Received: 24 March 2020 | Accepted: 26 June 2022

DOI: 10.1111/poms.13800

POMS

ORIGINAL ARTICLE

Optimization of active surveillance strategies for heterogeneous patients with prostate cancer

Zheng Zhang[1] | **Brian T. Denton**[2] | **Todd M. Morgan**[3]

[1]Department of Service Science and Operations Management, Zhejiang University, Hangzhou, Zhejiang, China

[2]Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, USA

[3]Division of Urologic Oncology, Urology, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Brian T. Denton, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA.
Email: btdenton@umich.edu

Handling Editor: Z. Max Shen

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 72001187, 71821002; Fundamental Research Funds for the Central Universities; National Science Foundation, Grant/Award Number: CMMI 0844511

Abstract

Prostate cancer (PCa) is common in American men with long latent periods, during which the disease is asymptomatic. Active surveillance is a monitoring strategy commonly used for patients diagnosed with low-risk PCa who may harbor latent high-risk PCa. The optimal monitoring strategy attempts to minimize the disutility of testing while ensuring that the patient is detected at the earliest time when the disease progresses. Unfortunately, guidelines for the active surveillance of PCa are often one-size-fits-all strategies that ignore the heterogeneity among multiple patient types. In contrast, personalized strategies based on partially observable Markov decision process (POMDP) models are challenging to implement in practice given the large number of possible strategies that can be used. This article presents a two-stage stochastic programming approach that selects a set of strategies for predefined cardinality based on patients' disease risks. The first-stage decision variables include binary variables for the selection of periods at which to test patients in each strategy and the assignment of multiple patient types to strategies. The objective is to maximize a weighted reward function that considers the need for cancer detection, missed detection, and cost of monitoring patients. We discuss the structure and complexity of the model and reformulate a logic-based Bender's decomposition formulation that can solve realistic instances to optimality. We present a case study for the active surveillance of PCa and show that our model results in strategies that vary in intensity according to patient disease risk. Finally, we show that our model can generate a small number of strategies that can significantly improve the existing "one-size-fits-all" guideline strategies used in practice.

KEYWORDS
active surveillance, health policy, optimization, stochastic programming

1 | INTRODUCTION

Many chronic diseases have latent periods during which there are no physical symptoms until the later stages. There are many examples of such diseases, including cardiovascular diseases and cancer, which together are the leading causes of death in most developed countries. Chronic diseases, such as these, have better treatment options and health outcomes when they are detected and treated early. For instance, for cancer, a localized tumor may be surgically removed; however, if cancer metastasizes, the treatment options are limited to chemotherapy or radiation therapy, which at best attempt to delay the progression of the disease. For this reason, it is important to identify the high-risk stages of the disease as soon as possible. However, diagnostic tests and procedures used to detect latent diseases can be painful, risky, and expensive. Therefore, strategies for early detection of latent diseases must balance the benefits of early detection with the disutility of diagnostic testing.

Prostate cancer (PCa) is an informative test case because it has clearly defined stages based on disease pathology and long latent periods during which the disease is asymptomatic. PCa is also an important public health challenge because it is the most common cancer in men and the second leading cause of cancer-related death among men in the United States (USCS, 2018). Previously, the norm was to treat all healthy

Accepted by Z. Max Shen, after three revisions.

patients diagnosed with PCa using radical prostatectomy (surgical removal of the prostate gland) or radiation therapy. Recently, evidence has surfaced that low-risk forms of the disease may not benefit from aggressive treatment (Bill-Axelson et al., 2018). As a result, the practice is shifting to the use of *active surveillance*, for men diagnosed with low-risk PCa. The purpose of surveillance is to minimize overtreatment and subsequent morbidity for men with the indolent disease while ensuring that men who progress to high-risk disease receive treatment when appropriate. Early studies suggested that active surveillance has a survival rate similar to that of immediate treatment for patients who are diagnosed with low-risk PCa (Hamdy et al., 2016). This is important because active surveillance can delay or avoid treatment via radical prostatectomy or radiation therapy, which have significant side effects such as sexual dysfunction, incontinence, and other impacts on a patient's health status (Guenther et al., 2019).

The fact that PCa is not directly observable the disease is stochastic, and diagnostic tests are imperfect, can cause disutility to patients, and raise important—and as yet unanswered—questions about the optimal surveillance strategy. The recent availability of data from multiple long-term clinical studies has made it possible to estimate stochastic models for the natural history of patients diagnosed with low-risk PCa (Barnett et al., 2018; Inoue et al., 2018). The approach of Barnett et al. (2018), based on the enumeration of a small list of surveillance strategies, suggested an incentive for optimizing the strategy.

PCa surveillance is a special case of disease monitoring. For example, PCa during surveillance only has two latent states, including low- and high-risk stages, the action to perform is either "test" or "wait," and the test result is dichotomous in which the chance of a false-positive result is zero (a complete description of the problem is given in Section 4). This example could be applied to other disease contexts in which risk can be dichotomized with respect to decisions about whether and when to collect information that comes with a cost or burden to patients. Currently, the commonly used partially observable Markov decision process (POMDP) model (see, e.g., Ayer et al., 2012; Zhang et al., 2012) generates a dynamic strategy based on a probability distribution over the set of core states defining the health status of each patient, and the resulting policies are at odds with the structure of many contemporary strategies used in medical practice. In contrast, all strategies endorsed by professional societies that we are aware of prescribe a predefined schedule of diagnostic tests and procedures (see, for example, Lawrentschuk & Klotz, 2011). This endorsement is partly due to the fact that such strategies are easy to interpret and provide patients with reasonable expectations in the future. Unfortunately, while such strategies are easy to interpret, they are "one-size-fits-all" strategies and lack the personalization of policies based on POMDPs that use patients' risks.

In this article, we study a stochastic programming model for the PCa surveillance problem to design easy-to-implement (static) surveillance strategies that can 1) provide patients with reasonable expectations, 2) allow the coordination of decisions among multiple patient types, and 3) balance the ease of implementation with the desire for individualized strategies for patients. We describe the stochastic programming framework that utilizes a *black box* simulation model based on the enumeration of sample paths of disease progression and detection. The decisions define (a) the selection of periods at which to test patients for each of a predefined number of surveillance strategies and (b) the assignment of patients based on disease risk factors to strategies to maximize a cumulative reward function that considers the goals of early detection and the harms of diagnostic tests and procedures. Figure 1 illustrates the shared decision-making process from perspectives of patients and physicians. We formulate the model as a two-stage stochastic integer program, and we develop a *logic-based Benders decomposition* formulation that exploits the structure of the model. We apply our approach to a case study based on a recently validated stochastic model for the active surveillance of PCa.

The remainder of this article is organized as follows. Section 2 provides a background on the active surveillance of PCa to provide a context for the application of our model. Section 3 reviews the most relevant literature on optimizing latent disease detection problems. A two-stage stochastic integer programming model is described in Section 4 along with some fundamental properties of the model. Section 5 compares the stochastic programming model with a benchmark POMDP approach. Section 6 presents a logic-based Benders decomposition (LBD) approach. Section 7 provides numerical results illustrating the performance of our model and evaluates the incremental benefit of increasing the number of surveillance strategies relative to a one-size-fits-all strategy. Section 8 concludes the article.

2 | BACKGROUND ON ACTIVE SURVEILLANCE FOR PROSTATE CANCER

The intent of active surveillance is to monitor patients diagnosed with a low-risk stage of PCa over time, to detect (possible) progression to a high-risk stage of the disease. Surveillance strategies dichotomize risk based on pathological information obtained from biopsies, clinical staging, and prostate-specific antigen (PSA) tests. Prostate biopsy is the most important method for definitively detecting PCa progression. The result of a biopsy, known as the *Gleason score*, is a discrete score assigned by a pathologist. Patients below the Gleason score threshold are assigned to the low-risk PCa category, while patients above the threshold are attached to the high-risk category. *Definitive treatments* such as radical prostatectomy and radiation therapy are recommended for patients detected with the high-risk stage of PCa who can tolerate these treatments. Active surveillance is recommended for patients with low-risk PCa as a means to delay or avoid treatment. In some cases, elderly patients may not be candidates for such treatments because of the risk of adverse outcomes.

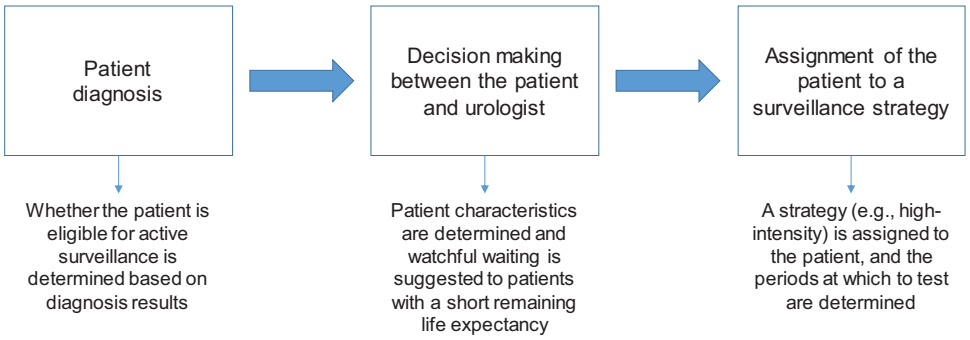

FIGURE 1 Shared decision-making process from the patient and physician perspectives [Color figure can be viewed at wileyonlinelibrary.com]

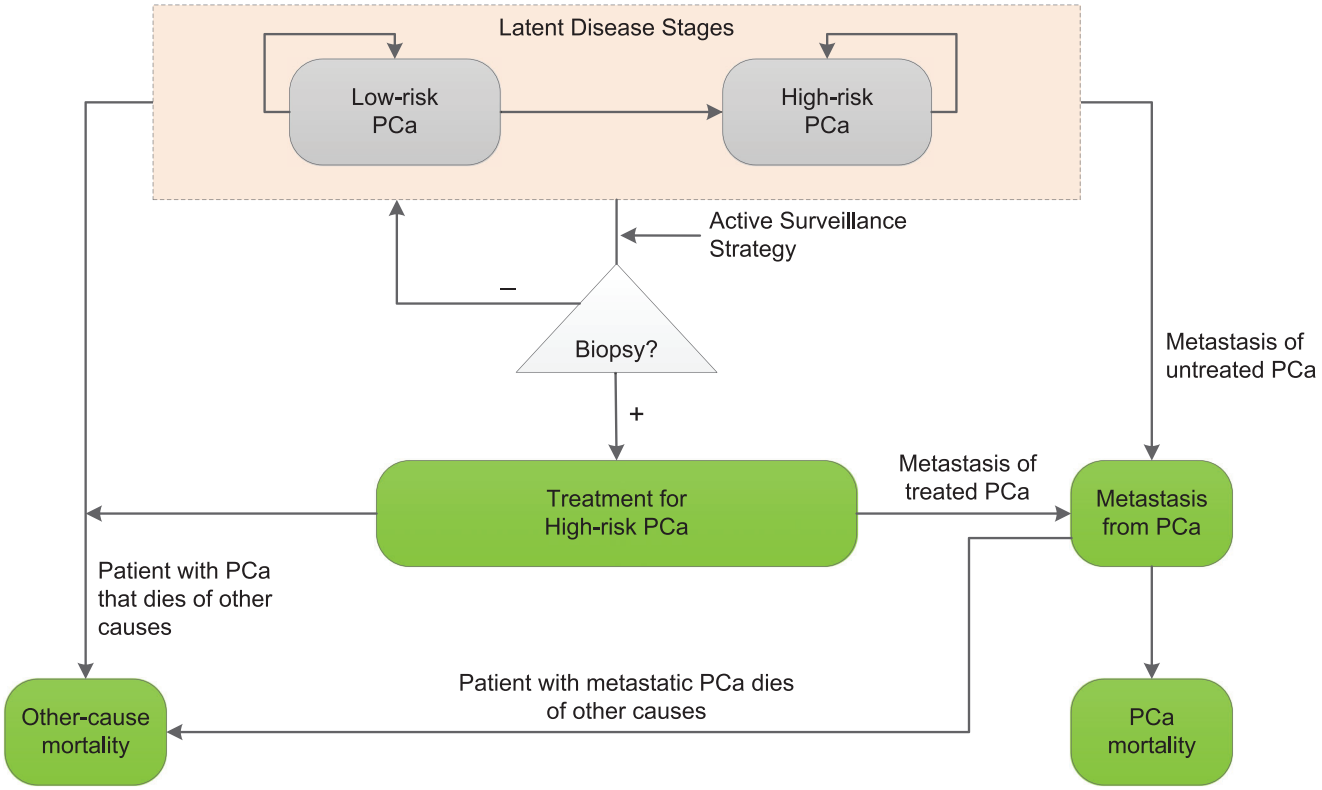

FIGURE 2 Illustration of disease progression and decision-making for active surveillance for PCa [Color figure can be viewed at wileyonlinelibrary.com]

A sketch of the active surveillance process (Figure 2) illustrates the system dynamics, including the progression from low- to high-risk PCa, treatment for high-risk cancer upon detection, and progression from treated PCa or latent high-risk PCa to *metastatic PCa*. Metastatic cancer frequently leads to *PCa mortality*. At the same time, patients in any health stage may also die from other competing diseases, known as *all-other-cause mortality*, where its probability changes with patient's age and it could also differ based on factors such as the presence of chronic diseases (e.g., diabetes, other types of cancer that may be present). Both low- and high-risk PCa stages are latent but may be detected via biopsy. Prostate biopsies are invasive tests that involve the insertion of needles into the patient's prostate gland to remove tissue samples. The tissue was then evaluated by a pathologist who reports the Gleason score. The Gleason score ranges from 6 to 10. Patients with a Gleason score below 7 are generally considered candidates for active surveillance, and those with a score of 7 or above are typically recommended for radical prostatectomy or radiation therapy. The true-negative rate of a biopsy (also known as the *specificity*) is close to 100%; however, the true-positive rate (also known as the *sensitivity*)—as estimated from a large cohort study—is 61% (Barnett et al., 2018). Although the specificity of biopsy is nearly perfect, and it is the gold standard test, it is painful and a source of significant anxiety for patients. In fact, many

TABLE 1 Well-known published strategies based on institutional studies of active surveillance for PCa

Existing strategy	Surveillance plan
Johns Hopkins	Annual biopsy
University of California, San Francisco	Biopsy 1 year after diagnosis, then every 2 years
University of Toronto	Biopsy 1 year after diagnosis, then every 3 years

patients have poor adherence to active surveillance, citing concerns regarding biopsies (Kinsella et al., 2018). Therefore, the frequency of biopsies is an important consideration in the effective implementation of active surveillance. On the other hand, the delay in detecting the high-risk stage of the disease is also an important consideration that is at odds with the goal of minimizing the use of biopsies.

The PSA test is a commonly used biomarker test for PCa progression. This is a simple blood test, in which high values or an increasing trend in PSA levels can be regarded as a signal of PCa progression. However, the test outcome has very poor sensitivity and specificity (Ankerst & Thompson, 2006). Nevertheless, PSA is considered part of the standard practice of care and it often serves as a trigger for additional biopsies (Bokhorst et al., 2015). Such biopsies are not anticipated in advance; therefore, we refer to them as "off-schedule" biopsies.

In our model and case study, we focus on the optimizing the schedule of biopsies that define the surveillance strategy. Several well-known strategies have been published (Table 1), including the Johns Hopkins (JH) strategy (Carter et al., 2007), the University of California, San Francisco (SF) strategy (Dall'Era et al., 2008), and the University of Toronto (UT) strategy (Van den Bergh, 2007). However, the guideline strategies vary significantly, and the best strategy remains unclear. Moreover, all of the guideline strategies are one-size-fits-all strategies, whereas the best policy may vary in intensity from one patient to another owing to the heterogeneity among multiple patient types.

3 | LITERATURE REVIEW

Active surveillance for PCa is a latent disease detection problem, which involves sequential decisions about whether and when to use a diagnostic test. The goal is to schedule diagnostic tests to detect the latent high-risk stage of the disease (e.g., high-risk PCa in our study) as early as possible while trading off this benefit against the disutility of testing. The risk of metastasis of high-risk disease promotes frequent testing, while the pain, risk of infection and bleeding, and cost of testing suggest infrequent testing. In this section, we review the literature related to latent disease detection problems and solution approaches. We close this section with a description of the main contributions of this article to the literature.

3.1 | Latent disease detection

The latent disease detection problem has been studied mainly in the area of chronic disease monitoring and infectious disease control (de Vries et al., 2021; Kamalzadeh et al., 2021; Nenova & Shang, 2022; Tunç et al., 2022). We provide a few representative examples here to illustrate the body of research in this area. Brandeau et al. (1993) studied an infectious disease control problem to analyze the cost and benefit of screening women of childbearing age for the human immunodeficiency virus (HIV). The authors proposed a dynamic compartmental model of the HIV epidemic that incorporates disease transmission and progression over time. Helm et al. (2015) studied a chronic disease monitoring problem to forecast and control glaucoma progression. The authors proposed a multivariate state-space model of disease progression based on a Kalman filter to sequentially forecast the likelihood of progression and determine the next time to test. Bertsimas et al. (2018) proposed a framework to find a decision-making process that works well under multiple models of latent disease progression and applied their method to the PCa-screening problem.

The above problems are similar in a number of aspects. First, there are multiple latent states, and at least one of them requires timely treatment. Second, the detection of the latent disease relies on tests that have measurement errors. Third, decisions are made sequentially based on previous test outcomes. Finally, the test is costly or harmful, and thus it is a trade-off between the cost to test patients and the reward associated with timely treatment.

3.2 | POMDP-based approaches

POMDPs generalize Markov decision processes (MDPs) to the case in which the underlying disease state is not directly observable and is instead denoted as a probability distribution over the set of possible states (also known as *a belief state*) that is updated over time based on observations.

Due to the nature of sequential decision-making, most latent disease detection problems are formulated as POMDP models. Steimle and Denton (2017) reviewed MDP and POMDP models that have been applied to the prevention, screening, and treatment of chronic diseases such as diabetes, heart disease, and cancer. In terms of applications of POMDP models in medicine, Hauskrecht and Fraser (2000) constructed a POMDP framework for the problem of ischemic heart disease management. The authors used a hierarchical Bayesian belief network to denote the disease dynamics, and they exploited regularities and specificity of the problem domain to improve the computational performance of the POMDP model. Vozikis et al. (2012) also considered an ischemic heart disease management problem using a POMDP model. The authors proposed a heuristic method to alleviate the computational obstacle of the POMDP model and trade-off for speed. Ayer et al. (2012) formulated a POMDP model

to design a personalized mammography screening policy for breast cancer screening based on the prior screening history and personal risk characteristics of women. Their POMDP model incorporated age-specific unobservable disease progression and age-specific mammography test characteristics. PCa screening has been the focus of much research using (descriptive) simulation models to evaluate PSA screening of healthy populations (Mühlberger et al., 2017; Underwood et al., 2012). From an optimization perspective, Zhang et al. (2012) formulated the first POMDP model to find an optimal prostate biopsy referral policy that depends on the patient's age and belief. However, since they considered disease detection as an endpoint, and they assumed a single biopsy, their model was a stochastic partially observable stopping time problem and thus differs from the surveillance setting that may perform multiple biopsies in different periods. Sandkç et al. (2013) proposed a POMDP framework for liver transplantation problems that incorporated wait list information into the patient's decision-making. The authors used this model to assess the loss in expected life days for patients under imperfect information as a price of privacy, and they concluded that the imperfect information that is typically available is nearly sufficient to result in a negligible price of privacy. Erenay et al. (2014) formulated a POMDP model to optimize colonoscopy screening policies for detection of colon cancer that incorporate age, gender, and other risk factors. Otten et al. (2017) extended the POMDP model of Ayer et al. (2012) for breast cancer screening by considering multiple risk categories based on the differentiation of the primary tumor. They suggested different screening strategies for different risk groups.

POMDPs are applicable for general latent disease detection problems, including multiple latent states, multiple actions, and general transition probabilities such as cases in colonoscopy screening (Erenay et al., 2014) and breast cancer screening (Otten et al., 2017). However, the POMDP model may have several drawbacks when applied to PCa surveillance practice, which involves two latent states and multiple patients. First, the POMDP solution is a dynamic strategy depending on a belief state that needs to be evaluated by a decision support tool based on historical data that may not be available for all patients. Second, POMDP models are often complex to formulate coordinating decisions among multiple patient types of heterogeneity. Ignoring this heterogeneity leads to a tractable model but also comes at a loss in accuracy; on the other hand, incorporating the coordinating decisions into the state space is associated with exponential increases in computational complexity. Finally, physicians prefer strategies that are defined a priori because they provide guidance on future expectations for patients and physicians.

The most related work to ours is that of Q. Chen et al. (2018), who proposed an MDP-inspired mixed-integer programming (MIP) model for liver cancer screening. They considered an *M-switch strategy* (i.e., a strategy that can switch M times during the time of the surveillance process). The MIP model can be solved efficiently, and thus overcomes the computational challenges of the POMDP model counterpart. However, the MIP model proposed by Q. Chen et al. (2018) is still based on the Markov assumption. Similar to POMDPs, their model generates an individualized strategy for each type of patient, and as the number of patients grows, the physician needs to document a growing number of strategies. Our model differs from Q. Chen et al. (2018) in several aspects: (1) it can constrain the number of strategies, which is critical for implementation; and (2) we develop a tailored solution approach for our model, while the MIP model in Q. Chen et al. (2018) was solved using standard MIP solvers.

3.3 | Our contributions

We formulate a stochastic integer programming (SIP) model to optimize the detection of latent diseases and apply it to active surveillance strategies for PCa. Compared with the aforementioned literature, the contributions of this article are as follows:

- Our approach generates easy-to-implement strategies (structurally consistent with current practice) that predefine a schedule of diagnostic tests, in contrast to the POMDP models commonly proposed in the literature. Using our model, we demonstrate the relative gains in increasing the number of strategies for PCa from a single "one-size-fits-all" strategy, ultimately proving that increasing a smaller number of strategies can achieve significant gains over the one-size-fits-all strategy.
- We develop an improved solution approach based on a tailored LBD to solve our generic SIP model, and we show that it significantly raises the computational efficiency compared to the standard methods using a commercial MIP solver.
- We allow for disease models with the possibility of coordinating decisions among multiple patient types, which would be difficult to address in POMDP models. Moreover, we provide a case study demonstrating the practical use of the model to determine optimal surveillance strategies for patients with PCa for the first time.

4 | STOCHASTIC INTEGER PROGRAMMING MODEL

This section describes a SIP model for the PCa surveillance strategy optimization problem with multiple patient types. This model applies to PCa surveillance with two latent states and biopsy result observations only when no information is updated. The following section will introduce a benchmark model based on the POMDP formulation that applies to problems with more latent states and observations. We then compare the two models regarding their applications to PCa surveillance practice at the end of Section 5.

The surveillance cohort consists of a set of heterogeneous patients of P types, where each patient type $p \in \mathcal{P} \equiv \{1, \dots, P\}$ consists of n_p patients, which correspond to a

TABLE 2 Notation used in the stochastic integer programming model

Notation	Description
Parameters	
p, k, t, s	Index for patient types, periods of PCa progression, periods of treatment (detection), and surveillance strategies, respectively
σ	True positive rate of the biopsy test
d_{pt}	Recommendation on early discontinuation of surveillance at period t for patient type p
ξ_{pk}	Probability of PCa progression at period k for patient type p
$\hat{\xi}_{pk}$	Number of patients who are in type p and PCa progress at period k
$\hat{\pi}_{pkt}$	Adjusted cost of a biopsy test at period t for patient type p when PCa progressed at period k
r_{pkt}	Expected lump-sum reward if PCa is treated at period t for patient type p when PCa progressed at period k
First-stage decision variables	
x_{st}	Whether a biopsy is scheduled at period t in strategy s
y_{ps}	Whether strategy s is assigned to patient type p
w_{pt}	Whether a biopsy is scheduled at period t for patient type p
Second-stage decision variables	
v_{pkt}	Probability of being detected via biopsy at period t for patient type p when PCa progressed at period k
u_{pkt}	Probability of receiving a biopsy test at period t for patient type p when PCa progressed at period k.
q_{pkt}	Probability of being undetected at the beginning of period t for patient type p when PCa progressed at period k

unique combination of patient's characteristics such as the diagnosis age and the enrollment health condition that may affect the value of early treatment and the risk of PCa progression, respectively. The notations used in this model are summarized in Table 2. We denote the probability of an event e by $\mathbb{P}(e)$, the cardinality of a set $\mathcal{X}$ by $|\mathcal{X}|$, and the vector or matrix form of a scalar x by $\mathbf{x}$.

The horizon for active surveillance is fixed at T. As with all published guidelines, we discretize the horizon into a set of periods, $\mathcal{T} \equiv \{1, \dots, T\}$, indexed by t. For PCa, each period corresponds to a year, which is the minimum interval between two consecutive biopsies proposed by all published guidelines. Because patients are treated immediately following a positive test outcome, t can also index the periods of treatment if we slightly extend set $\mathcal{T}$ to be set $\mathcal{T}^+ \equiv \{1, \dots, T+1\}$, where patients at period $T+1$ are no longer candidates for standard treatment (i.e., period of no treatment). Moreover, we include the decision to discontinue surveillance when treatment is no longer beneficial relative to other-cause mortality. We include this decision to ensure that all patients have the same length of surveillance, T, while some types of patients may discontinue earlier if surveillance is no longer beneficial. We define $d_{pt} \in \{0, 1\}$ such that $d_{p1} \geq d_{p2} \geq \cdots \geq d_{pT}$ for all p; $d_{pt} = 0$ denotes the discontinuation of surveillance at period t and beyond, and $d_{pt} = 1$ otherwise. Typically, we estimate $t^* = \arg\max_t \{d_{pt} = 1\}$ as the oldest age for patient type p beyond which the patient will not benefit from treatment, and t^* is independent of the first-stage decision variables. Therefore, d_{pt} can be predetermined based on the patient's expected life span, and it is regarded as a parameter in our model.

The maximum time for PCa progression is K. However, the period at which PCa progression occurs for patients of type p, K_p, is randomly distributed over a discrete set $\mathcal{K} \equiv \{0, 1, \dots, K+1\}$ indexed by k, where $\mathsf{K}_p = K+1$ if PCa does not progress by the end of K periods, and $\mathsf{K}_p = 0$ if PCa has already progressed at the time of diagnosis (indicating that the patient was misdiagnosed, due to a false-negative result for the diagnostic biopsy). The distribution of K_p is specific to p, because the patient type is associated with the rate of PCa progression through the classification of patients for active surveillance. We let $\xi_{pk} = \mathbb{P}(\mathsf{K}_p = k)$ denote the probability of PCa progression at period k for patient type p. Note that PCa may progress during the surveillance horizon, resulting in k and t having parallel indices of periods defined for PCa progression and surveillance, respectively, that is, k and t share the same definition of the period which corresponds to "one year," and they are addable. We assume that $K \geq T$ since patients may also experience PCa progression after T when surveillance is no longer recommended for all patient types.

A surveillance strategy is defined as a specified subset of periods in which a biopsy is scheduled. In contrast to the guideline (one-size-fits-all) strategy, we provide a set of S strategies for the surveillance cohort where each strategy $s \in \mathcal{S} \equiv \{1, \dots, S\}$ is implemented on a tailored subgroup of the patient types, aiming to improve the precision of surveillance with a limited number of strategies. We assume $S < P$ to avoid the triviality of the decision. The strategy design and the patient type assignment are determined by the following first-stage decision variables:

- $x_{st} \in \{0, 1\}$, which denotes whether a biopsy is scheduled at period t in strategy s ($x_{st} = 1$), or not ($x_{st} = 0$) for all $s = 1, \dots, S$ and $t = 1, \dots, T$.
- $y_{ps} \in \{0, 1\}$, which denotes whether patient type p is assigned to strategy s ($y_{ps} = 1$), or not ($y_{ps} = 0$) for all $p = 1, \dots, P$ and $s = 1, \dots, S$.
- $w_{pt} = \sum_{s \in \mathcal{S}} x_{st} y_{ps}$, which is an auxiliary variable introduced for ease of exposition, denoting whether a biopsy is scheduled at period t for patient type p or not.

In short, the strategy assigned to patient type p is defined as $\Pi_p = \{t \mid w_{pt} = 1\}$.

4.1 | Postprogression parameters and second-stage decisions

We make the following two assumptions (Assumptions 1 and 2) to ensure that the period of PCa progression, K_p, is an

exogenous random variable, which is helpful in formulating the problem as a two-stage stochastic integer program.

Assumption 1. The rates of PCa progression, PCa metastasis, PCa mortality, and other-cause mortality are independent of biopsy testing.

Several clinical studies such as a study of more than 2000 patients by the Mayo Clinic found that cancer biopsies do not promote cancer spread (Mayo Clinic, 2015), and similar assumptions (i.e., clinical tests do not affect the rate of disease progression) are common in the literature on cancer screening (Bertsimas et al., 2018; Zhang et al., 2012).

Assumption 2. The false-positive rate of the prostate biopsy test is 0.

The false-positive rate is negligible because once cancerous tissue is sampled, it is highly likely to be correctly diagnosed by a trained pathologist. As a result of Assumption 1, the treatment time (immediately after detection) is always after the progression time.

Based on Assumptions 1 and 2, if we observed $\mathsf{K}_p = k$ in a priori, then surveillance (and treatment, if any) does not affect the probability of PCa progression, that is, ξ_{pk} becomes independent of strategy Π_p for all p, implying that the health outcome for a given strategy is decomposable by k. Moreover, patients will not be treated before PCa progression, and thus the sequence of events leading to treatment is fixed as "low-risk PCa → high-risk PCa → treated PCa" (see Section EC.1.1 in the Supporting Information for illustration).

More specifically, since ξ_{pk} is independent of Π_p, the patient's health outcome is purely determined by the following factors: the patient type (i.e., p), the time of PCa progression (i.e., k), and the delay of treatment (i.e., $t-k$). Therefore, a sample path corresponds to the combination of p, k, and t (see Section EC.1.2 in the Supporting Information for illustration). Further, we predetermine the following sample-path parameters:

- $r_{pkt} \in \mathbb{R}$, which is the expected lump sum reward for patient type p being treated at period t when PCa progresses at period k (note that the lump sum reward can equivalently account for posttreatment states such as metastasis and mortality due to their independence of the strategy).
- $\pi_{pkt} \in [0,1]$, which is the proportion of patient type p staying in surveillance at period t when PCa progresses at period k (i.e., $1-\pi_{pkt}$ is the proportion of patients who stop active surveillance due to symptomatic diagnosis of metastatic cancer, or death from any cause, including PCas).

Moreover, patients of type p need to pay an instantaneous cost, c_p, for the disutility of one biopsy test. For ease of presentation, let $\hat{\pi}_{pkt} = c_p \pi_{pkt}$ denote the adjusted cost of a biopsy test incurred on the patients who stay in surveillance. The test always returns a negative outcome for unprogressed PCa, and it randomly returns a positive outcome for progressed PCa with probability σ (which is the true-positive rate of the biopsy test, and $0 < \sigma \leq 1$ because the biopsy test is based on tissue sampling that could miss the presence of high-risk PCa).

Note that $\hat{\boldsymbol{\pi}}$ and $\mathbf{r}$ are sample-path values estimated from a clairvoyant's view of the random variable K_p. Instead, the surveillance strategy determines the probability corresponding to each sample path, and the probability of biopsy for the patient, which can be expressed by the following second-stage decision variables:

- $v_{pkt} \in [0,1]$: the probability that PCa progression is detected via biopsy and treated at the end of period t for patient type p when PCa progressed at period k where $v_{pk,T+1}$ denotes the probability of undetected PCa at the end of surveillance.
- $u_{pkt} \in [0,1]$: the probability that a biopsy test is performed at period t for patient type p when PCa progressed at period k.
- $q_{pkt} = d_{pt}(1 - \sum_{t'=k}^{t-1} v_{pkt'})$: an auxiliary variable denoting the probability of being undetected at the beginning of period t for patient type p when PCa progressed at period k.

Note that the definition of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{q}$ as decision variables could be confusing at first glance; however, they represent probabilities that are defined as a function of the first- and second-stage variables in the second-stage recourse problem.

4.2 | Model formulation

Now, we formulate the complete problem as the following two-stage stochastic integer program based on the enumeration of all sample paths of PCa progression and detection:

$$\mathbf{P}: \max_{\mathbf{x},\mathbf{y},\mathbf{w}} \; Q(\mathbf{x},\mathbf{y},\mathbf{w}) = \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{K}} \hat{\xi}_{pk} \tilde{Q}(\mathbf{x},\mathbf{y},\mathbf{w},p,k), \tag{1a}$$

$$\text{s.t.} \sum_{s\in\mathcal{S}} y_{ps} = 1, \quad \forall p \in \mathcal{P}, \tag{1b}$$

$$-w_{pt} + \sum_{s\in\mathcal{S}} x_{st} y_{ps} = 0, \quad \forall p \in \mathcal{P}, t \in \mathcal{T}, \tag{1c}$$

$$x_{st}, y_{ps}, w_{pt} \in \{0,1\}, \tag{1d}$$

where $\hat{\xi}_{pk} = n_p \xi_{pk}$ denotes the expected number of patients who are of type p and for whom PCa progresses at period k. We include $\hat{\xi}_{pk}$ in the objective function (1a) because the goal is to maximize the total expected reward across all patients. Constraint (1b) ensures that each patient type is assigned to exactly one strategy, and constraint (1c) links x_{st}, y_{ps} with w_{pt} which determines whether a biopsy test is scheduled at period

t for patient type p. Moreover, $\tilde{Q}(\mathbf{x}, \mathbf{y}, \mathbf{w}, p, k)$ is the maximum reward associated with p and k, which is determined by the following second-stage recourse program:

$$\tilde{Q}(\mathbf{x}, \mathbf{y}, \mathbf{w}, p, k) = \max_{\mathbf{u},\mathbf{v},\mathbf{q}} \sum_{t=\bar{k}}^{T+1} r_{pkt} v_{pkt} - \sum_{t\in\mathcal{T}} \hat{\pi}_{pkt} u_{pkt}, \tag{2a}$$

$$\text{s.t. } q_{pkt} + d_{pt} \sum_{t'=k}^{t-1} v_{pkt'} = d_{pt}, \quad \forall 1 \le t \le T, \tag{2b}$$

$$-q_{pkt} w_{pt} + u_{pkt} = 0, \quad \forall 1 \le t \le T, \tag{2c}$$

$$-q_{pkt} w_{pt} \sigma + v_{pkt} = 0, \quad \forall k \le t \le T, \tag{2d}$$

$$v_{pk,T+1} + \sum_{t'=k}^{T} v_{pkt'} = 1, \tag{2e}$$

$$u_{pkt}, v_{pkt}, q_{pkt} \ge 0, \tag{2f}$$

where $\bar{k} = \min(k, T+1)$, $v_{pk0} = 0$ for all p, k, and the summation $\sum_{t'=\bar{k}}^{t-1}(\cdot) = 0$ if $t \le k$. Objective (2a) determines the reward associated with the detection of high-risk PCa minus the cost associated with the disutilities of biopsy testing. Constraint (2b) determines the probability of being undetected at the beginning of period t, which requires that PCa is not detected in previous periods and that the patient continues surveillance at period t. Constraint (2c) determines the probability of receiving a biopsy test at period t, which requires that PCa is undetected at the beginning of period t, and a biopsy is scheduled at period t. Constraint (2d) determines the probability of being detected via biopsy at period t, which requires that the patient is biopsied and the outcome is positive. Finally, constraint (2e) determines the probability of undetected PCa at the end of period T.

Remark 1. **P** is a stochastic program based on the enumeration of sample paths of PCa progression and detection, which differs from the traditional *sample average approximation* based on sampling the random variables. With straightforward modifications, **P** can be applied to optimize different criteria, such as minimizing the *expected delay of detection* of high-risk PCa and maximizing the *expected survival rate* (see Section EC.1.3 in the Supporting Information for illustration).

4.3 | Model linearization

We further show that **P** with a single patient type can be reduced to be a conjunctive normal form satisfiability problem, which is a well-known example of NP-completeness (see Section EC.1.4 in the Supporting Information for the proof). This motivates the methodological developments in the remainder of this section because there unlikely exists a polynomial-time algorithm for the optimal solution of **P**. As we observe, **P** has nonlinear constraint (1c) in the first-stage program and nonlinear constraints (2c) and (2d) in the second-stage program. In this subsection, we linearize the model, which is valid under the following assumption.

Assumption 3. If a patient is in the high-risk PCa stage and has not discontinued surveillance, we assume that treating the patient earlier is at least as good as treating later, that is, $r_{pkt} \ge d_{pt} r_{pkt'}$, $\forall t+1 \le t' \le T+1$ for all $p \in \mathcal{P}, k \in \mathcal{K}, k \le t \le T$.

This assumption is the fundamental underlying reason why treatment is recommended for high-risk PCa. Using Assumption 3, the second-stage recourse function, $\tilde{Q}(\mathbf{x}, \mathbf{y}, \mathbf{w}, p, k)$, is monotonic with respect to v_{pkt}, which is specified by the following lemma.

Lemma 1. *Under Assumption 3, when the first-stage decision variables* $\mathbf{x}$, $\mathbf{y}$, *and* $\mathbf{w}$ *are fixed, and* $d_{pt} = 1$, $\tilde{Q}(\mathbf{x}, \mathbf{y}, \mathbf{w}, p, k)$ *is nondecreasing as* v_{pkt} *increases for all* $p \in \mathcal{P}$, $k \in \mathcal{K}$, *and* $k \le t \le T+1$.

Using Lemma 1, we further obtain the following theorem (proofs of all theorems, lemmas, and propositions are relegated in Section EC.2 in the Supporting Information):

Theorem 1. *Under Assumption 3,* v_{pkt} *and* u_{pkt} *automatically achieve their upper bound and lower bounds, respectively, at the optimal solution for* **P**.

Theorem 1 is helpful for binding v_{pkt} and u_{pkt} when to linearize the formulation. Specifically, constraint (2c) can be reformulated as

$$u_{pkt} - q_{pkt} - \hat{M}_{pt} w_{pt} \ge -\hat{M}_{pt}, \quad \forall 1 \le t \le T, \tag{3}$$

which enforces $u_{pkt} \ge q_{pkt}$ if $w_{pt} = 1$ and $u_{pkt} \ge 0$ otherwise, where $\hat{M}_{pt}$ is a sufficiently large number (which is known as the *big-M* coefficient), and constraint (2d) can be reformulated as

$$v_{pkt} - \sigma q_{pkt} \le 0, \quad \forall k \le t \le T, \tag{4a}$$

$$v_{pkt} - M_{pt} w_{pt} \le 0, \quad \forall k \le t \le T, \tag{4b}$$

which enforces $v_{pkt} \le \sigma q_{pkt}$ if $w_{pt} = 1$, and $v_{pkt} \le 0$ otherwise, where M_{pt} is a *big-M* coefficient. We further linearize the first-stage program by removing its auxiliary decision variables w_{pt}, and we present an extensive linearized model of **P** in Section EC.1.5 in the Supporting Information.

4.4 | Practical properties for surveillance strategy assignment

We introduce valid inequalities by assuming that the assignments of patients in each subset of patient types, $\hat{\mathcal{P}}$, are continuous. Specifically, if patient types $p-1$ and $p+1$ are

assigned to strategy s, then patient type p must be assigned to the same strategy. This results in the following easy-to-assign constraints:

$$y_{p_j s} + y_{p_j,s+1} \geq y_{p_{j-1},s}, \quad \forall 2 \leq j \leq |\hat{\mathcal{P}}|,\ 1 \leq s \leq S-1, \tag{5a}$$

$$y_{p_j S} \geq y_{p_{j-1},S}, \quad \forall 2 \leq j \leq |\hat{\mathcal{P}}|, \tag{5b}$$

where p_j denotes the jth patient type in $\hat{\mathcal{P}}$. Constraints (5a) and (5b) naturally break the symmetry among the assignment decisions, potentially cutting off a large number of equally good solutions. As a result, constraints (5a) and (5b) can significantly improve the computational efficiency.

Finally, we refer to the linearized model of $\mathbf{P}$ (see Section EC.1.5 in the Supporting Information) combined with the easy-to-assign constraints (5a) and (5b) as model $\hat{\mathbf{P}}$. In Section 6, we develop solution methodologies based on Benders decomposition for model $\hat{\mathbf{P}}$ to efficiently solve realistic-scale problem instances.

5 | COMPARISON BETWEEN THE P AND POMDP MODEL

5.1 | Introduction to benchmark POMDP

We consider a discrete-time, finite-horizon POMDP model for surveillance of a single-patient type, in which we suppress the use of the subscript p. We present a POMDP model in Section 5.1.1 that allows for arbitrary numbers of latent states and observations. Then in Section 5.1.2, we present a special case of the POMDP for PCa surveillance with two latent states and biopsy result observations only when no information is updated.

5.1.1 | POMDP formulation

We consider an absorbing state, T, which corresponds to a postsurveillance state such as "treatment" or "mortality." During the surveillance, the disease is in a latent state $S_t \in \mathcal{S}$, where $\mathcal{S}$ is the set of all latent states (e.g., cancer-free, early-stage cancer, and late-stage cancer). Let $b_t(S_t)$ denote the probability of being in state S_t, and thus its vector form $\mathbf{b}_t \equiv \{b_t(S_1), \ldots, b_t(S_{|\mathcal{S}|})\}$ is the belief state at the beginning of period t. Based on $\mathbf{b}_t$, an action, A_t, is selected from the set of all candidate actions (e.g., biopsy and wait), $\mathcal{A}$. After taking A_t, the disease may transition to T at the end of period t with probability $\bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, A_t)$; otherwise, if the disease continues surveillance, the patient receives a routine test that gives an observation, $O_t \in \mathcal{O}$, with probability $\bar{p}_t(O_t \mid \mathbf{b}_t, A_t)$, where $\mathcal{O}$ is the set of all candidate observations. Based on O_t, one can update the belief state from $\mathbf{b}_t$ to $\mathbf{b}_{t+1}$ for the next period according to Bayesian inference (see Section EC.1.6 in the Supporting Information).

Let $V_{t+1}(\mathbf{b}_{t+1})$ denote the optimal value at period $t+1$ when the belief state is $\mathbf{b}_{t+1}$. The POMDP model that aims to maximize the expected value at period t can be formulated as follows:

$$V_t(\mathbf{b}_t) = \max_{A_t \in \mathcal{A}} R_t(\mathbf{b}_t, A_t) + \lambda \Bigg(\bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, A_t)\bar{R}_t(\mathsf{T}) + (1 - \bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, A_t)) \sum_{O_t \in \mathcal{O}} V_{t+1}(\mathbf{b}_{t+1})\bar{p}_t(O_t \mid \mathbf{b}_t, A_t) \Bigg), \tag{6}$$

where $\lambda \in (0, 1]$ is a discount factor, $R_t(\mathbf{b}_t, A_t)$ is the immediate reward if the belief state is $\mathbf{b}_t$ and the action chosen is A_t, and $\bar{R}_t(\mathsf{T})$ is the expected lump sum reward if the disease transitions to the absorbing state at the end of period t after taking A_t.

5.1.2 | POMDP for PCa surveillance

In PCa surveillance, the disease is either at low risk (L) or high risk (H), which constructs the state set as $\mathcal{S} \equiv \{\mathsf{L}, \mathsf{H}\}$. Based on a belief state $\mathbf{b}_t \equiv \{b_t(\mathsf{L}), b_t(\mathsf{H})\}$, the candidate action is either to perform (B) or defer (W) a biopsy, which constructs the action set as $\mathcal{A} \equiv \{\mathsf{B}, \mathsf{W}\}$. Since PCa is treated at the end of period t if the biopsy result is positive, the patient who continues surveillance in the next period must experience a negative biopsy result (if $A_t = \mathsf{B}$) or no result (if $A_t = \mathsf{W}$). Therefore, the Bayesian inference has no impact on updating the belief state since $|\mathcal{O}| = 1$ for all $A_t \in \mathcal{A}$.

Recall that PCa progression is independent of the surveillance strategy (Assumption 1), the belief state can be specified as

$$b_t(S_t) = \begin{cases} \sum_{k=0}^{t} \xi_{pk}, & \text{if } S_t = \mathsf{H}, \\ 1 - \sum_{k=0}^{t} \xi_{pk}, & \text{if } S_t = \mathsf{L}. \end{cases} \tag{7}$$

Thus, (7) updates the belief state as the cumulative probability of PCa progression up to period t, where ξ_{pt} is the probability of PCa progression at period k. Since $|\mathcal{O}| = 1$, we suppress O_t and reformulate the POMDP model (6) for the PCa surveillance setting, which we refer to as the *POMDP-surveillance* model, as follows:

$$V_t(\mathbf{b}_t) = \max \begin{cases} -d_{pt}\tilde{\pi}_{pt} + \lambda \left(\bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, \mathsf{B})\bar{R}_t(\mathsf{T}) + (1 - \bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, \mathsf{B}))V_{t+1}(\mathbf{b}_{t+1}) \right), & \text{if } A_t = \mathsf{B}, \\ \lambda V_{t+1}(\mathbf{b}_{t+1}), & \text{if } A_t = \mathsf{W}, \end{cases} \tag{8}$$

for $t = 1, \ldots, \bar{T}$, where $\bar{T} = \arg\max_{t \in \mathcal{T}} \{d_{pt} = 1\}$ denotes the recommended discontinuation age for patient type p. When PCa is biopsied, that is, $A_t = \mathsf{B}$, the reward, $R_t(\mathbf{b}_t, \mathsf{B}) =$

$-d_{pt}\hat{\pi}_{pt}$ where $\hat{\pi}_{pt}$ is the weighted cost of a biopsy test at period t, and the probability of detection $\bar{p}_t(\mathsf{T} \mid \mathbf{b}_t, \mathsf{B}) = b_t(\mathsf{H})d_{pt}\sigma$ because only high-risk PCa can be detected, and the sensitivity of biopsy is σ. When PCa is not biopsied, that is, $A_t = \mathsf{W}$, it receives no reward and has no chance of detection. The terminal value $V_{\bar{T}+1}(\mathbf{b}_{\bar{T}+1}) = \bar{R}_{\bar{T}+1}(\mathsf{T})$, denoting that all PCa leftovers are untreated. Finally, λ is set to 1, denoting that the decision-maker is risk-neutral.

5.2 | Conditions of equivalence between POMDP-surveillance and P

In the POMDP-surveillance model, the belief state update (7) is independent of the action chosen and the observation seen at each period, which makes its solution equivalent to a *static* policy in the sense that all actions can be made at the beginning period because the dynamic information observed in the following periods has no impact on the decision-making.

In the following theorem, we show that the solution of the POMDP-surveillance model is equivalent to the solution of **P** under a special case of **P** when there is only one patient type, and all parameters $\hat{\pi}_{pkt}$, r_{pkt} are independent of k.

Theorem 2. *Under a special case of* **P** *such that* $P = 1$, $\hat{\pi}_{pkt}$ *and* r_{pkt} *are independent of* k, *that is,* $\hat{\pi}_{pkt} = \hat{\pi}_{pt}$, $r_{pkt} = R_t(\mathsf{T})$, *and* $r_{pk,T+1} = R_{\bar{T}+1}(\mathsf{T})$ *for all* k, *the POMDP-surveillance model achieves the same total expected reward as* **P** *under the respective solutions corresponding to any given policy* Π *such that*

$$A_t^{\Pi} = \begin{cases} \mathsf{B}, & \text{if } w_{pt}^{\Pi} = 1, \\ \mathsf{W}, & \text{if } w_{pt}^{\Pi} = 0, \end{cases} \tag{9}$$

for all $1 \le t \le \bar{T}$, *where* A_t^{Π} *is the action chosen in the POMDP-surveillance model, and* w_{pt}^{Π} *is the first-stage decision of* **P**.

Theorem 2 applies to any feasible strategy including the optimal strategy Π^* that maximizes the total expected reward. Therefore, the solutions of the POMDP-surveillance model and **P** are equivalent when $P = 1$, $\hat{\pi}_{pkt}$ and r_{pkt} are independent of k, which means that the POMDP-surveillance model is a special case of **P**.

5.3 | Comparison between POMDP-surveillance and P

Based on (6) and Theorem 2, we summarize the advantage (A) and disadvantage (D) of **P** compared with the POMDP-surveillance model for multiple patient types and multiple states as follows:

- (A). When $P \ge S$, multiple patient types share the same surveillance strategy to maintain the ease of implementation. For **P**, the number of strategies can be controlled by the assignment decisions, which results in a factor of P additional decision variables. In contrast, for the POMDP-surveillance model, the number of strategies can be controlled through coordinating the actions of multiple patient types (Amato & Oliehoek, 2015), which results in a multiagent POMDP with a joint belief space of all patient types, and it requires expanding the belief space from $|\mathcal{S}|$ to $|\mathcal{S}|^P$. Therefore, as P increases, the belief space of the POMDP-surveillance model increases exponentially with respect to P versus a linear increase with respect to P for **P**.
- (D). **P** is suited for dichotomous health conditions ($|\mathcal{S}| = 2$) in clinical settings; however, it is less straightforward than the POMDP-surveillance model to scale up to account for multiple latent states. This is because, in **P**, ξ_{pk} only defines the probability of an irreversible state transition from low- to high-risk PCa. The model we proposed can be readily extended to the *left to right* Markov models with more than two states (see Section EC.1.7 in the Supporting Information); however, extensions to more general transition dynamics are less straightforward and potentially a topic for future research. On the other hand, from a practical point of view, dichotomous health conditions are quite common in clinical settings where surveillance is conducted to detect a change in health status requiring treatment.

6 | BENDERS DECOMPOSITION OF $\hat{\mathbf{P}}$

Benders decomposition (Benders, 1962) is a well-known technique for solving large-scale SIP models that often have a special block structure when the uncertainty is denoted by scenarios. A comprehensive description of Benders decomposition is provided in Rahmaniani et al. (2017) and van Ackooij et al. (2017).

Following the standard approach for Benders decomposition, we consider a relaxed master problem (**RMP**) of $\hat{\mathbf{P}}$ that determines the first-stage decision variables, $\mathbf{x}$, and $\mathbf{y}$, and sequentially bound the objective value, η, of **RMP** by accumulating optimality cuts over iterations until **RMP** finds a feasible solution. Specifically, the relaxed master problem is defined as follows:

$$\begin{aligned} \mathbf{RMP}: \quad & \max_{(\mathbf{x},\mathbf{y})\in\mathcal{X}} \quad \eta \\ & \text{s.t.} \quad \mathcal{F}_v(\eta, \mathbf{x}, \mathbf{y}) \quad \text{(optimality cut)} \end{aligned} \quad , \tag{10}$$

where $\mathcal{X} = \{(\mathbf{x}, \mathbf{y}) \mid (1b),\ (5a) \text{ and } (5b)\}$ defines the feasible region of $\mathbf{x}, \mathbf{y}$, and $\mathcal{F}_v(\eta, \mathbf{x}, \mathbf{y})$ defines the set of optimality cuts accumulated up to iteration v.

6.1 | Logic-based optimality cuts

Traditional Benders decomposition was unable to adequately solve practical problems owing to the existence of *big-M* coefficients such as those appearing in (3) and (4), which result in a poor linear programming (LP) relaxation of the subproblem and weaken the optimality cut. We consider a LBD that does not exploit the LP relaxation; thus, it may help to mitigate the effect of *big-M* coefficients (Hooker & Ottosson, 2003).

First, to implement the LBD, we consider a *single-strategy problem* (**SSP**) as defined on a given set of patient types, $\widehat{\mathcal{P}}$, in which all patients share the same strategy. Therefore, the first-stage decision variables in **SSP** are restricted to be determining the periods to perform biopsies for a single strategy, that is, $x_t \in \{0,1\}$ for all $t \in \mathcal{T}$. The problem can be formulated as a mixed-integer program as follows:

$$\textbf{SSP}: \ \hat{Q}(\widehat{\mathcal{P}}) = \max_{\mathbf{x},\mathbf{v},\mathbf{u}} \sum_{p\in\widehat{\mathcal{P}}} \sum_{k\in\mathcal{K}} \hat{\xi}_{pk} \left[\sum_{t=k}^{T+1} r_{pkt} v_{pkt} - \sum_{t\in\mathcal{T}} \hat{\pi}_{pkt} u_{pkt} \right], \tag{11a}$$

$$\text{s.t. } v_{pkt} + d_{pt}\sigma \sum_{t'=k}^{t-1} v_{pkt'} \le d_{pt}\sigma, \quad \forall p \in \widehat{\mathcal{P}}, k \in \mathcal{K}, k \le t \le T, \tag{11b}$$

$$\sum_{k\in\mathcal{K}} v_{pkt} \le M_{pt} x_t, \quad \forall p \in \widehat{\mathcal{P}}, 1 \le t \le T, \tag{11c}$$

$$\sum_{t=k}^{T} v_{pkt} = 1, \quad \forall p \in \widehat{\mathcal{P}}, k \in \mathcal{K}, \tag{11d}$$

$$-u_{pkt} - \sum_{t'=k}^{t-1} v_{pkt'} \le 1 - d_{pt} - x_t, \quad \forall p \in \widehat{\mathcal{P}}, k \in \mathcal{K}, 1 \le t \le T, \tag{11e}$$

$$x_t \in \{0,1\}, v_{pkt}, u_{pkt} \ge 0. \tag{11f}$$

Proposition 1. *We define $\delta_{\widehat{\mathcal{P}}}$ as the loss of reward from restricting patients in $\widehat{\mathcal{P}}$ to share the same strategy, where*

$$\delta_{\widehat{\mathcal{P}}} = \sum_{p\in\widehat{\mathcal{P}}} \hat{Q}(\{p\}) - \hat{Q}(\widehat{\mathcal{P}}) \ge 0, \tag{12}$$

where $\hat{Q}(\{p\})$ is the optimal reward corresponding to a single patient type p.

Proposition 1 implies that an increased number of strategies are always beneficial for the total reward associated with a given set of patient types. Therefore, the maximum value of the total expected reward, $\bar{\eta}$, is achieved when every patient type has an individual strategy, which can be determined as follows:

$$\bar{\eta} = \sum_{p\in\mathcal{P}} \hat{Q}(\{p\}). \tag{13}$$

Based on the incumbent solution of **RMP** at iteration v, we construct a subset of patients, $\widehat{\mathcal{P}}_s^v = \{p \in \mathcal{P} \mid y_{ps}^v = 1\}$, corresponding to the patient types who are assigned to strategy s. These patient types are forced to share the same strategy, which is a restriction that may cause a loss of rewards. Directly following Proposition 1, the reward loss in strategy s can be determined by the following constraint:

$$\hat{Q}(\widehat{\mathcal{P}}_s^v) + \delta_{\widehat{\mathcal{P}}_s^v} \le \sum_{p\in\widehat{\mathcal{P}}_s^v} \hat{Q}(\{p\}), \text{ if } y_{ps} = 1 \ \forall p \in \widehat{\mathcal{P}}_s^v, \tag{14}$$

where $\delta_{\widehat{\mathcal{P}}_s^v}$ is the loss of reward from restricting patients in $\widehat{\mathcal{P}}_s^v$ as defined in (12). It means that the loss of reward is enforced if all patient types are assigned to strategy s; otherwise, the constraint is relaxed. This constraint can be reformulated as follows:

$$\hat{Q}(\widehat{\mathcal{P}}_s^v) + \delta_{\widehat{\mathcal{P}}_s^v} \sum_{p\in\widehat{\mathcal{P}}_s^v} y_{ps} \le \sum_{p\in\widehat{\mathcal{P}}_s^v} \hat{Q}(\{p\}) + \delta_{\widehat{\mathcal{P}}_s^v}(|\widehat{\mathcal{P}}_s^v| - 1). \tag{15}$$

Constraint (15) can be applied simultaneously to all the strategies. The total reward loss is the sum of the losses across all strategies, that is, $\sum_{s\in S} \delta_{\widehat{\mathcal{P}}_s^v}$. Then, we add the following optimality cut to $\mathcal{F}_v(\eta, \mathbf{x}, \mathbf{y})$ at iteration v:

$$\eta + \sum_{s\in S} \left(\delta_{\widehat{\mathcal{P}}_s^v} \sum_{p\in\widehat{\mathcal{P}}_s^v} y_{ps} \right) \le \bar{\eta} + \sum_{s\in S} \left(\delta_{\widehat{\mathcal{P}}_s^v}(|\widehat{\mathcal{P}}_s^v| - 1) \right), \tag{16}$$

where η is the objective value of **RMP**, which is the total expected reward currently equal to $\sum_{s\in S} \hat{Q}(\widehat{\mathcal{P}}_s^v)$, and $\bar{\eta}$ is the maximum value of the total expected reward defined in (13).

Remark 2. We also consider a standard Bender's decomposition (SBD) based on a dual LP relaxation of the subproblem that includes many big-M coefficients. Using a computational study, we find that the LBD method significantly outperformed the SBD method and a standard branch-and-cut algorithm. The SBD model and detailed results are provided in Section EC.1.8 in the Supporting Information.

7 | CASE STUDY OF ACTIVE SURVEILLANCE FOR PROSTATE CANCER

In this section, we present a hypothetical case study based on two major surveillance studies conducted in the United States. We describe how we parameterize our model and show a series of numerical results of this case study to demonstrate the benefits of our model solution compared with previously published surveillance strategies. Finally, we present a sensitivity analysis to establish the robustness of our conclusions concerning the variation in model parameters.

TABLE 3 Summary of parameters used in the case study

Notation	Meaning	Value	Reference
c_B	Disutility in the year of biopsy	0.05	Zhang et al. (2012)
$\hat{c}_T$	Annual disutility for posttreatment	0.091[a] (first year), 0.05 (afterwards)	Litwin et al. (2001) Heijnsdijk et al (2012)
c_T	Disutility in the year of treatment	0.247	Heijnsdijk et al. (2012)
$\hat{c}_M$	Annual disutility for metastasis	0.4	Heijnsdijk et al. (2012)
σ	Sensitivity of biopsy	0.61	Barnett et al. (2018)
$\hat{\mathbf{w}}$	Misclassification rate[b]	0.044 (JH[c] cohort), 0.361 (SF[d] cohort)	Inoue et al. (2018)
$\mathbf{w}$	Annual progression rate[b]	0.024 (JH cohort), 0.06 (SF cohort)	Inoue et al. (2018)
f	Annual metastasis rate of treated high-risk PCa	0.006	Mayo Clinic Radical Prostatectomy Registry
e	Annual metastasis rate of untreated high-risk PCa	0.069	Ghani et al. (2005)
g_t	Annual death rate of metastatic PCa	0.074 (for age<65) 0.070 (for age≥65)	Zhang et al. (2012)
a_t	Annual death rate of all-other causes	Age-specific	US Life Tables (2012)

[a]Litwin et al. (2001) reported that the quality of life recovers to 82.4%, 96.3%, and 100% levels of the baseline at 12, 24, and 36 months, respectively. Based on these ratios, we estimate the disutility of posttreatment as $\hat{c}_T^1 = \hat{c}_T^\infty + \frac{1-96.3\%}{1-82.4\%}(c_T - \hat{c}_T^\infty) = 0.091$ in the first year of posttreatment where $\hat{c}_T^\infty$ is the annual disutility when the patient recovers to the baseline, which is estimated as 0.05 (Heijnsdijk et al., 2012).

[b]Our estimation is based on biopsy upgrading (i.e., the Gleason score of biopsy upgraded from 6 to 7 or higher) studies from Inoue et al. (2018) that include both the JH cohort and the SF cohort. For each cohort, we built a respective hidden Markov model to estimate $\hat{\mathbf{w}}$ and $\mathbf{w}$, with the sensitivity of biopsy testing being estimated as 0.61 (Barnett et al., 2018), based on the maximum likelihood estimation.

[c]JH = Johns Hopkins.

[d]SF = University of California, San Francisco.

TABLE 4 Age-specific death rate from all-other causes, a_t

Age	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85–89	90–94	95–99
Value	0.006	0.009	0.013	0.018	0.027	0.043	0.072	0.122	0.200	0.299

7.1 | Model parameterization

Our study focuses on quality adjusted life years (QALYs), a commonly used criterion for public health problems (Ayer et al., 2012; Erenay et al., 2014; Zhang et al., 2012). The QALYs are measured on a scale of 0 (death) to 1 (a year of perfect health). The quality of life adjustments is based on estimates of disutility associated with the biopsy test, treatment, and metastasis. Specifically, the disutility for biopsy and initial treatment, enduring disutility for metastasis, and long-term side effects of the treatment after initial recovery are provided in Table 3.

To estimate the parameters of the stochastic process, we collected data corresponding to transitions among the five stages of active surveillance for PCa (Figure 2), including PCa progression (from low- to high-risk PCa), detection (from high-risk to treated PCa), metastasis (from high-risk and treated to metastatic PCa), and death (from low-risk, high-risk, treated, and metastatic PCa to mortality). The sensitivity of biopsy, annual metastasis rate from treated and untreated PCa, and the annual death rate from metastatic PCa were drawn from public studies (see Table 3). We estimated the age-specific death rate due to all-other causes (Table 4) based on the difference between the death rate due to all diseases (Arias et al., 2016, Table 2) and the death rate due to PCa (USCS, 2018).

7.1.1 | Heterogeneity among multiple patient types

We consider the following parameters: diagnosis age, $\mathtt{A}_p$, and the cohort where patients are from, $\mathtt{H}_p$, to identify the heterogeneity among patient types. First, the diagnosis age is an important factor because it influences all-other-cause mortality which increases as patients age; as all-other-cause mortality increases the relative benefits of active surveillance for PCa decrease. Second, men in different active surveillance studies have different risks for PCa progression. Based on empirical studies by Inoue et al. (2018), we estimate the rate of misclassification at diagnosis, $\hat{\mathbf{w}}$, and the annual progression rate, $\mathbf{w}$, for both the JH study and the SF study cohorts (Table 3). We distinguish these studies because they show significant differences in risk estimates of PCa progression; typically, the JH cohort has a lower misclassification rate and lower progression rate than the SF cohort. This difference mainly contributes to the different eligibility criteria for enrollment in active surveillance for the two studies (Inoue et al., 2018). Generally, the JH study had more restrictive clinical criteria for enrollment than the SF study.

Remark 3. This case study is based on the design of PCa surveillance strategies for a hypothetical population based

on the SF and JH studies, in order to serve as a plausible example of the application of model **P**. The model optimizes the assignment of patient types to surveillance strategies to maximize overall long-term rewards, where the patients are classified into mutually exclusive patient types depending on their diagnosis age and a hypothetical label of either the "JH study" or "SF study." While we present this specific example for illustration purposes, the model can be generalized to generate surveillance strategies and optimize the assignment for any relevant choices of patient types a policymaker might define (e.g., based on the patient's family history, prognostic genetic tests).

7.2 | Parameter estimation

Once the patient type, progression time, and treatment time are fixed, the disease model can be decomposed into multiple Markov models corresponding to the disease dynamics model for a given combination of p, k, and t. We used the Markov process model based on the parameter estimates in Table 3 to estimate the following model parameters for the stochastic integer program in Section 4 (the results of the parameter estimation are available in Section EC.1.9 in the Supporting Information):

- The probability of PCa progression at period k for patient type p, ξ_{pk}, is estimated based on the rate of misclassification at diagnosis ($\hat{\mathbf{w}}$) and the annual progression rate ($\mathbf{w}$) as follows:

$$\xi_{pk} = \begin{cases} \hat{w}_{\mathtt{H}_p} & \text{if } k = 0, \\ (1-\hat{w}_{\mathtt{H}_p})(1-w_{\mathtt{H}_p})^{k-1} w_{\mathtt{H}_p} & \text{if } k \geq 1. \end{cases} \tag{17}$$

  where $\hat{w}_{\mathtt{H}_p}$ and $w_{\mathtt{H}_p}$ are the misclassification rate and the annual progression rate for cohort $\mathtt{H}_p$, respectively. Note that patients from the same cohort, even with different diagnosis ages, have the same ξ_{pk}, that is, $\xi_{pk} = \xi_{p'k}$ if $\mathtt{H}_p = \mathtt{H}_{p'}$ for all distinct $p, p' \in \mathcal{P}$ and $k \in \mathcal{K}$.
- The probability of patient type p staying in active surveillance at period t if PCa progresses at period k assuming that the patient is not treated, π_{pkt}, can be estimated as follows:

$$\pi_{pkt} = \begin{cases} 1, & \text{if } t = 1,\ k = 0, \\ 1-e, & \text{if } t = 1,\ k \geq t, \\ \pi_{pk,t-1}\left(1 - a_{\mathtt{A}_p+t-1}\right), & \text{if } t \geq 2,\ k \geq t, \\ \pi_{pk,t-1}\left(1 - a_{\mathtt{A}_p+t-1}\right)(1-e), & \text{if } t \geq 2,\ k < t. \end{cases} \tag{18}$$

  Note that patients with the same diagnosis age, even from different cohorts, have the same π_{pkt}, that is, $\pi_{pkt} = \pi_{p'kt}$ if $\mathtt{A}_p = \mathtt{A}_{p'}$ for all distinct $p, p' \in \mathcal{P}$, $k \in \mathcal{K}$, and $t \in \mathcal{T}$.
- The expected QALYs if patient type p is treated at period t if PCa progresses at period k, r_{pkt}, are estimated as follows:

$$r_{pkt} = \sum_{t'=\mathtt{A}_p}^{96} \left[\sum_{S \in \{\mathsf{L},\mathsf{H},\mathsf{T},\mathsf{M}\}} v_{t'}(S,t) \mathbb{P}_{t'}(S,p,k,t) \right], \tag{19}$$

  where $\{\mathsf{L},\mathsf{H},\mathsf{T},\mathsf{M}\}$ are a set of states, denoting "low-risk," "high-risk," "treated," and "metastatic" PCa, respectively. $v_{t'}(S,t)$ is the unit QALY associated with being in state S at period t' if the patient is treated at period t; for example, when $S = \mathsf{M}$, $v_{t'}(S,t) = 1 - \hat{c}_M = 1 - 0.4$ for all t'. $\mathbb{P}_{t'}(S,p,k,t)$ is the probability of being in state S for patient type p who is treated at period t when PCa progresses at period k, which is determined by forward induction based on the state transition probabilities (e, f, g_t, and a_t). For example, when $t' > t$ and $S = \mathsf{M}$, $\mathbb{P}_{t'}(\mathsf{M},p,k,t) = (1-a_{t'})(1-g_{t'})\mathbb{P}_{t'-1}(\mathsf{M},p,k,t) + (1-a_{t'})f\mathbb{P}_{t'-1}(\mathsf{T},p,k,t)$. Note that patients with the same diagnosis age, even from different cohorts, have the same r_{pkt}, that is, $r_{pkt} = r_{p'kt}$ if $\mathtt{A}_p = \mathtt{A}_{p'}$ for all $k \in \mathcal{K}$ and $t \in \mathcal{T}^+$. We assume that the termination age is 95, and let $v_{96}(S,t)$ denote the remaining expected QALY beyond 95 which is estimated by an infinite horizon Markov process model in the absence of biopsy testing. Furthermore, we estimate $r_{pk,T+1}$ by assuming that the patient has never been treated. Note that we can easily account for history-dependent disutility such as $\hat{c}_T$ when estimating r_{pkt} because there are no further actions given that the time of treatment is fixed at period t.
- Whether a patient should discontinue surveillance at age $\mathtt{A}_p + t$ and beyond (d_{pt}) is estimated as follows:

$$d_{pt} = \begin{cases} 0, & \text{if } t \geq 2 \text{ and } d_{p,t-1} = 0, \\ & \text{or } \exists k \text{ such that } r_{pkt} - c_B < r_{pk,T+1}, \\ & \text{or } \exists k, t' > t \text{ such that } r_{pkt} < r_{pkt'}, \\ 1, & \text{otherwise.} \end{cases} \tag{20}$$

  Note that we report the estimation of d_{pt} in Figure 4 as a part of our recommendation.
- The cost associated with the disutility of a unit biopsy for patient type p is estimated as $c_p = c_B$ for all $p \in \mathcal{P}$.

7.3 | Model validation

In this subsection, we validate the disease model based on estimated parameters from Section 6.2. Specifically, we run independent simulations for different patient types, predict their outcomes over different periods during and after active surveillance, and compare the model outputs with published sources on health outcomes.

We selected several health outcome indicators, including (1) the expected time between the beginning of surveillance

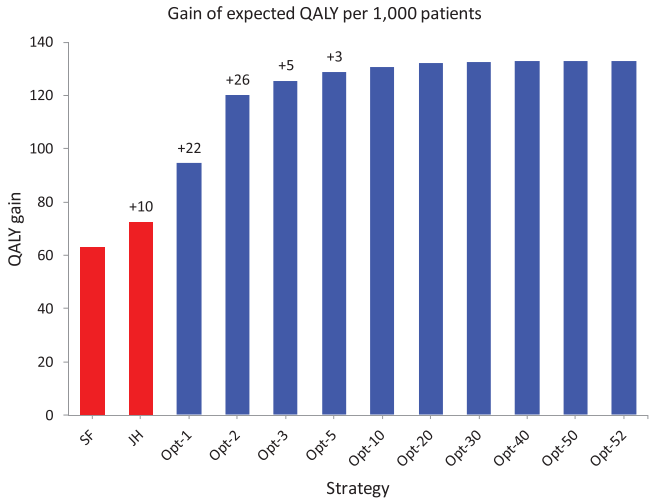

FIGURE 3 The performance comparison between the guideline strategies and the optimal solutions where *Opt-S* denotes the optimal solution with *S* strategies, and the label above each column denotes marginal QALY gain relative to the column directly to its left (labels with ≤ 2 QALY gain were omitted) [Color figure can be viewed at wileyonlinelibrary.com]

and detection and (2) the PCa-specific mortality rate. We compared the model outputs with recently published studies on PCa surveillance (Inoue et al., 2018; Tosoian et al., 2016). The results show that our model outputs are well aligned with the outcomes reported in the literature, suggesting that the disease model can accurately reflect the progression of PCa. More details on model validation are provided in Section EC.1.10 in the Supporting Information.

7.4 | Analysis of optimal surveillance strategies

For our analysis, we consider patients who are diagnosed among $\mathtt{A}_p \in \{50, \ldots, 75\}$ year-old, which is consistent with the U.S. Preventive Services Task Force recommendation (Moyer, 2012). We consider patients from the JH and SF cohorts, that is, $\mathtt{H}_p \in \{\text{JH, SF}\}$. As a result, there are $P = 2 \times 26 = 52$ patient types. We consider a weighted sum of rewards across the patient types, where the number of patients for each type, n_p, is estimated based on patients enrolled in the JH and SF studies in 1995–2014 as published in Inoue et al. (2018) (see Table EC.2 in the Supporting Information). We assume that PCa may progress within 45 years after diagnosis, that is, $k \in \{0, 1, \ldots, 46\}$, where $k = 0$ means that PCa was misclassified at diagnosis, and $k = 46$ means that PCa never progresses. Active surveillance may last for $t \in \mathcal{T} = \{1, 2, \ldots, 11\}$ years, where all patients in the first year of surveillance are scheduled for a biopsy, known as a *confirmatory biopsy*, which is intended to check for possible misdiagnosis. Our decision variables are to decide at which years ($t = 2, \ldots, 11$) to perform biopsies in each strategy and the assignment of patients to strategies.

The benchmarks are the published guideline strategies, as shown in Table 1, including JH, SF strategies, and the UT strategy, a well-known surveillance strategy in Canada. All guideline strategies are one-size-fits-all strategies.

We solve multiple instances of model $\hat{\mathbf{P}}$ with $S = \{1, \ldots, S\}$ where the number of strategies, S, varies from 1 to 52. We let *Opt-S* denote the optimal solution using S strategies. We choose the UT strategy as the bottom line because it performs the worst on the JH and SF cohorts in all the experiments. For each other strategy, we accounted for the gain in expected QALYs per 1000 patients compared with the UT strategy.

Figure 3 presents the results for the guideline strategies (JH and SF) and the optimal strategies for various choices of S. When considering a single strategy, both JH and SF strategies perform better than the UT strategy, by approximately 63 QALYs per 1000 patients, and the JH strategy slightly outperforms the SF strategy. The optimal solution with a single strategy, *Opt*-1, significantly improves upon the three guidelines, outperforming the JH strategy by 22 QALYs per 1000 patients. This improvement is due to the adjustment of the biopsy-testing periods compared with the JH and SF strategies. Specifically, the *Opt*-1 strategy performs biopsies at years 1–3, 5–6, 8, and 10–11 after diagnosis, which is adaptive to the patient's age.

The performance can be improved further if we allow for two or more strategies; however, benefits diminish as S increases (Figure 3), while the complexity of implementation increases in S. Figure 4 illustrates the case of two strategies at the bottom of the figure: one strategy has a higher number of biopsies than the other. The assignment of the high-intensity strategy is inclined towards the SF cohort and younger patients because the SF cohort is at a higher risk, and younger patients have a much longer time over which progression could occur. Both strategies are static, and the model assigns each patient type to the two strategies. Note that some elderly patients (diagnosed between 65 and 75 years of age) are recommended to discontinue active surveillance early due to the increased death rate from all-other causes, which negates the benefits of treatment and surveillance. There is no universally agreed upon age at which patients should discontinue active surveillance. de Carvalho et al. (2017) suggested varying discontinuation ages (from 65 to 82 years) for different categories of patients, which depends on disease risk, the age at diagnosis, and biopsy frequency. The only explicit recommendation by the American Society of Clinical Oncology (ASCO) endorsed that "serial biopsy should not continue past the age of 80" (R. C. Chen et al., 2016). Our numerical results show that patients are recommended to discontinue active surveillance at 76–77 years of age, depending on their age at diagnosis, which is slightly earlier than the ASCO recommendation. Thus, our results may help inform public policy decisions regarding when to terminate active surveillance.

7.5 | Sensitivity analysis

We further conducted a series of one-way sensitivity analyses to validate the robustness of the recommended solution

Diagnosis age	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
Discontinuation age	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	76	76	76	76	76	76	76	77	77	77	77
JH cohort	+	+	+	+	+	+	+	+	+	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SF cohort	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	○	○	○	○
Surveillance plans	Strategy +: perform biopsies at years of **1-5, 7-8, 10-11** after diagnosis Strategy ○: perform biopsies at years of **1, 5, 8** after diagnosis																									

FIGURE 4 The optimal model-based solution with two strategies (our recommendation) for the active surveillance of PCa (example of implementation: a man with a risk profile consistent with the JH cohort, diagnosed at age 65 would have a surveillance biopsy at age 66, 70, 73, according to strategy o and then transition to watchful waiting at age 76.) [Color figure can be viewed at wileyonlinelibrary.com]

to the parameter uncertainty. The selected data are allowed to vary within their ranges, as reported in Table EC.5 in the Supporting Information. We compare four solutions including the JH strategy, SF strategy, *Opt*-2 solution (our recommendation), and a perfect information solution, which is the optimal solution with two strategies solved with the updated parameters, thus providing an upper bound of the total expected QALYs. Similar to the results above, we report the gain in expected QALYs per 1000 patients compared with the UT strategy in Figure EC.6 in the Supporting Information. The results indicate the robustness of our claims of increasing QALY gains over published guidelines to model parameter uncertainty.

8 | CONCLUSION

This article addresses an active surveillance strategy optimization problem that arises in many diseases, including PCa. The problem involves sequential decisions under uncertainty about latent health states, which are often formulated as POMDPs. However, POMDPs are often difficult to solve and become computationally intractable when multiple patient types need to coordinate their decisions in active surveillance practice. We present an alternative approach based on a two-stage stochastic nonlinear integer program, which is more efficient in coordinating decisions among multiple patient types, generating a set of static strategies, and assigning them to patients of different types. We show that this problem is NP-complete, providing motivation for our analysis, which supports the relaxation of the stochastic nonlinear integer program. Moreover, we develop a customized LBD method that can improve the quality of surveillance strategies.

We conduct a case study based on validated data published in medical journals, and the results show that our model solution can significantly improve upon the published guideline strategies used in practice. Finally, we recommend an easy-to-implement solution using two strategies and provide sensitivity analyses to support the robustness of our claims. Based on our results, we briefly summarize the following observations:

- Our recommended solution (*Opt*-2) consistently outperformed all guideline strategies, even under parameter uncertainty, and it gained 67–141 QALYs per 1000 patients compared with the UT strategy. Considering that this solution is similarly easy to implement (static and suitable for all types of patients), the *Opt*-2 solution could be a strong competitor to existing guideline strategies. Compared with the existing strategies, the*Opt*-2 solution benefits from differentiating between young and elderly patients with respective biopsy schedules of different intensities. Moreover, the *Opt*-2 solution benefits from optimizing the timing of biopsy in each strategy and the assignment of patient types to strategies.
- The major sources of parameter uncertainty in performance are the annual metastasis rate from untreated high-risk PCa, f, and the annual death rate from all-other causes, a_t. Specifically, the value of surveillance strategy optimization is more significant for a larger f (i.e., surveillance optimization is more valuable for more progressive PCa that metastasizes quickly if treatment is missed) and smaller a_t (i.e., surveillance optimization is more valuable for younger and healthier patients who are less likely to die from all-other causes).
- The performance of our recommended solution is already close to that of the perfect solution, implying that the value of further considering the ambiguity in parameter estimation might be limited while simultaneously increasing the complexity for physicians and patients.

Our study has some limitations that present opportunities for future research. First, we do not consider the detection of PCa progression using PSA, digital rectal examination (DRE), and magnetic resonance imaging (MRI). We choose to omit PSA and DRE from the model for several reasons: (a) they have very poor accuracy in detecting PCa progression; and (b) PSA and DRE are often used as dynamic triggers for off-schedule biopsies, whereas we focus on designing

a predefined schedule to provide patients with reasonable expectations in the future. We choose to omit MRI because MRI-guided prostate biopsy is performed less frequently than a standard biopsy. It is worthwhile to note that our model can be directly applied to the MRI-guided prostate biopsy with updated parameter estimation. Second, we do not consider the treatment decision for progressive PCa because only radical prostatectomy data are readily available. However, it is worth noting that our model parameters, r_{pkt}, are sufficiently flexible to account for specific treatment for different types of patients and detection time. Therefore, the treatment decisions can be independently optimized without any changes to the structure of our model. Third, we only considered parameter estimates from the JH and the SF cohorts because of the ready availability of data for these studies; however, the patient distributions may not be representative of the population distribution for those particular regions. Nevertheless, our modeling framework can be applied to any population distribution that policymakers might define. These limitations notwithstanding, our findings and the model and methods we propose lay an important foundation for future studies of disease surveillance strategies.

ACKNOWLEDGMENTS

We thank the departmental editor Max Shen, an anonymous associate editor, and two anonymous referees for their constructive comments that helped improve the paper. This work was supported by the National Natural Science Foundation of China (No. 72001187, 71821002) and the Fundamental Research Funds for the Central Universities. This material is also based upon work supported by the National Science Foundation under Grant No. CMMI1536444. 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.

ORCID

Zheng Zhang https://orcid.org/0000-0002-6943-3723
Brian T. Denton https://orcid.org/0000-0002-6372-6066

REFERENCES

Amato, C., & Oliehoek, F. (2015). Scalable planning and learning for multiagent POMDPs. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*, 1995–2002.

Ankerst, D. P., & Thompson, I. M. (2006). Sensitivity and specificity of prostate-specific antigen for prostate cancer detection with high rates of biopsy verification. *Archivio Italiano di Urologia Andrologia*, *78*(4), 125–129.

Arias, E., Heron, M., & Xu, J. (2016). United States Life Tables, 2012. *National Vital Statistics Reports*, *65*(8), 1–68. https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_08.pdf

Ayer, T., Alagoz, O., & Stout, N. K. (2012). OR Forum-POMDP approach to personalize mammography screening decisions. *Operations Research*, *60*(5), 1019–1034. http://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1019

Barnett, C. L., Auffenberg, G. B., Cheng, Z., Yang, F., Wang, J., Wei, J. T., Miller, D. C., Montie, J. E., Mamawala, M., & Denton, B. T. (2018). Optimizing active surveillance strategies to balance the competing goals of early detection of grade progression and minimizing harm from biopsies. *Cancer*, *124*(4), 698–705.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, *4*(1), 238–252.

Bertsimas, D., Silberholz, J., & Trikalinos, T. (2018). Optimal healthcare decision making under multiple mathematical models: application in prostate cancer screening. *Health Care Management Science*, *21*, 105–118. https://doi.org/10.1007/s10729-016-9381-3

Bill-Axelson, A., Holmberg, L., Garmo, H., Taari, K., Busch, C., Nordling, S., Häggman, M., Andersson, S.-O., Andrén, O., Steineck, G., Adami, H.-O., & Johansson, J.-E. (2018). Radical prostatectomy or watchful waiting in prostate cancer—29-year follow-up. *New England Journal of Medicine*, *379*(24), 2319–2329.

Bokhorst, L. P., Alberts, A. R., Rannikko, A., Valdagni, R., Pickles, T., Kakehi, Y., Bangma, C. H., Roobol, M. J., & PRIAS study group. (2015). Compliance rates with the prostate cancer research international active surveillance (PRIAS) protocol and disease reclassification in noncompliers. *European Urology*, *68*(5), 814–821.

Brandeau, M. L., Owens, D. K., Sox, C. H., & Wachter, R. M. (1993). Screening women of childbearing age for human immunodeficiency virus: A model-based policy analysis. *Management Science*, *39*(1), 72–92.

Carter, H. B., Kettermann, A., Warlick, C., Metter, E. J., Landis, P., Walsh, P. C., & Epstein, J. I. (2007). Expectant management of prostate cancer with curative intent: An update of the Johns Hopkins experience. *The Journal of Urology*, *178*(6), 2359–2365.

Chen, Q., Ayer, T., & Chhatwal, J. (2018). Optimal m-switch surveillance policies for liver cancer in a hepatitis c–infected population. *Operations Research*, *66*(3), 673–696.

Chen, R. C., Bryan Rumble, R., D Loblaw, A., Finelli, A., Ehdaie, B., Cooperberg, M. R., Morgan, S. C., Tyldesley, S., Haluschak, J. J., Tan, W., Justman, S., & Jain, S. (2016). Active surveillance for the management of localized prostate cancer (Cancer Care Ontario Guideline): American Society of Clinical Oncology clinical practice guideline endorsement. *Journal of Clinical Oncology*, *34*(18), 2182–2190.

Dall'Era, M. A., Konety, B. R., Cowan, J. E., Shinohara, K., Stauf, F., Cooperberg, M. R., Meng, M. V., Kane, C. J., Perez, N., Master, V. A., & Carroll, P. R. (2008). Active surveillance for the management of prostate cancer in a contemporary cohort. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, *112*(12), 2664–2670.

de Carvalho, T. M., Heijnsdijk, E. A., & de Koning, H. J. (2017). When should active surveillance for prostate cancer stop if no progression is detected? *The Prostate*, *77*(9), 962–969.

de Vries, H., van de Klundert, J., & Wagelmans, A. (2021). Toward elimination of infectious diseases with mobile screening teams: Hat in the DRC. *Production and Operations Management*, *30*(10), 3408–3428.

Erenay, F. S., Alagoz, O., & Said, A. (2014). Optimizing colonoscopy screening for colorectal cancer prevention and surveillance. *Manufacturing & Service Operations Management*, *16*(3), 381–400.

Ghani, K. R., Grigor, K., Tulloch, D. N., Bollina, P. R., & McNeill, S. A. (2005). Trends in reporting Gleason score 1991 to 2001: Changes in the pathologist's practice. *European Urology*, *47*(2), 196–201.

Guenther, E., Klein, N., Zapf, S., Weil, S., Schlosser, C., Rubinsky, B., & Stehling, M. (2019). Prostate cancer treatment with irreversible electroporation (IRE): Safety, efficacy and clinical experience in 471 treatments. *PloS One*, *14*(4), e0215093.

Hamdy, F. C., Donovan, J. L., Lane, J. A., Mason, M., Metcalfe, C., Holding, P., Davis, M., Peters, T. J., Turner, E. L., Martin, R. M., Oxley, J., Robinson, M., Staffurth, J., Walsh, E., Bollina, P., Catto, J., Doble, A., Doherty, A., Gillatt, D., … ProtecT Study Group. (2016). 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *New England Journal of Medicine*, *375*(15), 1415–1424.

Hauskrecht, M., & Fraser, H. (2000). Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artificial Intelligence in Medicine*, *18*, 221–244.

Heijnsdijk, E. A., Wever, E. M., Auvinen, A., Hugosson, J., Ciatto, S., Nelen, V., Kwiatkowski, M., Villers, A., Páez, A., Moss, S. M., Zappa, M., Tammela, T. L. J., Mäkinen, T., Carlsson, S., Korfage, I. J., Essink-Bot, M.-L., Otto, S. J., Draisma, G., Bangma, C. H., … de Koning, H. J. (2012). Quality-of-life effects of prostate-specific antigen screening. *New England Journal of Medicine*, *367*(7), 595–605.

Helm, J. E., Lavieri, M. S., Van Oyen, M. P., Stein, J. D., & Musch, D. C. (2015). Dynamic forecasting and control algorithms of glaucoma progression for clinician decision support. *Operations Research*, *63*(5), 979–999.

Hooker, J. N., & Ottosson, G. (2003). Logic-based Benders decomposition. *Mathematical Programming*, *96*(1), 33–60.

Inoue, L. Y., Lin, D. W., Newcomb, L. F., Leonardson, A. S., Ankerst, D., Gulati, R., Carter, H. B., Trock, B. J., Carroll, P. R., Cooperberg, M. R., Cowan, J. E., Klotz, L. H., Mamedov, A., Penson, D. F., & Etzioni, R. (2018). Comparative analysis of biopsy upgrading in four prostate cancer active surveillance cohorts. *Annals of Internal Medicine*, *168*(1), 1–9.

Kamalzadeh, H., Ahuja, V., Hahsler, M., & Bowen, M. E. (2021). An analytics-driven approach for optimal individualized diabetes screening. *Production and Operations Management*, *30*(9), 3161–3191.

Kinsella, N., Stattin, P., Cahill, D., Brown, C., Bill-Axelson, A., Bratt, O., Carlsson, S., & Van Hemelrijck, M. (2018). Factors influencing men's choice of and adherence to active surveillance for low-risk prostate cancer: A mixed-method systematic review. *European Urology*, *74*(3), 261–280.

Lawrentschuk, N., & Klotz, L. (2011). Active surveillance for low-risk prostate cancer: An update. *Nature Reviews Urology*, *8*(6), 312.

Litwin, M. S., Melmed, G. Y., & Nakazon, T. (2001). Life after radical prostatectomy: A longitudinal study. *The Journal of Urology*, *166*(2), 587–592.

Mayo Clinic (2015). Cancer biopsies do not promote cancer spread, research finds. *Science Daily*. www.sciencedaily.com/releases/2015/01/150109093717.htm

Moyer, V. A. (2012). Screening for prostate cancer: US preventive services task force recommendation statement. *Annals of Internal Medicine*, *157*(2), 120–134.

Mühlberger, N., Boskovic, K., Krahn, M. D., Bremner, K. E., Oberaigner, W., Klocker, H., Horninger, W., Sroczynski, G., & Siebert, U. (2017). Benefits and harms of prostate cancer screening–predictions of the oncotyrol prostate cancer outcome and policy model. *BMC Public Health*, *17*(1), 596.

Nenova, Z., & Shang, J. (2022). Chronic disease progression prediction: Leveraging case-based reasoning and big data analytics. *Production and Operations Management*, *31*(1), 259–280.

Otten, J. W. M., Witteveen, A., Vliegen, I., Siesling, S., Timmer, J. B., & IJzerman, M. J. (2017). Stratified breast cancer follow-up using a partially observable MDP. In R. Boucherie & N. van Dijk (Eds.), *Markov Decision processes in practice* (pp. 223–244). Springer.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*(3), 801–817.

Sandkç, B., Maillart, L. M., Schaefer, A. J., & Roberts, M. S. (2013). Alleviating the patient's price of privacy through a partially observable waiting list. *Management Science*, *59*(8), 1836–1854.

Steimle, L. N., & Denton, B. T. (2017). Screening and treatment of chronic diseases. *Markov Decision Processes in Practice*, *248*, 1–33.

Tosoian, J. J., Carter, H. B., Lepor, A., & Loeb, S. (2016). Active surveillance for prostate cancer: Current evidence and contemporary state of practice. *Nature Reviews Urology*, *13*(4), 205–215.

Tunç, S., Alagoz, O., & Burnside, E. S. (2022). A new perspective on breast cancer diagnostic guidelines to reduce overdiagnosis. *Production and Operations Management*, *31*, 2361–2378. https://doi.org/10.1111/poms.13691

Underwood, D. J., Zhang, J., Denton, B. T., Shah, N. D., & Inman, B. A. (2012). Simulation optimization of PSA-threshold based prostate cancer screening policies. *Health Care Management Science*, *15*(4), 293–309.

USCS, C. S. W. G. (2018). U.S. cancer statistics data visualizations tool, based on November 2017 submission data (1999–2015). *U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute*. www.cdc.gov/cancer/dataviz

van Ackooij, W., de Oliveira, W., & Song, Y. (2017). Adaptive partition-based level decomposition methods for solving two-stage stochastic programs with fixed recourse. *Informs Journal on Computing*, *30*(1), 57–70.

Van den Bergh, R. (2007). Prospective validation of active surveillance in prostate cancer: The PRIAS study. *European Urology*, *52*, 1560–1563.

Vozikis, A., Goulionis, J. E., & Benos, V. K. (2012). The partially observable Markov decision processes in healthcare: An application to patients with ischemic heart disease (IHD). *Operational Research*, *12*(1), 3–14.

Zhang, J., Denton, B. T., Balasubramanian, H., Shah, N. D., & Inman, B. A. (2012). Optimization of prostate biopsy referral decisions. *Manufacturing & Service Operations Management*, *14*(4), 529–547.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhang, Z., Denton, B. T., & Morgan, T. M. (2022). Optimization of active surveillance strategies for heterogeneous patients with prostate cancer. *Production and Operations Management*, *31,* 4021–4037. https://doi.org/10.1111/poms.13800