The Optimal Time to Initiate HIV Therapy under Ordered Health States

Steven M. Shechter
Operations and Logistics Division, Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, British Columbia V6T 1Z2, Canada, steven.shechter@sauder.ubc.ca,

Matthew D. Bailey
Department of Industrial Engineering, University of Pittsburgh, 1048 Benedum Hall, Pittsburgh, Pennsylvania 15261, USA, mdbailey@pitt.edu,

Andrew J. Schaefer
Department of Industrial Engineering, Department of Medicine, University of Pittsburgh, 1048 Benedum Hall, Pittsburgh, Pennsylvania 15261, USA, schaefer@ie.pitt.edu,

Mark S. Roberts
Section of Decision Sciences and Clinical Systems Modeling, Department of Medicine, Department of Health Policy and Management, Department of Industrial Engineering, University of Pittsburgh, 200 Meyran Avenue, 2nd Floor, Pittsburgh, Pennsylvania 15213, USA, robertsm@max.upmc.edu

The question of when to initiate HIV treatment is considered the most important question in HIV care today. 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 use Markov decision processes to develop the first HIV optimization models that aim to maximize the expected lifetime or quality-adjusted lifetime of a patient. We prove conditions that establish structural properties of the optimal solution and compare them to our data and results. Model solutions, based on clinical data, support a strategy of treating HIV earlier in its course as opposed to recent trends toward treating it later.

Subject classifications: dynamic programming; Markov decision process; health care treatment; HIV/AIDS; stochastic modeling applications.
Area of review: Policy Modeling and Public Sector OR

1. 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 as of the end of 2005. In that year alone, 5 million people became infected with HIV while 3 million people died of AIDS. Unfortunately, the number of people living with HIV continues to grow (UNAIDS 2005).

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 (Stine 2003). Antiretroviral therapy involves the administration of antiretroviral drugs, which are designed to inhibit HIV replication so as to preserve the vital CD4 cells. Because HIV rapidly develops resistance to a single drug (Shernoff and Smith 2001), the standard of care in high-income countries is to administer highly active antiretroviral therapy (HAART, also referred to as “cocktail therapy”), which is the simultaneous use of three or more antiretroviral drugs to slow the development of resistance. HAART has led to significant reductions in HIV-related morbidity, mortality, and health care utilization (Palella Jr. et al. 1998, Tashima et al. 2001). However, despite considerable advances in HIV treatments, there is still debate about the optimal way to use them.

The best time to start a patient on HAART is an open question (Ahdieh-Grant et al. 2003, Cohen and Boyle 2004, Harrington and Carpenter 2000, Ho 1995, Hoffman and Mulcahy 2005, Holmberg et al. 2004, Jeffrey et al. 2003, Lepri et al. 2001, Mauskopf et al. 2005, O’Shaughnessy et al. 2000, Phillips et al. 2003, Schackman et al. 2002a, Tebas et al. 2001). 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” (Hoffman and Mulcahy 2005). 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 the need to find another effective therapy, and preserving a limited number of treatment options (National Institutes of Health 2005). 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 (National Institutes of Health 2005).

Decisions involving the initiation of HAART are currently based on clinical judgment and national guidelines for HIV care, which are influenced by the outcomes of clinical studies and expert opinion (National Institutes of Health 2005). The strategy for initiating HAART has changed over the years. In the late 1990s, the treatment paradigm was to “hit hard, hit early” (Ho 1995); however, that approach has since been challenged and current guidelines suggest delaying the initiation of HAART until the CD4 count (the number of CD4 cells per mm$^3$ of blood) falls between 200 and 350 (National Institutes of Health 2005). Therefore, the new paradigm may be to “hit hard, but only when necessary” (Harrington and Carpenter 2000). Clearly, the best time to initiate HAART is unresolved and evolving.

There have been a number of clinical studies that consider the optimal time to initiate therapy, such as observational cohort studies (Ahdieh-Grant et al. 2003, Anastos et al. 2002, Chaisson et al. 2000, Egger et al. 2002, Hogg et al. 2001, Lepri et al. 2001, O’Shaughnessy et al. 2000, Palella
et al. 2003, Phillips et al. 2001, Sterling et al. 2003b) and a randomized clinical trial (Community Programs for Clinical Research on AIDS 2002). Clinical studies are certainly necessary and final tests of the effectiveness of a proposed change to treatment guidelines; however, they are not practical for narrowing the field of potential policies to test. On the other hand, mathematical approaches have the benefit of evaluating many treatment strategies with little time, cost, or risk to patients. For example, Monte Carlo simulation models have been useful for evaluating various outcome measures based on preset strategies for initiating therapy (Braithwaite et al. 2006, Freedberg et al. 2001, Richter et al. 2002). They are limited, however, as means of seeking optimal treatment strategies, as it is necessary to run simulations for each potential treatment policy. Furthermore, because simulation results are point estimates rather than analytically derived solutions as obtained in an MDP, choosing the policy that yields the best outcome is subject to a probability of being the incorrect choice (see Law and Kelton (2000) for further discussion of this topic). Some theoretical studies sought to determine the starting time that maximizes the duration for which the CD4 count remains above a certain level (Berman 1994, Berman and Dubin 1992); however, these were performed in the pre-HAART treatment era and did not consider the possibility of patient death nor model quality of life. These latter limitations also apply to other HAART-era control-theoretic models which aimed to control the viral load (the amount of HIV virus in the blood) within certain ranges (Alvarez-Ramirez et al. 2000, Brandt and Chen 2001, Butler et al. 1997, Fister et al. 1998, Jeffrey et al. 2003, Joshi 2002, Kirschner et al. 1997, Kirschner and Webb 1996, Wein et al. 1997, 1998). Furthermore, these models did not consider stochastic progression of patient health.

2. Contributions

The purpose of this paper is to address the question of the optimal time to initiate HIV therapy, with a goal of maximizing a patient’s total expected lifetime or quality-adjusted lifetime. To our knowledge, ours is the first HIV optimization model to consider these objectives. We accomplish this by developing Markov decision process (MDP) models that capture essential aspects of HIV progression and treatment. Generally, an MDP 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 (Puterman 1994).
HIV therapy planning fits this framework. A physician sees an HIV patient periodically, observing the patient’s state of HIV through laboratory measurements of prognostic variables such as CD4 count or viral load. Based on these measurements, the physician decides (with the patient) whether to initiate therapy or wait 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 approach to the question of when to initiate therapy. First, MDPs are discrete-time stochastic dynamic programs, 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, 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. We feel that a model in which patients cannot die, which is an implicit assumption made in the control-theoretic models previously mentioned, lacks face validity. Third, patient variables such as CD4 count are known to vary considerably day to day (Stine 2003), so it is important to use a framework that considers the stochastic progression of the disease. 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 about how certain structure on the input may guarantee certain structure on the optimal solution. We refer the reader to Puterman (1994) for extensive coverage of MDPs and references to other literature on the topic. For a review of MDPs applied to health care, see Schaefer et al. (2004).

The rest of the paper is organized as follows. In Section 3, we describe the modeling framework of our problem. Section 4 shows structural properties that exist under different conditions. In Section 5, we discuss the use of clinical data to develop the key components of the MDP: a Markov model of the progression of CD4 count and patient survival prior to initiating therapy, along with the estimation of remaining patient lifetimes after initiating therapy. We then discuss the solution of the MDP under these models and various sensitivity analyses. Finally, we present conclusions, limitations, and future work in Section 7.
3. Modeling framework

We consider a patient in a chronic stage of HIV who must decide (with his or her physician) when to initiate therapy. We assume the patient visits a physician periodically, and at each visit, a measurement of the patient’s health is taken. The frequency of these visits may depend on the physician as well as the stage of HIV, but we assume they occur monthly (Freedberg et al. also considered a model with monthly cycles to reflect the setting for HIV care (Freedberg et al. 1998)). Because CD4 count is arguably the most important factor in considering when to initiate therapy (National Institutes of Health 2005, Sterling 2002), we develop the MDP as a function of this variable alone. We acknowledge that there are additional clinical characteristics that may aid in predicting patient outcomes (such as viral load); however, such a multidimensional state space model loses ordering properties and requires more intricate analysis beyond the scope of this paper. Although further patient characteristics could refine our policy recommendations, we feel that CD4 captures a sufficient amount of the patients’ risk and disease progression while providing a computationally tractable model. Based on the CD4 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 in the various health states. If the patient waits, a reward is accrued between physician visits, and the patient transitions to another health 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 health state. Our objective is to maximize the total expected lifetime reward for the patient.

Formally, the components of the MDP are as follows:

\[ T = \{1, 2, \ldots, \infty\} \]: 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 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 the 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 construction, \( v^*(0) = 0 \).

Our setup fits the framework of a stochastic longest path problem (Bertsekas and Tsitsiklis 1996), 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 (Bellman 1957, Puterman 1994):

\[
\begin{align*}
  v(s) &= \max \left\{ r(s) + \sum_{j \in S} p(j|s)v(j), \ R(s) \right\} \quad \text{for all } s \in \{1, 2, \ldots, N\}, \text{ and} \\
  v(0) &= 0.
\end{align*}
\]

(1)

For example, if the objective is to maximize a patient’s total expected lifetime, then we let \( r(s) = 1 \) month; \( R(s) \equiv L(s) \) = 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:

\[
\begin{align*}
  v(s) &= \max \left\{ 1 + \sum_{j \in S} p(j|s)v(j), \ L(s) \right\} \quad \text{for all } s \in \{1, 2, \ldots, N\}, \text{ and} \\
  v(0) &= 0.
\end{align*}
\]

(2)

An issue that often arises in the medical decision making literature is that of discounting future health outcomes (Drummond et al. 1997, Gold et al. 1996). 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) (Keeler and Cretin 1983, Weinstein and Stason 1977). 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 (Puterman 1994). 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 (Puterman 1994). 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 (Ross 1983). 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 (Puterman 1994, Ross 1983). 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 (Puterman 1994). 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. Structural properties

Before proceeding with solving the MDP based on clinical data, 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 for large-scale models. For further discussion of computational considerations, we refer the reader to Shechter (2006).

We first state the following definition (Barlow and Proschan 1965):

**Definition 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

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

is nondecreasing in $i$ for all $s \in \{0, \ldots, N\}$.

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).

Now consider the following assumptions:

- **(As1)** $r(s)$ and $R(s)$ are both nonnegative and nondecreasing in $s$.
- **(As2)** $P$ is IFR.

In the following intuitively appealing theorem, we show that lower states are “worse,” and higher health states are “better”. The proofs of all results are contained in the Appendix.
Theorem 1. Under assumptions \( A_{s1} \) and \( A_{s2} \), \( v^*(s) \) 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 5, our data satisfy this condition.

Theorem 2. \( a^*(s) = I \) for all \( s \) if and only if:

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

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

4.1. Patient-specific considerations

While national guidelines policies are based on studies from entire cohorts of patients, a case can be made for more patient-focused care (Henry 2000). Towards that end, we incorporate two important patient factors for deciding when to initiate HIV therapy: quality of life and adherence.

4.1.1. 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 (Chesney 2003, Duran et al. 2001, Lenert et al. 2002). 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 (AIDS Education Global Information System 2006). Tengs and Lin (2002) provide a meta-analysis of studies eliciting patient utilities for different stages of HIV. Estimates by CD4 strata can be found in Freedberg et al. (1998) and Schackman et al. (2002b). However, none of the papers we reviewed (Bayoumi and Redelmeir 1998, 1999, Cleary et al. 1993, Copfer et al. 1996, Freedberg et al. 1998, Gelber et al. 1992, Holtgrave and Qualis 1995, Lenert et al. 2002, Mrus et al. 2006, Owens et al. 1997, Paltiel et al. 1998, Revicki et al. 1995, Sanders et al. 1994, Schackman et al. 2002b, Tsevat et al. 1999, 1996, Wu et al. 1990, 1991) presents 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. (2002) 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.

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 \(s\). Applying these utilities to Bellman’s equations that maximize expected remaining lifetime (2) yields Bellman’s equations for the problem of maximizing quality-adjusted lifetime:

\[
v_u(s) = \max \left\{ u_w(s) + \sum_j p(j|s)v_u(j), \quad u_i(s)L(s) \right\} \quad \text{for all } s \in \{1, 2, \ldots, N\},
\]

where \(v_u(s)\) represents the expected remaining quality-adjusted lifetime from state \(s\), and \(L(s)\) is the expected remaining (unadjusted) life months after initiating therapy from state \(s\). In Section 5, we solve our MDP by performing sensitivity analyses around CD4-based and therapy-based utility estimates.

4.1.2. 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 (National Institutes of Health 2005). Various studies have demonstrated associations between lower adherence and worse outcomes such as higher viral load (Haubrich et al. 1999), lower CD4 count (Paterson et al. 2000), higher incidence of AIDS (Bangsberg et al. 2001), more days in the hospital (Paterson et al. 2000), and higher mortality rates (Carmona et al. 2000). Though by most conventions taking at least 80% of prescribed medication implies compliance with the therapy, one study exposed significant differences in outcomes even for adherence levels that differ in the 80-100% range (Paterson et al. 2000). However, adherence rates are often lower than this as various studies have reported 40-50% of patients taking less than 80% of their medication (Bennett et al. 1998, Eldred et al. 1998, Gir et al. 1998).

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 later.

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:

\[
v_a(s|m) = \max \left\{ 1 + \sum_{j \in S} p(j|s)v_a(j|m), \ mL(s) \right\} \quad \text{for all } s \in \{1, 2, ..., N\}, \quad \text{and (3)}
\]

where \( v_a(s|m) \) 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 from all states (i.e., the condition of Theorem 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:

\[
\Delta_s \equiv L(s) - [1 + \sum_{j \in S} p(j|s)L(j)] \quad \text{for all } s \in S,
\]

\[
\Delta_{\text{min}} \equiv \min_{s \in S} \Delta_s, \text{ and } \Delta_{\text{max}} \equiv \max_{s \in S} \Delta_s.
\]

In words, \( \Delta_s \) 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 3. For every state $s$ such that $1/m > 1 + \Delta_s$, it is uniquely optimal to wait.

It follows that if $1/m > 1 + \Delta_{\text{max}}$, 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 $s_{\text{min}}$ be a state that minimizes $\Delta_s$.

COROLLARY 2. $a^*(s_{\text{min}}) = W$, uniquely, if and only if $1/m > 1 + \Delta_{\text{min}}$.

Note that by Theorem 2, if the optimal policy in the expected lifetime problem without explicit consideration of patient adherence is to initiate therapy in every state, then $\Delta_{\text{min}} > 0$. In that case, it follows from Corollary 2 that a necessary condition for waiting to be uniquely optimal in state $s_{\text{min}}$ 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 intuitive result, which states that patients with greater levels of adherence have greater expected remaining lifetimes.

THEOREM 4. $v_a(s|m)$ is nondecreasing in $m$ for each $s$.

We note that other structural results, including conditions leading to optimal control-limit policies, can be found in Shechter (2006).

5. Implementation

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. (Veterans Aging Cohort Study 2005). 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 Section 7.

We focus our analysis on patients between 40 and 50 years old, as this ten year age bracket has the largest number of patients in the cohort. Also, we categorize the CD4 count into four distinct categories: $0 - 49$, $50 - 199$, $200 - 349$, and $\geq 350$. These strata can be found in other clinical studies and allow for results to be interpreted in the context of the guidelines categories of $<200$, $200-350$, and $>350$ (National Institutes of Health 2005).
5.1. Estimation of the Transition Probability Matrix

Although there is a general decline in CD4 count prior to initiating HAART, our data demonstrate a stochastic deterioration that includes upward fluctuations. Therefore, instead of considering just the mean decrease in CD4 counts over time (Mellors et al. 1997), we build a discrete-time Markov model that describes the monthly transition probabilities among the various CD4 categories and death, prior to initiating 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 (Beck and Pauker 1983, Roberts and Sonnenberg 2000, Sonnenberg and Beck 1993). 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 (Craig et al. 1999, Craig and Sendi 2002, Gentleman et al. 1994, Isaman et al. 2006, Welton and Ades 2005). Details of other discrete-time Markov models of CD4 progression can be found in (Aalen et al. 1997, Brundage et al. 1990, Freedberg et al. 1998, Longini 1990, Longini et al. 1991, Mauskopf et al. 2005, Sendi et al. 1999). Because our model also requires estimates of expected remaining lifetime upon initiating therapy (discussed in the next section), we develop both parts of the model from a single data source (VACS).

To build a stochastic model of the natural history of a patient’s CD4 count (progression prior to initiating therapy), we first fit curves to each patient’s sequence of CD4 measurements prior to initiating therapy. Patients with only one observation are eliminated, patients with two or three observations are fitted with a least-squares linear model, and patients with four or more observations are fitted using a cubic smoothing spline. A thorough treatment of smoothing splines can be found in Green and Silverman (1994). Briefly, 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 (the linear regression through the points being the smoothest of all splines). Because of the significant variability in a patient’s CD4 measurements over time, we use the splines that maximally capture the CD4 fluctuations; that is, we use the roughest of the splines, which goes through every data point (also referred to as the “interpolating spline”). A more detailed discussion of our use of smoothing splines can be found in Shechter (2006).

After fitting the curves, we partition the curves into monthly intervals and record the estimated CD4 counts at each interval. After categorizing the CD4 counts into the four groups mentioned above, we generate all the pairs of observations (CD4 category at the start of this month, CD4
category at the start of next month) represented in each curve across all the patients, count
the number of transitions from each category to the other categories, and construct a transition
probability matrix of movements between CD4 categories and death. A similar approach has been
used to model the progression of end-stage liver disease in the U.S. before patients receive a
transplant (Alagoz et al. 2005).

A challenge in estimating the transition probabilities is that as CD4 counts fall, patients eventu-
ally initiate therapy, thus censoring our natural history observations. In other words, as patients
get sicker and begin therapy, we lose the ability to see how their condition would have progressed
without therapy. Rarely does a patient die while in the pre-therapy phase of HIV, so ignoring
this issue would significantly underestimate patient risk of death. We considered different scenarios
for how these censored observations would have been distributed to death and the different CD4
categories had therapy not been initiated. More details on this can be found in Shechter (2006).

5.2. Estimation of Expected Remaining Lifetime after Initiating Therapy

Several authors have compared survival rates across different CD4 categories after patients initiated
et al. 2001, Sterling et al. 2003a). 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 are many patients still alive on therapy since starting it in the late
1990s. But since 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 mortality risks the 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 (Braithwaite et al. 2006, Freedberg et al. 2001,
King et al. 2003, Mauskopf et al. 2005). For example, King et al. (2003) report median lifetimes
of 5.5, 8.5, and 15.4 years for patients initiating HAART from CD4 categories ≤ 50, ≤ 200, and
> 200, respectively. We take a statistical approach. First, we use standard Cox proportional hazards
models (Cox 1972) to calculate hazard ratios between male VACS HIV+ patients and HIV- controls
(further controlling for age and race) across the various CD4 categories. 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, or in this case, between two
different types of VACS populations (Collett 1994). We then apply these hazard ratios to standard
life table data and use a life table construction method (Anderson 1999) to estimate the expected remaining lifetime for 45-year-old men. By working with life table estimates of remaining lifetimes, we consider the aging of the cohort. Table 1 shows the results, with an expected remaining lifetime 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 ≥ 350.

Note that the estimates of remaining survival 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. These estimates capture the effects of partial adherence and viral resistance as they are represented in patients who initiate therapy. In other words, we take a modeling perspective of determining the optimal time to initiate therapy as it is administered today, with the associated downstream decisions and effects. We leave for future work optimal switching and sequencing decisions after initiating therapy. Also note that we 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.

### 5.3. Results

As discussed in Section 5.1, we considered various scenarios for handling the censored observations in the development of the transition probability matrix. In the present analysis, we shall use three of those scenarios: a baseline weighting of redistributing censored observations towards death, a heavier weighting, and a lighter weighting (we refer the reader to Shechter (2006) for more detail on these scenarios). Table 1 of Section 5.2 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 2 gives the optimal policy along with the optimal value vector (in life years) for each combination of the natural history and survival models, under an objective of maximizing expected
Table 2  Optimal policy and value vector under various combinations of the natural history and survival models

<table>
<thead>
<tr>
<th>NH model</th>
<th>Survival Model</th>
<th>Output</th>
<th>0-49</th>
<th>50-199</th>
<th>200-349</th>
<th>≥350</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline</td>
<td>baseline</td>
<td>Optimal Policy</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>6.34</td>
<td>11.61</td>
<td>17.25</td>
<td>24.79</td>
</tr>
<tr>
<td>baseline</td>
<td>110%</td>
<td>Optimal Policy</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>6.97</td>
<td>12.77</td>
<td>18.98</td>
<td>27.27</td>
</tr>
<tr>
<td>baseline</td>
<td>90%</td>
<td>Optimal Policy</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>5.71</td>
<td>10.45</td>
<td>15.53</td>
<td>22.31</td>
</tr>
<tr>
<td>lighter</td>
<td>baseline</td>
<td>Optimal Policy</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>6.38</td>
<td>11.70</td>
<td>17.25</td>
<td>24.79</td>
</tr>
<tr>
<td>lighter</td>
<td>110%</td>
<td>Optimal Policy</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>6.97</td>
<td>12.77</td>
<td>18.98</td>
<td>27.27</td>
</tr>
<tr>
<td>lighter</td>
<td>90%</td>
<td>Optimal Policy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimal Values</td>
<td>5.97</td>
<td>10.76</td>
<td>15.56</td>
<td>22.31</td>
</tr>
</tbody>
</table>

lifetime. For the optimal policy vector, a value of “1” indicates to initiate therapy from that CD4 category, while a value of “0” 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 \( A_1 \) and \( A_2 \) 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 2, and hence, we obtain an optimal policy of initiating therapy from each of the CD4 categories. 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. For example, for the baseline survival model, the values associated with CD4 categories \(< 50 \) and \( 50 - 199 \), 6.38 and 11.70 life years, are similar to the values 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.

Our natural history model depends on the choice of the degrees of freedom (\( df \)), which can be
viewed as a smoothing parameter. As indicated in Section 5.1, we chose the parameter yielding the roughest curve: i.e., the one that maximally captures the data fluctuations. We also ran the model using the average spline instead (the spline with \( df \) half way between the value yielding the smoothest and roughest splines). With this change, every combination of natural history and survival model represented in Table 2 results in an optimal policy of initiating therapy from every CD4 category (Shechter 2006).

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 4.1.1, 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 Schackman et al. (2002b) and Freedberg et al. (1998) to estimate off-HAART utilities according to our CD4 categories (Table 3). 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 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 4 shows the on-HAART utilities associated with three different multiplicative factors of the off-therapy utilities (.9, .7, and .5).

Table 5 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. (2002b)) or .29 (for the set of utilities in Freedberg et al. (1998)) that we obtain a solution for which it is optimal to wait in some CD4 category. Using the utilities in Schackman et al. (2002b) for being off HAART and applying the .34 multiplier leads to on-therapy utility estimates of .30, .31, .33, and .33 for the CD4 categories 0−49, 50−199, 200−349, and ≥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 (Tengs and Wallace 2000). 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 (Tengs and Wallace 2000). Also, as discussed in Section 4.1.1, Lenert et al. (2002) 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 (Braithwaite 2006). 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.

6. Opportunities for Future OR Modeling in HIV Treatment

Because HIV treatment is such an important public health issue, this section outlines various research directions to expand upon the presented model. Broadly, the directions include more
detailed description of patient health and an examination of the Markovian assumption (Section 6.1); a more detailed consideration of patient adherence and viral evolution in response to therapy (Section 6.2); scheduling HAART beyond its initiation, including the sequencing and timing of therapy switches (Section 6.3); and the effect of newly developed therapies on optimal HAART strategies (Section 6.4). We believe that these areas will provide fruitful operations research opportunities for the foreseeable future as data evolve to better support the modeling of these extensions.

6.1. Increased Physiological Detail

Although CD4 count is the primary driver of decisions of when to initiate therapy, viral load is also considered in national guidelines. Shechter (2006) developed an MDP model, along with structural properties, of this two-variable state space. However, he did not calibrate the model with data, which would involve creating cubic splines for viral load and looking at the joint transitions of CD4 and viral load among their respective categories prior to treatment initiation. Furthermore, such a model would also need to estimate remaining survival as a function of both the patient’s CD4 and viral load at the time therapy is initiated. Hemoglobin level, another laboratory measurement that is thought to have prognostic value (independent of CD4 and viral load) (Justice et al. 2002), could be included in the model as well.

Our model assumed that the transition probability matrix is Markovian with respect to the CD4-based health state. An early 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 (De Gruttola et al. 1993); however, there is some evidence that health history may influence the transitions to other living health states (Dal Maso et al. 2000). Therefore, further research could explore the Markovian assumption. Our state space is small enough that there would be no computational difficulty to include some of the CD4 history. However, 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 NIH guidelines consider current laboratory values alone rather than including prior values.

6.2. Further Consideration of Adherence and Resistance

We mentioned in Section 5.2 that our model implicitly captures patients’ partial adherence to therapy and the development of viral resistance to the extent these effects are represented in the survival of patients in our dataset. However, it is plausible that the hazard ratio of HIV+ to HIV- patients in our less than 10 years of survival data increases over time, rather than stays constant
(which we assumed in obtaining our estimates of expected lifetimes). For example, if many patients experience viral resistance to all therapeutic options beyond the time limits of our data, or if patient adherence wanes over time, then our estimates of survival after initiating therapy should be lower. Therefore, further work may be spent on validating these drivers of our survival estimates. We note, however, that the significant decrements to quality of life for patients on therapy (presented in Table 5) can be viewed more generally as a sensitivity analysis on the survival estimates upon initiating therapy. For example, in Scenario 3 of Table 5, we took our baseline survival estimates and multiplied them by .5 to represent decreased quality of life. But the .5 could also represent downward pressure on the survival estimates resulting from the resistance and adherence issues just discussed. On the other hand, there would be upward pressure on the survival estimates once we incorporate the optimal downstream decisions of the timing and sequencing of therapy switches (see Section 6.3).

In Section 4.1.2, we considered a model that explicitly incorporates a particular patient’s likelihood to adhere better or worse than the typical patient in the cohort, and we evaluated it under the objective of maximizing expected lifetime. Another avenue for enhancing the clinical validity of our model is to consider adherence under an objective of maximizing expected quality-adjusted lifetime. Such a model would require understanding how patients’ utilities affect their adherence levels, which in turn affect their length and quality of life. Modeling such dynamics remains a challenge. However, such consideration may lead to patient-specific policies that recommend delaying therapy for certain states.

6.3. Scheduling HAART after Initiation

As indicated in Section 5.2, our estimates of survival are implicitly based on the current guidelines for sequencing and switching therapies after initiating therapy. An interesting area of further work is to consider optimizing the downstream decisions of sequencing and timing therapy switches, which would affect the survival estimates used in our model. Shechter (2006) explored various conceptual frameworks for these downstream scheduling decisions, but available data do not yet permit the calibration of empirical models. Moreover, the validity of the optimal scheduling model will require understanding the adherence and resistance dynamics associated with different sequences, as described in the previous section. We are unaware of any study that quantitatively models the effects of particular drugs in particular sequences.
6.4. The Effect of Newly Developed Therapies on Optimal Initiation Strategy

Another avenue for future research is to consider the possibility that some patients in the early stages of disease progression may delay the start of therapy with the hope that more potent therapies with fewer side-effects and toxicities and higher barriers to viral resistance may become available. Given the evolution of therapeutic options in just the past 10 years, this may be a defensible argument for delaying therapy. We could expand our model to consider the probability of new therapies becoming available, though much work would be needed to decide how best to forecast these events.

7. 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 chronic, yet manageable condition (Selwyn and Rivard 2003). 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 patients earlier in the course of disease as opposed to the more recent approach of treating them 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 2). Under the lighter redistribution of censors to death, we obtained some counter-intuitive optimal policies of waiting when CD4 counts are lower and initiating when they are higher. However, as discussed in Section 5.3, the values of those policies do not differ much from the policy of initiating therapy from every CD4 category. 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 NIH guidelines (National Institutes of Health 2005)). 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.

Although avoiding side effects of therapies is a common reason for delaying therapy, our results also support a policy of initiating therapy immediately under a variety of quality-of-life considerations. As shown in Table 5 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. (2002a) also found that it took a 70% decrease in on-therapy quality-of-life before a policy of delaying therapy (until CD4 falls below 200) appeared better than a policy of initiating therapy earlier (when CD4 falls below 350).

We note that our results of initiating therapy from each state are strengthened by 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.

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 $\geq 350$, respectively.

We captured the increased risk of death with age through our estimates of survival after initiating therapy (the terminal rewards), as described in Section 5.2. Additionally, we assumed that patients receive these terminal rewards regardless of when therapy is initiated. A time-based decrement to the terminal reward would require a significantly more complex model and significantly more data. As noted in the conclusions, however, our assumption biases 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 be careful in applying our results to non-VACS populations. However, we can generate easily the MDP components and solutions based on data from other cohorts.

Although our results are contrary to current treatment guidelines, other recent work also supports a strategy of treating HIV earlier in its course. For example, the study by Schackman et al. (2002a) discussed above suggests earlier treatment may be better. Also, Holmberg et al. (2004) 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. (2006) reported that concerns about toxicities associated with earlier treatment may be unfounded and suggested that earlier initiation of HAART may be better.
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 doctors periodically until the decision is made to initiate therapy, and there is uncertainty in how the patient’s health will progress. 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 support our results give to moving away from recent trends of initiating therapy later in the course of disease and to “hit earlier” instead.

Acknowledgments

This research was supported by a grant for Health Services Dissertation Research from the Agency for Healthcare Research and Quality (1 R36 HS15533-01), a grant from the National Institute of Alcohol Abuse and Alcoholism (1 K23 AA14483-01), and a grant from the National Science Foundation (CMI-0546960). We would also like to thank the Veterans Aging Cohort Study for giving us access to the data used in this paper, and we thank Scott Braithwaite, Jon Caulkins, Joyce Chang, David Paltiel, four anonymous referees, and one anonymous Associate Editor for their valuable insights and input.

Appendix. Proofs of Theorems

We will make use of the following lemma, which we state without proof (Puterman 1994):

**Lemma 1.** Suppose \( \{v_i\} \ (i=1,\ldots,N) \) is a sequence of numbers, and \( \{p_i\} \) and \( \{q_i\} \ (i=1,\ldots,N) \) are two discrete probability distributions such that
\[
\sum_{i=k}^{N} q_i \geq \sum_{i=k}^{N} p_i \quad \text{for all } k \in \{1,\ldots,N\}.
\]

Then, if \( \{v_i\} \) is nondecreasing (nonincreasing),
\[
\sum_{i=1}^{N} q_i v_i \geq (\leq) \sum_{i=1}^{N} p_i v_i.
\]

**Proof of Theorem 1**

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 \( As2 \) (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 \( As1 \) (i.e., \( r(s) \) is nonnegative and nondecreasing in \( s \) for all \( s \)), we have:

\[
r(s+1) + \sum_{j \in S} p(j|s+1)v^i(j) \geq r(s) + \sum_{j \in S} p(j|s)v^i(j) \geq 0 \quad \text{for all } s \in S. \tag{A1}
\]

We also know by assumption \( As1 \) that:

\[
R(s+1) \geq R(s) \geq 0 \quad \text{for all } s \in S. \tag{A2}
\]

Combining (A1) and (A2) yields:

\[
\max \left\{ r(s+1) + \sum_{j} p(j|s+1)v^i(j), \ R(s+1) \right\} \geq \max \left\{ r(s) + \sum_{j} p(j|s)v^i(j), \ R(s) \right\} \geq 0 \quad \text{for all } s \in S. \tag{A3}
\]

According to the value iteration algorithm,

\[
v^{i+1}(k) = \max \left\{ r(k) + \sum_{j} p(j|k)v^i(j), \ R(k) \right\} \quad \text{for all } k \in S. \tag{A5}
\]

Therefore, (A4) and (A5) 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 \( s \). \( \square \)

**Proof of Theorem 2**

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

\[
v(s) = \max \left\{ r(s) + \sum_{j} p(j|s)v(j), \ R(s) \right\} \quad \text{for all } s \in \{1,2,...,N\}, \text{ and } v(0) = 0,
\]

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 (1). Therefore, \( v(s) = R(s) \) is an optimal value function which is achieved by the policy \( a^*(s) = I \) for all \( s \). □

**Proof of Corollary 1**

Clearly \( v^*(s) \geq R(s) \) 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. □

**Proof of Theorem 3**

As a consequence of Corollary 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 \( \Delta_s \), the above is equivalent to:

\[ 1/m > 1 + \Delta_s. \] □

(A6)

**Proof of Corollary 2**

The sufficiency part of the proof follows directly from Theorem 3.

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 \( \Delta_s \), 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 2, this implies that \( a^*(s) = I \) for all \( s \), which contradicts our first assumption that \( a^*(s_{\min}) = W \), uniquely. Therefore, \( 1/m > 1 + \Delta_{\min} \). □

**Proof of Theorem 4** Let \( m_1 \leq m_2 \). We prove this by performing parallel iterations of the value iteration algorithm to solve for \( v_a(s|m_1) \) and \( v_a(s|m_2) \) for all \( s \). Suppose for some iteration, \( i \), of the algorithm \( v_a^i(s|m_1) \leq v_a^i(s|m_2) \) for all \( s \) (note that starting each problem with a vector of zeroes, satisfies this). Then for each \( s \geq 1 \) (the value associated with state \( s = 0 \) is always 0),

\[
 v_a^{i+1}(s|m_1) = \max \left\{ 1 + \sum_j p(j|s)v_a^i(j|m_1), \ m_1 L(s) \right\}, \quad \text{and} \quad (A7)
\]
\[
 v_a^{i+1}(s|m_2) = \max \left\{ 1 + \sum_j p(j|s)v_a^i(j|m_2), \ m_2 L(s) \right\}. \quad (A8)
\]

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

**References**


Braithwaite, R.S. 2006. Personal Correspondence.


Butler, S., D. Kirschner, S. Lenhart. 1997. Optimal control of chemotherapy affecting the infectivity of HIV. 


human immunodeficiency virus-infected women reporting protease inhibitor use. *Clinical Infectious Diseases* 33 2055–2060.


