effectiveness analysis examining the question of how often a patient should visit a physician for laboratory measurements.

Although there have been great strides in HIV therapies over the past 20 years, there is still considerable debate over the best way to use them. Questions regarding the optimal time to initiate therapy, the time to switch therapies, and the sequencing of therapies are central to effective HIV care today. The solutions to our model of when to initiate therapy support former strategies of administering HAART earlier in the course of HIV as opposed to recent trends of delaying therapy. While data are not mature enough to solve our models of the switching and sequencing of therapies, the models provide methodological frameworks for thinking about these difficult questions.

BIBLIOGRAPHY

[1] O.O. Aalen, V.T. Farewell, D. de Angelis, N.E. Day, and O.N. Gill. A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: Application to AIDS prediction in England and Wales. *Statistics in Medicine*, 16(19):2191–2210, 1997.

[2] L. Ahdieh-Grant, T.E. Yamashita, J.P. Phair, R. Detels, S.M. Wolinsky, J.B. Margolick, C.R. Rinaldo, and L.P. Jacobson. When to initiate highly active antiretroviral therapy: A cohort approach. *American Journal of Epidemiology*, 157(8):738–746, 2003.

[3] AIDS Education Global Information System. Opportunistic infections, 2006. Retrieved June 1, 2006, www.aegis.com/topics/oi.

[4] O. Alagoz, C.L. Bryce, S. Shechter, A. Schaefer, C.C.H. 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*, 25(6):620–632, 2005.

[5] O. Alagoz, L.M. Maillart, A.J. Schaefer, and M.S. Roberts. The optimal timing of living-donor liver transplantation. *Management Science*, 50(10):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. To appear in *Operations Research*, 2005.

[7] J. Alvarez-Ramirez, M. Meraz, and J.X. Velasco-Hernandez. Feedback control of the chemotherapy of HIV. *International Journal of Bifurcation and Chaos*, 10(9):2207–2219, 2000.

[8] K. Anastos, Y. Barrón, P. Miotti, B. Weiser, M. Young, N. Hessol, R.M. Greenblatt, M. Cohen, M. Augenbraun, A. Levine, and A. Muñoz for the Women's Interagency HIV Study Collaborative Study Group. Risk of progression to AIDS and death in women infected with HIV-1 initiating highly active antiretroviral treatment at different stages of disease. *Archives of Internal Medicine*, 162(17):1973–1980, 2002.

[9] R.N. Anderson. Method for constructing annual U.S. life tables. *National Center for Health Statistics: Vital and Health Statistics*, 2(129):1–36, 1999. Available at http://www.cdc.gov/nchs/data/series/sr_02/sr02_129.pdf.

[10] AVERT.org. The different stages of HIV infection, 2006. Retrieved June 20, 2006, www.avert.org/hivstages.htm.

[11] D.R. Bangsberg, S. Perry, E.D. Charlebois, R.A. Clark, M. Robertson, A.R. Zolopa, and A. Moss. Non-adherence to highly active antiretroviral therapy predicts progression to AIDS. *AIDS*, 15(9):1181–1183, 2001.

[12] R.E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. John Wiley & Sons, Inc., New York, New York, 1965.

[13] R.E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, Inc., New York, New York, 1975.

[14] A.M. Bayoumi and D.A. Redelmeir. Preventing mycobacterium avium complex in patients who are using protease inhibitors: A cost-effectiveness analysis. *AIDS*, 12(12):1503–1512, 1998.

[15] A.M. Bayoumi and D.A. Redelmeir. Economic methods for measuring the quality of life associated with HIV infection. *Quality of Life Research*, 8(6):471–480, 1999.

[16] J.R. Beck and S.G. Pauker. The Markov process in medical prognosis. *Medical Decision Making*, 3(4):419–458, 1983.

[17] R.E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, New Jersey, 1957.

[18] M. Bennett, D. Indyk, and S. Golub. Adherence re-framed in the BIG picture: A qualitative ecological perspective on HIV+ patients and protease inhibitors [abstract 32365]. Geneva, 1998. Marathon Multimedia. In: Program and Abstracts of the Twelfth World AIDS Conference.

[19] Z. Benyamini and U. Yechiali. Optimality of control limit maintenance policies under nonstationary deterioration. *Probability in the Engineering and Informational Sciences*, 13:55–70, 1999.

[20] S.M. Berman. Optimal timing of antiretroviral therapy in HIV infection. In J. Galambos and J. Gani, editors, *Studies in Applied Probability: Essays in Honour of Lajos Takács*, pages 3–15. Applied Probability Trust, Sheffield, England, 1994.

[21] S.M. Berman and N. Dubin. Is earlier better for AZT therapy in HIV infection? A mathematical model. In N.P. Jewell, K. Dietz, and V.T. Farewell, editors, *AIDS Epidemiology: Methodological Issues*, pages 366–383. Birkhäuser, Boston, Massachusetts, 1992.

[22] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, Belmont, Massachusetts, 1996.

[23] A. Biswas and A. Dewanji. Scheduling optimal examination times in a simple illness-death model. *Journal of Statistical Planning and Inference*, 126(1):173–187, 2004.

[24] R.S. Braithwaite, 2006. Personal Correspondence.

[25] R.S. Braithwaite, A.C. Justice, C.C.H. Chang, J.S. Fusco, S.R. Raffanti, J.B. Wong, and M.S. Roberts. Estimating the proportion of patients infected with human immunodeficiency virus who will die of comorbid disease. *American Journal of Medicine*, 118(8):890–898, 2005.

[26] R.S. Braithwaite, A.C. Justice, S.M. Shechter, A.J. Schaefer, J.B. Wong, and M.S. Roberts. A Monte Carlo simulation of HIV disease that explicitly represents genotypic resistance and adherence to therapy. Submitted for publication, 2006.

[27] R.S. Braithwaite, M.S. Roberts, J. Conigliaro, S.M. Shechter, A.J. Schaefer, K. McGinnis, and A.C. Justice. Estimating the impact of alcohol consumption on survival for HIV+ individuals in the Veterans Aging Cohort Study. Submitted for publication, 2006.

[28] R.S. Braithwaite, S.M. Shechter, M.S. Roberts, A.J. Schaefer, D.R. Bangsberg, P.R. Harrigan, and A.C. Justice. Explaining variability in the relationship between antiretroviral adherence and HIV mutation accumulation. Submitted for publication, 2006.

[29] M.E. Brandt and G. Chen. Feedback control of a biodynamical model of HIV-1. *IEEE Transactions on Biomedical Engineering*, 48(7):754–759, 2001.

[30] M. Brown and S.M. Ross. Optimal issuing policies. *Management Science*, 19(11):1292–1294, 1973.

[31] M. Brown and H. Solomon. Optimal issuing policies under stochastic field lives. *Journal of Applied Probability*, 10:761–768, 1973.

[32] J.F. Brundage, J.G. McNeil, R.N. Miller, L.I. Gardner Jr., S.M. Harrison, C. Hawks, D.B. Craig, R. Redfield, D.S. Burke, and the U.S. Army Retrovirus Research Group. The current distribution of CD4+ T-lymphocyte counts among adults in the United States with Human Immunodeficiency Virus infections: Estimates based on the experience of the U.S. Army. *Journal of Acquired Immune Deficiency Syndromes*, 3(1):92–94, 1990.

[33] S. Butler, D. Kirschner, and S. Lenhart. Optimal control of chemotherapy affecting the infectivity of HIV. In D. Arino, D. Axelrod, and M. Kimmel, editors, *Advances in Mathematical Populations Dynamics: Molecules, Cells, Man*, pages 557–569. World Scientific Press, Singapore, 1997.

[34] A. Carmona, H. Knobel, A. Guelar, S. Grau, J. Mateau de Antonio, J.L. Lopez-Colomes, and P. Saballs. Factors influencing survival in HIV infected patients treated with HAART [Abstract TuOrB417]. Durban, South Africa, 2000. Thirteenth International AIDS Conference.

[35] R.E. Chaisson, J.C. Keruly, and R.D. Moore. Association of initial CD4 cell count and viral load with response to highly active antiretroviral therapy. *JAMA*, 284(24):3128–3129, 2000.

[36] N. Cheonis. Immune reconstitution syndrome, 2005. Retrieved April 26, 2006, www.thebody.com/sfaf/winter05/irs.html#reactions.

[37] M. Chesney. Adherence to HAART regimens. *AIDS Patient Care and STDs*, 17(4):169–177, 2003.

[38] M.A. Chesney, J. Ickovics, F.M. Hecht, G. Sikipa, and J. Rabkin. Adherence: A necessity for successful HIV combination therapy. *AIDS*, 13:S271–S278, 1999.

[39] Y.S. Chow, H. Robbins, and D. Siegmund. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Company, Boston, Massachusetts, 1971.

[40] A.H. Christer and P.A. Scarf. A robust replacement model with applications to medical equipment. *Journal of the Operational Research Society*, 45(3):261–275, 1994.

[41] P.D. Cleary, F.J. Fowler Jr., J. Weissman, M.P. Massagli, I. Wilson, G.R. Seage III, C. Gatsonis, and A. Epstein. Health-related quality of life in persons with acquired immune deficiency syndrome. *Medical Care*, 31(7):569–580, 1993.

[42] C.J. Cohen and B.A. Boyle. Antiretroviral therapy: The "when to start" debates. *Clinical Infectious Diseases*, 39(11):1705–1708, 2004.

[43] D. Collett. *Modelling Survival Data in Medical Research*. Chapman & Hall, London, United Kingdom, 1994.

[44] Community Programs for Clinical Research on AIDS. SMART: Strategies for Management of Anti-Retroviral Therapy, 2002. Retrieved December 6, 2005, www.smart-trial.org.

[45] A.E. Copfer, N.M. Ampel, T.E. Hughes, K.J. Gregor, C.L. Dols, S.J. Coons, K. Colgan, and A.W. Wu. The use of two measures of health-related quality of life in HIV-infected individuals: A cross-sectional comparison. *Quality of Life Research*, 5(2):281–286, 1996.

[46] D.R. Cox. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34:187–220, 1972.

[47] B.A. Craig, D.G. Fryback, R. Klein, and B.E.K. Klein. A Bayesian approach to modelling the natural history of a chronic condition from observations with intervention. *Statistics in Medicine*, 18(11):1355–1371, 1999.

[48] B.A. Craig and P.P. Sendi. Estimation of the transition matrix of a discrete-time Markov chain. *Health Economics*, 11(1):33–42, 2002.

[49] L. Dal Maso, M. Gava, P. Pezzotti, N. Torelli, S. Franceschi, and R. Giovannia for the Italian Seroconversion Study. Markov models for HIV disease progression: An unverified assumption. *Journal of Acquired Immune Deficiency Syndromes*, 25(5):466–468, 2000.

[50] P. Dalgaard. *Introductory Statistics with R.* Springer-Verlag, New York, 2002.

[51] R.M. D'Amato, R.T. D'Aquila, and L.M. Wein. Management of antiretroviral therapy for HIV infection: Modelling when to change therapy. *Journal of Antiviral Therapy*, 3:147–158, 1998.

[52] R.M. D'Amato, R.T. D'Aquila, and L.M. Wein. Management of antiretroviral therapy for HIV infection: Analyzing when to change therapy. *Management Science*, 46(9):1200–1213, 2000.

[53] E.J. Dasbach, D.G. Fryback, P.A. Newcomb, R. Klein, and B.E. Klein. Cost-effectiveness of strategies for detecting diabetic retinopathy. *Medical Care*, 29:20–39, 1991.

[54] V. De Gruttola, M. W ulfsohn, M.A. Fischl, and A. Tsiatis. Modeling the relationship between survival and CD4 lymphocytes in patients with AIDS and AIDS-related complex. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 6:359–365, 1993.

[55] R. Dekker. Applications of maintenance optimisation models: A review and analysis. *Reliability Engineering and System Safety*, 51:229–240, 1996.

[56] C. Derman. On optimal replacement rules when changes of state are Markovian. In R. Bellman, editor, *Mathematical Optimization Techniques*, pages 201–210. University of California Press, Berkeley, California, 1963.

[57] C. Derman and M. Klein. Inventory depletion management. *Management Science*, 4(4):450–456, 1958.

[58] C. Derman, G.J. Lieberman, and S.M. Ross. On the use of replacements to extend system life. *Operations Research*, 32(3):616–627, 1984.

[59] M.F. Drummond, B. O'Brien, G.L. Stoddart, and G.W. Torrance. *Methods for the Economic Evaluation of Health Care Programmes.* Oxford University Press, Oxford, England, 1997.

[60] S. Duran, B. Spire, F. Raffi, V. Walter, D. Bouhour, V. Journot, V. Cailleton, C. Leport, J.P. Moatti, and the APROCO Cohort Study Group. Self-reported symptoms after initiation of a protease inhibitor in HIV-infected patients and their impact on adherence to HAART. *HIV Clinical Trials*, 2(1):38–45, 2001.

[61] M. Egger, M. May, G. Chene, A.N. Phillips, B. Ledergerber, F. Dabis, D. Costagliola, A. D'Arminio Monforte, F. de Wolf, P. Reiss, J.D. Lundgren, A.C. Justice, S. Staszewski, C. Leport, R.S. Hogg, C.A. Sabin, M.J. Gill, B. Salzberger, J.A.C. Sterne, and the ART Cohort Collaboration. Progonosis of HIV-1-infected patients starting highly active antiretroviral therapy: A collaborative analysis of prospective studies. *The Lancet*, 360(9327):119–129, 2002.

[62] W. El-Sadr and J. Neaton for the SMART Study Investigators. Episodic CD4-guided use of ART is inferior to continuous therapy: Results of the SMART study [abstract 106LB]. Denver, Colorado, 2006. Thirteenth Conference on Retroviruses and Opportunistic Iinfections, available at: http://www.retroconference.org.

[63] L.J. Eldred, A.W. Wu, R.E. Chaisson, and R.D. Moore. Adherence to antiretroviral and pneumocystis prophylaxis in HIV disease. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 18(2):117–125, 1998.

[64] E.G. Enns and P.F. Ehlers. Optimal issuing policies for non-identical units with stochastic lifetimes. *Journal of Applied Probability*, 13:183–189, 1976.

[65] T.S. Ferguson. Optimal stopping and applications, 2003. Retrieved December 6, 2005, http://www.math.ucla.edu/~tom/Stopping/sr1.pdf.

[66] K.R. Fister, S. Lenhart, and J.S. McNally. Optimizing chemotherapy in an HIV model. *Electronic Journal of Differential Equations (http://ejde.math.txstate.edu/)*, 32:1–12, 1998.

[67] K.A. Freedberg, E. Losina, M.C. Weinstein, A.D. Paltiel, C.J. Cohen, G.R. Seage, D.E. Craven, H. Zhang, A.D. Kimmel, and S.J. Goldie. The cost effectiveness of combination antiretroviral therapy for HIV disease. *The New England Journal of Medicine*, 344(11):824–831, 2001.

[68] K.A. Freedberg, J.A. Scharfstein, G.R. Seage, E. Losina, M.C. Weinstein, D.E. Craven, and A.D. Paltiel. The cost effectiveness of preventing AIDS-related opportunistic infections. *JAMA*, 279(2):130–136, 1998.

[69] R.D. Gelber, W.R. Lenderking, D.J. Cotton, B.F. Cole, M.A. Fischl, A. Goldhirsch, and M.A. Testa. Quality-of-life evaluation in clinical trials of zidovudine therapy in patients with mild symptomatic HIV infection. *Annals of Internal Medicine*, 116(12 Pt 1):961–966, 1992.

[70] R.C. Gentleman, J.F. Lawless, J.C. Lindsey, and P. Yan. Multi-state Markov models for analysing incomplete disease history data with illustrations for HIV disease. *Statistics in Medicine*, 13(8):805–821, 1994.

[71] E. Gir, R. Pratt, E.H. Bunch, and W.L. Holzemer. Adherence to anti-retroviral therapy: A four country comparison [abstract 60129]. Geneva, 1998. Marathon Multimedia. In: Program and Abstracts of the Twelfth World AIDS Conference.

[72] M.R. Gold, D.L. Patrick, G.W. Torrance, D.G. Fryback, D.C. Hadorn, M.S. Kamlet, N. Daniels, and M.C. Weinstein. Identifying and valuing outcomes. In M.R. Gold, J.E. Siegel, L.B. Russell, and M.C. Weinstein, editors, *Cost-Effectiveness in Health and Medicine*, pages 82–134. Oxford University Press, New York, New York, 1996.

[73] P.J. Green and B.W. Silverman. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman & Hall, London, United Kingdom, 1994.

[74] B. Hajek. Optimal control of two interacting service stations. *IEEE Transactions on Automatic Control*, AC-29(6):491–499, 1984.

[75] M. Harrington and C.C.J. Carpenter. Hit HIV hard, but only when necessary. *The Lancet*, 355(9221):2147–2152, 2000.

[76] R.H. Haubrich, S.J. Little, J.S. Currier, D.N. Forthal, C.A. Kemper, G.N. Beall, D. Johnson, M.P. Dube, J.Y. Hwanga, J.A. McCutchan, and the California Collaborative Treatment Group. The value of patient-reported adherence to antiretroviral therapy in predicting virologic and immunologic response. *AIDS*, 13(9):1099–1107, 1999.

[77] K. Henry. The case for more cautious, patient-focused antiretroviral therapy. *Annals of Internal Medicine*, 132(4):306–311, 2000.

[78] D.D. Ho. Time to hit HIV, early and hard. *New England Journal of Medicine*, 333:450–451, 1995.

[79] C. Hoffman and F. Mulcahy. When to start HAART. In C. Hoffman, J. Rockstroh, and B.S. Kamps, editors, *HIV Medicine*. Flying Publisher, Paris, France, 2005. www.hivmedicine.com.

[80] R.S. Hogg, B. Yip, K.J. Chan, E. Wood, K.J.P. Craib, M.V. O'Shaughnessy, and J.S.G. Montaner. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA*, 286(20):2568–2577, 2001.

[81] S.D. Holmberg, F.J. Palella, K.A. Lichtenstein, and D.V. Havlir. The case for earlier treatment of HIV infection. *Clinical Infectious Diseases*, 39:1699–1704, 2004.

[82] D.R. Holtgrave and N.L. Qualis. Threshold analysis and HIV prevention programs. *Medical Decision Making*, 15:311–317, 1995.

[83] W.J. Hopp and S.C. Wu. Machine maintenance with multiple maintenance actions. *IIE Transactions*, 22(3):226–232, 1990.

[84] C. Hu, W.S. Lovejoy, and S.L. Shafer. Comparison of some suboptimal control policies in medical drug therapy. *Operations Research*, 44(5):696–709, 1993.

[85] D.J.M. Isaman, W.H. Herman, and M.B. Brown. A discrete-state discrete-time model using indirect observation. *Statistics in Medicine*, 25(6):1035–1049, 2006.

[86] A.K.S. Jardine and J.A. Buzacott. Equipment reliability and maintenance. *European Journal of Operational Research*, 19(3):285–296, 1985.

[87] A.M. Jeffrey, X. Xia, and I.K. Craig. When to initiate HIV therapy: A control theoretic approach. *IEEE Transactions on Biomedical Engineering*, 50(11):1213–1220, 2003.

[88] D.W. Jorgenson, J.J. McCall, and R. Radner. *Optimal Replacement Policy*. Rand McNally & Company, Chicago, Illinois, 1967.

[89] H.R. Joshi. Optimal control of an HIV immunology model. *Optimal Control Applications and Methods*, 23(4):199–213, 2002.

[90] A.C. Justice, 2005. Personal Correspondence.

[91] A. Kamina. *Modeling the Emergence of Drug-Resistant Mutants: Implications for Treatment Strategies for HIV-1 Infection*. Ph.D. Dissertation, Yale University, 2000.

[92] E.P.C. Kao. Optimal replacement rules when changes of state are semi-Markovian. *Operations Research*, 21(6):1231–1249, 1973.

[93] E.L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481, 1958.

[94] E.B. Keeler and S. Cretin. Discounting of life-saving and other nonmonetary effects. *Management Science*, 29(3):300–306, 1983.

[95] J.T. King, A.C. Justice, M.S. Roberts, C.C.H. Chang, and J.S. Fusco. Long term HIV/AIDS survival estimation in the highly active antiretroviral therapy era. *Medical Decision Making*, 23(1):9–20, 2003.

[96] R.L.A. Kirch and M. Klein. Surveillance schedules for medical examinations. *Management Science*, 20(10):1403–1409, 1974.

[97] D. Kirschner, S. Lenhart, and S. Serbin. Optimal control of the chemotherapy of HIV. *Journal of Mathematical Biology*, 35(7):775–792, 1997.

[98] D. Kirschner and G.F. Webb. A model for treatment strategy in the chemotherapy of AIDS. *Bulletin of Mathematical Biology*, 58(2):367–390, 1996.

[99] J.P. Klein and M.L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York, 1997.

[100] M. Klein. Inspection-maintenance-replacement schedules under Markovian deterioration. *Management Science*, 9(1):25–32, 1962.

[101] V.G. Kulkarni. *Modeling and Analysis of Stochastic Systems*. Chapman and Hall, London, United Kingdom, 1995.

[102] Lab Tests Online. Lab Tests Online: CD4 count, 2006. Retrieved June 5, 2006, www.labtestsonline.org/understanding/analytes/cd4/test.html.

[103] A.M. Law and W.D. Kelton. *Simulation Modeling and Analysis, third edition*. McGraw-Hill, New York, New York, 2000.

[104] J.F. Lawless. *Statistical Models and Methods for Lifetime Data*. John Wiley, New York, 1982.

[105] L.A. Lenert, M. Feddersen, A. Sturley, and D. Lee. Adverse effects of medications and trade-offs between length of life and quality of life in Human Immunodeficiency Virus infection. *The American Journal of Medicine*, 113(3):229–232, 2002.

[106] A.C. Lepri, A.N. Phillips, A. d'Arminio Monforte, F. Castelli, A. Antinori, A. de Luca, P. Pezzotti, F. Alberici, A. Cargnel, P. Grima, R. Piscopo, T. Prestileo, G. Scalise, M. Vigevani, and M. Moroni for the ICONA study group. When to start highly active antiretroviral therapy in chronically HIV-infected patients: Evidence from the ICONA study. *AIDS*, 15(8):983–990, 2001.

[107] K. Lichtenstein, C. Armon, K. Buchacz, A. Moorman, K. Wood, J. Brooks, and the HIV Outpatient Study Group. Early, uninterrupted ART is associated with improved outcomes and fewer toxicities in the HIV Outpatient Study (HOPS) [abstract 769]. Denver, Colorado, 2006. Thirteenth Conference on Retroviruses and Opportunistic Infections, available at: http://www.retroconference.org.

[108] G.J. Lieberman. LIFO vs. FIFO in inventory depletion management. *Management Science*, 5(2):102–105, 1958.

[109] T.L. Lincoln and G.H. Weiss. A statistical evaluation of recurrent medical examinations. *Operations Research*, 12(2):187–205, 1964.

[110] I.M. Longini Jr. Modeling the decline of CD4+ T-lymphocyte counts in HIV-infected individuals. *Journal of Acquired Immune Deficiency Syndromes*, 3(9):930–931, 1990.

[111] I.M. Longini Jr., W.S. Clark, L.I. Gardner, and J.F. Brundage. The dynamics of CD4+ T-lymphocyte decline in HIV-infected individuals: A Markov modeling approach. *Journal of Acquired Immune Deficiency Syndromes*, 4(11):1141–1147, 1991.

[112] J. Martinez-Picado, E. Negredo, L. Ruiz, A. Shintani, C.R. Fumaz, C. Zala, P. Domingo, J. Vilaró, J.M. Llibre, P. Viciana, K. Hertogs, C. Boucher, R.T. D'Aquila, B. Clotet, and the SWATCH Study Team. Alternation of antiretroviral drug regimens for HIV infection. *Annals of Internal Medicine*, 139(2):81–89, 2003.

[113] J. Mauskopf, M. Kitahata, T. Kauf, A. Richter, and J. Tolson. HIV antitretoviral treatment: Early versus later. *Journal of Acquired Immune Deficiency Syndromes*, 39(5):562–569, 2005.

[114] J.J. McCall. Maintenance policies for stochastically failing equipment: A survey. *Management Science*, 11(5):493–524, 1965.

[115] D. Meichenbaum and D. C. Turk. Treatment adherence: Terminology, incidence, and conceptualization. In *Facilitating Treatment Adherence: A Practitioner's Guidebook*, pages 19–39. Plenum, New York, 1987.

[116] J.W. Mellors, A. Munoz, J.V. Giorgi, J.B. Margolick, C.J. Tassoni, P. Gupta, L.A. Kingsley, J.A. Todd, A.J. Saah, R. Detels, J.P. Phair, and C.R. Rinaldo Jr. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Annals of Internal Medicine*, 126(12):946–954, 1997.

[117] V. Miller, C. Sabin, K. Hertogs, S. Bloor, J. Martinez-Picado, R. D'Aquila, B. Larder, T. Lutz, P. Gute, E. Weidmann, H. Rabenau, A. Phillips, and S. Staszewski. Virological and immunological effects of treatment interruptions in HIV-1 infected patients with treatment failure. *AIDS*, 14(18):2857–2867, 2000.

[118] G.E. Monahan. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.

[119] J.M. Mrus, B.R. Schackman, A.W. Wu, K.A. Freedberg, J. Tsevat, M.S. Yi, and R. Zackin. Variations in self-rated health among patients with HIV infection. *Quality of Life Research*, 15(3):503–514, 2006.

[120] National Institute of Allergy and Infectious Diseases. International HIV/AIDS trial finds continuous antiretroviral therapy superior to episodic therapy, 2006. Retrieved April 20, 2006 www.smart-trial.com/members2/download/NIH-press-release_001.pdf.

[121] National Institute of Allergy and Infectious Diseases. Treatment of HIV infection, 2006. Retrieved April 21, 2006, www.niaid.nih.gov/factsheets/treat-hiv.htm.

[122] P.T. Nieuwkerk, M.A.G. Sprangers, D.M. Burger, R.M.W. Hoetelmans, P.W.H. Hugen, S.A. Danner, M.E. van der Ende, M.M.E. Schneider, G. Schrey, P.L. Meenhorst, H.G. Sprenger, R.H. Kauffmann, M. Jambroes, M.A. Chesney, F. de Wolf, and J.M.A. Lange. Limited patient adherence to highly active antiretroviral therapy for HIV-1 infection in an observational cohort study. *Archives of Internal Medicine*, 161:1962–1968, 2001.

[123] U.S. Department of Health and Human Services. Guidelines for the use of antiretroviral agents in HIV-1 infected adults and adolescents, 2005.

[124] M. O'Shaughnessy, J. Montaner, R. Hogg, B. Yip, K. Chan, and K. Craib. To start or not to start? Diminished effectiveness of anti-retroviral therapy among patients initiating therapy with CD4+ cell counts below 200/mm3 [abstract LbPeB7050]. Durban, South Africa, 2000. Thirteenth International AIDS Conference.

[125] D.K. Owens, A.B. Cardinalli, and R.F. Nease Jr. Physicians' assessments of the utility of health states associated with human immunodeficiency virus infection. *Quality of Life Research*, 6(1):77–86, 1997.

[126] F.J. Palella, M. Deloria-Knoll, J.S. Chmiel, A.C. Moorman, K.C. Wood, A.E. Greenberg, S.D. Holmberg, and the HIV Outpatient Study (HOPS) Investigators. Survival benefit of initiating antitretroviral therapy in HIV-infected persons in different CD4+ cell strata. *Annals of Internal Medicine*, 138(8):620–626, 2003.

[127] F.J. Palella Jr., K.M. Delaney, A.C. Moorman, M.O. Loveless, J. Fuhrer, G.A. Satten, D.J. Aschman, S.D. Holmberg, and the HIV Outpatient Study Investigators. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine*, 338(13):853–860, 1998.

[128] A.D. Paltiel, J.A. Scharfstein, G.R. Seage III, E. Losina, S.J. Goldie, M.C. Weinstein, D.E. Craven, and K.A. Freedberg. A Monte Carlo simulation of advanced HIV disease: Application to prevention of CMV infection. *Medical Decision Making*, 18(2 Suppl):S93–S105, 1998.

[129] D.L. Paterson, S. Swindells, J. Mohr, M. Brester, E.N. Vergis, C. Squier, M.M. Wagener, and N. Singh. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Annals of Internal Medicine*, 133(1):21–30, 2000.

[130] H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

[131] A.N. Phillips, A.C. Lepri, F. Lampe, M. Johnson, and C.A. Sabin. When should antiretroviral therapy be started for HIV infection? Interpreting the evidence from observational studies. *AIDS*, 17(3):1863–1869, 2003.

[132] A.N. Phillips, S. Staszewski, R. Weber, O. Kirk, P. Francioli, V. Miller, P. Vernazza, J.D. Lundgren, and B. Ledergerber for the Swiss HIV Cohort Study and the Frankfurt HIV Clinic Cohort and the EuroSIDA Study Group. HIV viral load response to antiretroviral therapy according to the baseline CD4 cell count and viral load. *JAMA*, 286(20):2560–2567, 2001.

[133] W. Pierskalla. Optimal issuing policies in inventory management-I. *Management Science*, 13(5):395–412, 1967.

[134] W.P. Pierskalla. Inventory depletion management with stochastic field life functions. *Management Science*, 13(11):877–886, 1967.

[135] W.P. Pierskalla and J.A. Voelker. A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23:353–388, 1976.

[136] M.L. Pinedo. *Scheduling: Theory, Algorithms, and Systems (2nd Edition)*. Prentice Hall, Upper Saddle River, New Jersey, 2002.

[137] PLATO Collaboration. Predictors of trend in CD4-positive T-cell count and mortality among HIV-1 infected individuals with virological failure to all three antiretroviral-drug classes. *The Lancet*, 364(9428):51–62, 2004.

[138] E.L. Porteus. *Foundations of Stochastic Inventory Theory*. Stanford University Press, Stanford, California, 2002.

[139] K.E. Poundstone, R.E. Chaisson, and R.D. Moore. Differences in HIV disease progression by injection drug use and by sex in the era of highly active antiretroviral therapy. *AIDS*, 15(9):1115–1123, 2001.

[140] M.L. Puterman. *Markov Decision Processes*. John Wiley and Sons, New York, New York, 1994.

[141] S.P. Raffanti, J.S. Fusco, B.H. Sherrill, N.I. Hansen, A.C. Justice, R. D'Aquila, W.J. Mangialardi, and G.P. Fusco for the Collaborations in HIV Outcomes Research/United States Project. Effect of persistent moderate viremia on disease progression during HIV therapy. *JAIDS*, 37(1):1147–1154, 2004.

[142] D.A. Revicki, A.W. Wu, and M.I. Murray. Change in clinical status, health status, and health utility outcomes in HIV infected patients. *Medical Care*, 33(4 Suppl):AS173–AS182, 1995.

[143] A. Richter, B. Hauber, K. Simpson, J.A. Mauskopf, and D. Yin. A Monte Carlo simulation of AIDS treatment regimens. *Pharmacoeconomics*, 20(4):215–224, 2002.

[144] G.K. Robbins, V. De Gruttola, R.W. Shafer, LM. Smeaton, S.W. Snyder, C. Pettinelli, M. Dubé, M.A. Fischl, R.B. Pollard, R. Delapenha, L. Gedeon, C. van der Horst, R.L. Murphy, M.I. Becker, R.T. D'Aquila, S. Vella, T.C. Merigan, and M.S. Hirsch for the AIDS Clinical Trials Group 384 Team. Comparison of sequential three-drug regimens as initial therapy for HIV-1 infection. *New England Journal of Medicine*, 349(24):2293–2303, 2003.

[145] M.S. Roberts and F.A. Sonnenberg. Decision modeling techniques. In G.B. Chapman and F.A. Sonnenberg, editors, *Decision Making in Health Care: Theory, Psychology, and Applications*, pages 20–65. Cambridge University Press, 2000.

[146] S.M. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, San Diego, California, 1983.

[147] D.L. Sackett and J.C. Snow. The magnitude of compliance and noncompliance. In D.L. Sackett and R.B. Haynes, editors, *Compliance with Therapeutic Regimens*, pages 11–27. Johns Hopkins University Press, Baltimore, Maryland, 1976.

[148] G.D. Sanders, D.K. Owens, N. Padian, A.B. Cardinalli, A.N. Sullivan, and R.F. Nease. A computer-based interview to identify HIV risk behaviors and to assess patient preferences for HIV-related health states. In *Proceedings of the Annual Symposium on Computer Applications in Medical Care, Washington, DC*, pages 20–24. Hanley & Belfus Inc., Philadelphia, 1994.

[149] P.A. Scarf. On the application of mathematical models in maintenance. *European Journal of Operational Research*, 99(3):493–506, 1997.

[150] B.R. Schackman, K.A. Freedberg, M.C. Weinstein, P.E. Sax, E. Losina, H. Zhang, and S.J. Goldie. Cost-effectiveness implications of the timing of antiretroviral therapy in HIV-infected adults. *Archives of Internal Medicine*, 162(21):2478–2486, 2002.

[151] B.R. Schackman, S.J. Goldie, K.A. Freedberg, E. Losina, J. Brazier, and M.C. Weinstein. Comparison of health state utilities using community and patient preference weights derived from a survey of patients with HIV/AIDS. *Medical Decision Making*, 22(1):27–38, 2002.

[152] A.J. Schaefer, M.D. Bailey, S.M. Shechter, and M.S. Roberts. Modeling medical treatment using Markov decision processes. In M. Brandeau, F. Sainfort, and W. Pierskalla, editors, *Operations Research and Health Care: A Handbook of Methods and Applications*, pages 597–616. Kluwer Academic, 2004.

[153] P.A. Selwyn and M. Rivard. Palliative care for AIDS: Challenges and opportunities in the era of Highly Active Antiretroviral Therapy. *Journal of Palliative Medicine*, 6(3):475–487, 2003.

[154] P.P. Sendi, H.C. Bucher, B.A. Craig, D. Pfluger, and M. Battegay for the Swiss HIV Cohort Study. Estimating AIDS-free survival in a severely immunosuppressed asymptomatic HIV-infected population in the era of antiretroviral triple combination therapy. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 20(4):376–381, 1999.

[155] R.W. Shafer, L.M. Smeaton, G.K. Robbins, V. De Gruttola, S.W. Snyder, R.T. D'Aquila, V.A. Johnson, G.D. Morse, M.A. Nokta, A.I. Martinez, B.M. Gripshover, P. Kaul, R. Haubrich, M. Swingle, S.D. McCarty, S. Vella, M.S. Hirsch, and T.C. Merigan for the AIDS Clinical Trials Group 384 Team. Comparison of four-drug regimens and pairs of sequential three-drug regimens as initial therapy for HIV-1 infection. *New England Journal of Medicine*, 349(24):2304–2315, 2003.

[156] M. Shaked and J.G. Shanthikumar. *Stochastic Orders and Their Applications*. Academic Press, Inc., San Diego, California, 1994.

[157] S.M. Shechter, M.D. Bailey, and A.J. Schaefer. Replacing nonidentical vital components to extend system life. Under review, 2006.

[158] S.M. Shechter, M.D. Bailey, A.J. Schaefer, and M.S. Roberts. The optimal time to initiate HIV therapy under ordered health states. Under review, 2006.

[159] S.M. Shechter, C.C.H. Chang, R.S. Braithwaite, A.J. Schaefer, and M.S. Roberts. Estimating the expected remaining lifetime after initiating highly active antiretroviral therapy. Working paper, University of Pittsburgh, 2006.

[160] S.M. Shechter, C.C.H. Chang, R.S. Braithwaite, A.J. Schaefer, and M.S. Roberts. A Markov model of the natural history of HIV. Working paper, University of Pittsburgh, 2006.

[161] Y.S. Sherif and M.L. Smith. Optimal maintenance models for systems subject to failure - A review. *Naval Research Logistics Quarterly*, 28(1):47–74, 1981.

[162] M. Shernoff and R.A. Smith. HIV treatments: A history of scientific advance, 2001. Retrieved April 26, 2006, www.thebody.com/bp/jul01/treatments.html.

[163] N. Singh, C. Squier, C. Sivek, M. Wagener, M. H. Nguyen, and V. L. Yu. Determinants of compliance with antiretroviral therapy in patients with human immunodeficiency virus: Prospective assessment with implications for enhancing compliance. *AIDS Care*, 8(3):261–269, 1996.

[164] F.A. Sonnenberg and J.R. Beck. Markov models in medical decision making: A practical guide. *Medical Decision Making*, 13(4):322–328, 1993.

[165] J. Stebbing, M. Nelson, C. Orkin, S. Mandalia, M. Bower, A. Pozniak, and B. Gazzard. A randomized trial to investigate the recycling of stavudine and didanosine with and without hydroxyurea in salvage therapy RESTART. *Journal of Antimicrobial Chemotherapy*, 53(3):501–505, 2004.

[166] T.R. Sterling. Antiretroviral therapy: When to start & what to start with. *The Hopkins HIV Report*, 14(5):1, 2002.

[167] T.R. Sterling, R.E. Chaisson, J. Keruly, and R.D. Moore. Improved outcomes with earlier initiation of highly active antiretroviral therapy among Human Immunodeficiency Virus-infected patients who achieve durable virologic suppression: Longer follow-up of an observational chort study. *Journal of Infectious Diseases*, 188(11):1659–1665, 2003.

[168] T.R. Sterling, R.E. Chaisson, and R.D. Moore. Initiation of highly active antiretroviral therapy at CD4 T lymphocyte counts of >350 cells/mm3: Disease progression, treatment durability, and drug toxicity. *Clinical Infectious Diseases*, 36(6):812–815, 2003.

[169] G.J. Stine. *AIDS Update, 2003: An Annual Overview of Acquired Immune Deficiency Syndrome*. Prentice Hall, Upper Saddle River, New Jersey, 2003.

[170] K.T. Tashima, J.W. Hogan, L.I. Gardner, C. Korkontzelou, E.E. Schoenbaum, P. Schuman, A. Rompalo, and C.C.J. Carpenter. A longitudinal analysis of hospitalization and emergency department use among human immunodeficiency virus-infected women reporting protease inhibitor use. *Clinical Infectious Diseases*, 33(12):2055–2060, 2001.

[171] P. Tebas, K. Henry, R. Nease, R. Murphy, J. Phair, and W.G. Powderly. Timing of antiretroviral therapy. Use of Markov modeling and decision analysis to evaluate long-term implications of therapy. *AIDS*, 15(5):591–599, 2001.

[172] T.O. Tengs and T.H. Lin. A meta-analyses of utility estimates for HIV/AIDS. *Medical Decision Making*, 22(16):475–481, 2002.

[173] T.O. Tengs and A. Wallace. One thousand health-related quality-of-life estimates. *Medical Care*, 38(6):583–637, 2000.

[174] J. Tsevat, S.N. Sherman, J.A. McElwee, K.L. Mandell, L.A. Simbartl, F.A. Sonnenberg, and F.J. Fowler Jr. The will to live among HIV-infected patients. *Annals of Internal Medicine*, 131(3):194–198, 1999.

[175] J. Tsevat, J.G. Solzan, K.M. Kuntz, J. Ragland, J.S. Currier, R.L. Sell, and M.C. Weinstein. Health values of patients infected with human immunodeficiency virus: Relationship to mental health and physical functioning. *Medical Care*, 34(1):44–57, 1996.

[176] UNAIDS. UNAIDS/WHO AIDS epidemic update: December 2005. Retrieved April 19, 2006, http://www.unaids.org/epi/2005/doc/report.asp.

[177] C. Valdez-Flores and R.M. Feldman. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, 36(4):419–446, 1989.

[178] W.N. Venables and D.M. Smith and the R Development Core Team. *An Introduction to R*. Network Theory Limited, Bristol, United Kingdom, 2004.

[179] Veterans Aging Cohort Study, 2005. Retrieved December 6, 2005, http://www.vacohort.org/.

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

[181] L.M. Wein, R.M. D'Amato, and A.S. Perelson. Mathematical analysis of antiretroviral therapy aimed at HIV-1 eradication or maintenance of low viral loads. *Journal of Theoretical Biology*, 192(1):81–98, 1998.

[182] L.M. Wein, S.A. Zenios, and M.A. Nowak. Dynamic multidrug therapies for HIV: A control theoretic approach. *Journal of Theoretical Biology*, 185:15–29, 1997.

[183] M.C. Weinstein and W.B. Stason. Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine*, 296(13):716–721, 1977.

[184] N.J. Welton and A.E. Ades. Estimation of Markov chain transition probabilities and rates from fully and partially observed data: Uncertainty progagation, evidence synthesis, and model calibration. *Medical Decision Making*, 25(6):633–645, 2005.

[185] A.W. Wu, W.C. Matthews, L.T. Brysk, J.H. Atkinson, I. Grant, I. Abramson, C.J. Kennedy, J.A. McCutchan, S.A. Spector, and D.D. Richman. Quality of life in a placebo-controlled trial of zidovudine in patients with AIDS and AIDS-related complex. *Journal of Acquired Immune Deficiency Syndromes*, 3(7):683–690, 1990.

[186] A.W. Wu, H.R. Rubin, W.C. Mathews, J.E. Ware Jr., L.T. Brysk, W.D. Hardy, S.A. Bozzette, S.A. Spector, and D.D. Richman. A health status questionnaire using 30 items from the Medical Outcomes Study. Preliminary validation in persons with early HIV infection. *Medical Care*, 29(8):786–798, 1991.

[187] C.H. Wu, M.E. Lewis, and M. Veatch. Dynamic allocation of reconfigurable resources in a two-stage queueing system with reliability considerations. *IEEE Transactions on Automatic Control*, 51(2):309–314, 2006.

[188] P.W. Zehna. Inventory depletion policies. In K.J. Arrow, S. Karlin, and H. Scarf, editors, *Studies in Applied Probability and Management Science*, pages 263–287. Stanford University Press, Stanford, California, 1962.

[189] M. Zelen. Optimal scheduling of examinations for the early detection of disease. *Biometrika*, 80(2):279–293, 1993.