Optimization of Prostate Biopsy Referral Decisions

Jingyu Zhang
Graduate Program in Operations Research, North Carolina State University,
Raleigh, NC 27695, jzhang2@ncsu.edu

Brian T. Denton
Edward P. Fitts Department of Industrial & Systems Engineering, NC State University,
Raleigh, NC 27695, bdenton@ncsu.edu

Hari Balasubramanian
Department of Mechanical & Industrial Engineering, University of Massachusetts,
Amherst, MA 01003, hbalasubraman@ecs.umass.edu

Nilay D. Shah
Division of Health Care Policy and Research, Mayo Clinic,
Rochester, MN 55905, Shah.Nilay@mayo.edu

Brant A. Inman
Department of Surgery, School of Medicine, Duke University,
Durham, NC 27710, brant.inman@ duke.edu

Abstract

Prostate cancer is the most common solid tumor in American men and is screened for using prostate-specific antigen (PSA) tests. We report on a non-stationary partially observable Markov decision process (POMDP) for prostate biopsy referral decisions. The core states are the patients’ prostate cancer related health states, and PSA test results are the observations. Transition probabilities and rewards are inferred from the Mayo Clinic Radical Prostatectomy Registry (MCRPR) and the medical literature. The objective of our model is to maximize expected quality-adjusted life years (QALYs). We solve the POMDP model to obtain an age and belief (probability of having prostate cancer) dependent optimal biopsy referral policy. We also prove a number of structural properties including the existence of a control-limit type policy for the biopsy referral decision. Our empirical results demonstrate a non-decreasing belief threshold in age, and we provide sufficient conditions under which PSA screening should be discontinued for older patients. Finally, the benefits of screening under the optimal biopsy referral policy are estimated, and sensitivity analysis is used to prioritize the model parameters that would benefit from additional data collection.
Keywords: partially observable Markov decision process, PSA screening, biopsy, control-limit policy, stopping time problem

1 Introduction

Prostate cancer is the most common solid tumor in American men. The American Cancer Society (2010) estimated that 217,730 new cases of prostate cancer would be diagnosed, and 32,050 deaths would occur in the United States in 2010. At its current level of prevalence, it is estimated that 16 percent of men will be diagnosed with prostate cancer during their lifetimes (National Cancer Institute 2009). While the direct health impact is felt by men, the indirect social effects are felt by all. Therefore prostate cancer is an important societal issue and preventive screening is an important consideration as part of the health service system.

Prostate cancer screening relies heavily on the prostate-specific antigen (PSA) test, a simple blood test that indicates the amount of PSA in the blood serum. PSA varies in a continuous range. Patients with higher than normal PSA have a greater risk of prostate cancer; however, higher than normal PSA levels may occur for other reasons including natural variation over time, prostatic infections and benign enlargement of the prostate gland. If a patient’s PSA test result is classified as “suspicious”, he may be referred for biopsy, which has a negligible false positive rate but a non-trivial false negative rate (about 20%). In addition to providing imperfect information, biopsies are painful and carry the possibility of side-effects. Therefore, from the patient’s perspective, it is necessary to decide how best to use these tests to trade off the impact on quality of life from screening, with the long-term potential benefits of early detection and treatment of prostate cancer.

When should a patient be referred for a biopsy? How should a patient’s age and PSA history influence the referral decision? Surprisingly, there has been very little research on determining optimal decisions related to these questions. In this article, we address these and other important questions related to prostate cancer screening. We focus on a population-based model and investigate biopsy referral policies which maximize a patient’s expected quality-adjusted life years (QALYs). Our model trades off the potential rewards from early detection and treatment of prostate cancer (additional QALYs) with the side effects of biopsy. To this end we formulate a partially observable Markov decision process (POMDP) model for the biopsy referral decision process. The
health states in our model are not directly observable, but can be probabilistically inferred from PSA test results and biopsy results. We use several data sources to calibrate our model including a large regional data set from Olmsted County in Minnesota.

We begin by presenting some relevant background on prostate cancer screening and a detailed mathematical formulation of our POMDP model. Next, we present several theorems illustrating generalizable insights into the structure of the optimal biopsy referral policy. For instance, we demonstrate under reasonable assumptions that the optimal biopsy referral decision follows a control-limit type policy with respect to a belief (probability of having prostate cancer) threshold. We show that the expected QALYs are nonincreasing with respect to belief which is an important factor in computing bounds on the optimal policy. We provide conditions under which there is a finite age at which screening should be discontinued. Finally, we present empirical results for the optimal age dependent biopsy referral policy based on a data set for a large population from Olmsted County, Minnesota. Sensitivity analysis is used to identify which parameters most significantly affect the optimal biopsy policy and to provide guidance about how to prioritize further data collection through patient surveys and randomized controlled trials (RCTs) to develop more detailed future versions of our model.

The remainder of this paper is organized as follows. Section 2 provides some background on prostate cancer. Section 3 reviews the related literature on prostate cancer screening and relevant literature on medical decision making in other contexts. Our POMDP model is described in Section 4, and in Section 5 a number of structural properties are presented. Section 6 provides a detailed description of the data used to populate our model, and presents numerical results illustrating the optimal policy and sensitivity analysis. In Section 7 we discuss the results, and in Section 8 we summarize the most significant conclusions that can be drawn from our study.

2 Prostate Cancer Background

Prostate cancer is a disease in which malignant cancer cells form in the prostate gland. Since prostate tumors progress slowly, and at early stages there are usually no physical symptoms, screening with the PSA blood test is common. This blood test quantifies the amount of PSA that escapes into the blood from the prostate, measured typically in ng/mL. Patients with higher than normal
PSA values have a greater risk of prostate cancer. However, a patient’s PSA varies in a continuous range, and higher than normal levels may occur for a variety of other reasons. As a result, the definition of a suspicious test result vs. a likely benign PSA test result is an open question. Figure 1 illustrates the imperfect nature of PSA testing using a receiver operating characteristic (ROC) curve based on our population data set described in Section 6.1. The sensitivity of a PSA test decreases as a function of the specificity as the PSA threshold increases.

Figure 1: An ROC curve that illustrates the imperfect nature of PSA tests for diagnosing prostate cancer. The different points on the curve correspond to different PSA thresholds used to distinguish a suspicious and likely benign test. The curve was generated using the dataset described in Section 6.1.

A typical screening process is illustrated in Figure 2 where the patient receives routine PSA tests at regular intervals (typically annually). If the test result is suspicious the patient is typically referred for biopsy, and if the biopsy indicates cancer the patient is referred for treatment.

The biopsy procedure is ultrasound-guided. Hollow needles are typically passed 12 times into the peripheral zone of the prostate gland to extract tissue. Each needle extricates a core of tissue to be analyzed by a pathologist. Biopsy is a fairly accurate detection method in the sense that the false positive rate is nearly zero. However, there is a significant probability of a false negative biopsy, since biopsy involves sampling only a small portion of the prostate. Although minimally invasive, biopsy is painful and carries non-negligible short and long-term risks for the patient.

The imperfect nature of the PSA test and biopsy, as well as the QALY decrements of biopsy and subsequent treatment, have raised questions about the most effective and efficient policies for prostate cancer screening. In some cases it is clearly not logical to perform such tests (e.g.,
screening a 90 year old male, since his probability of dying from prostate cancer is generally much smaller than his risk of death from other competing causes such as heart disease). Furthermore, the imperfect sensitivity and specificity of the test illustrated in Figure 1 has led to criticism of the use of the PSA test for population screening (Holmstrom et al. 2009) and the concern of over-diagnosis (Etzioni et al. 2002, Welch and Black 2010). The definition of a suspicious PSA test result, and the decision to refer a patient for biopsy must trade off the potential benefits from early detection with the side effects of biopsy and subsequent treatment.

Once detected, there are multiple options for treating prostate cancer. One option is radical prostatectomy (surgical removal of the prostate gland) which is one of the most common forms of treatment in the U.S. (Welch and Albertsen 2009). Other treatment options include active surveillance (monitoring prostate cancer progression through regular biopsies), brachytherapy (implantation of radioactive seeds in the prostate), and external beam radiation therapy. All of these treatment options can have serious side effects (e.g. urinary, sexual and gastrointestinal dysfunction) which impact the patient’s future long-term quality of life. In the case of active surveillance, which has become more common in recent years, the patient may be subjected to hormone therapy and multiple future biopsies.

In summary, there are many decisions involved in the design of screening and treatment policies for prostate cancer. These are complicated by the probabilistic progression of prostate cancer, and the imperfect nature of the tests used to detect it. There have been recent advances in data
collection and analysis for the progression of prostate cancer and its relationship to the biomarker PSA (see Section 3). There is also a growing literature on the study of treatment effects using RCTs. However, as we point out in the next section, the optimal policies for whether and when to biopsy, the topic of this article, are not yet well understood. To our knowledge, ours is the first optimization study of prostate cancer screening that explicitly treats the imperfect nature of PSA tests and prostate biopsies.

3 Literature Review

The medical community has recently focused considerable attention on the use of PSA tests for prostate cancer screening. The majority of family physicians and urologists in the U.S. use PSA tests to screen their patients, commonly initiating annual screening at age 50 (Woolf and Rothemich 1999). However, recently some have suggested that PSA screening should not be done routinely since it can result in unnecessary biopsies, potential harm to the patient, and increased treatment costs. The U.S. Preventive Services Task Force (USPSTF), for instance, recommended a guideline for prostate cancer screening, stating that people older than 75 years should not be screened (U.S. Preventive Services Task Force 2008). The guideline made no specific recommendation for people younger than 75 years, citing insufficient evidence. Another guideline from the American Urological Association (2009) recommended PSA screening starting at age 40, followed by future screening intervals based on previous results.

There have been a number of recent studies of the value of PSA screening for detecting prostate cancer (often with conflicting conclusions). Fall et al. (2007) evaluated the accuracy of changes in PSA as predictors of lethal prostate cancer outcomes. They analyzed a cohort of men with localized prostate cancer and found that PSA is a poor predictor of the number that will develop lethal cancer. The controversy over PSA screening continued when the results of two large clinical trials were reported in 2009. The U.S. Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial (Andriole et al. 2009) concluded that screening does not reduce mortality. On the other hand, the European Randomized Study of Screening for Prostate Cancer (ERSPC) trial (Schroder et al. 2009) provided evidence of benefits from PSA screening, concluding that PSA screening reduced prostate cancer mortality by approximately 20%. Each of these studies reported
serious problems in terms of bias calling into question the results of the trials (Barry 2009).

Uncertainty also exists about the optimal frequency of screening over the course of a patient’s lifetime. van der Cruijsen-Koeter et al. (2003) performed an RCT comparing screened and unscreened patients as part of the European Randomized Study of Screening for Prostate Cancer. They concluded that sensitivity of the PSA test is high, and PSA testing detects the majority of cancer cases. Roobol et al. (2007) compared the incidence of prostate cancer between a Swedish and Dutch population that were screened at 2-year and 4-year intervals, respectively. They found no statistically significant difference in the incidence rate, suggesting that both screening intervals were of equal merit from a health outcomes perspective.

Some authors have developed simulation models to evaluate alternative prostate cancer screening policies. Ross et al. (2000) report on a simulation model to compare simple strategies (e.g. no screening, and screening intervals of 1, 2, and 5 yrs) based on performance measures including the number of PSA tests per 1000 men and prostate cancer deaths prevented. Their model is based on a Markov process for progression of the disease. They consider two competing criteria, prostate cancer deaths prevented per 1000 men and number of PSA tests per 1000 men. Etzioni et al. (2008) used a simulation model to attempt to determine if declines in advanced stage prostate cancer can be attributed to PSA screening. They concluded that PSA screening has contributed in part to declines in incidence and resulting mortality.

The prostate biopsy process has received some attention in the operations research literature. For instance, the physical placement of needles based on imaging information is an important and complex decision that is part of the overall screening process. Haas et al. (2007) estimated the diagnostic accuracy of needle biopsy, to be approximately 80%. Sofer et al. (2003) studied the optimal number of samples, and placement of the individual needles. They formulated and solved a nonlinear integer programming problem to determine optimal placement of needles assuming the objective was to maximize the probability of cancer detection.

Markov decision processes (MDP) have been applied to several types of medical decisions. For instance, Alagoz et al. (2004) used an MDP model to study the structure of optimal policies for the timing of living-donor liver transplantation; Denton et al. (2009) used an MDP model to study optimal policies for initiating cholesterol lowering medication for patients with diabetes; and Shechter et al. (2005) used an MDP model to study the optimal time to initiate HIV therapy.
Chhatwal et al. (2010) used a MDP to study the optimal policy for breast cancer biopsy based on mammography observations. In their model, states are defined by the probability of breast cancer at each epoch. A Bayesian Network is used to estimate the probabilities. In all of these cases the authors defined a discrete set of health states for a Markov process, and they computed the optimal policy to maximize expected future QALYs. QALYs are commonly employed in the health policy literature and are based on a decrement (from a nominal life year of 1.0) to represent the patient’s perceived value of a year of life in a particular health state. Gold et al. (2002) provide a review of QALYs, and Packer (1968) and Fanshel and Bush (1970) are early references on the use of QALYs in the operations research literature.

POMDPs have been used to study medical decisions in other contexts. Hu et al. (1993) considered the problem of choosing an appropriate drug infusion plan for the administration of anesthesia. Hauskrecht and Fraser (2000) applied a POMDP formulation to the problem of treating patients with ischemic heart disease. Maillart et al. (2008) used a partially observable Markov process to study breast cancer screening policies using mammography; they evaluated age-dependent screening policies and studied the tradeoff between lifetime mortality risk of breast cancer and the expected number of mammograms. Ivy (2008) further studied a POMDP model of breast cancer screening; they considered the patient and third-party payer perspective by computing cost optimal screening policies subject to a patient-based utility constraint at each decision epoch.

Our study differs from the above referenced work in several respects. First, in contrast to the medical literature which uses simulation to evaluate a small number of alternative screening policies we study an optimization model which seeks to find the policy that maximizes expected QALYs for the patient. Second, we are unaware of any existing work addressing partial observability resulting from two stages of imperfect medical tests with a continuous serum biomarker such as PSA. For instance, for the breast cancer screening articles referenced above mammogram results are discrete (normal or abnormal) and biopsy is perfect resulting in significantly different models. Third, we identify a number of interesting structural properties of our model which provide insight into the optimal screening policy. These properties are not obvious based on well known sufficient conditions (see, for example, Albright 1979, White 1980, and Lovejoy 1987) due to the non-stationary fact of the core process, effects of competing cause mortality, and lack of a stochastic ordering of the core states. Finally, we calibrate our model with a data set based on a large regional population that
includes all screening and treatment events. Other studies using standard statistical methods or simulation have used datasets for high risk patients or patients under study in RCTs which can result in selection bias.

4 Partially Observable Markov Decision Process

In this section we describe our POMDP model for prostate biopsy referral decisions. The objective in our model is to maximize expected QALYs for the patient. QALYs are estimated by decrementing a normal life year as a result of various events including (a) occurrence of biopsy, (b) treatment upon detection of cancer, and (c) long-term complications resulting from treatment. The optimal policy for biopsy therefore trades off the long-term benefits from early detection of prostate cancer with the short term negative impact of biopsy and long-term side effects of treatment.

In our model patients progress through (unobservable) health states and (observable) PSA intervals. PSA intervals are defined by clinically relevant ranges. The PSA intervals determine a conditional probability that the patient has prostate cancer. At each decision epoch the patient’s PSA is measured, and a decision is made to refer the patient for biopsy, or to defer the referral decision until the next decision epoch. If a patient receives a positive biopsy result he is assumed to be treated. Radical prostatectomy, active surveillance, external beam radiation therapy and brachytherapy (Lee et al. 1999) are all common forms of treatment that can be considered in our model. If a patient receives a negative biopsy result then he awaits PSA screening in the next decision epoch. Following is a description of our model.

Time Horizon: PSA screening is performed annually, \( t = \{40, 41, 42, \ldots, \infty \} \). In this infinite horizon problem, biopsy decisions are made until an old-enough age, \( \hat{N} \), using a finite horizon non-stationary POMDP. \( \hat{N} \) corresponds to a liberal upper bound on when screening is discontinued due to the risk of treatment being greater than the benefits. An infinite horizon Markov process is used beyond age \( \hat{N} \). Note that we provide support for the assumption of annual screening by an upperbound on the age of screening in Section 5.

Actions: Action, \( a_t \in \{B, W\} \), denotes the decision to perform a biopsy (B) or defer the biopsy decision (W) until the next decision epoch, \( t + 1 \).

States: At each decision epoch a patient is in one of several health states including no cancer (NC),
prostate cancer present but not detected (C), organ confined cancer detected (OC), extraprostatic cancer detected (EP), lymph node-positive cancer detected (LN), metastasis (M), and death from prostate cancer and all other causes (D). The states NC and C are not directly observable, but the other health states are assumed to be completely observable. The possible transitions among states are illustrated in Figure 3(a). Figure 3 illustrates how the model can be simplified by aggregating the three observable cancer stages (OC, EP and LN) into a single state in which non-metastatic prostate cancer has been detected and treated (T). Note that the state aggregation does not cause a loss of accuracy in our model, since the reward for state $T$ is the expected discounted future rewards, which are independent of the actions. Therefore, in our POMDP model, we use the aggregated set of core states $S = \{NC, C, T, M, D\}$ illustrated in Figure 3(c).

Figure 3: POMDP model simplification: aggregating the three non-metastatic prostate cancer stages after detection into a single core state $T$. Solid lines denote the transitions related to prostate cancer; dotted lines denote the action of biopsy and subsequent treatment; dashed lines in (c) denote death from other causes (for simplicity these are omitted from (a) and (b)).

**Observations:** At each decision epoch the patient is observed in one of a set of observable states including PSA intervals, non-metastatic cancer detected and treated (T), metastasis (M) or death (D), indexed by $\ell_t \in O = \{1, 2, 3, ..., m, T, M, D\}$.

**Information Matrix:** Conditional probabilities relate the underlying core states to the observations. We let $q_t(\ell_t|s_t)$ denote the probability of observing $\ell_t \in O$ given he is in health state $s_t \in S$.

**Belief States:** The belief state (or belief vector), $\pi_t = (\pi_t(NC), \pi_t(C), \pi_t(T), \pi_t(M), \pi_t(D))$, defines the probability the patient is in one of the four health states at epoch $t$. $\pi_t \in \Pi \equiv \{\pi_t \in$
\[ \mathbb{R}^5 \mid \sum_{i \in S} \pi_t(i) = 1, \pi_t(i) \geq 0, i \in S \]. Note that for a patient without a positive biopsy result, his belief state can be represented as \( \pi_t = (1 - \pi_t(C), \pi_t(C), 0, 0, 0) \). For this case we use \( \pi_t(C) \), the first component of vector \( \pi_t \), to denote \( \pi_t \) for short in the remainder of this paper.

**Transition Probabilities:** \( p_t(s_{t+1}|s_t, a_t) \) denote the core state transition probability from health state \( s_t \) to \( s_{t+1} \) at epoch \( t \) given action \( a_t \).

**Rewards:** \( r_t(s_t, a_t) \) is the immediate reward (measured in QALYs) given the patient is in core state \( s_t \) and action \( a_t \) is taken at decision epoch \( t \). Thus, the belief state immediate reward is \( r_t(\pi_t, a_t) = \sum_{s_t \in S} r_t(s_t, a_t) \pi_t(s_t) \).

The goal of our model is to determine the biopsy referral policy that maximizes expected discounted QALYs over the patient’s lifetime. It is well known that POMDPs can be converted into an equivalent completely observable Markov decision process on the continuous belief states \( \pi_t \) (Astrom 1965, Sondik 1971, Monahan 1982). The optimal value function and the corresponding optimal action for our model can be written as

\[
v_t(\pi_t) = \max_{a_t \in \{W,B\}} \left\{ r_t(\pi_t, a_t) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1}) p_t(\ell_{t+1}|\pi_t, a_t) \right\}, \forall \pi_t \in \Pi, \tag{1}
\]

and

\[
a_t^*(\pi_t) = \arg \max_{a_t \in \{W,B\}} \left\{ r_t(\pi_t, a_t) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1}) p_t(\ell_{t+1}|\pi_t, a_t) \right\}, \forall \pi_t \in \Pi,
\]

where

\[
p_t(\ell_{t+1}|\pi_t, a_t) = \sum_{s_{t+1} \in S} q_{t+1}(\ell_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \pi_t(s_t),
\]

and \( \lambda \in [0, 1] \) is the discount factor. Bayesian updating is used to revise the patient’s belief state over time as PSA observations are obtained. Bayesian updates are defined by the following transformation of the belief state:

\[
\pi_{t+1}(s_{t+1}) = \frac{q_{t+1}(\ell_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \pi_t(s_t)}{\sum_{s_{t+1} \in S} q_{t+1}(\ell_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \pi_t(s_t)}, \tag{2}
\]

where \( \pi_{t+1}(s_{t+1}) \), the component of the belief vector, \( \pi_{t+1} \), is a function of \( \ell_{t+1}, a_t \) and \( \pi_t \). Thus (2) updates the belief state of a patient based on their prior belief state and their most recent observed
PSA interval. The sequence of probabilities \( \{\pi_t, t = 1, \cdots, \infty\} \) has been shown to follow a Markov process (Monahan 1982), and therefore (1) defines a continuous state MDP.

4.1 Transition Probability Matrices and Reward Vectors

Table 1 defines the parameters used to construct the core state transition probability matrices and the rewards in our model. All the parameters in Table 1 are non-negative and not greater than 1 by definition.

Table 1: Detailed description of model parameters defining transition probabilities and rewards for the core state process.

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_t ) Annual prostate cancer incidence rate</td>
</tr>
<tr>
<td>( d_t ) Annual death rate from other causes</td>
</tr>
<tr>
<td>( b_t ) Annual metastasis rate for patients in state ( T )</td>
</tr>
<tr>
<td>( e_t ) Annual metastasis rate for patients in state ( C )</td>
</tr>
<tr>
<td>( z_t ) Annual prostate cancer death rate excluding death from other causes for patients in state ( M )</td>
</tr>
<tr>
<td>( f ) Biopsy detection rate for patients with prostate cancer</td>
</tr>
<tr>
<td>( \mu ) One-time utility decrement associated with biopsy</td>
</tr>
<tr>
<td>( \epsilon ) Annual utility decrement of living in state ( T )</td>
</tr>
<tr>
<td>( \gamma ) Annual utility decrement of living in state ( M )</td>
</tr>
</tbody>
</table>

We denote the transition probability matrix at epoch \( t \) given the decision to wait, \( W \), as \( P_t(s_{t+1}|s_t, W) \) consisting of elements \( p_t(s_{t+1}|s_t, W) \), \( \forall s_{t+1} \in S, s_t \in S \), where the non-zero elements are:

\[
p_t(\text{NC}|\text{NC}, W) = (1 - d_t)(1 - w_t),
\]

\[
p_t(\text{C}|\text{NC}, W) = (1 - d_t)w_t,
\]

\[
p_t(\text{D}|\text{NC}, W) = d_t,
\]

\[
p_t(\text{C}|\text{C}, W) = (1 - d_t)(1 - e_t),
\]

\[
p_t(\text{M}|\text{C}, W) = e_t(1 - d_t),
\]

\[
p_t(\text{D}|\text{C}, W) = d_t,
\]
\[ p_t(T|T, W) = (1 - d_t)(1 - b_t), \]
\[ p_t(M|T, W) = b_t(1 - d_t), \]
\[ p_t(D|T, W) = d_t, \]
\[ p_t(M|M, W) = (1 - d_t)(1 - z_t), \]
\[ p_t(D|M, W) = d_t + z_t(1 - d_t), \]
\[ p_t(D|D, W) = 1. \]

We denote the transition probability matrix given the decision to biopsy, \( B \), as \( P_t(s_{t+1}|s_t, B) \) consisting of elements \( p_t(s_{t+1}|s_t, B) \), \( \forall s_{t+1} \in S, s_t \in S \), where the non-zero elements are:

\[ p_t(NC|NC, B) = (1 - d_t)(1 - w_t), \]
\[ p_t(C|NC, B) = (1 - d_t)w_t, \]
\[ p_t(D|NC, B) = d_t, \]
\[ p_t(C|C, B) = (1 - f)(1 - d_t)(1 - e_t), \]
\[ p_t(T|C, B) = f(1 - b_t)(1 - d_t), \]
\[ p_t(M|C, B) = fb_t(1 - d_t) + e_t(1 - f)(1 - d_t), \]
\[ p_t(D|C, B) = d_t, \]
\[ p_t(T|T, B) = (1 - d_t)(1 - b_t), \]
\[ p_t(M|T, B) = b_t(1 - d_t), \]
\[ p_t(D|T, B) = d_t, \]
\[ p_t(M|M, B) = (1 - d_t)(1 - z_t), \]
\[ p_t(D|M, B) = d_t + z_t(1 - d_t), \]
\[ p_t(D|D, B) = 1. \]
Note that the last 6 elements of $p_t(s_{t+1}|s_t, W)$ and $p_t(s_{t+1}|s_t, B)$ are identical because they are independent of the action. From this point forward, we write them as $p_t(T|T)$, $p_t(M|T)$ $p_t(D|T)$, $p_t(M|M)$, $p_t(D|M)$ and $p_t(D|D)$, respectively.

There are a number of parameters that define the reward function in our model. We define the following rewards given the decision to wait, $W$:

$$\bar{r}_t(NC, W) = 1,$$
$$\bar{r}_t(C, W) = 1,$$
$$\bar{r}_t(T) = 1 - \epsilon,$$
$$\bar{r}_t(M) = 1 - \gamma,$$
$$\bar{r}_t(D) = 0.$$

The rewards given the decision to biopsy, $B$, are as follows:

$$\bar{r}_t(NC, B) = 1 - \mu,$$
$$\bar{r}_t(C, B) = 1 - \mu - f\epsilon,$$
$$\bar{r}_t(T) = 1 - \epsilon,$$
$$\bar{r}_t(M) = 1 - \gamma,$$
$$\bar{r}_t(D) = 0.$$

The rewards defined above reflect the following assumptions. All the parameters, $w_t$, $d_t$, $b_t$, $e_t$, $z_t$, $f$, $\mu$, $\epsilon$, $\gamma$ and $\lambda$, have values in [0, 1]. A patient who has a biopsy suffers a loss of $\mu$ QALYs in the year of biopsy to represent pain, anxiety, and short term procedure side effects such as infection. A non-metastatic prostate cancer patient who has a positive biopsy, and is treated, suffers a loss of $\epsilon$ QALYs for all future life years, i.e., $\epsilon$ reflects quality of life decrement due to permanent side effects of treatment. Note that we are implicitly assuming that the utility decrements, $\mu$, $\epsilon$ and $\gamma$ are based on population averages, and the decision maker is risk neutral.
5 Structural Properties

In this section we discuss the structure of our model and prove several structural properties that give insights into the optimal policy for biopsy referral. In our model we focus on primary screening, assuming that patients have at most one biopsy. In reality, about 7–12% of men undergoing biopsy have had a previous negative biopsy (Nguyen et al. 2010, Thompson et al. 2006). This is likely because the 10-12 cores obtained in a standard office prostate biopsy miss cancer in some men (Haas et al. 2007). Importantly, however, a prior negative prostate biopsy is an indicator of the absence of prostate cancer and therefore decreases the probability that the patient will have cancer detected at re-biopsy (Thompson et al. 2006, Ashley et al. 2008). In re-biopsy situations, the technique of biopsy also changes, with more cores being sampled to ensure that a “hidden” cancer is not missed again (Rabets et al. 2004, Chon et al. 2002). Unfortunately, when a cancer is found at re-biopsy, it is often a very low risk tumor and frequently clinically insignificant (Epstein et al. 2005, Bastian et al. 2004, Master et al. 2005). It can be surmised from these data that the decision to re-biopsy a man is inherently different than the decision to biopsy him for the first time, and for this reason we have opted to focus on the first biopsy decision since it informs about 90% of biopsy decisions. From the perspective of our model, this can be interpreted to mean that once patients are biopsied they leave the system. This assumption has also been made in previous cancer screening studies (Chhatwal et al. 2010).

Assuming a single biopsy, the optimality equation (1) can be rewritten as

\[
v_t(\pi_t) = \max \left\{ r_t(\pi_t, W) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1})p_{t}(\ell_{t+1}\pi_t, W), R_t(\pi_t) \right\}, \forall t, \forall \pi_t \in \Pi, (3)
\]

where \( R_t(\pi_t) \) is the expected discounted future reward given \( a_t = B \) at decision epoch \( t \), which can be written as:

\[
R_t(\pi_t) = -\mu + (1 - \pi_t(C))\tilde{R}_t(NC) + \pi_t(C)((1 - f)\tilde{R}_t(C) + f\tilde{R}_t(T)), (4)
\]

where \( \mu \) is the QALY decrement of biopsy; \( f \) is the biopsy detection rate; \( \pi_t(C) \), a component of vector \( \pi_t \), is the probability the patient is in state \( C \). We let \( \tilde{R}_t(NC) \), \( \tilde{R}_t(C) \), \( \tilde{R}_t(T) \) and \( \tilde{R}_t(M) \)
denote the expected discounted future rewards under the policy of never referring the patient for biopsy after age \( t \) for states \( NC, C, T \) and \( M \), respectively, which can be written as:

\[
\begin{align*}
\bar{R}_t(NC) &= \bar{r}_t(NC, W) + \lambda p_t(NC|NC, W)\bar{R}_{t+1}(NC) + p_t(C|NC, W)\bar{R}_{t+1}(C), \\
\bar{R}_t(C) &= \bar{r}_t(C, W) + \lambda p_t(C|C, W)\bar{R}_{t+1}(C) + \lambda p_t(M|C)\bar{R}_{t+1}(M), \\
\bar{R}_t(T) &= \bar{r}_t(T) + \lambda p_t(T|T)\bar{R}_{t+1}(T) + \lambda p_t(M|T)\bar{R}_{t+1}(M), \\
\bar{R}_t(M) &= \bar{r}_t(M) + \lambda p_t(M|M)\bar{R}_{t+1}(M).
\end{align*}
\] (5)

We let \( \bar{R}_t(\pi_t) \) denote the expected discounted future QALYs given the patient is never referred for biopsy, and is in belief state \( \pi_t \), in decision epoch \( t \). It can be written as:

\[
\bar{R}_t(\pi_t) = (1 - \pi_t(C))\bar{R}_t(NC) + \pi_t(C)\bar{R}_t(C).
\] (6)

To simplify the notation in later proofs related to (3), we define

\[
v_t(\pi_t, W) = v_t(\pi_t, W) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\ell_{t+1}|\pi_t, W).
\] (7)

Formulation (3) can be viewed as a partially observable optimal stopping time problem. At each decision epoch, the decision maker selects between the expected reward associated with biopsy, \( \bar{R}_t(\pi_t) \), or deferral of the decision to biopsy for one more decision epoch.

When there is no utility decrement associated with a PSA test, it can be shown that any additional PSA test provides non-negative benefit in total expected QALYs. We formalize this concept with the following proposition.

**Proposition 1.** The incremental benefit of an additional PSA test is non-negative.

Proof: see the appendix in the online supplement.

An immediate corollary to the above proposition is the following.

**Corollary 1.** Annual PSA screening is optimal when decisions to perform PSA tests are made not more frequent than annually.

Proof: this corollary is a direct result of the above proposition.
Proposition 1 and Corollary 1 imply that PSA screening should be done as frequently as possible in order to maximize the expected QALYs from the patient perspective. This is intuitively reasonable given we assume no disutility of PSA testing. Thus, the more PSA testing that is done, the greater the ability to discriminate between patients with and without prostate cancer. Since annual screening is the highest frequency of screening generally supported in the medical literature (U.S. Preventive Services Task Force 2008, American Urological Association 2009, American Cancer Society 2010), in the results that follow in Section 6 we assume patients are screened at each of a set of annual decision epochs.

Several structural properties can be proved about the optimal biopsy referral policy under the assumption of annual screening and reasonable assumptions about the model parameters. Lemmas, detailed proofs for the lemmas, propositions, and theorems, can be found in the appendix in the online supplement. We begin by stating several important assumptions and propositions, before presenting our main theoretical results.

**Assumption 1.** $w_t \leq Kb_t, \forall t$, where $K = (\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(M))/(\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(C))$.

Assumption 1 means that the annual prostate cancer incidence rate is not more than $K$ times the prostate cancer metastasis rate for patients in state $T$. Lemmas 1, 2, and 3 in the appendix in the online supplement together imply that the coefficient $K$ is greater than or equal to 1. Therefore, Assumption 1 is consistent with published mortality and incidence data for prostate cancer, which we discuss in Section 6.

**Assumption 2.** $Pr(PSA_t \geq PSA^*|C) \geq Pr(PSA_t \geq PSA^*|NC), \forall PSA^* \geq 0, \forall t$.

Assumption 2 means that for any given PSA threshold, $PSA^*$, and any age, $t$, patients in state $C$ are more likely to have PSA test results, $PSA_t$, higher than or equal to the threshold than patients in $NC$. In other words, a patient with cancer is likely to have higher PSA test results than a patient with no cancer. This is accepted as a basic clinical fact of PSA screening and is well established in the medical literature.

Before stating the next proposition, we provide the definition of the increasing failure rate (IFR) property for a matrix. This property plays an important role in the proof of our main results.

**Definition 1.** A $m \times n$ matrix $A$ has the IFR property if and only if $\sum_{j=k}^{n} A_{ij} \leq \sum_{j=k}^{n} A_{i'j}, \forall i < i' \in \{1, \cdots, m\}, \forall k \in \{1, \cdots, n\}$. 

17
Proposition 2. The information matrix $Q(t|s_t)$ has the IFR property.

Proof: see the appendix in the online supplement.

To say that $Q_t(l_t|s_t)$ is IFR means $\sum_{l_t=k}^D q_t(l_t|s_t^{(1)}) \leq \sum_{l_t=k}^D q_t(l_t|s_t^{(2)}), \forall k \in O$, for all $s_t^{(1)} \in S$ and $s_t^{(2)} \in S$ such that $s_t^{(1)} \leq s_t^{(2)}$. Here the order of $s_t$ is the same as the natural order of $S$, i.e., $NC < C < T < M < D$. Proposition 2 is a direct result of Assumption 2 and central to the proof of our main theoretical results presented below.

Assumption 3. If there exists a $\bar{t}$ such that $a^*_t(\pi_t(C) = 1) = W$ then $a^*_t(\pi_t(C) = 1) = W$ for all $t > \bar{t}$.

Assumption 3 means that if it is not optimal to treat a patient known to have prostate cancer at a given age, then it is not optimal to treat the patient in the future. In other words, patients who are detected with prostate cancer are treated immediately or not at all. Alternatively, this can be interpreted as follows. Patients who are not treated upon detection of cancer leave the screening process.

In the remainder of this section we present our main theoretical results that provide general insight into the optimal policy for biopsy referral decisions. Theorems 1 and 2 provide monotonicity results for our POMDP model.

Theorem 1. Under Assumptions 1 and 3, the optimal biopsy referral policy is of control-limit type with $\pi^*_t(C)$ such that

$$a^*_t(\pi_t) = \begin{cases} W, & \text{if } \pi_t(C) < \pi^*_t(C) \\ B, & \text{if } \pi_t(C) \geq \pi^*_t(C). \end{cases}$$

Proof: see the appendix in the online supplement.

Theorem 1 states that the optimal policy is monotonically nonincreasing in $\pi_t(C)$ under the above assumptions. The existence of a control-limit type policy means the optimal decisions on the belief space are separated by the threshold, $\pi^*_t(C)$. This is important for two reasons. First, the introduction of partial observability and the fact that the underlying health states are not ordered (state $C$ and $T$ are not ordered in that the reward of being in state $C$ is greater than in state $T$, but the prostate cancer metastasis rate of state $C$ is higher than state $T$) makes the proof of the structural properties novel relative to existing literature (e.g., Albright 1979, White...
Second, such policies are intuitive and much easier to implement in practice, which is particularly important in an already challenging clinical environment. Although the result in Theorem 1 seems intuitive there are counter examples demonstrating that intuition can be misleading. For instance, Lovejoy (1987) presents an example of a simple POMDP, which counter to intuition, does not have a threshold policy.

**Theorem 2.** Under Assumptions 1, 2 and 3, \( v_t(\pi_t) \) is nonincreasing in \( \pi_t(C) \) for any \( t \).

Proof: see the appendix in the online supplement.

Theorem 2 states that, consistent with intuition, the value function is monotonically nonincreasing in the belief of being in state \( C \) under the above assumptions. Intuitively this means that the patients expected QALYs do not increase as their probability of having prostate cancer increases. This property is useful in estimating bounds on the value function resulting from discretizing the continuous belief states, which is done in the application of certain solution methods.

Next, we provide a sufficient and necessary condition for the existence of the stopping time policy in our POMDP model in Theorem 3 and Corollary 2.

**Theorem 3.** Under Assumption 3, there exists a finite age, \( N \), at which it is optimal to discontinue biopsy referral if and only if the following condition is satisfied:

\[
R_N(T) - R_N(C) \leq \mu/f.
\]

Proof: see the appendix in the online supplement.

Theorem 3 provides a general result for a partially observable stopping time problem which is potentially relevant to other medical decision making problems in which the patient’s health state is not known without invasive and imperfect testing. Discontinuing biopsy referral at age \( N \) means \( a_t^*(\pi_t) = W, \forall \pi_t \in \Pi, \forall t \geq N \). Note that this implies that PSA screening should be discontinued which we state formally as the following corollary.

**Corollary 2.** If \( a_t^*(\pi_t) = W, \forall \pi_t \in \Pi, \forall t \geq N \), PSA screening should be discontinued.

Proof: this corollary is a direct result of the above proposition.

In words, Corollary 2 states that if the incremental benefit of treatment is not greater than the ratio of disutility of biopsy to the biopsy detection rate, it is on longer optimal to screen. Intuitively
this means that reducing the negative impact of biopsy or increasing the biopsy detection rate will increase the age at which screening should be discontinued. Furthermore, improving the benefit of treatment will also increase the age at which screening is discontinued.

Theorem 3 and Corollary 2 provide an insight into published clinical recommendations that prostate cancer screening should be terminated for older patients when their risk of dying of prostate cancer becomes low relative to other causes of death (e.g. heart disease). Estimating the stopping time $N$ is useful for two reasons. First, it provides a foundation for guidelines such as those of the U.S. Preventive Services Task Force (2008) which recommend terminating screening for older patients. Second, it provides a means to improve computational efficiency in solving the POMDP since it defines a finite horizon beyond which the policy is fixed.

6 Computational Results

In this section we present results based on our POMDP model. We describe the data used to estimate our model parameters and details about how we estimated model parameters based on the medical literature. We present sensitivity analysis based on variation of the model parameters, and we present estimates of the benefits of annual PSA screening. Finally, we discuss insights that can be drawn from the results of our numerical experiments.

6.1 Data Description

The data we used for parameter estimation in our model consists of 11,872 patients from Olmsted County, MN. It includes PSA values, biopsy information (if any), diagnosis information (if any), and the corresponding ages for patients recorded from 1983 through 2005. This regional data set includes all patients in Olmsted County irrespective of their prostate cancer risk. We use it to estimate prostate cancer probabilities conditional on PSA level for a general population. To our knowledge, it is the largest dataset of its kind for a North American population.

Among the patients in our data set, 1,140 patients have at least one biopsy prior to detection during their lifetime. 739 patients (81.4%) have exactly one biopsy. Since we focus on the prostate biopsy referral policy for primary screening we do not consider PSA records after cancer treatment. Based on expert opinion we removed abnormal PSA records such as the following. Suppose there
are three consecutive PSA tests for any patient; the first one and the third one are lower than 2.5 ng/ml; the second one is greater than 4.0 ng/ml and 2 times higher than the first and third. Then we assume the second PSA value is abnormal (likely caused by infections or data entry error), and remove it from the data set. This method eliminates a total of 94 abnormal PSA records.

6.2 Estimating Parameters

Since some people who have prostate cancer that has not yet been detected, the information matrix, $Q_t(\ell_t|s_t)$, is subject to bias. We used the methods proposed by Begg and Greenes (1983) to correct for this bias. We use biopsy as the confirmative test; thus, we assume that patients who have positive biopsies are true cancer patients and those who have negative biopsy are true no cancer patients.

We first separate the patients into different groups according to their PSA values ([0, 1), [1, 2.5), [2.5, 4), [4, 7), [7, 10) and \( \geq 10 \)) and ages ([40, 50), [50, 60), [60, 70), [70, 80) and \( \geq 80 \) ). Within each group, we assume patients without a confirmative test (biopsy) have the same probability of prostate cancer as patients who have had a confirmative test. The probability of having prostate cancer based on patients with confirmative tests is used to infer the cancer state of patients without confirmative tests. The resulting information matrix is

$$Q_t(\ell_t|s_t) = \begin{pmatrix}
0.471 & 0.337 & 0.101 & 0.059 & 0.015 & 0.017 & 0 & 0 & 0 \\
0.138 & 0.229 & 0.192 & 0.226 & 0.103 & 0.112 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}, \forall t.$$

The rows of $Q_t(\ell_t|s_t)$ correspond to states $NC$, $C$, $T$, $M$ and $D$, respectively; the columns correspond to PSA intervals $[0, 1)$, $[1, 2.5)$, $[2.5, 4)$, $[4, 7)$, $[7, 10)$, $[10, \infty)$, $T$, and $D$, respectively. $Q_t(\ell_t|s_t)$ is fixed for all the ages in this empirical study since our preliminary numerical experiments showed that changes in $Q_t(\ell_t|s_t)$ with respect to age do not significantly influence the optimal policy.

In the results we present, we assume patients detected with prostate cancer are treated by radical prostatectomy. Radical prostatectomy is historically the most common treatment (National Comprehensive Cancer Network 2007, Burkhardt et al. 2002) and reported to be the best treatment.
in terms of expected QALYs for all the ages (Sommers et al. 2007). It is also the only form of treatment for which the patients’ cancer stages can be understood by pathological examination of the removed organ. In order to estimate the annual metastasis rate for patients in state $T$, $b_t$, we use the weighted average of the metastasis rate of three non-metastatic prostate cancer stages using Mayo Clinic Radical Prostatectomy Registry (MCRPR) survival data. In our base case $b_t = 0.006$.

We estimated the annual death rate for metastasis from the 5 year death rates for patients’ age < 65 and ≥ 65 from the SEER data (National Cancer Institute 2008). Based on our estimates the disease specific annual death rate of metastatic prostate cancer is $b_t = 0.074$ for $t < 65$ and $b_t = 0.070$ for $t ≥ 65$. The disease specific metastasis rate from cancer not detected is estimated using the weighted sum of the grade-specific metastasis rate (Scardino et al. 1994) and the probabilities of grades upon detection (Ghani et al. 2005). The base case estimate is $e_t = 0.069$ for all ages.

In our base case, we use a decrement of $\mu = 0.05$ in the year of biopsy to estimate quality adjustment in the year a patient has a biopsy. Since no estimates of utility decrement exist yet for prostate biopsy, this is an estimate based on a similar choice of parameters for a recent bladder cancer study for the occurrence of surveillance cystoscopy (Kulkarni et al. 2009) and a breast cancer biopsy study (Chhatwal et al. 2010). In our base case, the disutility of metastasis is $\gamma = 0.24$ (Brenner et al. 2007). We assume that the annual QALY in years after treatment via prostatectomy is the mean of two extremes: (a) the most severe (metastasis) and (b) minor (mild sexual disfunction) symptoms according to patient surveys reported in Brenner et al. (2007). Hence annual QALYs of being in state $C$, $1 - \epsilon$, equals 0.855, which is the midpoint of 0.76 and 0.95.

The prostate cancer incidence rate, $w_t$ (shown in Table 2), is estimated from an autopsy review study (Bubendorf et al. 2000) that provides estimates of prostate cancer prevalence in the general population in ten-year age intervals. The mortality rate from other causes, $d_t$, is age specific and

Table 2: The age-specific values of the prostate cancer incidence rate, $w_t$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$40 - 50$</th>
<th>$50 - 60$</th>
<th>$60 - 70$</th>
<th>$70 - 80$</th>
<th>$\geq 80$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_t$</td>
<td>$2.32 \times 10^{-3}$</td>
<td>$4.70 \times 10^{-3}$</td>
<td>$7.02 \times 10^{-3}$</td>
<td>$6.17 \times 10^{-3}$</td>
<td>$1.17 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

based on the general mortality rate from the National Vital Statistics Reports (Arias 2010) minus the prostate cancer mortality rate from the National Cancer Institute (2008). Note that because
the National Cancer Institute reports a single prostate cancer incidence rate for ages greater than 95 and the National Vital Statistics Reports (Arias 2010) reports a single all cause mortality rate for ages greater than 95, we assume that \(d_t\) are fixed after the age of 95, i.e., \(\hat{N} = 95\) in our numerical experiment. Our base case biopsy detection rate is 0.8 (Haas et al. 2007).

A summary of all of the parameter values and their sources are provided in Table 3. It is worth noting that these parameter estimates satisfy Assumptions 1, 2, and 3 in Section 5. Thus, they validate our assumptions empirically. In Section 6.3 we use sensitivity analysis to evaluate the influence of changes in each of these parameters on the optimal biopsy referral policy.

Table 3: Parameters, their sources and specific values used in our base case analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Source</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w_t)</td>
<td>Bubendorf et al. (2000)</td>
<td>Age Specific</td>
</tr>
<tr>
<td>(d_t)</td>
<td>National Cancer Institute (2008), Arias (2010)</td>
<td>Age Specific</td>
</tr>
<tr>
<td>(b_t)</td>
<td>MCRPR</td>
<td>0.006</td>
</tr>
<tr>
<td>(e_t)</td>
<td>Scardino et al. (1994), Ghani et al. (2005)</td>
<td>0.069</td>
</tr>
<tr>
<td>(z_t)</td>
<td>National Cancer Institute (2008)</td>
<td>Age Specific</td>
</tr>
<tr>
<td>(f)</td>
<td>Haas et al. (2007)</td>
<td>0.8</td>
</tr>
<tr>
<td>(\mu)</td>
<td>Kulkarni et al. (2009), Chhatwal et al. (2010)</td>
<td>0.05</td>
</tr>
<tr>
<td>(\epsilon)</td>
<td>Bremner et al. (2007)</td>
<td>0.145</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>Bremner et al. (2007)</td>
<td>0.24</td>
</tr>
</tbody>
</table>

6.3 Computational Experiments and Sensitivity Analysis

POMDP models such as we propose are often computationally intractable. However, due to the low dimensionality of the belief state instances of our model, the POMDP can be solved in reasonable time using incremental pruning (Zhang and Liu 1996, Cassandra et al. 1997). All the experiments in this study was completed done on a 64-bit Intel Xeon 2.5 GHz CPU with 6MB of cache. All instances were solved in less than 6 minutes.

Results based on the base-case parameter estimates are presented in Figure 4. The optimal policy, the belief threshold between biopsy and wait, is illustrated in the figure. There are several interesting properties of the optimal policy. First, as expected from Theorem 1, the optimal policy is control-limit type. Second, there is a stopping time for screening at age 74, and thus the optimal policy is consistent with Theorem 3. Finally, we note that the threshold is decreasing in age; thus,
as patients age, their probability of having prostate cancer must be higher in order for a biopsy referral to be optimal. This is consistent with increasing all other cause mortality and the general consensus in the medical community that due to the low mortality rate of prostate cancer treatment becomes less beneficial as age increases.

Some of the parameters in our model are subject to variation either because they differ among patients due to differences in preferences (e.g. the perceived impact of biopsy, $\mu$) or physiology (e.g. anticipated recovery from surgery, $\epsilon$). Furthermore, there is variation in reported estimates such as the annual prostate cancer incidence rate, $w_t$ (Bubendorf et al. 2000). Therefore we present the results of a one-way sensitivity analysis for the model parameters that define the core process and reward function.

Figure 5(a) is a one-way sensitivity analysis for $w_t$, in which the upper and lower bounds on $w_t$ (shown in Table 4) are based on the lowest and highest estimates reported in autopsy studies (Bubendorf et al. 2000). We assume that people younger than 40 do not have prostate cancer. Figure 5(a) depicts the optimal policy for the base case and the lower and upper bounds on $w_t$. Because of the very low sensitivity of the optimal threshold to $w_t$, the three lines are nearly overlapping.

In Figure 5(b) the all other cause mortality, $d_t$, was perturbed by 20%. Note that increasing and decreasing 20% from the baseline values provided in Section 6.2 are the default variations of all the parameters if not otherwise specified. Figure 5(c) illustrates the sensitivity with respect to the annual prostate cancer metastasis rate, $b_t$, of patients in state $T$. Figure 5(d) shows the one-way
Figure 5: One-way sensitivity analysis for parameters: \( w_t, d_t, b_t, e_t, z_t, f, \mu, \epsilon, \gamma \) and \( \lambda \). Solid lines denote the base-case optimal policy and dashed lines denote the optimal policies of the upper and lower bounds.

(a) \( w_t \) changes in bounds from Bubendorf et al. (2000)

(b) \( d_t \) changes ±20%

(c) \( b_t \) changes ±20%

(d) \( e_t \) changes ±20%

(e) \( z_t \in (0.07, 0.37) \)

(f) \( f \in (0.64, 0.96) \)

(g) \( \mu \in (0.01, 0.1) \)

(h) \( \epsilon \in (0.05, 0.24) \)

(i) \( \gamma \in (0.15, 0.46) \)

(j) \( \lambda \in (0.97, 1) \)
sensitivity analysis of the annual prostate cancer metastasis rate of patients in state $C$, $e_t$. One-way sensitivity analysis of the annual death rate from metastasis excluding death from other causes, $z_t$, is illustrated in Figure 5(e), where the lower bound, 0.07, an estimate from Messing et al. (2006) and the upper bound, 0.37, is an estimate from Aus et al. (2005). Figure 5(f) shows the one-way sensitivity analysis for the biopsy detection rate, $f$.

Figure 5(g) is an one-way sensitivity analysis of varying the disutility of biopsy, $\mu$, from 0.01 to 0.1. Figure 5(h) shows the one-way sensitivity analysis for the utility decrement after prostatectomy, $\epsilon$, which has lower bound $\epsilon = 0.05$ (consistent with mild urinary problem after prostatectomy) and upper bound $\epsilon = 0.24$ (consistent with metastasis) which are taken from Bremner et al. (2007). Figure 5(i) is the one-way sensitivity analysis for the disutility of being in state $M$, $\gamma$, with lower bound $\gamma = 0.15$ from Krahn et al. (2003) and upper bound $\gamma = 0.46$ from Sandblom et al. (2004). Figure 5(j) shows the one-way sensitivity analysis on the discount factor, $\lambda$, with a baseline value of 1. In the sensitivity analysis, the lower bound of $\lambda$ takes 0.97, the most commonly used value in the health economics literature (Gold et al. 2002). No upper bound is provided in this case.

From Figure 5 we can see that the optimal policy is most sensitive to the utility decrements, $\epsilon$, $\mu$, and $\gamma$, which are the factors affecting the reward function. The optimal policy is also quite sensitive to $d_t$, $e_t$, $z_t$ and $\lambda$. On the other hand, it is not very sensitive to the prostate cancer incidence rate, $w_t$, the annual metastasis rate, $b_t$, or the biopsy detection rate, $f$.

Figure 5 illustrates the sensitivity of the optimal threshold to changes in model parameters. Figure 6 illustrates the changes to the optimal value function at age 40 given that the parameters are varied in the same ranges used in Figure 5. When we consider the sensitivity analysis on the value function for the optimal policy, we find it is most sensitive to the prostate cancer incidence
rate, $w_t$. It is also sensitive to the prostate cancer metastasis rate for patients in state $C$, $e_t$, and the utility decrement of treatment, $\epsilon$.

Although not illustrated in Figure 6, we found that changing the discount factor $\lambda$ from 1 to 0.97 will make the optimal value function decrease more than 10 QALYs. In this sense, the optimal value is the most sensitive to $\lambda$. The use of a discount factor, particularly over long time frames, has been a highly debated topic for decades (see Gold et al. 2002 for a discussion of this). From Figure 6: One-way sensitivity analysis on optimal values for model parameters: $d_t$, $w_t$, $\epsilon$, $z_t$, $\mu$, $\gamma$, $e_t$, $f$ and $b_t$.

![Sensitivity Analysis Diagram]

Figures 6, $d_t$ and $w_t$ are the parameters affecting expected QALYs the most; From Figure 5, the parameters defining the reward function, $\mu$, $\epsilon$ are among the ones with greatest influence on the optimal policy.

6.4 Benefits of Prostate Cancer Screening

We measured the total estimated benefit of prostate cancer screening by estimating how much the value function at age 40, i.e., the expected QALYs for a 40 year old patient with no prostate cancer, improves when the optimal policy is adopted versus no PSA screening ($a_t = W, \forall t, \forall \pi_t \in \Pi$). In Table 5 the optimal objective values are provided for our base case along with several choices of
model parameters, $\lambda$, $\epsilon$, and $\mu$. The benefits of prostate cancer screening are most significant for cases in which $\mu$ and $\epsilon$, the factors that define the effect of screening and treatment on the patients quality of life, are minimized. For $\lambda = 1$ the base-case benefit of screening is 0.167 QALYs per person for the male population regardless of their risk of prostate cancer. For the case that is most favorable for the benefits of prostate cancer screening and treatment ($\epsilon = 0.05$, $\mu = 0.01$) the benefit is 0.312 QALYs per person.

Table 5: Sensitivity analysis for expected QALYs for a 40-year-old patient assuming $\pi_{40}(NC) = 1$ comparing the optimal policy to the case of no screening, allowing at most one biopsy over one’s lifetime. Base-cases values are shown in bold.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\epsilon$</th>
<th>$\mu$</th>
<th>Expected QALYs under optimal biopsy referral policy</th>
<th>Improvements over no screening (QALYs)</th>
<th>Percentage improvement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.97</td>
<td>0.05</td>
<td>0.01</td>
<td>21.893</td>
<td>0.101</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>21.882</td>
<td>0.090</td>
<td>0.413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>21.872</td>
<td>0.080</td>
<td>0.367</td>
</tr>
<tr>
<td>0.145</td>
<td>0.05</td>
<td>0.01</td>
<td>21.843</td>
<td>0.051</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>21.833</td>
<td>0.041</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>21.825</td>
<td>0.033</td>
<td>0.153</td>
</tr>
<tr>
<td>0.24</td>
<td>0.01</td>
<td>0.01</td>
<td>21.804</td>
<td>0.012</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>21.799</td>
<td>0.007</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>21.796</td>
<td>0.004</td>
<td>0.020</td>
</tr>
<tr>
<td>0.05</td>
<td>0.1</td>
<td>0.01</td>
<td>37.784</td>
<td>0.312</td>
<td>0.824</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>37.762</td>
<td>0.290</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>37.741</td>
<td>0.269</td>
<td>0.711</td>
</tr>
<tr>
<td>0.145</td>
<td>0.01</td>
<td>0.01</td>
<td>37.659</td>
<td>0.187</td>
<td>0.495</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td><strong>37.639</strong></td>
<td><strong>0.167</strong></td>
<td><strong>0.442</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>37.621</td>
<td>0.149</td>
<td>0.395</td>
</tr>
<tr>
<td>0.24</td>
<td>0.01</td>
<td>0.01</td>
<td>37.555</td>
<td>0.083</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>37.539</td>
<td>0.067</td>
<td>0.177</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>37.528</td>
<td>0.056</td>
<td>0.148</td>
</tr>
</tbody>
</table>

7 Discussion

The U.S. Preventive Services Task Force (2008) provides the following recommendation: “Current evidence is insufficient to assess the balance of benefits and harms of screening for prostate cancer in men younger than age 75 years. Do not screen for prostate cancer in men age 75 years or older”. Our theoretical results in Section 5 provide a foundation for the concept of stopping screening
at older ages. Our base case results from our empirical study in Section 6 estimate the optimal stopping time at 74 years of age, which is surprisingly close to the USPSTF recommended stopping time. However, based on our sensitivity analysis we find the decision is highly dependent on a number of factors that may vary among patients. Based on our sensitivity analysis we find it may be optimal to discontinue screening as early as age 63 if the quality of life after treatment is expected to be very low.

The partially observable nature of cancers, such as a prostate cancer, make it challenging to estimate model parameters. Our sensitivity analysis provides a basis to help prioritize research on the collection of data to better estimate model parameters. For example, from our sensitivity analysis, we observed that the referral threshold is most sensitive to utility decrements of QALYs, but not very sensitive to the biopsy detection rate. Thus, it appears that improvements in prostate biopsy technology are not likely to significantly influence the expected QALYs of the optimal biopsy referral policy. On the other hand, methods for estimating disutility of biopsy and treatment, an area in which very little research has been done, are very important for measuring the benefit of PSA screening in individual patients. Furthermore, reducing the disutility of treatment and biopsy could significantly influence the optimal policy, and incorporating individual characteristics of a patient (body mass index and medical history) could help better inform biopsy and treatment decisions.

Our results in Section 6 shed light on the recent controversy about PSA screening (Andriole et al. 2009, Schroder et al. 2009). For instance, we quantify the benefits of annual PSA screening. Our results indicate that there is a potentially significant benefit from PSA screening. For example, the results in Table 5 illustrate the benefits may be significantly greater than the benefit per person of some well-known population-based prevention programs such as vaccination against measles and rubella, which has an estimated benefit per person of 0.008 QALYs (Wright and Weinstein 1998)).

8 Conclusions

We presented a non-stationary POMDP model to estimate optimal biopsy referral policies. Theoretical results indicate the existence of a control-limit policy, and the conditions under which there exists a finite stopping time for screening. We reported numerical results for the age-specific belief
threshold for biopsy based on a population using data for Olmsted County, MN, and several secondary data sources from the literature. Our results for the optimal biopsy referral policy provide a number of insights into the potential value of an annual PSA screening program. They also provide a means for prioritizing the collection of new data needed to better understand prostate cancer screening decisions.

There are some limitations of our study. First, prostatectomy was assumed to be the only treatment because it is one of the most common treatments, and because prostatectomy data was readily available for our study. Our model could be easily adapted to consider other treatment options such as radiation therapy, brachytherapy and active surveillance, if data on expected quality adjusted survival after such treatments becomes available. Second, we have not considered physical screening through digital rectal examination (DRE). This is due primarily to the lack of data for DRE. If DRE data was available, we could factor it into our model by considering it as an additional observation. Third, we focus on primary screening and the case of a first biopsy referral decision; however, some patients undergo additional biopsies if they continue to have high PSA test results following an initial biopsy. Finally, due to the demographic makeup of the population in Olmsted county, MN, our model was based on data for a largely Caucasian population. Future work based on a more diverse population could reveal insights about the role of race as a risk factor in prostate biopsy referral decisions. We believe our model provides a foundation for these future studies.

Proofs of the structural properties in Section 5

**Proposition 1.** The incremental benefit of an additional PSA test is non-negative.

**Proof.** A PSA test is assumed to be an instantaneous action. Let \( t_- \) be the time immediately prior to a PSA test, and \( t_+ \) be the time immediately following a PSA test. Similarly, \( \pi_{t_-}(C) \) denotes the belief prior to the PSA test, and \( \pi_{t_+}(C) \) denotes the belief following the PSA test. Let \( v^N_{t_-}(\pi_{t_-}(C)) \) be the value function when no PSA test is done at \( t \). Let \( v^P_{t_-}(\pi_{t_-}(C)) \) be the value function when a PSA test is done at \( t \). The proposition requires \( v^P_{t_-}(\pi_{t_-}(C)) \geq v^N_{t_-}(\pi_{t_-}(C)) \) for all \( t \), which can
be proved as follows:

\[
q^P_{t-}(\pi_{t-}(C)) = \sum_{\ell_t \in O} \Pr(\ell_t|\pi_{t-}(C))v_{t-}(\pi_{t-}(C)) \\
= \sum_{\ell_t \in O} \Pr(\ell_t|\pi_{t-}(C))v^N_{t-}(\pi_{t+}(C)) \\
= \sum_{\ell_t \in O} \left(q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))\right)v^N_{t-}(\pi_{t+}(C)) \\
= \sum_{\ell_t \in O} \left(q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))\right)\frac{q_t(\ell_t|C)\pi_{t-}(C)}{q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))}v^N_{t-}(\pi_{t+}(C)) \\
\geq v^N_{t-}\left(\sum_{\ell_t \in O} \left(q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))\right)\right)\frac{q_t(\ell_t|C)\pi_{t-}(C)}{q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))} = v^N_{t-}(\pi_{t-}(C)),
\]

where the second equality holds because a PSA test is assumed to be an instantaneous action, so that the value function of a patient immediately before and after \(t\) is the same when no PSA test is done; the fourth equality holds because the core state does not change from \(t_-\) to \(t_+\); the inequality holds due to the convexity of the POMDP value function and \(\sum_{\ell_t \in O} \left(q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))\right) = 1\); the last equality holds because \(\sum_{\ell_t \in O} \left(q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))\right)\frac{q_t(\ell_t|C)\pi_{t-}(C)}{q_t(\ell_t|C)\pi_{t-}(C) + q_t(\ell_t|\text{NC})(1 - \pi_{t-}(C))} = 1\).

**Proposition 2.** The information matrix \(Q_t(l_t|s_t)\) has the IFR property.

**Proof.** From Assumption 2, for any set of ordered PSA intervals, \(\{1, \cdots, m\}\), which is also a set partition on PSA values \([0, \infty)\), we have \(\sum_{l_t = k}^m q_t(l_t|C) \geq \sum_{l_t = k}^m q_t(l_t|\text{NC})\), \(\forall k \in \{1, \cdots, m\}\), \(\forall t\). Thus the \(2 \times m\) submatrix on the upper left corner of \(Q_t\) with rows in \(\text{NC}\) and \(C\) and columns in \(\{1, \cdots, m\}\) has the IFR property. Since \(q_t(T|T) = 1\), \(q_t(l_t|T) = 0\), \(\forall l_t \neq T\), \(q_t(M|M) = 1\), \(q_t(l_t|M) = 0\), \(\forall l_t \neq M\), \(q_t(D|D) = 1\) and \(q_t(l_t|D) = 0\), \(\forall l_t \neq D\), it follows that \(Q_t\) has the IFR property.

**Lemma 1.** \(\bar{R}_t(C) \geq 0\), \(\forall t\).
Proof. From (5) we have

\[
\bar{R}_t(M) = \bar{r}_t(M) + \lambda p_t(M|\bar{M}, \bar{W})\bar{R}_{t+1}(M) \\
= 1 - \gamma + \lambda(1 - z_t)(1 - d_t)\bar{R}_{t+1}(M) \\
= 1 - \gamma + \sum_{i=1}^{\infty} (1 - \gamma) \prod_{j=1}^{i} \lambda(1 - z_{t-1+j})(1 - d_{t-1+j}) \\
\geq 0,
\]

where the third equality follows from using (5) for \( t + 1, \cdots, \infty \); the inequality follows from the fact that all the parameters have values in \([0,1]\).

Again from (5) we have

\[
\bar{R}_t(C) = \bar{r}_t(C, \bar{W}) + \lambda p_t(C|\bar{C}, \bar{W})\bar{R}_{t+1}(C) + \lambda p_t(M|C, \bar{W})\bar{R}_{t+1}(M) \\
= 1 + \lambda(1 - e_t)(1 - d_t)\bar{R}_{t+1}(C) + \lambda e_t(1 - d_t)\bar{R}_{t+1}(M) \\
\geq 1 + \lambda(1 - e_t)(1 - d_t)\bar{R}_{t+1}(C) \\
= 1 + \sum_{i=1}^{\infty} \prod_{j=1}^{i} \lambda(1 - e_{t-1+j})(1 - d_{t-1+j}) \\
\geq 0,
\]

Where the first inequality follows from (8); the third equality follows from using (5) for \( t+1, \cdots, \infty \); the second inequality follows from the fact that all the parameters have values in \([0,1]\).

Lemma 1 means the expected QALY of a patient in \( C \) is nonnegative.

Lemma 2. \( \bar{R}_t(C) \geq \bar{R}_t(M), \ \forall t. \)
Proof. From (5) we have

\begin{align*}
\bar{R}_t(C) - \bar{R}_t(M) &= \bar{r}_t(C, W) + \lambda p_t(C|C, W) \bar{R}_{t+1}(C) + \lambda p_t(M|C, W) \bar{R}_{t+1}(M) \\
&\quad - (\bar{r}_t(M) + \lambda p_t(M|M) \bar{R}_{t+1}(M)) \\
&= 1 + \lambda (1 - d_t)(1 - e_t) \bar{R}_{t+1}(C) + \lambda (1 - d_t) e_t \bar{R}_{t+1}(M) \\
&\quad - (1 - \gamma + \lambda (1 - d_t)(1 - z_t) \bar{R}_{t+1}(M)) \\
&= \gamma + \lambda (1 - d_t)(1 - e_t) \bar{R}_{t+1}(C) - \lambda (1 - d_t)(1 - e_t - z_t) \bar{R}_{t+1}(W) \\
&= \gamma + \lambda (1 - d_t)(1 - e_t)(\bar{R}_{t+1}(C) - \bar{R}_{t+1}(M)) + \lambda (1 - d_t) z_t \bar{R}_{t+1}(C) \\
&\geq \gamma + \lambda (1 - d_t)(1 - e_t)(\bar{R}_{t+1}(C) - \bar{R}_{t+1}(M)),
\end{align*}

(9)

where the inequality holds by dropping a non-negative term based on Lemma 1. Expanding (9) for \( t + 1, \cdots, \infty \), we have

\[
\bar{R}_t(C) - \bar{R}_t(M) \geq \sum_{i=1}^{\infty} \gamma \prod_{j=1}^{i} \lambda (1 - d_{t-1+j})(1 - e_{t-1+j}) \geq 0,
\]

which implies \( \bar{R}_t(C) \geq \bar{R}_t(M) \), \forall t. \qed

Lemma 2 means that patients in state \( C \) have expected QALYs no less than those in state \( M \) given \( a_t^*(\pi_t) = W, \forall t \). In other words the expected discounted future QALYs are higher for a patient with cancer not detected than for a patient with metastasis.

Lemma 3. \( \bar{R}_t(NC) \geq \bar{R}_t(C), \forall t. \)
Proof. From (5) we have

\[
\bar{R}_t(NC) - \bar{R}_t(C) = \bar{r}_t(NC, W) + \lambda p_t(NC|NC, W) \bar{R}_{t+1}(NC) + \lambda p_t(C|NC, W) \bar{R}_{t+1}(C) \\
- (\bar{r}_t(C, W) + \lambda p_t(C|C, W) \bar{R}_{t+1}(C) + \lambda p_t(M|C, W) \bar{R}_{t+1}(M))
\]

\[
= 1 + \lambda (1-d_t)(1-w_t) \bar{R}_{t+1}(NC) + \lambda (1-d_t) w_t \bar{R}_{t+1}(C) \\
- (1 + \lambda (1-d_t)(1-e_t) \bar{R}_{t+1}(C) + \lambda (1-d_t) e_t \bar{R}_{t+1}(M))
\]

\[
\geq 1 + \lambda (1-d_t)(1-w_t) \bar{R}_{t+1}(NC) + \lambda (1-d_t) w_t \bar{R}_{t+1}(C) \\
- (1 + \lambda (1-d_t) \bar{R}_{t+1}(C))
\]

\[
= \lambda (1-d_t)(1-w_t) \bar{R}_{t+1}(NC) - \lambda (1-d_t)(1-w_t) \bar{R}_{t+1}(C)
\]

\[
= \lambda (1-d_t)(1-w_t)(\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(C)),
\]

where the inequality holds by dropping a non-negative term based on Lemma 2. Expanding (10) for \( t+1, \cdots, \infty \), we have

\[
\bar{R}_t(NC) - \bar{R}_t(C) \geq \prod_{j=1}^{\infty} \lambda (1-d_{t-1+j})(1-w_{t-1+j}) \geq 0,
\]

which implies \( \bar{R}_t(NC) \geq \bar{R}_t(C), \forall t. \]

Lemma 3 means that patients in state \( NC \) have expected QALYs no less than those in state \( C \) given \( a^*_t(\pi_t) = W, \forall t \). In other words the expected discounted future QALYs are higher for a patient without cancer than for a patient with cancer in the absence of screening.

Lemma 4. \( \bar{R}_t(NC) \geq \bar{R}_t(T) \forall t. \)

Proof. From Assumption 1, we have

\[
w_t \leq \frac{\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(M)}{\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(C)} b_t \Rightarrow b_t(\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(M)) \geq w_t(\bar{R}_{t+1}(NC) - \bar{R}_{t+1}(C)) \quad \text{(by Lemma 3)},
\]

\[
\Rightarrow (b_t - w_t) \bar{R}_{t+1}(NC) + w_t \bar{R}_{t+1}(C) - b_t \bar{R}_{t+1}(M) \geq 0.
\]

(11)
From (5) we have
\[ R_t(\text{NC}) - R_t(T) = (\bar{r}_t(\text{NC}, W) + \lambda p_t(\text{NC}|\text{NC}, W) R_{t+1}(\text{NC}) + \lambda p_t(\text{C}|\text{NC}, W) R_{t+1}(\text{C})
- (\bar{r}_t(T) + \lambda p_t(T|T) R_{t+1}(T) + \lambda p_t(M|T) R_{t+1}(M))
\]
\[ = 1 + \lambda (1 - d_t)(1 - w_t) R_{t+1}(\text{NC}) + \lambda (1 - d_t) b_t R_{t+1}(M)
- (1 - \epsilon + \lambda (1 - d_t)(1 - b_t) R_{t+1}(T) + \lambda (1 - d_t) b_t R_{t+1}(M)) \] (12)
\[ = \epsilon + \lambda (1 - d_t)(1 - b_t) (\bar{R}_{t+1}(\text{NC}) - \bar{R}_{t+1}(T))
+ \lambda (1 - d_t)(b_t - w_t) R_{t+1}(\text{NC}) + w_t R_{t+1}(\text{C}) - b_t \bar{R}_{t+1}(M))
\geq \epsilon + \lambda (1 - d_t)(1 - b_t)(\bar{R}_{t+1}(\text{NC}) - \bar{R}_{t+1}(T)),
\]
where the inequality results from dropping the nonnegative term in (11). Using (12) for \(t+1, \cdots, \infty\), we have
\[ \bar{R}_t(\text{NC}) - \bar{R}_t(T) \geq \epsilon + \sum_{i=1}^{\infty} \epsilon \prod_{j=1}^{i} \lambda (1 - d_{t-1+j})(1 - b_{t-1+j}) \geq 0,
\]
which implies \(\bar{R}_t(\text{NC}) \geq \bar{R}_t(T), \forall t.\)

Lemma 4 states that the expected discounted future QALYs for a patient in state \text{NC} is not less than a patient in state \text{T} in the absence of screening.

**Lemma 5.** \(R_t(\pi_t)\) is nonincreasing in \(\pi_t(\text{C})\) for any \(t\).

**Proof.** \(R_t(\pi_t)\) denotes the cumulative reward for \(\pi_t\) given \(a_t^*(\pi_t) = B\). From (4) it follows that \(R_t(\pi_t)\) is nonincreasing in \(\pi_t(\text{C})\) if \(\bar{R}_t(\text{NC}) \geq (1 - f) \bar{R}_t(\text{C}) + f \bar{R}_t(T)\), which follows from Lemmas 3 and 4.

**Theorem 1.** Under Assumptions 1 and 3, the optimal biopsy referral policy is of control-limit type with \(\pi_t^*(\text{C})\) such that
\[ a_t^*(\pi_t) = \begin{cases} W, & \text{if } \pi_t(\text{C}) < \pi_t^*(\text{C}) \\ B, & \text{if } \pi_t(\text{C}) \geq \pi_t^*(\text{C}). \end{cases} \]

**Proof.** This theorem can be categorized and proved by considering three different cases, which are mutually exclusive and collectively exhaustive:
Case 1 \((a^*_t(\pi_t(C) = 1) = B)\): By the condition of this case, \(v_t(\pi_t(C) = 1, W) < R_t(\pi_t(C) = 1)\). From Lemma 5, \(R_t(\pi_t(C))\) is linear decreasing in \(\pi_t(C)\) and \(v_t(\pi_t(C), W)\) is convex in \(\pi_t(C)\) (Smallwood and Sondik 1973, Sondik 1978 proved this for general PODMP maximization problem).

It follows that \(v_t(\pi_t(C), W)\) and \(R_t(\pi_t(C))\) have at most one intersection of \(\pi_t(C)\), which implies \(a^*_t(\pi_t) = W\), if \(\pi_t(C) < \pi^*_t(C)\) and \(a^*_t(\pi_t) = B\), if \(\pi_t(C) \geq \pi^*_t(C)\), where \(\pi^*_t(C)\) is the intersection.

Case 2 \((a^*_t(\pi_t(C) = 1) = W\) and \(\bar{R}_t(T) < R_t(C)\)): From (6) and (4), it is straightforward to show \(\bar{R}_t(\pi_t(C)) > R_t(\pi_t(C))\), \(\forall \pi_t(C)\). And from (3) we have \(v_t(\pi_t(C)) \geq \bar{R}_t(\pi_t(C)) > R_t(\pi_t(C))\), i.e., \(a^*_t(\pi_t) \neq B\), \(\forall \pi_t(C)\). It can be written as \(a^*_t(\pi_t) = W\), if \(\pi_t(C) < \pi^*_t(C)\) and \(a^*_t(\pi_t) = B\), if \(\pi_t(C) \geq \pi^*_t(C)\), where \(\pi^*_t(C) = 1\).

Case 3 \((a^*_t(\pi_t(C) = 1) = W\) and \(\bar{R}_t(T) \geq \bar{R}_t(C)\)): From (4), we have
\[
\frac{dR_t(\pi_t(C))}{d\pi_t(C)} = \bar{R}_t(NC) - (1 - f)\bar{R}_t(C) - f\bar{R}_t(T).
\]

From (6), we have
\[
\frac{dR_t(\pi_t(C))}{d\pi_t(C)} = \bar{R}_t(NC) - \bar{R}_t(C).
\]

By (2) \(p_t(NC|C, W) = p_t(NC|C, B) = 0\) and \(\pi_t(C) = 1\), then \(\pi_{t+1}(C) = 1\). Furthermore \(\pi_t(C) = 1\) implies \(\pi_t(C) = 1\) for all \(t \geq \bar{t}\). By Assumption 3 we have \(v_t(\pi_t(C) = 1) = \bar{R}_t(\pi_t(C) = 1)\). By the convex property of \(v_t(\pi_t(C))\), we have
\[
\frac{dv_t(\pi_t(C))}{d\pi_t(C)}|_{\pi_t(C)=1} \geq \frac{d\bar{R}_t(\pi_t(C))}{d\pi_t(C)}.
\]

By assumption for this case we have
\[
\frac{dv_t(\pi_t(C))}{d\pi_t(C)}|_{\pi_t(C)=1} > \frac{dR_t(\pi_t(C))}{d\pi_t(C)},
\]

and by convexity
\[
\frac{dv_t(\pi_t(C))}{d\pi_t(C)} > \frac{dR_t(\pi_t(C))}{d\pi_t(C)}, \ \forall \pi_t(C).
\]

Thus \(a^*_t(\pi_t(C)) = W\) for all \(\pi_t(C)\), which implies \(a^*_t(\pi_t) = W\), if \(\pi_t(C) < \pi^*_t(C)\) and \(a^*_t(\pi_t) = B\), if \(\pi_t(C) \geq \pi^*_t(C)\), where \(\pi^*_t(C) = 1\).
Lemma 6. If \( v_{t+1}(\pi_{t+1}) \) is nonincreasing in \( \pi_{t+1}(C) \), then \( v_t(\pi_t, W) \) is nonincreasing in \( \pi_t(C) \) for any \( t \).

Proof. By definition in (7), we have \( v_t(\pi_t, W) = r_t(\pi_t, W) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1})p_t(\ell_{t+1} | \pi_t, W) \). Since \( r_t(\pi_t, W) = (1 - \pi_t(C))\bar{r}_t(NC, W) + \pi_t(C)\bar{r}_t(C, W) \) and \( \bar{r}_t(NC, W) \geq \bar{r}_t(C, W) \), then \( r_t(\pi_t, W) \) is nonincreasing in \( \pi_t(C) \). Next we prove that \( \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1})p_t(\ell_{t+1} | \pi_t, W) \) is nonincreasing in \( \pi_t(C) \) for any \( t \):

Since \( \Pr(C | \pi_t(C)) = \pi_t(C) \) and \( \Pr(NC | \pi_t(C)) = 1 - \pi_t(C) \), therefore

\[
\Pr(NC | \pi_t(1)(NC)) \leq \Pr(NC | \pi_t(2)(NC)),
\]

for all \( \pi_t(1)(NC) \) and \( \pi_t(2)(NC) \) such that \( \pi_t(1)(NC) \leq \pi_t(2)(NC) \). Let \( S = \{C, NC\} \) and write (13) as:

\[
\sum_{s_t \in C, NC} \Pr(s_t | \pi_t(1)(C)) \leq \sum_{s_t \in C, NC} \Pr(s_t | \pi_t(2)(C)), \quad \forall k \in S
\]

for all \( \pi_t(1)(C) \) and \( \pi_t(2)(C) \) such that \( \pi_t(1)(C) \leq \pi_t(2)(C) \). The transition probability matrix among core states \( NC, C, M \) and \( D \),

\[
p_t(s_{t+1} | s_t, W) = \begin{pmatrix}
(1 - d_t)(1 - w_t) & (1 - d_t)w_t & 0 & d_t \\
0 & (1 - d_t)(1 - e_t) & e_t(1 - d_t) & d_t \\
0 & 0 & (1 - z_t)(1 - d_t) & d_t + z_t(1 - d_t) \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

can easily be shown to have the IFR property. That is, \( \sum_{s_t+1 = k} p_t(s_{t+1} | s_t, W) \) is nondecreasing in \( s_t \).

Therefore, by Lemma 4.7.2 in Puterman (1994) and (14), \( \sum_{s_t+1 = k} \sum_{s_t = 1}^{[S]} p_t(s_{t+1} | s_t, W) \Pr(s_t | \pi_t(C)) \) is nonincreasing in \( \pi_t(C) \) for all \( k \in S \). It can be rewritten as

\[
\sum_{s_t+1 = k} \sum_{s_t = 1}^{[S]} p_t(s_{t+1} | s_t, W) \Pr(s_t | \pi_t(1)(C)) \leq \sum_{s_t+1 = k} \sum_{s_t = 1}^{[S]} p_t(s_{t+1} | s_t, W) \Pr(s_t | \pi_t(2)(C)), \quad \forall k \in S
\]

37
for all $\pi_t^{(1)}(C)$ and $\pi_t^{(2)}(C)$ such that $\pi_t^{(1)}(C) \leq \pi_t^{(2)}(C)$. By Proposition 2, $\sum_{\ell_{t+1}=k}^{[O]} q_{t+1}(\ell_{t+1}|s_{t+1})$ is nondecreasing in $s_{t+1}$. Therefore, by Lemma 4.7.2 in Puterman (1994) and (15),

$$\sum_{s_{t+1}=1}^{[S]} \sum_{\ell_{t+1}=k}^{[O]} q_{t+1}(\ell_{t+1}|s_{t+1}) \sum_{s_{t}=1}^{[S]} p_{t}(s_{t+1}|s_{t}, W) \Pr(s_{t}|\pi_{t}(C))$$

is nonincreasing in $\pi_{t}(C)$ for any $k \in O$. And since

$$p_{t}(\ell_{t+1}|\pi_{t}, W) = \sum_{s_{t+1}=1}^{[S]} q_{t+1}(\ell_{t+1}|s_{t+1}) \sum_{s_{t}=1}^{[S]} p_{t}(s_{t+1}|s_{t}, W) \Pr(s_{t}|\pi_{t}(C))$$

therefore

$$\sum_{\ell_{t+1}=k}^{[O]} p_{t}(\ell_{t+1}|\pi_{t}^{(1)}(C), W) \leq \sum_{\ell_{t+1}=1}^{[O]} p_{t}(\ell_{t+1}|\pi_{t}^{(2)}(C), W), \ \forall k \in O$$ (16)

for all $\pi_t^{(1)}(C)$ and $\pi_t^{(2)}(C)$ such that $\pi_t^{(1)}(C) \leq \pi_t^{(2)}(C)$. Since $v_{t+1}(\pi_{t+1})$ is nondecreasing in $\pi_{t+1}(C)$, then $v_{t+1}(\pi_{t+1})$ is nondecreasing in $\pi_{t}(C)$ following a proof similar to that of Lemma 1 in Albright (1979). Therefore from (16) and Lemma 4.7.2 of Puterman (1994) $\sum_{\ell_{t+1}=0}^{[O]} v_{t+1}(\pi_{t+1})p_{t}(\ell_{t+1}|\pi_{t}, W)$ is nondecreasing in $\pi_{t}(C)$. Therefore $r_{t}(\pi_{t}, W) + \lambda \sum_{\ell_{t+1}=0}^{[O]} v_{t+1}(\pi_{t+1})p_{t}(\ell_{t+1}|\pi_{t}, W)$, i.e., $v_{t}(\pi_{t}, W)$ is nondecreasing in $\pi_{t}(C)$. \qed

**Theorem 2.** Under Assumptions 1, 2 and 3, $v_{t}(\pi_{t})$ is nonincreasing in $\pi_{t}(C)$ for any $t$.

**Proof.** This theorem can be categorized and proved by two different cases:

**Case 1** ($a_{t}^{*}(\pi_{t}(C) = 1) = B, \ \forall t$): Then for any $t$, $v_{t}(\pi_{t}(C) = 1) = R(\pi_{t}(C) = 1)$ and $\frac{dv_{t}(\pi_{t})}{d\pi_{t}(C)}|_{\pi_{t}(C)=1} \geq \frac{dR(\pi_{t})}{d\pi_{t}(C)} \leq 1$ which follows from (4). Therefore $\frac{dv_{t}(\pi_{t})}{d\pi_{t}(C)} \leq 1, \ \forall \pi_{t}(C), \ \forall t$ by the convexity of $v_{t}(\pi_{t})$, i.e., $v_{t}(\pi_{t})$ is nonincreasing in $\pi_{t}(C)$ for any $t$.

**Case 2** (Otherwise): There exists an age $\bar{t}$ such that $a_{\bar{t}}^{*}(\pi_{\bar{t}}(C) = 1) = W$. This case is proved by induction. By the case assumption that there exists an age $\bar{t}$ such that $a_{\bar{t}}^{*}(\pi_{\bar{t}}(C) = 1) = W$, then we have $a_{\bar{t}}^{*}(\pi_{t}) = W, \ \forall \pi_{t} \in \Pi, \ \forall t \geq \bar{t}$ which follows from Assumption 3 and Theorem 1. Thus $v_{t}(\pi_{t}) = R_{t}(\pi_{t})$ and by (6), $v_{t}(\pi_{t})$ is nonincreasing in $\pi_{t}(C)$.

Assuming $v_{t+1}(\pi_{t+1})$ is nonincreasing in $\pi_{t+1}(C)$ for some $t + 1$, by Lemma 5 and 6, we have
\( v_t(\pi_t) = \max \left\{ r_t(\pi_t, W) + \lambda \sum_{\ell_{t+1} \in O} v_{t+1}(\pi_{t+1}) p_t(\ell_{t+1} | \pi_t, W), R_t(\pi_t) \right\} \) is nonincreasing in \( \pi_{t+1}(C) \).

Therefore \( v_t(\pi_t) \) is nonincreasing in \( \pi_t(C) \) for any \( t \). □

**Theorem 3.** Under Assumption 3, there exists a finite age, \( N \), at which it is optimal to discontinue biopsy referral if and only if the following condition is satisfied:

\[
\bar{R}_N(T) - \bar{R}_N(C) \leq \mu/f.
\]

**Proof.** We prove the necessity and sufficiency separately in the following two cases. 

**Case 1 (Necessity of \( \mu/f \geq \bar{R}_N(T) - \bar{R}_N(C) \)):** Since \( N \) is a stopping time for biopsy referral, \( a^*_t(\pi_t) = W, \forall \pi_t \in \Pi, \forall t \geq n \). Therefore \( v_N(\pi_N(C) = 1) = \bar{R}_N(\pi_t(C) = 1) \geq R_N(\pi_t(C) = 1) \). Then by (6) and (4) we have \( \bar{R}_N(C) \geq -\mu + (1 - f)\bar{R}_N(C) + f\bar{R}_N(T) \). Therefore \( \mu/f \geq \bar{R}_N(T) - \bar{R}_N(C) \).

**Case 2 (Sufficiency of \( \mu/f \geq \bar{R}_N(T) - \bar{R}_N(C) \)):** \( \frac{\mu}{f} \geq \bar{R}_N(T) - \bar{R}_N(C) \) implies \( \bar{R}_N(C) \geq -\mu + (1 - f)\bar{R}_N(C) + f\bar{R}_N(T) \). By (6) and (4) we have \( \bar{R}_N(\pi_N(C) = 1) \geq R_N(\pi_N(C) = 1) \), i.e., \( a^*_N(\pi_N(C) = 1) = W \). By Assumption 3

\[
a^*_t(\pi_t(C) = 1) = W, \forall t \geq N,
\]

and by Theorem 1 we have

\[
a^*_t(\pi_t) = W, \forall \pi_t \in \Pi, \forall t \geq N.
\]

Therefore \( N \) is a stopping time for biopsy referral. □

**Acknowledgements**

This project was funded in part by grant CMMI 0844511 from the National Science Foundation.

**References**


American Cancer Society. 2010. Cancer facts and figures 2010. Atlanta, GA.


