Screening and Treatment of Chronic Diseases

Lauren N. Steimle and Brian T. Denton

Abstract In recent years, Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs) have found important applications to medical decision making in the context of prevention, screening, and treatment of diseases. In this chapter, we provide a review of state-of-the-art models and methods that have been applied to chronic diseases. We provide a tutorial about how to formulate and solve these important problems emphasizing some of the challenges specific to chronic diseases such as diabetes, heart disease, and cancer. Then, we illustrate important considerations for model formulation and solution methods through two examples. The first example is an MDP model for optimal control of drug treatment decisions for managing the risk of heart disease and stroke in patients with type 2 diabetes. The second example is a POMDP model for optimal design of biomarker-based screening policies in the context of prostate cancer. We end the chapter with a discussions of the challenges of using MDPs and POMDPs for medical contexts and describe some important future directions for research.

1 Introduction

Chronic diseases are the leading cause of death and disablement in many countries [1]. Although these diseases cannot be cured, they can be controlled by screening and treatment. Clinicians are tasked with deciding which screening and treatment options are most beneficial for a patient. These decisions are made sequentially over long periods of a patient's life and are made in an uncertain environment. Although

Lauren N. Steimle
Department of Industrial and Operations Engineering, University of Michigan
1205 Beal Ave, Ann Arbor, MI 48109, e-mail: steimle@umich.edu

Brian T. Denton
Department of Industrial and Operations Engineering, University of Michigan
1205 Beal Ave, Ann Arbor, MI 48109, e-mail: btdenton@umich.edu

clinicians can observe a patient's current health or test results, there is uncertainty in the future progression of the disease, the effect of treatment on the patient, and even the correctness of test results. Medical decisions have grown even more complicated due to the aging patient population. Older patients often have multiple chronic conditions, and treatment for one condition may worsen another. Health care providers have recognized these growing problems and have responded with increased expenditures on data collection and tracking systems. With the growth in medical data comes the need for analytical methodology to convert these data into information. Recently, operations research methods have proven to be powerful tools to analyze these data to guide screening and treatment decisions.

Markov decisions processes (MDPs) are increasingly being used in the analysis of medical decisions, especially chronic disease screening and treatment decisions. Both screening and treatment decisions are characterized by large state spaces that define the severity of the disease, patient-specific clinical risk factors, and medication histories, and these decisions have uncertain outcomes due to differences among patients such as genetic, environmental, and dietary factors. The framework of MDPs lends itself well to these decisions since they are made sequentially over time in a fundamentally stochastic environment. Further, partially observable MDPs (POMDPs) are useful for studying systems in which the true state of the system is not known exactly, which is usually the case when screening for a chronic disease.

Modeling screening and treatment decisions using MDPs is not without its challenges. These clinical decisions take place over long time horizons (sometimes decades) under constraints due to medication conflicts, clinical practice requirements, or budget constraints. Furthermore, the influence of patient's treatment and screening history on future decisions leaves these models subject to the *curse of dimensionality* due to dramatic increase in the size of the state space that can be caused by this history dependence. As a result, optimization of the stochastic and sequential decision making process gives rise to computationally-intensive problems that are difficult to solve, even with state-of-the-art algorithms and computing resources. Fortunately, many of these problems have promising structural properties that can be exploited to achieve meaningful theoretical insights and lead to efficient exact and/or approximation methods.

The remainder of this chapter is organized as follows: in Section 2, we discuss some applications of MDPs to chronic diseases. In Section 3, we discuss how to formulate an MDP/POMDP model in the context of chronic disease and solution methods that can be used to determine optimal policies for these models. In Sections 4 and 5, we give in-depth descriptions of an MDP used for the treatment of type 2 diabetes and a POMDP model used for screening of prostate cancer. We end the chapter with discussion of the open challenges that need to be addressed when using MDP/POMDP models for chronic diseases and some concluding remarks.

2 Background on Chronic Disease Modeling

Surveys of operations research applications in healthcare can be found in [2, 3, 4]. Many of the examples are in the context of healthcare operations management, which has been an important application area for decades. In contrast to operations management, the study of disease screening and treatment policies has a shorter history and is confined to a relatively small, but fast growing, number of topic areas including liver and kidney transplant decisions [5, 6, 7, 8, 9, 10, 11], breast cancer screening [12, 13], intensity modulated radiation therapy [14, 15, 16] and brachy-therapy [17] for cancer treatment, the treatment of HIV [18], and public policy decisions related to the transmission of communicable diseases [19, 20].

MDPs can be used to study sequential decisions made in uncertain environments, which is why they are so powerful for the analysis of chronic disease screening and treatment. Before describing how these models are formulated, we provide some motivation for the study of MDPs in the context of chronic diseases by giving the following examples of clinically-relevant questions that have been answered:

- *At what point should a patient with HIV initiate highly active antiretroviral therapy (HAART)?*
  Human Immunodeficiency Virus (HIV) is a virus that attacks the CD4 white blood cells to the point the body can no longer protect itself against infections and disease. Acquired Immune Deficiency Syndrome (AIDS) is caused by HIV and eventually leads to death. Once someone acquires HIV, the virus will remain in the body for the remainder of that person's life. Highly active antiretroviral therapy (HAART) prevents the virus from multiplying and is the standard treatment for HIV patients. However, it was debated whether to "hit early, hit hard" with HAART, as was the treatment model in the late 1990s, or to wait until the CD4 count falls between 200 and 350 as suggested by more recent guidelines. The authors of [18] used an infinite-horizon MDP with the objective of maximizing a patient's total expected lifetime or quality-adjusted lifetime to answer this open question. The states of the MDP were defined by a patient's CD4 count, and at each monthly decision epoch, the decision was to *"initiate HAART"* or *"wait to initiate HAART"*. The authors proved that the optimal policy prescribes initiating therapy if and only if the CD4 count falls below a certain threshold. The optimal policy suggested that HAART should be initiated earlier supporting the "hit early, hit hard" approach to HIV treatment.

- *When should women receive mammograms to screen for breast cancer?*
  Breast cancer is the second leading cause of cancer death for women in the United States [21]. Detecting breast cancer in its early stages allows for treatment and decreases the risk of a breast cancer mortality. A mammogram is an x-ray image of the breast that can be used to detect breast cancer before a woman develops symptoms. If a mammogram shows a suspicious area, a biopsy can be performed to determine if the abnormality is cancer. While these tests are useful in determining if a patient has cancer, they are not perfect. Mammograms

can lead to radiation exposure and pain, and biopsies are an invasive procedure associated with pain and anxiety. Further, mammograms can give false negative and false positive results. The authors of [22] created a finite-horizon POMDP model to determine personalized mammography screening policies that depend on a woman's personal risk factors and past screening results. The unobservable states represent which stage of cancer the patient has such as no cancer, non-invasive cancer, invasive cancer, invasive cancer under treatment, or death. The actions of this POMDP are *"wait"* and *"mammography"*. If the action chosen is mammography, the decision maker can observe a positive or negative mammogram result. If the action is to wait, the patient can give a self-detection result that is either positive or negative. If a mammogram is positive, the patient will get a biopsy, and if a self-detection is positive, the patient will get a mammogram. With these observations in mind, the decision maker can update her *belief state* which describes the probability that a patient is in any given state given the history of mammogram results. The authors find that a control-limit policy exists that depends on the risk of noninvasive and invasive cancers and that a patient's screening history may affect the decision of whether to perform a mammography or not.

- *When should a patient with end-stage liver disease accept a living-donor transplant?* For patients with end-stage liver diseases such as primary biliary cirrhosis, hepatitis C, and acute failure (fulminants) disease, organ transplantation is the only treatment option. Provided that a patient with end-stage liver disease has a willing living donor, it might seem the patient should receive a transplant as soon as possible. However, depending on the quality of the match with the donor and the current health of the patient, this decision might give a lower expected total lifetime for the patient compared with the decision to wait. To analyze this situation, the authors of [5] create an infinite-horizon MDP model in which the state space is represented by a patient's "Model For End-Stage Liver Disease"(MELD) score. The MELD score quantifies the severity of end-stage liver disease based on laboratory results and is used for the purpose of transplant decisions. Higher MELD scores are associated with more severe liver disease. At each daily decision epoch, the actions are *"transplant"* and *"wait"*. If the decision is to wait, the patient will receive a reward of one life day and then progress probabilistically among the health states or die. Once the decision to transplant is made, the patient transitions into an absorbing state and receives a reward corresponding to the expected life days associated with the health of the patient at the time of the transplantation and the quality of the match with the donor. The authors prove that the optimal policy has a control-limit structure in which the patient will only accept a liver of a given quality if her MELD score is worse than the control-limit. For example, a MELD score of 20 is the control-limit given that the quality of the match has a score of 4. Therefore, a patient with a MELD score of 25 should accept this liver to transplant while a patient with a MELD score of 15 should wait to transplant.

These examples illustrate some treatment and screening decisions that can be analyzed using MDPs. More examples of MDPs used in medicine can be found in the reviews by [23, 24]. This chapter differs from these previous reviews in that we provide an in-depth discussion of how to formulate MDP models for chronic disease screening and treatment problems. We also provide detailed examples that illustrate MDP model formulation, validation, solutions, and interpretation of results. Finally we compare and constrast perfectly observable and imperfectly observable contexts. With this motivation, we will proceed to more formally describe how MDPs can be formulated to generate insights for screening or treating a chronic disease.

3 Modeling Framework for Chronic Diseases

The remainder of this chapter will focus on the modeling framework for MDPs specifically in the context of screening and treatment applications. This section will provide a tutorial on how to formulate, solve, and validate these models. In the following sections, we will provide several examples to illustrate the development of the formulation and potential challenges faced by researchers.

3.1 MDP and POMDP Model Formulation

To build an MDP model of a chronic disease treatment process, one must define the *decision epochs*, *time horizon*, *state space*, *action space*, *transition probabilities*, and *rewards* as they relate to the specific disease and screening/treatment options being considered.

Decision Epochs: Treatment and screening decisions are made at each decision epoch. The length of time between decision epochs for a chronic disease model usually corresponds to the time between treatment and/or screening decisions made by the clinician. For instance, in the case of liver transplantation, decisions about whether to transplant or not could be made daily, while in the case of type 2 diabetes, decisions about which medications to initiate are more likely to be made less frequently (e.g. every 6 or 12 months based on clinical guidelines). Determining the ideal time interval requires some understanding of the disease context and clinical practice.

Time Horizon: Another modeling choice is whether to consider a finite-horizon formulation, in which there are a finite number of decision epochs, or an infinite-horizon formulation. While the patient will die in a finite amount of time, some researchers use an infinite-horizon approach for treatment decisions when the time between epochs is short relative to the length of the horizon over which decisions are made. For example, in organ transplantation, if the decision epochs are daily, it

may be a suitable approximation to use an infinite-horizon. Usually infinite-horizon problems are associated with an absorbing state that is reached with probability 1, such as a post-treatment absorbing state. Moreover, infinite-horizon models are often *stationary*, i.e., model parameters do not vary over time.

State Space: The state space of the model represents the information that would be useful to a clinician when making decisions regarding a patient. A state vector typically includes the patient's health status, demographic information, and relevant medical history.

A patient's health status is usually defined by a number of clinical risk factors or a risk score that can be used by clinicians to predict the severity of a disease or the likelihood of developing a disease. For example, when determining whether or not to transplant a liver, clinicians consider a patient's MELD score which depends on a number of laboratory values that are useful in determining the severity of liver disease. While MELD scores are integer-valued, other metabolic risk factors, such as body mass index (BMI), are continuous. Most MDP models used for medical decisions discretize the true continuous state space to reduce the computation needed to solve the model. A finer discretization may be more representative of the true continuous state space, but it also increases the size of the state space and therefore the computation required to solve the model. Further, a finer discretization will decrease the number of observed transitions for some state-action pairs introducing more sampling error into the estimates of the transition probabilities. [25] provides a discussion of the trade-off between the model error introduced with a more coarse discretization and the sampling error that is associated with a finer discretization.

A patient's demographic information can be important for defining the state space of a model. The dynamics of some diseases vary depending on the demographics of the patient such as age and race. For example, [12] considers age because older women are at higher risk for developing breast cancer, but breast cancer is less aggressive in these women. These dynamics may be important in determining the optimal treatment or screening policies, but incorporating this information might require formulation and validation of unique models for these different populations.

Information about a patient's medical history, such as medication history or history of adverse events, may affect treatment decisions. For example, once a patient has had one heart attack, she is at increased risk to have a second heart attack. Although this history is important, MDP models require that the transitions among states must maintain the Markov property, i.e, the next state may only depend on the current state and the action taken. To maintain this property, it is necessary to incorporate any necessary history of the patient into the state definition. For example, the state definition may include which medications a patient has already initiated or how many adverse events the patient has already had.

In most MDP models of chronic disease, there is an absorbing state representing major complication and/or death. In some models, there are separate death states depending on the cause of death (e.g. death from a heart attack, death from other causes). It may be necessary to use more than one absorbing state when absorbing states that are reachable from a given health state vary or when rewards vary depend-

ing on the absorbing state that is reached. Defining the state space is closely tied to what sources exist to estimate transition probabilities, such as statistical survival models or patient data.

POMDPs are a generalization of MDPs in which the decision maker does not know the state of the system with certainty. This generalization is particularly useful within the context of chronic disease, because often clinicians cannot be 100% sure of the health state of their patients. While screening and diagnostic tests provide valuable information, these tests sometimes give false positive and false negative test results which leaves the true health state of the patient uncertain. In a POMDP, the state space is defined by a *core process* and an *observation process* (also referred to as a *message process*). With respect to chronic diseases, the core process corresponds to the true health of a patient, such as cancer-free, has non-invasive cancer, has invasive cancer, in treatment, or dead. To a clinician, the first three states are unobservable, meaning that the clinician cannot know with certainty the true state of the patient. The observation process corresponds to observable test results, such as a mammogram. The core process and the observation process are tied together probabilistically through an *information matrix* with elements that define probabilities of a particular observation given a particular core state. For example, the decision maker may know the true and false positive and negative rates of a biopsy based on clinical studies. Using Bayesian updating, the relationship between the core and observation processes and the observed test result can be used to continually update a *belief state* over the time horizon of the POMDP. The belief state is a probability distribution describing the believed true state of the system based on the decision maker's past observations. For additional details specific to POMDPs, the reader is referred to [26, 27].

Action Space: To identify the action space of the MDP, one must identify which screening or treatment options to consider. In the case where there is a clear "best" treatment option, the action space might be only two actions: treat the patient with the best therapy or wait. These are typically referred to as *optimal stopping-time problems* in the literature, because the decision maker aims to choose the optimal time to stop the process and enter the absorbing post-treatment state. For instance, deciding when to transplant an organ is usually a stopping-time problem with the action space being transplant or wait to transplant.

For some diseases, it is not clear which therapy is the best or how different therapies should be used together to treat the patient. In these cases, the action space can grow quite large because of the combinatorial nature of the actions. For example, if $M = \{m_1, m_2, ..., m_n\}$ is a set of different drugs that can be used in any combination to treat a patient, the size of the action space is $2^{|M|}$ and thus grows exponentially in the number of treatments considered.

In a POMDP model, the decision maker can take actions to gain information about the state of the system. For example, screening decisions can be modeled using POMDP models where the action space might represent the different types of screening tests available. Performing a screening test may not change the natural progression of the disease, but it can provide the decision maker with valuable in-

formation about the true health state of the patient, which in turn may be used to decide whether to do more invasive testing such as biopsy or radiologic imaging.

Transition Probabilities: The transition probabilities in an MDP model of a chronic disease usually describes the progression of the disease with and without treatment, the probability of an adverse event, and the probability of death. To describe the progression of a disease, a key step is to build a *natural history model*. The natural history model describes how the disease progresses under no treatment. Creating this model can be challenging because medical records will only contain data about patients who have been diagnosed and treated for the disease. To build a natural history model, one can use longitudinal data to estimate the effects of treatment by observing measurements of risk factors before and after a patient starts the treatment. In this way, one could estimate how the disease would progress if no treatment was given to a patient. It is important to note that these measures can be affected by bias associated with patterns that influence which patients are referred for treatment. For example, patients who initiate blood pressure lowering medications would typically have higher than normal blood pressure and may exhibit greater relative reduction in blood pressure than the general population.

When there is a clear "best" therapy, as is the case in optimal stopping-time problems, the modeler is not concerned with the effect of treatment on the transition probabilities. Upon initiating treatment, the patient will transition to an absorbing state representing post-treatment with probability 1. In other cases, the modeler must consider how treatment affects the transition probabilities. Presumably, initiating treatment will lower the probability of having an adverse event or dying from the disease. A recent proliferation of statistical models for estimating the risk of chronic disease complications can provide these inputs for MDPs. For instance, statistical models for type 2 diabetes include: the *Framingham* model [28, 29, 30], the *UKPDS* model [31, 32, 33, 34], and the *ACC/AHA* pooled risk calculator [35]. These models predict the probability of diabetes complications such as cardiovascular events (stroke and coronary heart disease), kidney failure, and blindness. Inputs include gender, race, family history, and metabolic factors like cholesterol, blood pressure, and blood glucose. Treatment can affect some of the inputs to these models and therefore can affect the transition probability to an adverse event state.

Another key input to an MDP model is the probability associated with transitioning to the death state. The probability of death caused by something other than the disease of interest is called *all other cause mortality*. All other cause mortality can have a large impact on treatment decisions. As all other cause mortality increases, treatment can become less beneficial since the probability of a complication or dying from the particular disease of focus for the MDP is not as likely. This is particularly important for chronic diseases that progress slowly. For example, the American Urology Association recommends not screening men for prostate cancer after age 75 because men who have not been diagnosed with prostate cancer by this age are not likely to die from this slowly progressing disease. Estimates for all other cause mortality can typically be found using mortality tables from the Centers for

Disease Control and Prevention (CDC) [36].

Rewards: The rewards and costs in a chronic disease MDP model may be associated with the economic and health implications associated with treatment and screening policies. To determine the relevant rewards and costs, one must identify the perspective of the decision maker: patient, third-party payer (e.g. Blue Cross Blue Shield, Medicare), or a societal perspective that combines these different perspectives. Treating or screening a patient for a chronic disease will offer some reward to the patient, such as a potentially longer life. However, these benefits come at some "cost" to the patient, whether it be a reduction in quality of life, such as side effects due to medication or discomfort due to a screening test, or a financial cost, such as medication or hospitalization expenses. Health services researchers typically use quality-adjusted life years (QALYs) to quantify the quality of a year of life with the discomfort due to medical interventions. A QALY of 1 represents a patient in perfect health with no disutility due to medical interventions and side effects of treatment. As the patient's quality of life decreases, whether from medication side effects or disablement from a disease, the patient's QALY value will tend towards zero. (The reader is referred to [37] for a review of QALYs and other quality of life measures.) Some MDP models are only concerned with maximizing a patient's QALYs. Other models take a societal perspective and attempt to balance the health benefits of treatment with the corresponding monetary costs of medical interventions. To balance these competing objectives, a common approach is to use a *willingness to pay* factor, which assigns a monetary value to a QALY. Values of $50,000 and $100,000 per QALY have commonly been used in the literature; however, the exact value to use is often debated [38].

MDPs are rather data-intensive due to the need for transition probabilities and rewards for each state-action pair. However, after gleaning these inputs from the literature or longitudinal patient data, solving these MDPs can generate meaningful insights into how and when to screen for and treat chronic diseases.

3.2 Solution Methods and Structural Properties

Various algorithms have been developed for solving MDPs. The appropriate method for solving an MDP depends on whether the MDP is formulated as an infinite-horizon or finite-horizon problem and the size of the state and action spaces. Methods such as *policy iteration*, *value iteration*, and *linear programming* have been used to solve infinite-horizon problems, while *backwards induction* is typically used to solve finite-horizon problems. One problem with MDP formulations is that they are subject to the curse of dimensionality. This is seen in MDPs for chronic disease where the size of the state space grows exponentially with the number of health risk factors defining the state. To circumvent this problem, approximation algorithms can be used. There has been a great amount of research on approximate dynamic

programming in general, but these approaches tend to be highly context dependent and very little work has been done in the context of chronic disease. [39, 40] provide a thorough review of approximation methods of MDPs.

Many MDP models for chronic diseases have certain structural properties that can be exploited for computational gains. One such property is the *increasing failure rate* (IFR) property describing the transition probability matrices. In the context of chronic diseases, the IFR property means that the worse the health status of the patient is, the more likely that the health status will become even worse. Usually this ordering naturally follows the severity of the chronic disease, with the ordering of the states defined by a patient's health status. For certain optimal stopping-time problems, it has been shown that the IFR property together with some additional (and nonrestrictive) conditions guarantees an optimal threshold policy (see Chapter 4 of [41]). These conditions have be used in the context of HIV [18], liver disease [5], and type 2 diabetes [42] to prove the existence of an optimal *control-limit policy*. A control-limit policy is one in which one action is used for all states below a certain value (e.g. wait to transplant if the MELD score is below 25) and another action for all states above a certain value (e.g. transplant if the MELD score is at least 25). Proving the existence of a control-limit policy can decrease the computational effort required to solve the MDP model, since the value function does not need to be explicitly calculated for every state-action pair.

POMDPS are generally much more challenging to solve than MDPs. Early methodological studies focused on exact methods that exploit the fact that the optimal value function for a POMDP is convex, and in the finite-horizon case it is piecewise linear and expressible using a finite set of supporting hyperplanes. The first exact method was provided by [43]. The authors proposed an iterative approach to generate supporting hyperplanes at each decision epoch. Due to exponential growth in the number of hyperplanes with respect to the number of decision epochs and observations and the fact that many of the hyperplanes are dominated, the authors further proposed an approach to reduce the number of hyperplanes to a minimal set using a linear programming formulation to identify dominated hyperplanes. Many authors have built on this early approach by developing more efficient ways of pruning unnecessary hyperplanes, including *incremental pruning* [44] and the *witness method* [45]. Exact methods are generally limited to small POMDPs. A well-known approximation approach for moderate-sized POMDPs is based on discretizing the continuous belief state to obtain an approximate finite state MDP. One of the first proposed approaches was the *fixed-grid algorithm* proposed by [46]. Many enhancements, including variable grid based approaches have built on this early idea. The reader is referred to [47] for discussion of finite grid based approaches. Grid based methods are limited in their applicability to large-scale POMDPs. For this reason, it is often necessary to develop approximation methods tailored to particular applications.

3.3 Model Validation

Once an MDP has been solved, it is critical to determine whether the results of the model are valid. Below are some common ways to validate MDP models for chronic diseases.

Expert Opinion: After the MDP has been solved, one can seek the opinion of an expert in the field, such as a clinician or a health services researcher, to determine if the results of the model are realistic. This form of validation is not very strong since it is subjective. Some experts may have differing opinions of whether the model results are actually valid. However, this form of validation is probably the easiest to use and can be a first step in validating the model before turning to more objective procedures.

Independent Study: To validate an MDP, one could compare the results to a model developed independently. For instance, an alternative stochastic model could be compared to the MDP using a reference policy (e.g. an existing screening or treatment guideline.) This approach is particularly useful if there is an existing *gold standard* model to compare to.

Retrospective validation: Retrospective validation compares the results of the MDP to past observed outcomes of an existing patient cohort. If this method of validation is used, one should use a different cohort for calibration of the model and for validation of the model. Using the same cohort to calibrate and validate the model could lead to *optimism bias*.

Prospective Validation: Prospective validation, the gold standard of validation, involves using the model to predict outcomes and comparing the predictions to the actual outcomes. This form of validation is considered very strong, because there is no contamination between data used to calibrate the model and the data used to validate it. However, the outcomes of interest in chronic disease modeling are long-term, which can lead to long periods of time between the obtainment of the results and the validation of the model. As a result, this form of validation is almost never done.

Validating the model is an important step to ensure that the results from the model are useful to clinicians. If the model cannot be validated, the modeler should carefully consider whether the assumptions of the model are justified, if the model parameters are accurate and generalizable to other patient populations, and if the model was implemented without errors. Sensitivity analysis often plays an important role in addressing concerns about inaccuracy of model parameters.

4 MDP Model for Cardiovascular Risk Control in Patients with Type 2 Diabetes

This section presents model formulation, solutions, and analysis of results for an MDP in the context of type 2 diabetes. Advances in medical treatments have extended the average lifespan of individuals and transformed many diseases from life-threatening in the near term to chronic conditions in need of long-term management. Diabetes is a good example. With 9.3% of the U.S. population estimated to have diabetes, it is recognized as a leading cause of mortality and morbidity [48]. The disease is associated with many serious complications such as coronary heart disease (CHD), stroke, blindness, kidney disease, limb amputation, and neurological disorders.

Patients with diabetes are at much higher risk of stroke and CHD events than those without diabetes. The risk of having one of these adverse events is affected by a number of risk factors including gender, race, height, weight, glucose, total cholesterol, high density lipids (HDL - often referred to as "good cholesterol"), and blood pressure (systolic and diastolic). Several medications now exist that can control cholesterol and blood pressure for patients with type 2 diabetes. However, there is considerable disagreement in the health care community about how best to use these medications [49, 50, 51]. Risk models exist to predict an individual patient's probability of complications related to diabetes [29, 30, 31, 32]; but alone they cannot provide optimal treatment decisions. Further, these risk models often give conflicting estimates of patient's risk, which adds another challenge to the decision-making process.

Historically, guidelines for the treatment of cholesterol and blood pressure have been "one size fits all" guidelines that do not account for the different risk profiles of the heterogeneous population. The guidelines for cholesterol treatment and the guidelines for blood pressure treatment in the United States were created by two independent committees. This artificial separation of guidelines for treating risk factors that both influence the risk of CHD and stroke could potentially lead to over-treatment of patients and increases in medical costs. These issues provide great motivation for an MDP approach to treatment planning that combines decisions for cholesterol and blood pressure control.

Recently, MDPs have been used to study the optimal treatment of patients with type 2 diabetes. [42] and [52] analyze the optimal time to initiate statins, the most common drug for managing cholesterol. [53] extends this work to study the effect of imperfect adherence on the optimal policy. [54] uses an MDP to determine the optimal simultaneous management of blood pressure and cholesterol. For the remainder of this section, we use the model in [54] as an example of model formulation, the effect of model parameters, and how the optimal compares to the guidelines. Additionally, we provide new results based on more recent data including a new risk model for estimating the probability of cardiovascular events [35].

4.1 MDP Model Formulation

In this MDP model, patients with type 2 diabetes progress between states defined by blood pressure and cholesterol levels. At every decision epoch, a clinician observes the patient's risk factors (i.e. cholesterol and blood pressure levels) and decides which medications (if any) to prescribe to the patient. This model takes a societal perspective and uses a bi-criteria objective, which balances the goal of having a low discounted medication cost with the goal of primary prevention (i.e. delaying the first occurrence of a CHD event or a stroke). Figure 1 gives a schematic representation of this decision process.

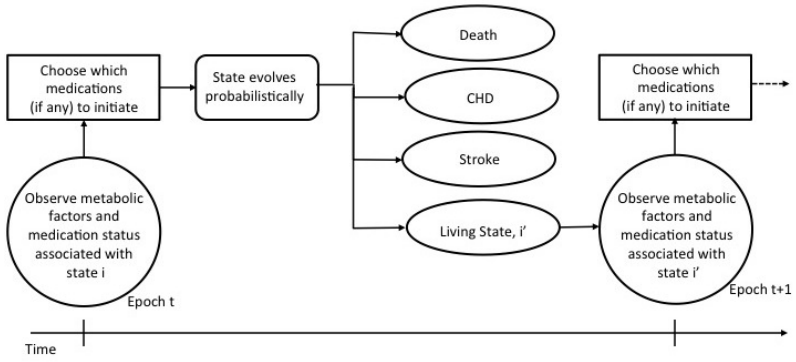

Fig. 1 The treatment decision process for managing cholesterol and blood pressure for patients with type 2 diabetes.

Decision Epochs / Time Horizon: The decision of which medications to initiate is revisited periodically within a finite horizon with N (yearly) decision epochs, with non-stationary rewards and transition probabilities. The set of decision epochs is $\mathscr{T} = \{0,1,2,...,N\}$. It is possible that some patients may live beyond decision epoch N and it is necessary to estimate the rewards that they will accrue in these additional years of life. To do this, an infinite-horizon approximation is used beyond epoch N in which treatment is held constant. This formulation is consistent with regular annual primary care visits for most adults. We assume $N = 100$ because physicians will not typically prescribe new medications to patients after they have reached a certain age.

State Space: The state space is made up of *living states* and *absorbing states*. The set of living states is denoted $S_{\mathscr{L}}$ and the states in this set are defined by a number of factors that characterize a patient's level of cardiovascular risk. Some of these factors, such as metabolic levels and medication status, change over time. Because changes in these values affect the cardiovascular risk, it is important to incorpo-

rate these values into the state space. Other relevant information such as race and gender, is incorporated into the model through the transition probability and reward parameters.

When considering R metabolic factors and M medications, a living state is represented by a vector $\mathbf{s} = \{s_1, ..., s_R, s_{R+1}, ..., s_{R+M}\} \in S_{\mathscr{L}}$. In this model, the first R components of $\mathbf{s}$ correspond to measurements of a patient's total cholesterol, HDL, and systolic blood pressure, and each of the next M components are binary indicators specifying whether or not the patient is on the corresponding medication.

In practice, measurements of cholesterol and blood pressure are continuous. To create a discrete state space, these continuous values are discretized according to clinically-relevant thresholds and then labeled low (L), medium (M), high (H), and very high (V). For metabolic risk factor k, we have $s_k \in \{L, M, H, V\}$. The actual cutpoints for these states were based on a combination of clinical considerations and availability of data for estimating transition probabilities (please see [54] for a detailed discussion of this).

As stated in Section 3, MDPs must maintain the Markov property, and thus any necessary information from the past must be incorporated into the state space. In this model, it is necessary to know whether a patient is already on a medication or not, and therefore this information must be added to the state space. Consider the j^{th} medication: if $s_{R+j} = 0$, the patient is not using medication j and if $s_{R+j} = 1$, the patient is using medication j. Notice that, in this model, the size of the living state space is $|S_{\mathscr{L}}| = 4^R \cdot 2^M$ and therefore the size of the living state space grows exponentially in R and M. Also, if a finer discretization of the metabolic risk factors was used, this growth would be even faster.

The model also has a set of absorbing states $S_{\mathscr{D}}$. These state vectors take on values that represent having a CHD event ($\mathbf{d}^C$), having a stroke ($\mathbf{d}^S$), or dying from a cause other than CHD or stroke ($\mathbf{d}^O$). The set of all absorbing states will be represented as $S_{\mathscr{D}} = \{\mathbf{d}^C, \mathbf{d}^S, \mathbf{d}^O\}$. Because primary prevention is the goal of the model, $\mathbf{d}^S$ and $\mathbf{d}^{CHD}$ are treated as absorbing states and no rewards are accrued after entering these states.

Action Space: Initiating a cholesterol or blood pressure lowering medication is assumed to be an irreversible decision, which is consistent with the clinical practice in which the intention is for the patient to remain on the medication permanently. For each medication j, at each decision epoch, we either initiate this medication (I_j) or wait at least one period to initiate the medication (W_j). Therefore, for a living state, the action space is represented by

$$A(\mathbf{s}) = A_1(\mathbf{s}) \times ... \times A_M(\mathbf{s}) \quad \forall \mathbf{s} \in S_{\mathscr{L}}$$

where M is the total number of medications considered and

$$A_j(\mathbf{s}) = \begin{cases} \{I_j, W_j\} & \text{if } s_{R+j} = 0 \text{ and } \mathbf{s} \in S_{\mathscr{L}}, \\ \{W_j\} & \text{if } s_{R+j} = 1 \text{ and } \mathbf{s} \in S_{\mathscr{L}} \end{cases}$$

This simply means that there is a choice of whether to start medication j or not, provided that the patient is not already on medication j. At every decision epoch t, that decision maker selects an action $a_t \in A(s_t)$ when the system is in state s_t.

Initiating a medication is assumed to have a proportional change on each metabolic factor. Cholesterol medications are designed to lower total cholesterol and raise HDL, while blood pressure medications lower systolic blood pressure. It is assumed that cholesterol medications have negligible effect on blood pressure and vice versa since there is no evidence to the contrary. The estimates of the effects of these drugs on the metabolic values were obtained from [54].

Transition Probabilities There are four types of probabilities in this MDP: the probability of non-diabetes-related death, probability of a CHD event, probability of a stroke, and the transition probabilities among living states. The estimates of these probabilities come from a combination of sources, including the literature and longitudinal patient data, as described below.

At epoch $t \in \mathscr{T}$, a non-diabetes-related death occurs with fixed probability p_t^O for every state $\mathbf{s} \in S_{\mathscr{L}}$. The probability p_t^O depends only on a patient's age and demographic information and can be estimated from mortality tables such as [36]. Note that we assume that p_t^O is independent of the risk factors for type 2 diabetes. If the patient is in state $\mathbf{s} \in S_{\mathscr{L}}$, a CHD or stroke event occurs with probability $p_t^C(\mathbf{s}, a_t)$ and $p_t^S(\mathbf{s}, a_t)$, respectively. These probabilities depend on the patient's age, metabolic state, current and initiated medications, as well as other attributes that affect risk such as race and gender. Estimates of these values can be obtained from risk models such as the *Framingham* model [28, 29, 30], the *UKPDS* model [33, 34], and the *ACC/AHA Pooled ASCVD risk calculator* [35]. These models fit risk equations to observational data for large cohorts of patients followed over many years to predict the probability of having an event within a certain time frame. Some models take the length of the time frame as an input to the equation, which gives an easy way to calculate the probability that corresponds to the time between epochs of the model. However, some models only give 10-year probabilities which must be adjusted to a 1-year probability to be used as an input to the MDP model. [55] provides a discussion of converting the time-interval of transition probabilities to an adverse event or death state under the assumption that the rate of these events is constant. This assumption likely leads to some degree of over-estimation of the yearly transition probability, since the model suggests that as a patient ages, they are more likely to have an adverse event.

If the patient was in state $\mathbf{s} \in S_{\mathscr{L}}$ and did not enter an absorbing state, she will transition probabilistically among the living states, entering state $\mathbf{s}' \in S_{\mathscr{L}}$ with probability $p(\mathbf{s}'|\mathbf{s})$, which is given by

$$p(\mathbf{s}'|\mathbf{s}) = \left(\Pi_{r=1}^{R} p(s'_r|s_r)\right)\left(\Pi_{m=1}^{M} \mathbb{1}(s'_{R+m}|s_{R+m}, a_t)\right) \quad \forall \mathbf{s}, \mathbf{s}' \in S_{\mathscr{L}} \tag{1}$$

The first product in (1) indicates the probability of having the metabolic levels of state $\mathbf{s}'$ given the patient had the metabolic levels of state $\mathbf{s}$. This model assumes that HDL, total cholesterol, and blood pressure progress independently so that the

transition probability of all metabolic factors is simply the product of the transition probabilities within each metabolic factor. For a given metabolic factor, one can estimate the transition probabilities from a longitudinal patient data set (please see [42] for a detailed description of an estimation procedure). After segmenting the continuous values of the factor into discrete groups, one can count the total number of transitions from each group to every other group for the metabolic factor of interest. Dividing through by the total number of transitions out of the given group gives the transition probability. The model used in [54] estimated transition probabilities from observational data on 663 diabetes patients from the Mayo Electronic Records and Diabetes Electronic Management System, which is a medical record containing detailed results of laboratory results such as blood pressure, cholesterol, and HbA1c for diabetes patients at the Mayo Clinic. Note that relaxing the independence assumption of the progression of the metabolic factors would decrease the number of observed samples and therefore the method described above would not be desirable due to large sampling error. This is what motivates the independences assumption, which is supported by relatively low correlation between these risk factors.

In (1), the product of the indicator functions, $\mathbb{1}\{s'_{R+m}|s_{R+m},a_t\}$ is used to distinguish between feasible transitions where $\mathbb{1}\{s'_{R+m}|s_{R+m},a_t\}=1$ if the transition from the medications used in state $\mathbf{s}$ to the medications used in state $\mathbf{s}'$ is valid given the actions taken in time t and 0 otherwise. For example, if a patient in state $\mathbf{s}$ was not on statins and the decision maker did not prescribe statins, then a transition to state $\mathbf{s}'$ in which statins are used is not possible. Since this is not a valid transition, the transition probability will be 0.

The complete set of transition probabilities are summarized in the following equation:

$$p_t(\mathbf{j}|\mathbf{s},a_t)=\begin{cases}[1-p_t^S(\mathbf{s},a_t)-p_t^C(\mathbf{s},a_t)-p_t^0]\cdot p(\mathbf{j}|\mathbf{s}) & \text{if } \mathbf{s},\ \mathbf{j}\in S_{\mathscr{L}}\\ p_t^S(\mathbf{s},a_t) & \text{if } \mathbf{j}=\mathbf{d}^S \text{ and } \mathbf{s}\in S_{\mathscr{L}}\\ p_t^{CHD}(\mathbf{s},a_t) & \text{if } \mathbf{j}=\mathbf{d}^C \text{ and } \mathbf{s}\in S_{\mathscr{L}}\\ p_t^O & \text{if } \mathbf{j}=\mathbf{d}^O \text{ and } \mathbf{s}\in S_{\mathscr{L}}\\ 1 & \text{if } \mathbf{s}=\mathbf{j}\in S_{\mathscr{D}}\\ 0 & \text{otherwise}\end{cases}$$

Rewards: As mentioned above, this model has a bi-criteria objective of maximizing the life years before the first CHD event or stroke while minimizing the discounted medication costs. To balance these competing objectives, we weight a life year (LY) by the willingness to pay factor, β. At epoch t, if the patient is in a living state, one life year is accrued with to give a reward of $r^{a_t}(\mathbf{s})=\beta$. The decision maker also incurs a cost $c^{a_t}(\mathbf{s})$ which is the total yearly cost of the current medications of the patient in state $\mathbf{s}$ as well as any medications initiated by the selected action a_t at epoch t. In other words, the patient continues to accumulate rewards until she incurs a cardiovascular event or dies from other causes.

Solution Method: For a patient in state $\mathbf{s}$ in epoch t, let $V_t(\mathbf{s})$ denote the patient's maximum total expected dollar reward prior to her first CHD or stroke event or death. The following optimality equations define the optimal value function $V_t^*(\mathbf{s})$ and the optimal action in each state based on the optimal value function:

$$V_t^*(\mathbf{s}) = \max_{a_t \in A(\mathbf{s})} \left\{ r_t^{a_t}(\mathbf{s}) - c_t^{a_t}(\mathbf{s}) + \alpha \sum_{\mathbf{j} \in S} p_t(\mathbf{j}|\mathbf{s}, a_t) V_{t+1}^*(\mathbf{j}) \right\} \tag{2}$$

and

$$a_t^*(\mathbf{s}) \in \arg\max_{a_t \in A(\mathbf{s})} \left\{ r_t^{a_t}(\mathbf{s}) - c_t^{a_t}(\mathbf{s}) + \alpha \sum_{\mathbf{j} \in S} p_t(\mathbf{j}|\mathbf{s}, a_t) V_{t+1}^*(\mathbf{j}) \right\} \tag{3}$$

where $\alpha \in [0,1)$ is the discount factor corresponding to the length between epochs, which is commonly set to 0.97 in health economic evaluations involving monetary costs and QALYs (see Chapter 7 of [56] for justification). $V_{N+1}^*(\mathbf{s})$ is assumed to be the expected discounted dollar reward accrued from period $N+1$ if the patient were to remain on the same medications given by state $\mathbf{s}$. Using $V_{N+1}^*(\mathbf{s})$ as a boundary condition, backward induction can be used to solve the MDP for the optimal decisions for each state and epoch. First, evaluate (2) at $t = N$ and proceed backwards until $t = 1$. The actions $a_t^*(\mathbf{s})$ that define the optimal policy are found by solving (3). Then, one can compare the optimal value function V_1^* to the value function V_1^π for any given policy π, which is of special interest when π is a common guideline used for cholesterol and blood pressure management.

4.2 Results: Comparison of Optimal Policies Versus Published Guidelines

In this section, we compare results for MDP-based policies with published treatment guidelines. In the United States, the guidelines for treatment of blood pressure and cholesterol are published by two independent committees. The Joint National Committee (JNC) is responsible for the American blood pressure guideline, while the Adult Treatment Panel (ATP) is responsible for the cholesterol guidelines [57, 58]. These guidelines have historically been "one size fits all" for diabetes patients and have not taken into account the individual risk profile of a patient. The action space of the model is consistent with the medications that these panels recommend. In this model, we consider statins and fibrates for cholesterol medications, and we consider the following blood pressure medications: thiazides, ACE-inhibitors, beta-blockers, and calcium-channel blockers.

The model in [54] used retrospective validation by comparing the results of the MDP with the outcomes of the patient cohort in the Framingham Heart Study (FHS) [59]. The different outcomes are shown in Table 1. Most of the FHS diabetes patients were diagnosed after age 40 and so these patients provide a lower bound for the

outcomes of patients diagnosed at age 40. The overall patient population of the FHS likely provide an upper bound on the outcomes of diabetic patients.

MDP Model/ Patient Cohort	LYs Before First Event (after age 50)
FHS: Diabetes Patients	14.2 (12.3 - 16.1)
FHS: Overall	21.2 (20.5 - 22.0)
Mason et. al (2014), MDP: No Treatment	18.9
Mason et. al (2014), MDP: U.S. Guideline	21.2

Table 1 Comparison of the expected LYs until the first event after age 50 from the MDP model presented with the model presented in [54] and the Framingham Heart Study (FHS). Confidence intervals are shown for the FHS.

Differences between the FHS and this model could be due to imperfect adherence to guidelines, all other cause mortality, and differences in the underlying risk of the patient population. For example, the risk associated with heart disease and stroke has decreased significantly since the start of the Framingham study in 1948. Differences between the results we present below and those in the earlier model [54] differ because we have updated the model with data, such as all other cause mortality, that has been released since the publication of [54].

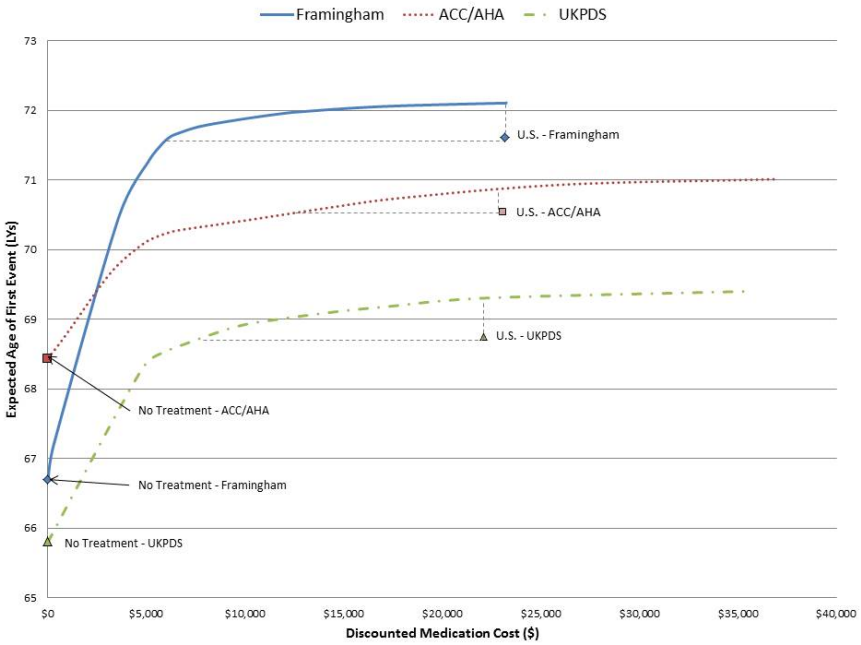

Fig. 2 Comparison of the expected life years until first event and discounted medication costs for optimal treatment policies and U.S. guidelines under different risk model assumptions.

Figure 2 shows the optimal trade-off curve between the expected life years before the first event and the expected discounted medication costs. To obtain each curve, first we specified a risk model to estimate p_t^S and p_t^{CHD}. Then, we solved the corresponding MDP with different values of the willingness to pay factor, β. The labeled points on the vertical axis correspond to a β value of \$0/LY and the optimal policy is to never initiate treatment because there is no weight on the expected life years in this case. As the value of β increases, more medications tend to be initiated leading to increases in life years.

The U.S. guidelines are also shown on the graph. At the time of publication of [54], JNC 7 [60] and ATP III [61] were the guidelines in the United States. We used policy evaluation to determine how well these guidelines performed. Under each risk model assumption, the optimal policy can increase the expected time until the first event for the same medication cost used in the U.S. guidelines. Alternatively, the same life years until the first event that are achieved using these guidelines could be achieved at much lower cost with the optimal policy. See [54] and [62] for additional results and comparison to international guidelines.

Figure 3 shows the optimal initiation of statins under different assumptions of the underlying risk model. The different risk models are functions of the patient's age, systolic blood pressure, HDL, and total cholesterol. The structure of these functions affects the optimal decisions associated with state.

Figure 2 shows that coordinating the treatment of blood pressure and cholesterol could be beneficial for patients with type 2 diabetes under each of the three risk model assumptions. Because the underlying risk of complications is a function of both cholesterol and blood pressure, treating each risk factor separately, as recommended by the U.S. guidelines, could lead to higher cost and lower age of a first complication. This is supported by the outcomes of the U.S. guidelines which give high expected LYs and high discounted medication costs. This work shows that the optimal coordinated treatment of blood pressure and cholesterol depends on the underlying risk of the patient. However, as mentioned above, the risk models used to determine the probability of a complication often conflict with each other. For this reason, it would be beneficial to develop MDP methodology that provides policies that perform well despite disparities between the assumed risk model and the true underlying risk. Methods for achieving this goal remain an open research question.

5 POMDP for Prostate Cancer Screening

Diagnosing chronic diseases is a challenge because most medical tests have some chance of *false positive* or *false negative* results. The former occurs when a test indicates a disease is present, when in fact it is not; the latter indicates a disease is not present, when in fact it is present. Successful diagnosis is critical to starting treatment early, and many chronic diseases, if detected early, have excellent outcomes. Prostate cancer is a good example. It is the most common cancer (excluding skin cancer) that affects men in many countries [21]. It is estimated that one in every

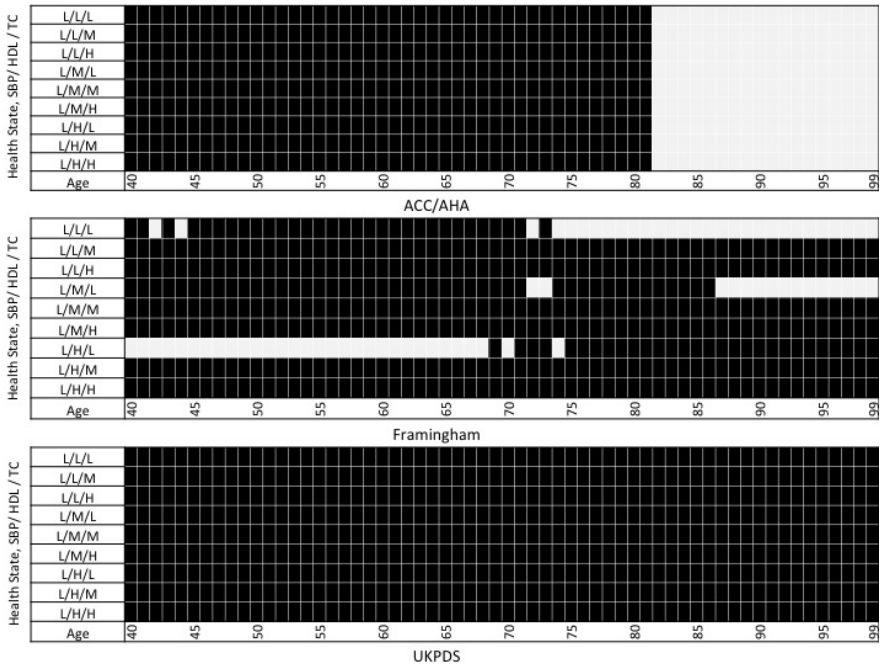

Fig. 3 A comparison of the optimal statin initiation actions under different risk model assumptions for $\beta = \$50{,}000$ per life year for selected blood pressure and cholesterol states. Black boxes indicate that the optimal decision is to initiate statins for this state and a white box indicates that the optimal decision is to wait to initiate statins. L/H/L is the healthiest state shown and L/L/H is the least healthy state shown.

seven U.S. men will be diagnosed with prostate cancer during his lifetime. Diagnosis is often based in part on a Prostate Specific Antigen (PSA) test that measures the amount of PSA in the blood. PSA varies from near zero to potentially high values (e.g. > 20 ng/ml). Men with prostate cancer often have elevated levels of PSA, but this can also be caused by other non-cancerous conditions. A commonly used threshold for asserting that a biopsy is warranted is 4ng/ml; however, this is subjective and it is has been observed to be associated with high false positive and false negative outcomes in the biopsy referral process. Figure 4 illustrates the imperfect nature of PSA testing using a receiver operating characteristic (ROC) curve. An ROC curve is generated by computing the true positive rate (also called *sensitivity* of the test) and one minus the false positive rate (also called *specificity* of the test) for various choices of the test threshold. Thus, the curve in Figure 4 illustrates that, as the PSA threshold for biopsy increases, the true positive rate of biopsy referral based on the PSA test increases and the false positive rate decreases (a perfect test would have a true positive rate of one a false positive rate of zero). Different points on the curve correspond to different choices of the threshold at which to recommend biopsy.

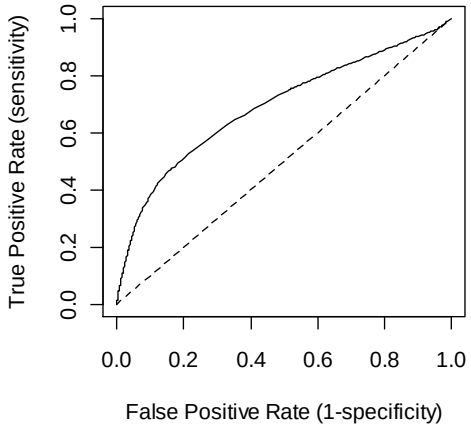

Fig. 4 Receiver operating characteristic (ROC) curve illustrating the imperfect nature of PSA tests for diagnosing prostate cancer.

Given the invasive nature of prostate biopsies, the optimal threshold at which to recommend biopsy is debated. Moreover, the decision for when to biopsy must consider the fact that screening tests are often done multiple times over an individual's lifetime. An example of a screening process is illustrated in Figure 5 where the patient receives routine PSA tests at regular intervals (often every year or every two years). If the PSA test result is over the biopsy threshold then the patient is typically referred for biopsy, and if the biopsy indicates cancer then the patient is referred for treatment. In practice, some clinicians consider the history of PSA test results for a patient, such as the rate of change with respect to time (often referred to as *PSA velocity*) because PSA is expected to increase with respect to tumor volume.

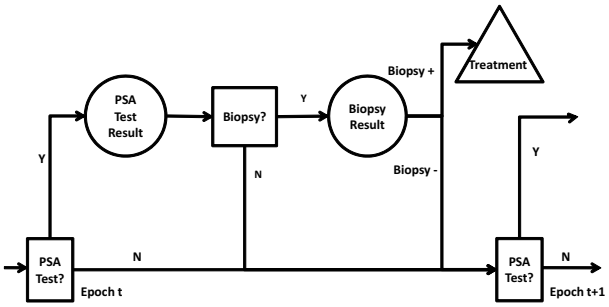

Fig. 5 Illustration of the typical stages of prostate cancer screening and treatment including PSA screening, biopsy, and treatment.

In this section, we present a POMDP model that uses an alternative approach for making screening decisions based on a patient's PSA history. In the model formulation that follows, Bayesian updating is used to estimate the probability that a patient has prostate cancer based on the complete history of PSA results. These probabilities are in turn used to decide when to perform a PSA test, and when to perform a biopsy. The model and the summary results we present are based on work presented in [63, 64] which together provide a complete description of the POMDP model, theoretical analysis of properties of the optimal policies, and a more complete description of the model parameters, results, and conclusions that can be drawn from the model.

5.1 POMDP Model Formulation

In the POMDP model, patients progress through (unobservable) prostate cancer states and (observable) PSA states. PSA states are treated as discrete, based on clinically relevant intervals, and estimated using a large observational data set. We assume that decision epochs occur annually, and the patient's PSA is measured at each epoch and a decision is made about whether to refer the patient for biopsy or defer the decision until the next epoch. Similar to the MDP model of the previous section, the POMDP model is also bi-criteria with the objective of maximizing a measure of life span minus the cost of screening and treatment for prostate cancer. To balance these competing objectives, we use a willingness to pay factor, β. In contrast to the application in Section 4, in this model QALYs are estimated by decrementing a normal life year due to the occurrence of biopsy (which is painful and has some, though low, probability of serious complications due to infection), side effects of treatment, and long term complications resulting from treatment. Costs are based on the cost of PSA tests, biopsies, and subsequent treatment. If a patient receives a positive biopsy result, he is assumed to be treated by prostatectomy (surgical removal of the prostate) which is a common form of treatment. If a patient receives a negative biopsy result, then screening discontinues, an assumption that is motivated by the fact that most men have at most one biopsy in their lifetime unless other symptoms arise warranting additional biopsies. Following is a mathematical description of the model.

Decision Epochs: PSA screening is performed annually, typically starting at age 40, and thus the set of decision epochs is $\mathscr{T} \in \{40, 41, 42, \cdots, N\}$, where N corresponds to a liberal upper bound on when screening is discontinued due to the risk of treatment being greater than the benefits.

Time Horizon: This is a finite horizon model. The rewards accrued for patients that live beyond the last decision epoch N are estimated based on expected future survival given that the patient is not screened beyond epoch N ($N = 95$ in our nu-

merical results).

State Space: 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 detected (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 5.1. Note that state T in part (c) of Figure 5.1 is an aggregate of treatment states *OC*, *EP*, and *LN*. This aggregation is possible since there are no actions in these states.

Action Space: The action at epoch t, $a_t \in \{B, DB, DP\}$, denotes the decision to perform a biopsy (B), defer biopsy and obtain a new PSA test result in epoch $t+1$ (DB), or defer the biopsy decision and PSA testing in decision epoch $t+1$ (DP). Combinations of these three actions over the decision horizon determine the PSA test and biopsy schedule. For instance, $a_{40} = DB$, $a_{41} = DP$, $a_{42} = DB$ and $a_{43} = B$ imply PSA testing at age 41 and 43, and followed by biopsy at age 43. Note that decisions are made sequentially and in this model decisions are based on the probability of prostate cancer at each decision epoch.

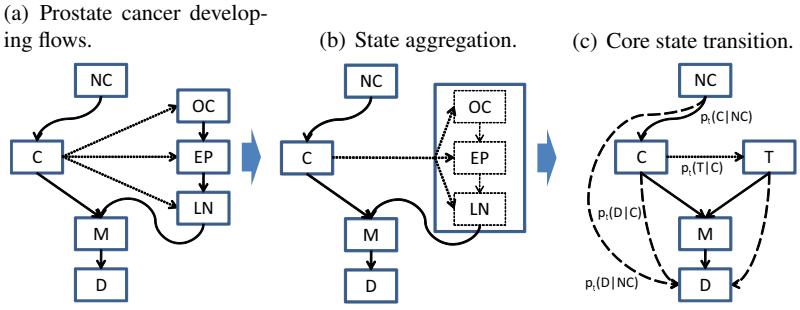

Fig. 6 POMDP model simplification: aggregating the three non-metastatic prostate cancer stages after detection into a single core state T. Solid lines denote the transitions related to prostate cancer; dotted lines denote the action of biopsy and subsequent treatment; dashed lines in (c) denote death from other causes (for simplicity these are omitted from (a) and (b)).

Observations: At each decision epoch, the patient is observed to be in one of a discrete set of observable PSA states based on clinically relevant ranges of PSA, non-metastatic cancer detected and treated (T), metastasis (M), or death (D). These observable states are indexed by $o_t \in O = \{1, 2, 3, ..., m, T, M, D\}$, where the first m states correspond to PSA states for patients either in state *C* or state *NC*. Note that the exact state, *C* or *NC*, cannot be known with uncertainty in this POMDP frame-

work. The *observations* are a unique aspect of POMDP models that arise in the context of imperfect observability, as opposed to the completely observable context of Section 4.

Transition Probabilities: The transition probability $p_t(s_{t+1}|s_t, a_t)$ denotes the core state transition probability from health state s_t to s_{t+1} at epoch t given action a_t. These represent the probability of a change in the patient's health status from one decision epoch to the next. By the nature of partially observable problems, such data is often difficult or impossible to estimate exactly. In the context of prostate cancer these estimates can be obtained using autopsy studies, in which all all fatalities within a given region, regardless of cause of death, are investigated to determine the presence and extent of prostate cancer [65]. This provides estimates of the true incidence of disease that are not biased by the fact that diseases like prostate cancer may be latent for an extended period of time before diagnosis.

Information Matrix: A unique part of POMDP models, compared to MDP models, is the set of conditional probabilities that relate the underlying core states (e.g. *C* or *NC*) to the observations (e.g. PSA states). We let $u_t(o_t|s_t)$ denote the probability of observing $o_t \in O$ given health state $s_t \in S$. Collectively, these transition probabilities define the elements of the *information matrix*, which we denote by U_t. The estimation of these probabilities requires data that can link the observations to the cores states. Often this is one of sets of model parameters that is the most difficult to estimate, because problems that are ideally modeled as partially observable are naturally ones in which limited data is available for the underlying core state of the system. Estimation of the information matrix is often made possible by a systematic randomized trial that evaluates the presence of disease independent of whether a patient has symptoms. In the case of prostate cancer, the *Prostate Cancer Prevention Trial (PCPT)* [66] had a protocol in which all men were biopsied independent of their PSA level. Based on data from this trial [67] fit a statistical model that can be used to estimated the probability a man has a given PSA level conditional on whether or not they are in state *C* or *NC*.

Belief States: The belief vector is a vector with elements each corresponding to one of the number of core states. In this model, each element corresponds to the probability the patient is in the corresponding core state. We denote the vector by $\mathbf{b}_t = (\mathbf{b}_t(NC), \mathbf{b}_t(C), \mathbf{b}_t(T), \mathbf{b}_t(M), \mathbf{b}_t(D))$, where $\mathbf{b}_t \in B \equiv \{\mathbf{b}_t \in \Re^5 \mid \sum_{i \in S} \mathbf{b}_t(i) = 1, \mathbf{b}_t(i) \geq 0, i \in S\}$. The optimal policy maps the belief states to the action space.

Rewards: $r^{a_t}(s_t)$ is the reward of living for a year given the patient is in health state s_t and decision a_t. The expected reward of living for a year is the average over possible health states: $r^{a_t}(\mathbf{b}_t) = \sum_{s_t \in S} r^{a_t}(s_t)\mathbf{b}_t(s_t)$. In this model, the reward is the product of QALYs and a willingness to pay factor minus the cost of a PSA test, biopsy or treatment, depending on the action a_t. The terminal reward at the end of the horizon, at period N is denoted $r_N(\mathbf{b}_t)$.

The overall objective of the model is to determine the optimal screening policy that maximizes the product of willingness to pay and QALYs minus the costs of screening and treatment over the patient's lifetime. The optimal value function and the corresponding optimal action for the model can be written as follows:

$$V_t^*(\mathbf{b}_t) = \max_{a_t \in \{B,DB,DP\}} \left\{ r^{a_t}(\mathbf{b}_t) + \alpha \sum_{o_{t+1} \in O} V_{t+1}^*(\mathbf{b}_{t+1}) p_t(o_{t+1}|\mathbf{b}_t, a_t) \right\}, \forall \mathbf{b}_t \in B, \quad (4)$$

and the boundary condition at the end of horizon is $V_N(\mathbf{b}_t) = r_N(\mathbf{b}_t), \forall \mathbf{b}_t \in B$. The optimal decision at epoch t in belief state $\mathbf{b}_t$ is

$$a_t^*(\mathbf{b}_t) \in \underset{a_t \in \{B,DB,DP\}}{\arg\max} \left\{ r^{a_t}(\mathbf{b}_t) + \alpha \sum_{o_{t+1} \in O} V_{t+1}^*(\mathbf{b}_{t+1}) p_t(o_{t+1}|\mathbf{b}_t, a_t) \right\}, \forall \mathbf{b}_t \in B,$$

where

$$p_t(o_{t+1}|\mathbf{b}_t, a_t) = \sum_{s_{t+1} \in S} u_{t+1}(o_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \mathbf{b}_t(s_t),$$

and $\alpha \in [0,1)$ is the previously defined 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:

$$\mathbf{b}_{t+1}(s_{t+1}) = \frac{u_{t+1}(o_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \mathbf{b}_t(s_t)}{\sum_{s_{t+1} \in S} u_{t+1}(o_{t+1}|s_{t+1}) \sum_{s_t \in S} p_t(s_{t+1}|s_t, a_t) \mathbf{b}_t(s_t)}, \quad (5)$$

where $\mathbf{b}_{t+1}(s_{t+1})$, the component of the belief vector, $\mathbf{b}_{t+1}$, is a function of o_{t+1}, a_t, and $\mathbf{b}_t$ (for simplicity, this dependence is omitted). Thus (5) updates the belief state of a patient based on the prior belief state and his most recent observed PSA interval. The sequence of probabilities $\{\mathbf{b}_t, t = 1, \cdots, \infty\}$ has been shown to follow a Markov process [26], and therefore (4) defines a continuous state MDP.

5.2 Results: Optimal Belief-Based Screening Policy

In this section, we present examples based on the above POMDP model (complete details about model parameter estimates and numerical results can be found in [64]). The data used for parameter estimation in the model consisted of 11,872 patients from Olmsted County, Minnesota. 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. Prostate cancer probabili-

ties conditional on PSA level were estimated from this data set to obtain the information matrix, U_t. In the results we present, we assume patients detected with prostate cancer were treated by radical prostatectomy. To estimate the annual transition probability from the treatment state, T, to the metastatic cancer state, M, a weighted average of the metastasis rate of three non-metastatic prostate cancer stages based on the Mayo Clinic Radical Prostatectomy Registry (MCRPR) were used. The disease specific transition probability from C to M was based on the metastasis rates reported by [68]. The transition probability from state NC to state C was based on the prostate cancer incidence rate estimated from an autopsy review study reported in [69] that provides estimates of prostate cancer prevalence in the general population in 10-year age intervals. The transition probability from all non-cancer states to state D is age specific and was based on the general mortality rate from the National Vital Statistics Reports [70] minus the prostate cancer mortality rate from the [71]. 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 [70] reports a single all cause mortality rate for ages greater than 95, we assume transition probabilities were fixed after the age of 95, i.e., $N = 95$ in the numerical experiments. The biopsy detection rate was 0.8 based on a study by [65]. To estimate the reward function we assumed an annual reward of 1 for each epoch minus disutilities for biopsy and treatment. Since no estimates of disutility exist yet for prostate biopsy, an estimate based on a bladder cancer study for the occurrence of surveillance cystoscopy [72] was used. We assumed patients treated by prostatectomy experience disutility due to side effects as reported in [73]. It is well known that POMDPs can be converted into an equivalent completely observable MDP on the continuous belief states $\mathbf{b}_t$ [43]. Even so, as noted earlier, POMDP models are typically much more computationally challenging to solve than completely observable MDPs, owing to the continuous nature of the belief state space. However, due to the low dimensionality of the belief state instances of this POMDP, it can be solved exactly using *incremental pruning* [44]. Incremental pruning is an exact solution method for POMDPs that builds upon early work of Monahan [74]. Monohan's approach enumerates the complete set of supporting hyperplanes (referred to as α-vectors) that define the optimal value function for the POMDP. Incremental pruning attempts to reduce the computational effort required by decomposing the set of alpha vectors and pruning unnecessary (dominated vectors). See [75] for a review of POMDP properties and methods including incremental pruning.

The model was validated using a combination of expert opinion, based on feedback from practicing urologists, and comparison of the model results to independent studies. For the latter validation, the POMDP model was used to estimate mean lifespan, proportion of men diagnosed with prostate cancer, and prostate cancer mortality. These results were compared to published estimates from the CDC mortality tables and from longitudinal studies of diagnosis and mortality rates. See Chapter 3 of [76] for complete details of the validation of the model including comparison of the model results to those of a randomized trial.

Figure 7 illustrates the optimal prostate cancer screening policy based on the validated POMDP. Two examples are presented. In the first, only quality-adjusted life

span is considered, and the costs of screening and treatment are not part of the reward function; this can be viewed as the *patient perspective*. In the second example, the reward function defines QALYs, weighted using a societal willingness to pay of $\beta = \$50,000$/QALY [38] minus the cost of screening and treatment; as mentioned earlier, this can be viewed as the *societal perspective* since it weights the benefits of additional quality-adjusted lifespan against the cost of screening and treatment. The belief threshold between the three decisions is illustrated by the lines in the figure. From the figure, it can be observed that the optimal policy is control-limit type, a property that can be proven to hold under certain conditions for this POMDP ([63]). A stopping time for screening occurs when the threshold for biopsy in the figures reaches 1.0. It is notable that there is a stopping time for screening at age 76 for the patient perspective. For the case of the societal perspective the stopping age is 71. The five year difference can be attributed to the cost of screening and treatment in the societal case. Finally, the policies in both examples demonstrates the optimal belief-based thresholds are age-dependent; as patients age, their probability of having prostate cancer must be higher in order for a biopsy referral to be optimal (action *B*). Moreover, the same is true for the decision to perform a PSA test (action *DB*). 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.

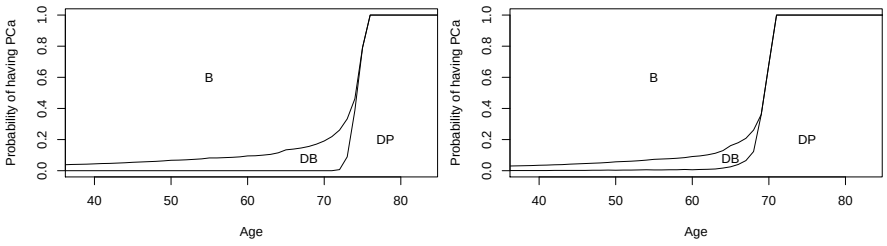

Fig. 7 Optimal prostate cancer screening policies from the patient perspective (left) and the societal perspective (right).

6 Open Challenges in MDPs for Chronic Disease

While MDPs serve as a powerful tool for developing screening and treatment policies for chronic diseases, there exist open challenges in terms of formulating the MDPs and implementing the results from MDPs into clinical settings. We reflect on some of the challenges that were faced in the examples of the previous two sections:

1. *Parameter Uncertainty:* Many estimates of the parameters used in chronic disease MDPs are subject to error. Transition probabilities among living states are

usually estimated from observational data and therefore are subject to sampling error. Transitions to death states and adverse event states are estimated using risk models found in the literature, but usually there is no "gold standard" model. Further, estimates of disutilities due to medications are based on patient surveys and will vary patient-to-patient. As seen in Section 4, MDPs can be sensitive to changes in the model parameters, which is problematic when the model parameters cannot be known with certainty. For this reason, it will be important to develop MDP models that are *robust* in the sense that they perform well under a variety of assumptions of the model parameters, while not being overly conservative. The reader is referred to [77] and [78] for more about robust dynamic programming.

2. *State Space Size and Transition Probability Estimates:* As discussed in Section 3, the continuously-valued metabolic risk factors are usually discretized to reduce the size of the state space. While a finer discretization of the state space might be more representative of the continuous-valued process, this will decrease the sample size of the transitions available for estimating transition probabilities. There is a natural trade-off between the fineness of the discretization of the state space and the error introduced in the transition probabilities due to sampling. Methods for determining the best discretization of the continuous state-space would reduce this barrier to formulating MDPs.
3. *Adjusting the Time Frame of Event Probabilities:* Many risk models provide the probability of an event or death within a fixed time (e.g. 10 years). While this information is useful to clinicians, MDP formulation requires converting these long-term probabilities into transition probabilities between epochs. As mentioned in Section 4, these probabilities can be converted under the assumption that the rate of events is constant, but this may not be realistic in all cases. Determining a method for converting probabilities under different assumptions about the rate of events would improve the accuracy of MDP models that use these risk probabilities.
4. *Solution Methods for a Large-Scale MDPs:* Chronic disease MDPs grow especially large because of the need to incorporate some history-dependence into the state space. Additionally, future models may incorporate risk factors for multiple, coexisting conditions which will cause the state space to grow ever larger. Because MDPs are subject to the curse of dimensionality, these models can be computationally-intensive to solve exactly. To provide support to clinicians in real-time, optimal policies should be able to be solved for quickly. This will not be possible in many chronic disease models, in which case fast approximation algorithms that provide near-optimal solutions will be necessary.
5. *Implementation of Optimal Policies:* The goal of these MDPs is to guide screening and treatment decisions made by the clinician. This requires that optimal policies can be made easily understood to clinicians. However, if the optimal policies are complicated, this could hinder the ability of the clinician to use the MDP results. Therefore, methods for designing structured policies that are near-optimal could potentially improve the likelihood of the policy being implemented in practice.

Tackling these challenges could make MDPs an even more useful tool for guiding clinicians and policy-makers in treatment and screening decisions.

7 Conclusions

Screening and treatment decisions for chronic disease are complicated by the long time periods over which these decisions are made and the uncertainty in the progression of disease, effects of medication, and correctness of test results. Throughout this chapter, we discussed a number of challenges that arise when modeling these decisions using MDPs, such as parameter uncertainty and the rapid growth in the size of the state space. Thus, there are still opportunities to study new application areas and develop new methodology, such as robust and approximate dynamic programming methods, for solving models in this context. These challenges notwithstanding, MDPs have recently found important applications to chronic diseases because they provide an analytical framework to study the sequential and dynamic decisions of screening and treating these diseases that develop stochastically over time.

Acknowledgements This material is based in part on work supported by the National Science Foundation under grant number CMMI 0969885 (Brian T. Denton). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. World Health Organization et al. The top 10 causes of death. July 2013. *Available at: who. int/mediacentre/factsheets/fs310/en/*, 2014.
2. M.L. Brandeau, F. Sainfort, and W.P. Pierskalla. *Operations Research and Health Care*. Kluwer Academic Publishers, Boston, 2004.
3. B.T. Denton. *Handbook of Healthcare Operations Management: Methods and Applications*. Springer, New York, 2013.
4. G. S. Zaric. *Operations Research and Health Care Policy*. Springer, 2013.
5. O. Alagoz, L. M. Maillart, A. J. Schaefer, and M.S. Roberts. The optimal timing of living donor liver transplantation. *Management Science*, 50(10):1420–1430, 2004.
6. O. Alagoz, C. L. Bryce, S. M. Shechter, A. J. Schaefer, C.-C. H. Chang, D. C. Angus, and M. S. Roberts. Incorporating biological natural history in simulation models: Empiric estimates of the progression of end-stage liver disease. *Medical Decision Making*, 25:620–632, 2005.
7. O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. Which waiting lists should an end-stage liver disease patient join? Technical report, University of Pittsburgh, 2006.
8. O. Alagoz, L.M. Maillart, A.J. Schaefer, and M.S. Roberts. Choosing among living-donor and cadaveric livers. *Management Science*, 53(11):1702–1715, 2007.
9. D.L. Segev, Gentry S.E., Warren D.S., Reeb B., and R.A. Montgomery. Kidney paired donation and optimizing the use of liver donor organs. *Journal of the American Medical Association*, 295:1655–1663, 2005.
10. S.A Zenios, G. M. Chertow, and L.M. Wein. Dynamic allocation of kidneys to candidates on the trasplant waiting list. *Operations Research*, 48(4):549–569, 2000.

11. X. Su and S. Zenios. Patient choice in kidney allocation: The role of the queuing discipline. *Manufacturing and Service Operations Management*, 6(4):280–301, 2005.
12. L.M. Maillart, J.S. Ivy, D. Kathleen, and S. Ransom. Assessing dynamic breast cancer screening policies. *Operations Research*, 56(6):1411–1427, 2008.
13. J. Chhatwal, O. Alagoz, and E.S. Burnside. Optimal breast biopsy decision-making based on mammographic features and demographic factors. *Operations research*, 58(6):1577–1591, 2010.
14. E.K. Lee, T. Fox, and I. Crocker. Integer programming applied to intensity-modulated radiation therapy treatment planning. *Annals of Operations Research*, 119:165–181, 2003.
15. A. Holder. Designing radiotherapy plans with elastic constraints and interior point methods. *Health Care Management Science*, 6:5–16, 2003.
16. F. Preciado-Walters, R. Rardin, M. Langer, and V. Thai. A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer. *Mathematical Programming*, 101:319–338, 2004.
17. E.K. Lee, R.J. Gallagher, D. Silvern, C. Wu, and M. Zaider. Treatment planning for brachytherapy: An integer programming model, two computational approaches, and experiments with permanent prostate implant planning. *Phys. Med. Biol.*, 44:145–165, 1999.
18. S.M. Shechter, M.D. Bailey, A.J. Schaefer, and M.S. Roberts. The optimal time to initiate HIV therapy under ordered health states. *Operations Research*, 56(1):20–33, 2008.
19. E.H. Kaplan. Probability models of needle exchange. *Operations Research*, 43(4):558–569, 1995.
20. G. Zaric and M.L. Brandeau. Optimal investment in a portfolio of HIV prevention programs. *Medical Decision Making*, 21:391–408, 2001.
21. Rebecca Siegel, Jiemin Ma, Zhaohui Zou, and Ahmedin Jemal. Cancer statistics, 2014. *CA: a cancer journal for clinicians*, 64(1):9–29, 2014.
22. T. Ayer, O. Alagoz, and N.K. Stout. A POMDP approach to personalize mammography screening decisions. *Operations Research*, 60(5):1019–1034, 2012.
23. O. Alagoz, H. Hsu, A. J. Schaefer, and M. S. Roberts. Markov decision processes: A tool for sequential decision making under uncertainty. *Medical Decision Making*, 30(4):474–483, 2010.
24. A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts. Modeling medical treatment using Markov decision processes. In M. Brandeau, F. Sainfort, and W. Pierskalla, editors, *Handbook of Operations Research/Management Science Applications in Health Care*, pages 597–616. Kluwer Academic Publishers, 2004.
25. E. Regnier and S. M. Shechter. State-space size considerations for disease-progression models. *Statistics in medicine*, 32(22):3862–3880, 2013.
26. G.E. Monohan. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.
27. W. S. Lovejoy. A survey of algorithmic methods for partially observed Markov decision processes. *Annals of Operations Research*, 28:47–66, 1991.
28. K.A. Anderson, P.M. Odel, P.W.F. Wilson, and W.B. Kannel. Cardiovascular disease risk profiles. *American Heart Journal*, 121:293–298, 1991.
29. P.A. Wolf, R.B. D'Agostino, A.J. Belanger, and W.B. Kannel. Probability of stroke: a risk profile from the Framingham study. *Stroke*, 22(3):312–318, 1991.
30. P.W.F Wilson, R. B. DAgostino, D. Levy, A. M. Belanger, H. Silbershatz, and W. B. Kannel. Prediction of coronary heart disease using risk factor categories. *Circulation*, 97(18):1837–1847, 1998.
31. R. C. Turner. The uk prospective diabetes study - a review. *Diabetes Care*, 21:C35–C38, 1998.
32. T.M.E Davis, H. Millns, I.M Stratton, R.R. Holman, and R.C. Turner. Risk factors for stroke in type 2 diabetes mellitus - United Kingdom Prospective Diabetes Study (UKPDS) 29. *Archives of Internal Medicine*, 159(10):1097–1103, 1999.
33. R. J. Stevens, V. Kothari, A. I. Adler, I. M. Stratton, and R. R. Holman. The UKPDS risk engine: a model for the risk of coronary heart disease in type ii diabetes (UKPDS 56). *Clinical Science*, 101(6):671–679, 2001.

34. V. Kothari, R.J. Stevens, A. I. Adler, I. M. Stratton, S.E. Manley, H. A. Neil, R. R. Holman, et al. UKPDS 60 risk of stroke in type 2 diabetes estimated by the UK Prospective Diabetes Study risk engine. *Stroke*, 33(7):1776–1781, 2002.
35. D.C. Goff, D. M. Lloyd-Jones, G. Bennett, C.J. O'Donnell, S. Coady, and J. Robinson. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *J Am Coll Cardiol*, 2014.
36. S.L. Murphy, J. Xu, and K.D. Kochanek. Deaths: final data for 2010. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 61(4):1–117, 2013.
37. D. Gold, M.R. andStevenson and D.G. Fryback. Halys and qalys and dalys, oh my: Similarities and differences in summary measures of population health. *Annu. Rev. Public Health*, 23:115–134, 2002.
38. K.L. Rascati. The $64,000 question – what is a quality-adjusted life year worth? *Clinical Therapeutics*, 28(7):1042–1043, 2006.
39. D.P. Bertsekas and J.N. Tsitsiklis. Neuro-dynamic programming: an overview. In *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, volume 1, pages 560–564. IEEE, 1995.
40. W.B. Powell. *Approximate Dynamic Programming*. John Wiley & Sons, Inc, Hoboken, New Jersey, 2007.
41. M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc, Hoboken, New Jersey, 1994.
42. M. Kurt, B.T. Denton, A.J. Schaefer, N.D. Shah, and S.A. Smith. The structure of optimal statin initiation policies for patients with type 2 diabetes. *IIE Transaction on Healthcare Engineering*, 1:49–65, 2011.
43. R.D. Smallwood and E.J. Sondik. The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.
44. A. Cassandra, M.L. Littman, and N.L. Zhang. Incremental pruning: a simple, fast, exact method for partially observable Markov decision processes. *Proceedings Thirteenth Annual Conference on Uncertainty in Artificial Intelligence, San Francisco, CA*, pages 54–61, 1997.
45. M.L. Littman. The witness algorithm: Solving partially observable Markov decision processes. *Brown University, Providence, RI*, 1994.
46. J.E. Eckles. *Optimum replacement of stochastically failing systems*. Department of Electrical Engineering, Stanford University., 1966.
47. W.S. Lovejoy. Computationally feasible bounds for partially observed Markov decision processes. *Operations research*, 39(1):162–175, 1991.
48. Centers for Disease Control, Prevention, et al. National diabetes statistics report: estimates of diabetes and its burden in the united states, 2014. *Atlanta, ga: US Department of health and human services*, 2014.
49. V. Snow, M. D. Aronson, E. R. Hornbake, C. Mottur-Pilson, and K. B. Weiss. Lipid control in the management of type 2 diabetes mellitus: A clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine*, 140(8):644–649, 2004.
50. D. G. Manuel, K. Kwong, P. Tanuseputro, J. Lim, C. A. Mustard, G. M. Anderson, S. Ardal, D. A. Alter, and A. Laupacis. Effectiveness and efficiency of different guidelines on statin treatment for preventing deaths from coronary heart disease: modelling study. *British Medical Journal*, 332(7555):1419–1422, 2006.
51. P.N. Durrington, H. Prais, D. Bhatnagar, M. France, V. Crowley, J. Khan, and J. Morgan. Indications for cholesterol-lowering medication: comparison of risk-assessment methods. *Lancet*, 353(9149):278–281, 1999.
52. B.T. Denton, M. Kurt, N.D. Shah, S.C. Bryant, and S.A. Smith. Optimizing the start time of statin therapy for patients with diabetes. *Medical Decision Making*, 29(3):351–367, 2009.
53. J.E. Mason, D.A. England, B.T. Denton, S.A. Smith, M. Kurt, and N.D. Shah. Optimizing statin treatment decisions for diabetes patients in the presence of uncertain future adherence. *Medical Decision Making*, 32(1):154–166, 2012.
54. J.E. Mason, B.T. Denton, N.D. Shah, and S.A. Smith. Optimizing the simultaneous management of blood pressure and cholesterol for type 2 diabetes patients. *European Journal of Operational Research*, 233(3):727–738, 2014.

55. D.K. Miller and S. M. Homan. Determining transition probabilities confusion and suggestions. *Medical Decision Making*, 14(1):52–58, 1994.
56. M.R. Gold, J.E. Siegel, L.B. Russell, and M.C. Weinstein. *Cost-Effectiveness in Health and Medicine*. Oxford University Press, New Rork, NY, 1996.
57. Paul A James, Suzanne Oparil, Barry L Carter, William C Cushman, Cheryl Dennison-Himmelfarb, Joel Handler, Daniel T Lackland, Michael L LeFevre, Thomas D MacKenzie, Olugbenga Ogedegbe, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the eighth joint national committee (jnc 8). *Jama*, 311(5):507–520, 2014.
58. NJ Stone, JG Robinson, and AH Lichtenstein. 2013 acc/aha guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the american college of cardiology/american heart association task force on practice guidelines. j am coll cardiol 2013 nov 12 [e-pub ahead of print]. correction. *Journal of the American College of Cardiology*, 63(25):3024–3025, 2014.
59. O.H. Franco, E. W. Steyerberg, F. B. Hu, J. Mackenbach, and W. Nusselder. Associations of diabetes mellitus with total life expectancy and life expectancy with and without cardiovascular disease. *Archives of internal medicine*, 167(11):1145–1151, 2007.
60. A.V. Chobanian, G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo Jr, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright Jr, et al. The seventh report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *Jama*, 289(19):2560–2571, 2003.
61. National Cholesterol Education Program NCEP Expert Panel et al. Third report of the National Cholesterol Education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report. *Circulation*, 106(25):3143, 2002.
62. J. Shah, N.D.and Mason, M. Kurt, B.T. Denton, A. Schaefer, V. Montori, and S. Smith. Comparative effectiveness of guidelines for the management of hyperlipidemia and hypertension for type 2 diabetes patients. *Plos One*, 6(1), 2011.
63. J. Zhang, H. Balasubramanian, B.T. Denton, N. Shah, and B. Inman. Optimization of prostate cancer screening decisions: A comparison of patient and societal perspectives. *Medical Decision Making*, DOI: 10.1177/0272989X11416513, 2011.
64. Jingyu Zhang, Brian T Denton, Hari Balasubramanian, Nilay D Shah, and Brant A Inman. Optimization of prostate biopsy referral decisions. *Manufacturing & Service Operations Management*, 14(4):529–547, 2012.
65. G.P. Haas, R.F. Delongchamps, V. Jones, V. Chandan, A.M. Seriod, A.J. Vickers, M. Jumbelic, G. Threatte, R. Korets, H. Lilja, and G. De la Roza. Needle biopsies on autopsy prostates: Sensitivity of cancer detection based on true prevalence. *Journal of the National Cancer Institute*, 99:1484–1849, 2007.
66. Ian M Thompson, Donna Pauler Ankerst, Chen Chi, Phyllis J Goodman, Catherine M Tangen, M Scott Lucia, Ziding Feng, Howard L Parnes, and Charles A Coltman. Assessing prostate cancer risk: results from the prostate cancer prevention trial. *Journal of the National Cancer Institute*, 98(8):529–534, 2006.
67. R. Gulati, L. Inoue, J. Katcher, W. Hazelton, and R. Etzioni. Calibrating disease progression models using population data: a critical precursor to policy development in cancer control. *Biostatistics*, 11(4):707–719, 2010.
68. W. J. Catalona, P. T. Scardino, J. R. Beck, B. J. Miles, G. W. Chodak, and R. A. Thisted. Conservative management of prostate cancer. *New England Journal of Medicine*, 330(25):1830–1832, 1994.
69. L. Bubendorf, A. Schöpfer, U. Wagner, G. Sauter, H. Moch, N. Willi, T. C. Gasser, and M. J. Mihatsch. Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Human pathology*, 31(5):578–583, 2000.
70. M. Heron. Deaths: Leading causes for 2004. *National Vital Statistics Reports*, 56(5):1–96, 2007.
71. National Cancer Institute. Surveillance epidemiology and end results.

72. Girish S Kulkarni, Shabbir MH Alibhai, Antonio Finelli, Neil E Fleshner, Michael AS Jewett, Steven R Lopushinsky, and Ahmed M Bayoumi. Cost-effectiveness analysis of immediate radical cystectomy versus intravesical bacillus calmette-guerin therapy for high-risk, high-grade (t1g3) bladder cancer. *Cancer*, 115(23):5450–5459, 2009.
73. K.E. Bremner, C.A.K.Y. Chong, G. Tomlinson, S.M.H Alibhai, and M.D Krahn. A review and meta-analysis of prostate cancer utilities. *Med Decis Making*, 27:288–298, 2007.
74. G.E. Monohan. A survey of partially observable Markov decision processes: Theory, mondels, and algorithms. *Management Science*, 28(1):1–16, 1982.
75. Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1):99–134, 1998.
76. D Underwood. *Risk-based Simulation Optimization of PSA-based Prostate Cancer Screening*. PhD thesis, North Carolina State University, 2015.
77. G.N. Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2):257–280, 2005.
78. A. Nilim and L.E. Ghaoui. Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, 55(5):780–798, 2005.