Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Optimizing Colonoscopy Screening for Colorectal Cancer Prevention and Surveillance

Fatih Safa Erenay, Oguzhan Alagoz, Adnan Said

To cite this article:
Fatih Safa Erenay, Oguzhan Alagoz, Adnan Said (2014) Optimizing Colonoscopy Screening for Colorectal Cancer Prevention and Surveillance. Manufacturing & Service Operations Management 16(3):381-400. http://dx.doi.org/10.1287/msom.2014.0484

M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 16, No. 3, Summer 2014, pp. 381–400
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

informs

http://dx.doi.org/10.1287/msom.2014.0484

Optimizing Colonoscopy Screening for Colorectal Cancer Prevention and Surveillance

Fatih Safa Erenay
Department of Management Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada,
ferenay@uwaterloo.ca

Oguzhan Alagoz
Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, Wisconsin 53706,
alagoz@engr.wisc.edu

Adnan Said
Department of Medicine, University of Wisconsin–Madison, Madison, Wisconsin 53792,
axs@medicine.wisc.edu

Millions of Americans undergo colonoscopy screening for colorectal cancer (CRC) prevention and surveillance every year. The efficiency of colonoscopy operations depends on how often patients are screened, which is a complex and controversial decision, as reflected by the discrepancy between clinical practice and guidelines. We develop a partially observable Markov decision process to optimize colonoscopy screening policies for the objective of maximizing total quality-adjusted life years. Our model incorporates age, gender, and risk of having CRC into the screening decisions and therefore provides a novel framework for personalized CRC screening. In addition to deriving the maximum attainable benefit from colonoscopy screening, which reflects the opportunity cost of following current guidelines, our results have several policy implications. Using clinical data, we show that the optimal colonoscopy screening policies may be more aggressive than the guidelines under some conditions. Optimal screening policies recommend that females with CRC history undergo colonoscopy more frequently than males. In contrast, females without CRC history should be screened less frequently than males. This result, which was not recognized before, signifies the role of gender in optimal CRC screening decisions.

Keywords: stochastic modeling; partially observable Markov decision processes; operations research applications in healthcare; cancer screening; colorectal cancer prevention and surveillance; colonoscopy
History: Received: April 10, 2012; accepted: January 3, 2014. Published online in *Articles in Advance* May 27, 2014.

1. Introduction

Approximately 9% (141,210 per year) of new cancer cases and 8.6% (49,380 per year) of cancer deaths are associated with colorectal cancer (CRC), making it the third most common cause of cancer death (according to number of deaths) among males and females (Siegel et al. 2011). Most CRC cases originate from adenomatous polyps, which are benign growths on the inner surface of the colon and rectum that may progress to CRC (Loeve et al. 2004). This pattern—from polyp to cancer—provides an effective method for the prevention of CRC: detecting polyps using CRC screenings, such as colonoscopy, and removing them before they become malignant. As a result, many gastroenterologists suggest their patients undergo colonoscopy screening after age 50 (USPSTF 2008).

The choice of the screening policy determines the efficiency and quality of colonoscopy operations since the increase in life expectancy and quality of life, the reduction in CRC risk and mortality, and the disutility from colonoscopy screening depend on the screening frequency specified by the policy. Considering that there are more than 99 million U.S. citizens older than 50 years recommended to undergo periodic CRC screening, and the annual demand for colonoscopy operations is around 2.21–7.96 million (Vijan et al. 2004); gastroenterologists' choice of colonoscopy screening policy obviously has a significant effect on the welfare of U.S. citizens and the U.S. economy. However, there is significant controversy within the clinical community about the best colonoscopy screening policy, as reflected by the high variation in the screening policy recommendations of practitioners (Klabunde et al. 2009). Therefore, determining the optimal colonoscopy screening policies is imperative but is also a complex task.

In this paper, we propose an analytical framework for optimizing colonoscopy screening policies for CRC prevention and surveillance while considering static and dynamic risk factors such as gender, age, and history of CRC or adenomatous polyp. Our purposes are (1) to provide a modeling framework for clinicians and policy makers to develop optimal personalized CRC screening protocols, and (2) to apply this model

to various patient groups and provide insights about the optimal policies. A brief background on CRC and colonoscopy screening is necessary to duly appreciate our contributions.

The estimated lifetime risk of CRC is 5.08% based on the Surveillance, Epidemiology, and End Results (SEER) data (SEER 2012). Gastroenterologists do not only suggest that their patients undergo preventative CRC screening, but also recommend undergoing follow-up and surveillance screening after adenomatous polyp removal (polypectomy) and CRC treatment to prevent recurrence and secondary occurrence (metachronous) of CRC (Winawer et al. 2003). Although the overall five-year survival rate for CRC is approximately 64%, CRC is very deadly in the case of a late diagnosis; i.e., the five-year survival rate for metastasized CRC is less than 12% (SEER 2012), which illustrates the importance of early diagnosis and hence the value of CRC screening. Patients may undergo several screening tests, such as colonoscopy and sigmoidoscopy, each having a different accuracy and disutility due to the pain, uneasiness, and anxiety associated with the preparations, complications, and delay before obtaining pathology results.

The literature on colorectal cancer discusses many subtle issues about CRC screening: the utility of screening for older patients (Schoen et al. 2006), the timing of CRC screening termination (Maheshwari et al. 2008), and the impact of newer screening modalities such as computed tomography (CT) colonography on CRC prevention (Regueiro 2005). In particular, clinicians need more guidance on choosing the best screening schedule for their patients (Pfister et al. 2004). Several clinical organizations, such as the American Gastroenterology Association (AGA) and the U.S. Preventive Services Task Force (USPSTF), have developed CRC screening guidelines (Winawer et al. 2003, USPSTF 2008) that generally classify patients into different risk levels based on personal history, as shown in Table 1. Yet recent surveys report that many clinicians do not follow the screening guidelines. For example, one study reports that 43% of clinicians recommend more frequent colonoscopy screening than the guidelines for low-risk patients (Klabunde et al. 2009). In addition, "a substantially larger proportion of physicians reported that their recommendations are more influenced by published evidence than practice guidelines" (Mysliwiec et al. 2004, p. 268).

Cost-effectiveness analyses, which evaluate the performance of a number of screening policies using simulation modeling, put effort into resolving this controversy (Frazier et al. 2000, Vijan et al. 2007). However, cost-effectiveness analyses cannot determine the optimal screening policies because of the computational difficulty of enumerating all possible policies. Therefore, analytical optimization models are necessary to determine the optimal CRC screening policies.

Table 1 Patient Risk Levels and Screening Guidelines

Risk level	Description	AGA guidelines for colonoscopy screening
Low	Asymptomatic patients without personal or family history of CRC	Every 10 years
High	Patients with a history of adenomatous polyp	First screening should be 3 years after the polypectomy. If it finds nothing significant, every 5 years afterward.
Post-CRC	Patients with history of CRC	First screening should be 1 year after CRC treatment. If it finds nothing, next screening should be after 3 years. If it finds nothing, every 5 years.

Source. Winawer et al. (2003).

In this context, we develop a finite-horizon partially observable Markov decision process (POMDP) model that determines the optimal screening policies for CRC prevention and surveillance. The recent update of USPSTF guidelines highlighted that "the goal of a screening program should be to maximize the number of life-years gained while minimizing the harms" (USPSTF 2008, p. 629). In line with this, we consider the patient's perspective and maximize the total expected quality-adjusted life-years (TQALYs), the difference between the expected total lifetime and total disutility due to screening and developing CRC. We consider only colonoscopy screening in this study, because it is the most accurate and commonly recommended screening test (Klabunde et al. 2009). We optimally solve our POMDP model using clinical data from a simulation study based on Mayo Clinic-Rochester patient records (Erenay et al. 2011), the SEER database (SEER 2012), and the literature.

Using a POMDP model allows us to incorporate a patient's personal risk of having CRC and adenomatous polyp and other factors that are ignored by the guidelines, such as age and gender, into the screening decisions. Therefore, our model provides insights to help clinicians develop more personalized screening recommendations. Although this study has methodological contributions for proposing a POMDP framework for CRC screening and surveillance, our main contribution lies in the implementation of the model using clinical data, determining the performance gap between the current guidelines and the optimal policies, and identifying the trends in the optimal policies to provide insights for the clinicians to improve their practice.

Some of these insights are as follows: (1) We find that the optimal colonoscopy screening policies are generally more aggressive than the current guidelines; i.e., the proposed optimal policies recommend more frequent colonoscopy screening. (2) Although low-risk women should undergo colonoscopy screening less frequently than low-risk men, women with a

personal history of CRC should undergo colonoscopy more often than men with a personal history of CRC, which was not recognized by the clinical community before. (3) Younger patients should be screened more frequently than older patients. (4) Compared to clinical guidelines, the optimal screening policies are associated with up to 3.2% more TQALYs, 86% less CRC risk, and 89% less CRC mortality, reflecting the opportunity cost of following guidelines instead of the optimal policies. (5) We also provide an algorithm based on our model to determine the stopping age for screening, a highly debated issue in CRC screening. (6) Under certain conditions, periodic policies with shorter screening intervals than the guidelines may perform acceptably well compared to the optimal policies.

2. Literature Review

Operations research applications of cancer screening are extensively reviewed in Pierskalla and Brailer (1994) and Alagoz et al. (2011). Parmigiani et al. (2002) propose a stochastic model to determine the optimal timing of once-in-a-lifetime colonoscopy screening, and Roberts et al. (2007) develop a simulation model to measure the performances of different CRC screening policies. Among operations research models in the literature, the hidden Markov chain of Leshno et al. (2003) is the most relevant model to ours, because they consider a similar health state space with two states for different polyp sizes and three states for CRC stages. We do not incorporate the size and stages of colorectal lesions in our definition; however, our model still represents the dynamics of CRC progression reasonably (see §5) and is compact enough to solve optimally. Our proposed model has the following advantages over that of Leshno et al. (2003): (1) They do not define a dynamic programming mechanism; i.e., they only evaluate the performances of six screening policies (e.g., colonoscopy every 10 years). In contrast, we define the dynamic programming mechanism for our proposed model and solve it optimally. Thus, unlike the model of Leshno et al. (2003), our model has the ability to provide the optimal screening recommendations based on a particular patient's individual risk of having a colorectal lesion. (2) Leshno et al. (2003) do not account for personal history of CRC, whereas our model ensures more aggressive CRC progression after polypectomy and successful CRC treatment. This is important because polyp or CRC diagnosis is a likely event; e.g., there is up to 25%–60% polyp prevalence after age 50 (Winawer et al. 1997). (3) Our model derives screening policies for high-risk and post-CRC patients as well, whereas Leshno et al. (2003) consider only the low-risk patients. (4) Finally, we use detailed data from clinical sources for our numerical experiments, whereas, the numerical experiments in Leshno et al. (2003) are mainly based on published population-level data.

In this context, this is the first analytical model on the problem of optimizing CRC screening and surveillance. Our modeling approach provides a novel framework for developing the optimal CRC screening policies personalized to individual CRC lesion risk and considering static (gender) and dynamic factors (history of CRC and polyp). Although our modeling framework is still basic, considering only some of the actual risk factors, it can be easily extended to consider a more realistic set of risk factors given the necessary data. However, even as it is, it provides valuable insights about the optimal colonoscopy screening policies and appropriateness of actual screening practices of clinicians as well as effects of age, gender, and personal history on the optimal CRC screening.

There are also related studies that consider other cancers. Ayer et al. (2012) and Maillart et al. (2008) develop a POMDP and hidden Markov chain for mammography screening for breast cancer, and Zhang et al. (2012b) propose a POMDP model for screening for prostate cancer. In addition, Zhang et al. (2012a) develop another POMDP application to determine the optimal timing of biopsy based on annual prostate-specific antigen test results. There are no biopsy decisions in colonoscopy screening because the lesions detected by colonoscopy are always removed. Among them, Ayer et al. (2012) and Zhang et al. (2012b) need further elaboration. Ayer et al. (2012) consider only the objective of maximizing TQALYs, as we do, whereas Zhang et al. (2012b) also consider maximizing the net benefit (willingness to pay $\times$ TQALYs $-$ costs). Both models consider only up to three partially observable health states and assume that individuals move to absorbing states after the diagnosis of cancer; i.e., these models terminate when cancer is diagnosed. In this context, our model is more complicated because we consider a more complex set of health states and do not terminate the model after CRC diagnosis. Our POMDP model has some similarities with those of Ayer et al. (2012) and Zhang et al. (2012b) since all of these models are on the cancer screening problem. In particular, this study and those of Ayer et al. (2012) and Zhang et al. (2012b) report common performance measures (e.g., TQALYs, cancer risk), compare our proposed strategies with the guidelines, and discuss the effect of aging on the optimal screening policies.

Nevertheless, in addition to focusing on CRC screening and providing unique intuitions for the optimal colonoscopy screening frequency, our POMDP approach is methodologically different from Ayer et al. (2012) and Zhang et al. (2012b). Both of these models assume that cancer progression depends only on age. Therefore, their models are only capable of analyzing the effect of static risk factors on the optimal policies. In fact, a random event (e.g., becoming obese or increasing breast density) can also cause an increase/decrease in

the rate of cancer progression. That is, all cancers are associated with both dynamic and static risk factors; however, Ayer et al. (2012) and Zhang et al. (2012b) do not provide a mechanism to incorporate the dynamic risk factors that could be crucial for a realistic model, as it is in our case. Current CRC screening guidelines provide different recommendations for patients who already had polypectomy and CRC; therefore, unlike Ayer et al. (2012) and Zhang et al. (2012b), our model considers both static (gender) and dynamic risk factors (having polypectomy or CRC treatment before) besides the age of the patient. We achieve this by defining completely observable risk levels (low-risk, high-risk, and post-CRC) in our state space where a patient moves from one risk level to another after a completely observable event occurs (screening result). Incorporating dynamic risk factors generates a complex optimal value function system; i.e., we have separate optimality equations for each risk level, which are interrelated because risk levels dynamically change as patients age. We show that the corresponding optimal value functions are piecewise linear and convex (PW&C) and we characterize the alpha vectors needed to efficiently solve our model. Our model, the PW&C proofs and alpha vector definitions can be easily extended to incorporate more dynamic risk factors. Thus, we provide a basic framework that is more realistic than the previous POMDP applications for developing personalized CRC screening considering both dynamic and static risk factors.

The consideration of these risk factors required us to collect more intensive data than Ayer et al. (2012) and Zhang et al. (2012b) and to conduct several meta-analyses to synthesize information from several clinical sources. Unlike the studies by Ayer et al. (2012) and Zhang et al. (2012b), having a precursor state in addition to cancer states adds flexibility to our model because many cancers such as stomach and head and neck arise from benign precursors whose removal delays or prevents the disease. Furthermore, both Ayer et al. (2012) and Zhang et al. (2012b) consider x-ray or blood-test-based noninvasive screening methods; however, invasive screening methods are associated with several details such as fatal complications like perforation, bowel obstruction, etc. Our approach also shows how such details can be incorporated into a cancer screening model.

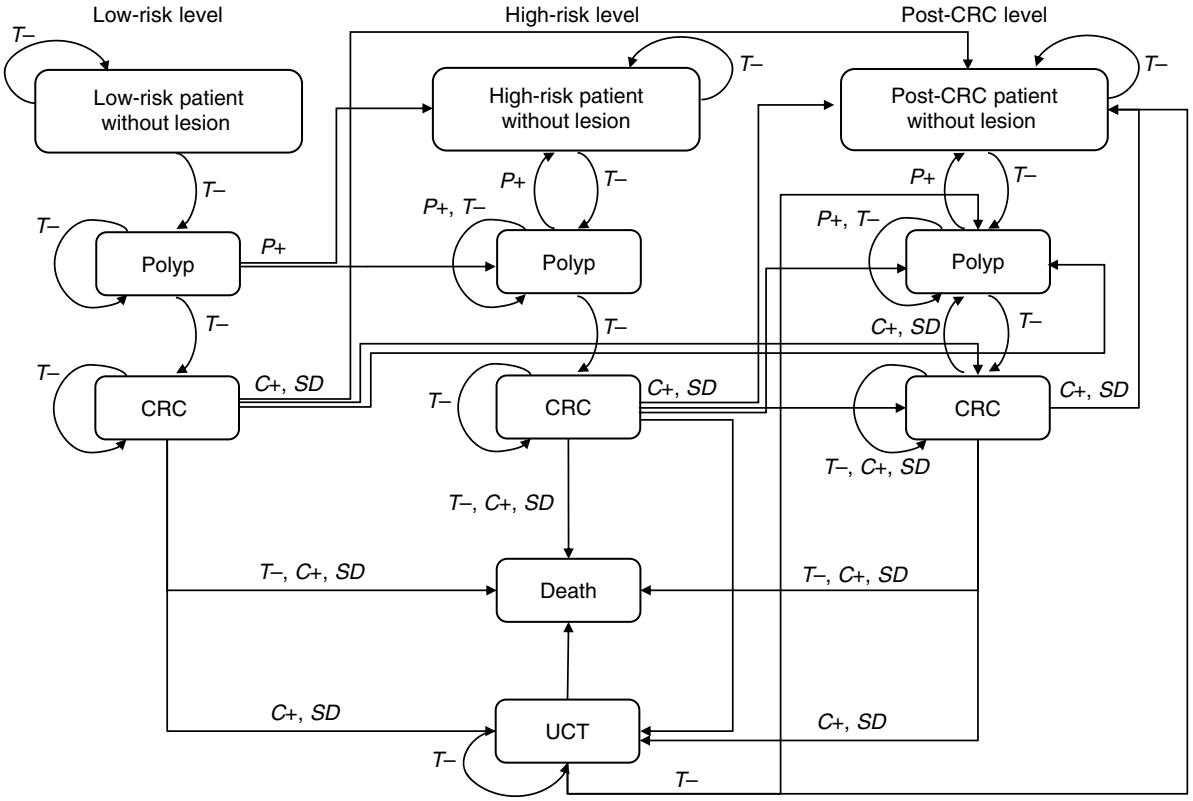
3. CRC Progression Under Colonoscopy Screenings

We represent the current health state of a patient using three health states: *without lesion*, (*having adenomatous*) *polyp*, (*having*) *CRC*. The health state transitions depend on current health state, screening decision (to undergo colonoscopy this year or not), and the results observed after colonoscopy. A colonoscopy may detect an adenomatous polyp, a CRC lesion, or nothing in the colon or rectum ($T-$). In addition, a patient may experience severe CRC symptoms and undergo a diagnostic colonoscopy screening, which is referred to as self-diagnosis (*SD*). We do not consider the self-diagnosis of polyps, since colorectal polyps are generally asymptomatic (Bond 2000). This is mainly because polyps rarely cause major symptoms such as gross bleeding and partial obstruction, and the more common minor symptoms such as minor bleeding are hard to detect or can be confused with the symptoms of other gastrointestinal problems such as hemorrhoid (Bond 2000).

Figure 1 shows how colorectal lesions progress and health states change based on colonoscopy results. There are three levels representing the disease progression for low-risk, high-risk, and post-CRC patients. Screening results $T-$, $P+$, $C+$, and *SD* refer to test $-$, detection of adenomatous polyp via colonoscopy, detection of CRC via colonoscopy, and self-diagnosis of CRC, respectively. Each object in the figure represents a core health state. The arrows illustrate possible core state transitions under various screening results. Transitions denoted by $T-$ and *SD* result from the natural progression of the disease and they can occur at any year. However, those denoted by $P+$ and $C+$ can only occur when the patient undergoes colonoscopy with positive findings. The transitions from mortality in *without lesion* and *polyp* states are omitted for clarity of presentation. We assume that all CRC cases follow polyp-to-cancer progression because nonpolyposis CRCs generally are caused by rare disorders such as hereditary nonpolyposis colorectal cancer (Winawer et al. 1997). We describe the health state transitions for low-risk patients, and the transitions for patients in the other risk levels follow the same pattern.

In Figure 1, the top node of the low-risk (LR) level represents a low-risk patient who has no lesion. The screening result for this state is always $T-$, because the specificity of colonoscopy (probability of accurately identifying the patients with no colorectal lesions) is equal to 1 (Frazier et al. 2000). Given that a patient is lesion free at the beginning of a year, he can stay in the same state or develop an adenomatous polyp and move to the *polyp* state within the year. If the patient has an adenomatous polyp at the beginning of a year, then the polyp can be missed ($T-$) with a probability of 1-sensitivity of colonoscopy (sensitivity: the probability of accurately detecting colorectal lesions) and this polyp either stays as an adenomatous polyp or turns into a CRC within the year. The adenomatous polyp can also be detected and removed ($P+$) at the beginning of the current year with a probability that is equal to the sensitivity of colonoscopy for polyps. In this case, this patient either develops a new adenomatous polyp and moves to the *polyp* state in the high-risk

Figure 1 Health State Transitions According to the Screening Results



(HR) level or he moves to the *high-risk patient without lesion* state. From this point on, this patient's health state transitions follow the transition scheme in the high-risk level.

If a patient is diagnosed with CRC via screening ($C+$) or self-diagnosis (SD) at the beginning of a year, then he undergoes standard CRC treatment. If the treatment is successful, the patient may develop a new adenomatous polyp and move to the *polyp* state in the post-CRC level, may develop a recurrent CRC and move to the *CRC* state in the post-CRC level, or may move to the *post-CRC patient without lesion* state within the year. In all cases, further transitions follow the transition scheme in the post-CRC level. If the treatment is not successful, the patient either dies or undergoes further cancer treatment (UCT) during that year. If the patient is in the UCT state, the health state transitions follow a similar route, depending on the success of the treatment in the current year. Finally, if a patient has undetected CRC at the beginning of a year because the CRC is missed by the previous screening tests and not self-diagnosed ($T-$), then two transitions may occur: the patient may die from CRC or natural reasons, or he may still have CRC next year.

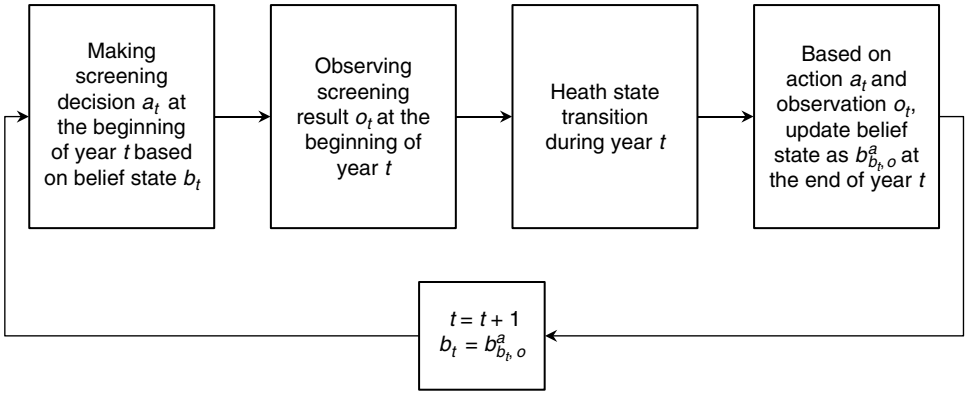
4. The POMDP Model

Our model combines the CRC progression paradigm explained above with annual colonoscopy screening decisions. Our discrete-time adaptive model keeps track of the patient's risk of having adenomatous polyp and CRC, which we refer to as belief state. Our model keeps the information gained from previous screening results by updating the belief state at every decision epoch based on the screening observations. The order of events in our model is illustrated in Figure 2.

In our model, the current decision epoch is denoted by $t \in \mathbf{T} = \{0, 1, 2 \ldots N\}$, which represents the years since age 50. This implies that our analysis computes the optimal screening policies under the assumption that screening starts at age 50, which is reasonable because almost all guidelines suggest initiating CRC screenings at age 50 (Winawer 2007). We choose annual decision epochs as the probability of developing a colorectal lesion within a year is very small (Winawer et al. 1997).

The health state of a patient at the beginning of year t is represented by $s_t \in \mathbf{S} = \{LR0, LR1, LR2, HR0, HR1, HR2, PC0, PC1, PC2, UCT, D\}$, where $\{LR0, LR1, LR2\}$, $\{HR0, HR1, HR2\}$, and $\{PC0, PC1, PC2\}$ represent *without lesion*, *polyp*, and *CRC* states for low-risk, high-risk, and post-CRC patients, respectively. In addition, UCT and D represent *under cancer treatment* and *death* states, respectively. The UCT and D states are completely observable; the others are partially observable. For brevity, we match each state $s \in \mathbf{S}$ with an integer, $i \in \{0, 1, \ldots, 10\}$ to represent the health states in analytical expressions.

Figure 2 Order of Events in the POMDP Model



The action at the beginning of year t is denoted by $a_t \in \mathbf{A} = \{DN, Co\}$, where DN and Co represent doing nothing and undergoing a colonoscopy, respectively. We assume that patients perfectly comply with the screening decisions. Although this is a valid assumption for high-risk and post-CRC patients, as around 80%–100% of such patients comply with follow-up screening, it may not hold for the low-risk patients; their compliance rate is around 50%–70% for preventative screening (Frazier et al. 2000, Vijan et al. 2007). However, under this assumption, we determine the maximal quality-adjusted life-years (QALY) benefits that can be attained via colonoscopy screening in the most ideal scenario. Actually, several cost-utility or cost-effectiveness studies on CRC screening make the same assumption (Frazier et al. 2000, Wilschut et al. 2011, Lansdorp-Vogelaar et al. 2009a). Nevertheless, we relax this assumption in our numerical studies to investigate the performance of the optimal policies under imperfect compliance (see Online Appendix I[1]). Table 2 summarizes the notation.

4.1. Observation Probabilities

As a result of the selected action, an observation $o_t \in \mathbf{O} = \{T-, P+, C+, SD\}$, referring to the screening result in year t, is seen. The observation probability in year t is denoted by $f_t \in \mathbf{F}$: $\mathbf{S} \times \mathbf{A} \times \mathbf{O} \to [0, 1]$. Specifically, $f_t(o \mid s, a)$ represents the probability of observing a screening result o given that the patient is in state s and action a is selected at the beginning of year t. Table 3 shows the $f_t(o \mid s, a)$ values for different states and observations.

Colonoscopy screening can detect colorectal lesions with probabilities equal to its sensitivity for adenomatous polyps and CRC denoted by τ^a_P and τ^a_C, respectively. To simplify the notation, we assume that, without loss of generality, a self-diagnosis of CRC, $o_t = SD$, may occur at the beginning of year t with probability ω_a. We define $f_t(T- \mid UCT, a) = f_t(T- \mid D, a) = 1$ because the UCT and D states are not associated with any screening decision. We also define $\tau^{DN}_P = \tau^{DN}_C = 0$, i.e., colorectal lesions can be only self-diagnosed if the patient does not undergo colonoscopy.

4.2. Health State Transition Probabilities

The health state transition probability function, $g_t \in \mathbf{G}$: $\mathbf{S} \times \mathbf{S} \times \mathbf{A} \to [0, 1]$, denotes the probability that a patient will be in state s' in year $t+1$, given that he is in state s and action a is selected in year t. That is,

$$g_t(s' \mid s, a) = \sum_{o \in \mathbf{O}} p_t(s' \mid s, a, o) f_t(o \mid s, a) \quad \forall s, s' \in \mathbf{S},\ a \in \mathbf{A}, \text{ and } t < N, \quad (1)$$

where $p_t(s' \mid s, a, o)$ represents the probability that a patient will be in state s' in year $t+1$ given that he is in state s, action a is selected, and screening result o is observed in year t. For example, let $\rho^t_{i,i'}$ for $i, i' \in \{1, \ldots, 9\}$ denote the transition probability from state i to state i', given that the patient does not die during year t, and let $\delta^t_{i,o}(a)$ denote the annual death probability in state i after selecting action a and seeing observation o. Then a patient who already has an adenomatous polyp in year t may still have an adenomatous polyp with probability $\rho^t_{1,1}[1 - \delta^t_{1,T-}(a)]$, or he may develop CRC with probability $\rho^t_{1,2}[1 - \delta^t_{1,T-}(a)]$ or die with probability $\delta^t_{1,T-}(a)$ by the next year (see Table 7 in Online Appendix A). A table form representation of the $p_t(s' \mid s, a, o)$ values is given in Online Appendix A; it also illustrates how we incorporate the effect of fatal colonoscopy complications on mortality probabilities into our model to account for the permanent future reward loss.

4.3. Immediate Rewards

The immediate rewards of our model, $q_t \in \mathbf{Q}$: $\mathbf{S} \times \mathbf{A} \to \mathcal{R}$, refer to the expected difference between the lifetime and the disutility (in life years) accrued in year t. Because the expected lifetime and disutility depend

[1] The online appendix is available as supplemental material at http://dx.doi.org/10.1287/msom.2014.0484.

Table 2 Model Notation

Notation	Description
$t \in \mathbf{T}$, $s_t \in \mathbf{S}$, $a_t \in \mathbf{A}$, and $o_t \in \mathbf{O}$	Current decision epoch, health state, colonoscopy decision (action), and observation; and their corresponding sets
$LR0$, $HR0$, and $PC0$	Without lesion states for low-risk, high-risk, and post-CRC patients
$LR1$, $HR1$, and $PC1$	Polyp states for low-risk, high-risk, and post-CRC patients
$LR2$, $HR2$, and $PC2$	CRC states for low-risk, high-risk, and post-CRC patients
D and UCT	Death and under cancer treatment states
DN and Co	Doing nothing and undergoing colonoscopy
$T-$, $P+$, $C+$, and SD	Test negative, a polyp is found, CRC is found, and CRC is self-diagnosed
$f_t(o \mid s, a)$	Probability of observing result o given current state s and action a
$g_t(s' \mid s, a)$	Transition probability from state s to s' given action a
$p_t(s' \mid s, a, o)$	Transition probability from state s to s' given action a and observation o
α_P^a and α_C^a	Sensitivity of colonoscopy for polyp and CRC given action a
ω_a and β	Probability of CRC self-detection given action a and detecting non-CRC lesion
$\rho_{i,j}^t$	Lesion progression probability from state i to j in year t
$\delta_{i,o}^t(a)$	Probability of mortality in state i given current action a and observation o
γ_i^t	Probability that treatment initiated at state i will be completed within year t
$\mu_{i,o}^t$	Probability of mortality from colonoscopic complications
$q_t(s, a)$ and $q_t(s, a, o, s')$	Immediate reward (expected QALYs in year t) for state s given action a and for a patient who moves from state s to state s' given action a and observation o
$q_N(s)$ and λ_t	Terminal reward (QALYs after age 100) for state s and discount factor in year t
d_C, d_{CT}, and d_{UCT}	Disutilities of undetected CRC, CRC treatment, and being in the UCT state
$\underline{d}(Co)$ and $\bar{d}(Co)$	Disutility of undergoing colonoscopy with and without polypectomy
$d^1(Co)$ and $d^2(Co)$	Disutility of undergoing colonoscopy with and without lesion removal
$\kappa_{s,o}^t(a)$, κ_{UCT}^t	Probability of immediate mortality from treatment, and screening complications given that the patient dies in decision epoch t
b_t and b_{UCT}^t	Belief state in year t, and after a completed treatment in the UCT state
$\tilde{b}_{b_t,o}^a$	Posterior belief state after action a and observation o given prior belief state b_t
$b_{b_t,o}^a$	Updated belief state in year $t+1$ (it is the normalized form of $\tilde{b}_{b_t,o}^a(s)$)
$V_t^*(b_t)$, $W_t^*(b_t)$, $Y_t^*(b_t)$, and $V_t^*(UCT)$	Optimal value functions for low-risk, high-risk, and post-CRC patients in year t given b_t and optimal value function for a patient in the UCT state in year t
$l_t(b_t, a, o)$	Probability of not occupying a completely observable state in year $t+1$

also on the observation o and next state s', we define $q_t(s, a)$ as follows:

$$q_t(s,a) = \sum_{s' \in \mathbf{S}} \sum_{o \in \mathbf{O}} p_t(s' \mid s, a, o) f_t(o \mid s, a) q_t(s, a, o, s') \quad \forall s \in \mathbf{S},\ a \in \mathbf{A}, \text{ and } t < N, \quad (2)$$

where $q_t(s, a, o, s')$ represents the expected QALYs of a patient between years t and $t+1$ if he moves to state $s' \in \mathbf{S}$ in year $t+1$, given that he is in state $s \in \mathbf{S}$, action $a \in \mathbf{A}$ is selected, and observation $o \in \mathbf{O}$ is seen in year t. We also define $q_N(s, a) = q_N(s)$ for $s \in \mathbf{S}$, $a \in \mathbf{A}$, where $q_N(s)$ represents the terminal rewards, i.e., the expected remaining QALYs of the patient at age $50+N$.

Table 3 $f_t(o \mid s, a)$ **Values When** $a \in \mathbf{A}$

s/o	$T-$	$P+$	$C+$	SD
$LR0(0)$	1	0	0	0
$LR1(1)$	$1-\tau_P^a$	τ_P^a	0	0
$LR2(2)$	$(1-\tau_C^a)(1-\omega_a)$	0	τ_C^a	$(1-\tau_C^a)\omega_a$
$HR0(3)$	1	0	0	0
$HR1(4)$	$1-\tau_P^a$	τ_P^a	0	0
$HR2(5)$	$(1-\tau_C^a)(1-\omega_a)$	0	τ_C^a	$(1-\tau_C^a)\omega_a$
$PC0(6)$	1	0	0	0
$PC1(7)$	$1-\tau_P^a$	τ_P^a	0	0
$PC2(8)$	$(1-\tau_C^a)(1-\omega_a)$	0	τ_C^a	$(1-\tau_C^a)\omega_a$
$UCT(9)$	1	0	0	0
$D(10)$	1	0	0	0

In addition to adenomatous polyps, colonoscopy may detect and remove other non-CRC-related lesions (e.g., hyperplastic polyps), which increases the likelihood of colonoscopic complications such as perforation (Levin et al. 2006). Therefore, we consider the effect of removing such lesions in calculating the disutility of colonoscopy. We define the disutility of colonoscopy with and without polypectomy as $\bar{d}(Co)$ and $\underline{d}(Co)$, respectively. Let $d^1(Co)$ and $d^2(Co)$ denote the disutility of colonoscopy with and without lesion removal (adenomatous polyps or non-CRC-related lesions), where $\bar{d}(Co) = d^1(Co)$. Furthermore, we define β as the probability that a colonoscopy detects a non-CRC-related lesion. Then $\underline{d}(Co) = \beta d^1(Co) + (1-\beta) d^2(Co)$. Note that $\underline{d}(DN) = \bar{d}(DN) = 0$; i.e., the disutility for not undergoing colonoscopy is equal to 0.

We also assign a disutility for having undetected CRC, d_C, because undetected CRC may weaken the patient and/or cause complications including bowel obstruction (Biondo et al. 2004). In addition, let d_{CT} and d_{UCT} denote the disutility of CRC treatment in states $s \in \{LR2, HR2, PC2\}$ and UCT, respectively. We assume that the disutility of colonoscopy is incurred at the

beginning of the year, whereas d_C, d_{CT}, and d_{UCT} accrue uniformly during the year. Finally, let $\kappa^t_{s,o}(a)$ and κ^t_{UCT} represent the conditional probability of immediate mortality from the complications of colonoscopy and CRC treatment, given that the patient dies in this decision epoch. Then

$$
\begin{aligned}
&q_t(s,a,o,s')\\
&=1-\underline{d}(a) \quad \text{if } o=T-,\ s\in\mathbf{S}\backslash\{LR2,HR2,PC2,UCT,D\},\\
&\qquad\qquad s'\in\mathbf{S}\backslash\{UCT,D\},\ a\in\mathbf{A};\\
&=1-\bar{d}(Co) \quad \text{if } o=P+,\ s\in\{LR1,HR1,PC1\},\\
&\qquad\qquad s'\in\mathbf{S}\backslash\{UCT,D\},\ a=Co;\\
&=1-d_C \quad \text{if } o=T-,\ s,s'\in\{LR2,HR2,PC2\},\ a\in\mathbf{A};\\
&=1-d_{UCT} \quad \text{if } o=T-,\ s=UCT,\ s'\in\mathbf{S}\backslash D,\ a=DN;\\
&=1-d_{CT} \quad \text{if } o\in\{C+,SD\},\ s\in\{LR2,HR2,PC2\},\\
&\qquad\qquad s'\in\mathbf{S}\backslash D,\ a\in\mathbf{A};\\
&=(0.5-\underline{d}(a))(1-\kappa^t_{s,T-}(a))\\
&\qquad \text{if } o=T-,s\in\mathbf{S}\backslash\{LR2,HR2,PC2,UCT,D\},\\
&\qquad\qquad s'=D,\ a\in\mathbf{A};\\
&=(0.5-\bar{d}(Co))(1-\kappa^t_{s,P+}(Co))\\
&\qquad \text{if } o=P+,\ s\in\{LR1,HR1,PC1\},\\
&\qquad\qquad s'=D,\ a=Co;\\
&=0.5(1-d_C)(1-\kappa^t_{s,T-}(a))\\
&\qquad \text{if } o=T-,\ s\in\{LR2,HR2,PC2\},\\
&\qquad\qquad s'=D,\ a\in\mathbf{A};\\
&=0.5(1-d_{UCT})(1-\kappa^t_{UCT})\\
&\qquad \text{if } o=T-,\ s=UCT,\ s'=D,\ a=DN;\\
&=0.5(1-d_{CT})(1-\kappa^t_{s,o}(a))\\
&\qquad \text{if } o\in\{C+,SD\},s\in\{LR2,HR2,PC2\},\\
&\qquad\qquad s'=D,\ a\in\mathbf{A};\\
&=0 \quad \text{otherwise.}
\end{aligned}
$$

The immediate rewards for the first five cases of $q_t(s,a,o,s')$ are equal to one year minus the annual disutility accrued since the patient survives through the year. The remaining cases consider the event that the patient dies during year t. For instance, if a low-risk patient dies after undergoing a colonoscopy with polypectomy (seventh case), then there are two possibilities: the patient may die from complications immediately after the colonoscopy (with probability $\kappa^t_{LR1,P+}(Co)=\mu^t_{1,P+}/(\delta^t_{1,P+}(Co))$) and the immediate reward is 0. The patient may also die during year t, i.e., average lifetime is equal to 0.5 years, and the immediate reward is equal to $(0.5-\bar{d}(Co))$ with probability $1-\kappa^t_{LR1,P+}(Co)$. A table form representation of the $q_t(s,a,o,s')$ values is available in Online Appendix B.

4.4. Belief States

The state space of our POMDP model includes completely observable health states (*UCT* and *D*) and all possible belief states that are probability distributions defined over the remaining health states. Each belief state, b_t, is a probability vector where $b_t(s)$, $s\in\mathbf{S}'\equiv\mathbf{S}\backslash\{UCT,D\}$ denotes the probability of being in health state s in year t. That is,

$$b_t=[b_t(LR0),b_t(LR1),\ldots,b_t(PC2)],$$

where $b_t(s)=P(s_t=s)\geq 0$ and $\sum_{s\in\mathbf{S}\backslash\{UCT,D\}}b_t(s)=1$.

For example, [0.918, 0.08, 0.002; 0, 0, 0; 0, 0, 0] refers to a low-risk patient for whom probability of having adenomatous polyp and CRC is 0.08 and 0.002, respectively. Because the risk level of a patient changes only after a completely observable test result ($P+$, $C+$, or SD), the following equation holds for only one j value: $\sum_{s\in\mathbf{S}_j}b_t(s)=1$, where $j\in\{1,2,3\}$, $\mathbf{S}_1=\{LR0,LR1,LR2\}$, $\mathbf{S}_2=\{HR0,HR1,HR2\}$, and $\mathbf{S}_3=\{PC0,PC1,PC2\}$. We define $\mathbf{B}$ as the set of all such belief states. The belief state is updated each year using a Bayesian update scheme after observing the screening result. The belief state updates in our POMDP model are different than those in classical POMDPs because of a different order of events in our approach (see Figure 2).

As noted by Smallwood and Sondik (1973), the belief state is a sufficient statistic for the past sequence of observations (screening results in our case). That is, by constantly updating the belief state after each observation, we save the information about the previous screening decisions and screening results, i.e., the personal screening history of the patient. For example, consider two patients who received CRC treatment at age 50. Assume that the first patient underwent two colonoscopies with $T-$ result within five years after the treatment; the second patient did not have any colonoscopy surveillance. The updated belief state of the first patient indicates less risk of having CRC and adenomatous polyp at age 55 than that of second patient.

We define $b^a_{b_t,o}$ as the updated belief state in year $t+1$, which is computed as follows, where $\tilde{b}^a_{b_t,o}(s)$ refers to the probability of occupying state $s\in\mathbf{S}$ at $t+1$ given the prior belief state (b_t), action (a), and observation (o) (see Online Appendix C). Note that $b^a_{b_t,o}(s)$ is the normalization of $\tilde{b}^a_{b_t,o}(s)$ with $(1-\tilde{b}^a_{b_t,o}(D)-\tilde{b}^a_{b_t,o}(UCT))$ because belief states are defined only over partially observable health states.

$$
\begin{aligned}
b^a_{b_t,o}(s)&=P(s_{t+1}=s\,|\,b_t,a,o)\\
&=\frac{\tilde{b}^a_{b_t,o}(s)}{1-\tilde{b}^a_{b_t,o}(D)-\tilde{b}^a_{b_t,o}(UCT)},\\
&\qquad \forall s\in\mathbf{S},a\in\mathbf{A},\ \text{and } o\in\mathbf{O} \text{ given that } s_{t+1}\in\mathbf{S}',\\
\text{where}\quad \tilde{b}^a_{b_t,o}(s)&=\frac{\sum_{s'\in\mathbf{S}'}p_t(s\,|\,s',a,o)f_t(o\,|\,s',a)b_t(s')}{\sum_{s''\in\mathbf{S}'}f_t(o\,|\,s'',a)b_t(s'')},\\
&\qquad \forall s\in\mathbf{S},\ a\in\mathbf{A},\ \text{and } o\in\mathbf{O}. \quad (3)
\end{aligned}
$$

The *UCT* state is not associated with screening decisions; therefore, once a patient leaves the *UCT* state and survives during year t, the new belief state, b^t_{UCT}, in year $t+1$ is computed as follows:

$$b^t_{UCT}(s) = P(s_{t+1}=s \mid s_t = UCT, a = DN, o = T-) = \frac{p_t(s \mid UCT, DN, T-)}{1-\sum_{s\in\{UCT,D\}} p_t(s \mid UCT, DN, T-)}, \quad \forall s \in \mathbf{S}_3. \quad (4)$$

4.5. Optimality Equations

The objective of the POMDP model is to maximize the expected TQALYs, the sum of the expected immediate rewards. We define the *optimal value functions* $V^*_t(b_t)$, $W^*_t(b_t)$, and $Y^*_t(b_t)$ as the maximum expected TQALYs from year t through N for post-CRC, high-risk, and low-risk patients, respectively, and $V^*_t(UCT)$ as the optimal value function for a patient in the *UCT* state. We derive these recursive optimality equations for any $b_t \in \mathbf{B}$ as follows:

$$V^*_t(b_t) = \max_{a\in\mathbf{A}}\Bigg[\sum_{s\in\mathbf{S}'} b_t(s)q_t(s,a) + \lambda_t \sum_{s\in\mathbf{S}'}\sum_{o\in\mathbf{O}} b_t(s)f_t(o\mid s,a)\big(l_t(b_t,a,o)V^*_{t+1}(b^a_{b_t,o}) + \tilde{b}^a_{b_t,o}(UCT)V^*_{t+1}(UCT)\big)\Bigg], \quad \text{for } t<N, \quad (5)$$

where $l_t(b_t, a, o) = (1-\tilde{b}^a_{b_t,o}(D)-\tilde{b}^a_{b_t,o}(UCT))$ denotes the probability of not occupying a completely observable state and $0 < \lambda_t \le 1$ represents the discount factor. Similarly,

$$V^*_t(UCT) = q_t(UCT,DN) + \lambda_t p_t(UCT \mid UCT, DN, T-)V^*_{t+1}(UCT) + \lambda_t\Bigg[\Bigg(1-\sum_{s'\in\mathbf{S}\setminus\mathbf{S}'} p_t(s' \mid UCT, DN, T-)\Bigg)V^*_{t+1}(b^t_{UCT})\Bigg], \quad \text{for } t<N, \quad (6)$$

$$W^*_t(b_t) = \max_{a\in\mathbf{A}}\Bigg[\sum_{s\in\mathbf{S}'} b_t(s)q_t(s,a) + \lambda_t\sum_{s\in\mathbf{S}'}\Bigg[\sum_{o\in\{T-,P+\}} b_t(s)f_t(o\mid s,a)l_t(b_t,a,o)W^*_{t+1}(b^a_{b_t,o}) + \sum_{o\in\{C+,SD\}} b_t(s)f_t(o\mid s,a)\big(l_t(b_t,a,o)V^*_{t+1}(b^a_{b_t,o}) + \tilde{b}^a_{b_t,o}(UCT)V^*_{t+1}(UCT)\big)\Bigg]\Bigg] \quad \text{for } t<N, \quad (7)$$

$$Y^*_t(b_t) = \max_{a\in\mathbf{A}}\Bigg[\sum_{s\in\mathbf{S}'} b_t(s)q_t(s,a) + \lambda_t\sum_{s\in\mathbf{S}'} b_t(s)\Bigg[f_t(T-\mid s,a)l_t(b_t,a,T-)Y^*_{t+1}(b^a_{b_t,T-}) + f_t(P+\mid s,a)l_t(b_t,a,P+)W^*_{t+1}(b^a_{b_t,P+}) + \sum_{o\in\{C+,SD\}} f_t(o\mid s,a)\big(l_t(b_t,a,o)V^*_{t+1}(b^a_{b_t,o}) + \tilde{b}^a_{b_t,o}(UCT)V^*_{t+1}(UCT)\big)\Bigg]\Bigg] \quad \text{for } t<N, \quad (8)$$

$$V^*_N(b_N) = W^*_N(b_N) = Y^*_N(b_N) = \sum_{s\in\mathbf{S}'} b_N(s)q_N(s), V^*_N(UCT) = q_N(UCT).$$

The first component of each optimal value function represents the expected immediate reward in year t, and the other components represent the total expected future rewards of a patient when observation o is seen in year t. The optimality equations for V^*_t, W^*_t, and Y^*_t are different because the risk level of the patient may change based on the current observation o.

4.6. Structural Properties

We show that the optimal value functions are PW&C in belief state b_t, which allows us to use Eagle's variant of Monahan's (E&M) algorithm for solving our POMDP model (Cassandra 1994). This property implies that the optimal value functions for a given belief state can be expressed as the maximum of multiplication of the belief state and a finite set of vectors, which are called α-vectors. Although Smallwood and Sondik (1973) prove that the optimal value functions are PW&C in the general POMDP framework, their results are not directly applicable to our model, since our model is different than the conventional POMDP models. First, our model includes completely observable states and dynamically changing risk levels. In addition, the observation scheme and belief update formula differ from those of the general POMDP framework. Therefore, we prove that the optimal value functions in our model are PW&C, as presented in Theorem 1. The proof of Theorem 1 is available in Online Appendix D. In addition, Online Appendix E characterizes the α-vectors defined in Theorem 1.

THEOREM 1. *$V^*_t(b_t)$, $W^*_t(b_t)$, and $Y^*_t(b_t)$, are PW&C in $b_t \in \mathbf{B}$. That is,*

$$V^*_t(b_t) = \max_{0\le k\le|\alpha_t|}\sum_{s\in\mathbf{S}'} b_t(s)\alpha^k_t(s), \quad W^*_t(b_t) = \max_{0\le k\le|\dot{\alpha}_t|}\sum_{s\in\mathbf{S}'} b_t(s)\dot{\alpha}^k_t(s), \quad \text{and} \quad Y^*_t(b_t) = \max_{0\le k\le|\ddot{\alpha}_t|}\sum_{s\in\mathbf{S}'} b_t(s)\ddot{\alpha}^k_t(s) \quad \forall b_t \in \mathbf{B},\ t \le N, \quad (9)$$

Table 4 Input Parameters

Notation	Parameter description	Estimation method and data source
$\rho^t_{0,1}$, $\rho^t_{3,4}$	Polyp onset probabilities	Reported in the National Polyp Study (Loeve et al. 2004)
$\rho^t_{1,2}$, $\rho^t_{4,5}$	Polyp-to-CRC progression probabilities	Calibration based on the SEER database (SEER 2012) and the literature
$\rho^t_{6,7}$, $\rho^t_{7,8}$	Lesion progression probabilities for post-CRC patients	Derived from MCRC-NH simulation using data from Mayo Clinic (Erenay et al. 2011)
$\rho^t_{i,6}$, $\rho^t_{i,7}$, $\rho^t_{i,8}$ for $i \in \{2, 5, 8, 9\}$	Lesion progression probabilities after CRC treatment	Meta-analysis based on Longo et al. (2000), Ohlsson and Pålsson (2003), Yun et al. (2008), and Erenay et al. (2011)
$\delta^t_{i,T-}(DN)$ for $i \in \{0, 1, 3, 4\}$	Probabilities of mortality for CRC-free patients	Reported in the U.S. life tables (Arias 2007)
$\delta^t_{2,T-}(DN)$, $\delta^t_{5,T-}(DN)$	Probabilities of mortality from undetected CRC for low- and high-risk patients	Calibration based on the SEER database (SEER 2012) and the literature
$\delta^t_{i,T-}(DN)$ for $i \in \{6, 7, 8\}$	Probabilities of mortality for post-CRC patients	Derived from MCRC-NH simulation using data from Mayo Clinic (Erenay et al. 2011)
$\delta^t_{9,T-}(DN)$	Probability of mortality in the *UCT* state	Reported in the SEER database as survival rate for distant CRC (SEER 2012)
$\delta^t_{i,T-}(Co)$ for $i \leq 8$, $\delta^t_{i,P+}(Co)$ for $i \in \{1, 4, 7\}$	Probabilities of mortality after a colonoscopy with $T-$ and $P+$ results	Calculated according to Equations (10) and (11) (see Online Appendix A)
$\delta^t_{2,o}(Co)$, $\delta^t_{5,o}(Co)$, $\delta^t_{8,o}(Co)$ for $o \in \{C+, SD\}$	Probabilities of mortality after CRC treatment	Reported in the SEER database as CRC survival rates (SEER 2012)
γ^t_i for $i \in \{2, 5, 8, 9\}$	Probabilities that CRC treatment will be completed within one year	Meta-analysis based on Longo et al. (2000), Ohlsson and Pålsson (2003), Yun et al. (2008), and Erenay et al. (2011)
$\bar{\mu}^t_{i,T-}$, $\underline{\mu}^t_{i,T-}$, $\mu^t_{i,P+}$	Probabilities of mortality from colonoscopy complications	Meta-analysis based on Vijan et al. (2007), Gatto et al. (2003), Levin et al. (2006), and Arias (2007)
d_C	Disutility of experiencing undetected CRC, 0.134 years	Meta-analysis based on Syngal et al. (1998) and SEER (2012)
d_{CT}, d_{UCT}	Disutilities of having CRC treatment and being in the *UCT* state, 0.4 and 0.75 years	Meta-analysis based on Ness et al. (2000) and SEER (2012)
$\underline{d}(Co)$, $\bar{d}(Co)$	Disutility of colonoscopy, 0.8 and 2.4 weeks	Meta-analysis based on data from the literature (for details, see Erenay 2010)
$q_N(s)$ for $s \in S \backslash D$	Terminal rewards	Markov process analysis based on disutility values and extrapolated transition probabilities
α^a_P, α^a_C	Sensitivity of colonoscopy for polyp and CRC	Reported in Frazier et al. (2000), Vijan et al. (2007) as 85% and 90%, respectively
ω_a	Probability of CRC self-detection	Meta-analysis based on Frazier et al. (2000) and SEER (2012)
Costs	Costs associated CRC screening and treatment	Reported in Lansdorp-Vogelaar et al. (2009a) and Vijan et al. (2007)

for some sets of vectors $\alpha_t = \{\alpha^1_t, \alpha^2_t, \ldots\}$, $\dot{\alpha}_t = \{\dot{\alpha}^1_t, \dot{\alpha}^2_t, \ldots\}$, *and* $\ddot{\alpha}_t = \{\ddot{\alpha}^1_t, \ddot{\alpha}^2_t, \ldots\}$, *where* $\alpha^i_t, \dot{\alpha}^i_t, \ddot{\alpha}^i_t \in R^{|\mathbf{S}'|}$ *for all* $i \in \mathbf{Z}$.

5. Parameter Estimation and Validation

We derive most of the input parameters for post-CRC patients using the metachronous CRC natural history (MCRC-NH) simulation model that uses data from the Mayo Clinic at Rochester (Erenay et al. 2011). We estimate the transition probabilities using data from the MCRC-NH simulation, the SEER database, the National Polyp Study (Loeve et al. 2004), U.S. life tables (Arias 2007), and clinical literature. The annual probability of polyp-to-CRC progression and mortality from undetected CRC are not available in any clinical database; thus, we estimate them via a novel calibration method, as described in Online Appendix F. We derive most of the cost inputs from the cost-effectiveness analysis of Lansdorp-Vogelaar et al. (2009a) because they conducted a detailed cost analysis using Medicare cost data and theirs is the only study that estimated CRC treatment and continuing care costs separately. Finally, for the base-case analysis, we set the discount factor λ to 1, as many medical decision makers prefer not discounting health benefits (Severens and Milne 2004); however, we conduct a sensitivity analysis on the discount factor (see §I.2 in Online Appendix I).

Table 4 summarizes our parameter estimation scheme. We do not include the details of some of the input parameters in this table because they vary by age and health state. In addition, we derive some of the input parameters by synthesizing information from multiple sources, thus providing multiple citations for a single parameter. We refer the readers to Erenay (2010) for more details about the input analysis and synthesis of data from multiple sources. Detailed explanations about the parameter estimation and calibration process are provided in Online Appendix F.

We validate our model using three steps. First, we validate our model based on expert opinion by collaborating with Dr. Adnan Said, the chief of the Gastroenterology and Hepatology department in the Veterans Affairs Medical Center in Madison, Wisconsin. Second, we compare our model outputs with the statistics from the SEER database (the largest cancer registry in the United States (SEER 2012)) and data from the Mayo Clinic as they appear in Erenay et al. (2011). For example,

Figure 3 **Calibrated vs. SEER's Risk of CRC and CRC Mortality for 50-Year-Old-Males**

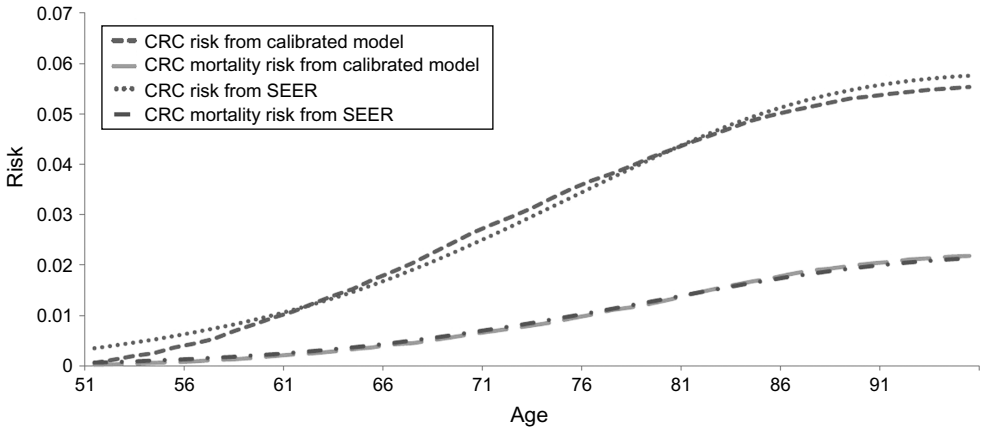


Table 5 **Comparison of the Proposed Model and MISCAN Simulation Model Based on Saved Life-Years**

	Every 20 years	Every 15 years	Every 10 years	Every 5 years
MISCAN	0.096 years	0.104 years	0.113 years	0.123 years
POMDP	0.084 years	0.100 years	0.118 years	0.142 years

we estimate the cumulative risk of developing CRC and risk of CRC mortality between ages 50 and 95 for 50-year-old low- and high-risk males using our model. As illustrated in Figure 3, we observe that these results are very close to those derived from the SEER database. We also compare our model outputs for post-CRC patients with MCRC incidence, polyp incidence, and survival rates after CRC treatment from the Mayo Clinic, the SEER database, and the literature. Therefore, we are content that our model complies with the existing clinical information about CRC progression in the United States.

Finally, we compare our model outputs with that of an existing clinically validated simulation model. Lansdorp-Vogelaar et al. (2009a) used the microsimulation screening analysis (MISCAN) CRC simulation model to estimate the life-year savings for 50-year-old asymptomatic patients through different screening policies (e.g., colonoscopy every 10 years) compared to no screening option under particular assumptions. We estimate the lifetime savings for the same conditions and observe that our estimations are very close to those of Lansdorp-Vogelaar et al. (2009a), as shown in Table 5. These results illustrate that our model represents CRC screening in the U.S. population well. More details about the validation process are available in Online Appendix G.

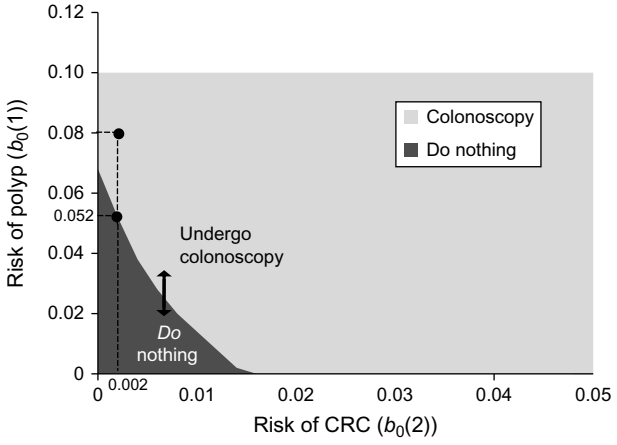
6. Numerical Results

For our numerical experiments, we set the planning horizon for ages 50–100, i.e., $N = 50$. Although we have a large POMDP model with 50 decision epochs and 11 core states, and POMDPs are known to be difficult to solve since the state space is infinite, we optimally solve our POMDP model within several minutes using the E&M algorithm. More details about the E&M algorithm and other POMDP solution methods are available in Cassandra (1994).

6.1. The Optimal Annual Colonoscopy Decisions

Figure 4 shows the optimal colonoscopy decisions of 50-year-old low-risk males for different combinations of CRC and adenomatous polyp risks (probability of having CRC and adenomatous polyp). The light-colored area shows the risk combinations for which it is optimal to undergo colonoscopy ($a_t^*(b_t) = Co$), and the dark-colored area shows those for which it is optimal not to undergo colonoscopy ($a_t^*(b_t) = DN$). The optimal colonoscopy screening decisions in Figure 4 are of control-limit type. For instance, if the CRC risk is equal to 0.002, the optimal colonoscopy decision is *colonoscopy* for all polyp risk values higher than 0.052 and *do nothing* for those lower than 0.052. We verify

Figure 4 **Optimal Colonoscopy Screening Decisions for 50-Year-Old Low-Risk Males**



that such optimal thresholds exist for all age, gender, and risk-level categories (see Online Appendix H).

Although gender-based differences in the optimal screening decisions for low- and high-risk patients are very small, such differences are significant for post-CRC patients. For example, the do nothing areas—i.e., the belief states for which the optimal colonoscopy decision is *DN*—for 60-year-old high-risk males and females (Figure 5) are slightly different because low- and high-risk females are slightly less prone to CRC than males (SEER 2012). However, as shown in Figure 5, the optimal threshold risks for post-CRC females are significantly lower than those for males, which promotes more aggressive screening for females after CRC treatment. This finding is not recognized by the current practice and may appear to be counterintuitive. However, this is because CRC progression is comparable between low-risk females and males, whereas it is more aggressive in post-CRC females (Erenay et al. 2011).

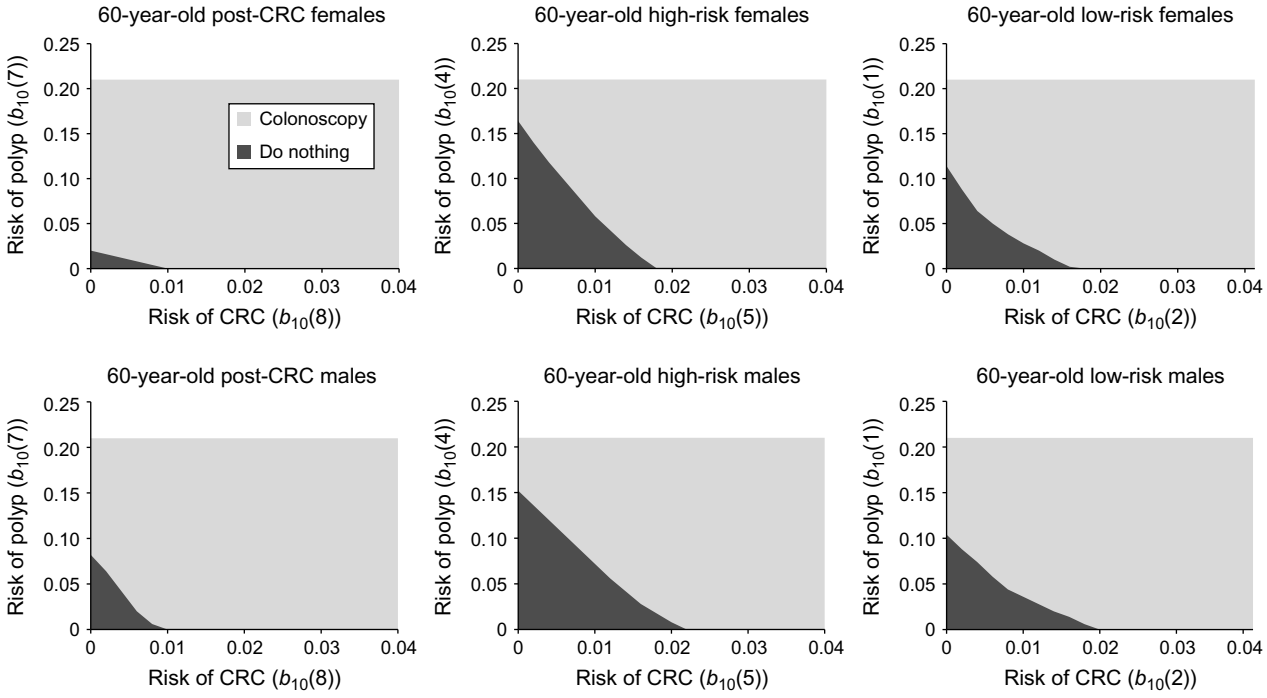
Risk level also affects the optimal risk thresholds (Figure 5). In general, the optimal thresholds for post-CRC patients are lower than those for low- and high-risk patients, which implies more aggressive screening. This may be because the CRC progression is more aggressive for post-CRC patients than it is for low- and high-risk patients (Winawer et al. 1997). Although the do nothing areas for low-risk patients are smaller than those for high-risk patients, the optimal screening policies recommend less frequent screening for low-risk patients than for high-risk patients, as shown in §6.3, which appears to be counterintuitive. However, note that polyp onset probabilities for high-risk patients are at least two times higher than those for low-risk patients (Loeve et al. 2004), whereas the proportion of the do nothing areas for high-risk and low-risk patients is less than two at each age. That is, the risk of having adenomatous polyp for a high-risk patient increases faster than that for a low-risk patient; therefore, it takes less time for a high-risk patient to reach a belief state outside of the do nothing area.

Finally, we observe that the optimal thresholds typically increase as age increases, which implies that the optimal screening decisions suggest more aggressive screening for younger patients (see Online Appendix H) because as patients age, their expected remaining lifetimes and benefits from CRC screening decrease. For example, clinicians currently hesitate to screen 88-year-old patients since they are likely to die before developing CRC (Schoen et al. 2006).

6.2. Optimal Colonoscopy Screening Policies

We refer to the process of deriving the optimal colonoscopy policies from the optimal annual colonoscopy decisions as *path enumeration*, which is summarized below. Algorithm 1 computes the optimal policy, a_t^* for $t_f \le t \le t_l$, given the current belief state (b), the future screening results ($o_t(\pi)$ where π is a scenario), and the first and last specified decision epochs (t_f and t_l). For example, the optimal preventative screening policy for low-risk patients between ages 50 and 60 can be found by analyzing the scenario where the screening result is $o_t(\pi) = T-$ for $0 \le t \le 10$.

Figure 5 Optimal Colonoscopy Screening Decisions and Gender



Algorithm 1 (Path Enumeration Algorithm)

Step 1. Let $t = t_f$ and $b_t = b$.

Step 2. Set $a_t^* = DN$ if $a_t^*(b_t) = DN$. Otherwise, set $a_t^* = Co$.

Step 3. Calculate the updated belief $b^{a_t^*}_{b_t,\, o_t(\pi)}$ using Equations (3) and (4).

Step 4. Set $t = t + 1$. Set $b_t = b^{a_t^*}_{b_t,\, o_t(\pi)}$.

Step 5. If $t > t_l$ stop. Otherwise, return to Step 2.

We estimate the initial belief states using the adenomatous polyp and CRC prevalence values reported in the literature and the SEER database. For instance, we estimate the belief state for 50-year-old low-risk males as [0.918, 0.08, 0.002; 0, 0, 0; 0, 0, 0]. The National Polyp Study reports the polyp prevalence at age 50–59 as 15%, which can be obtained only if the probability of having polyp at age 50 is 0.08 based on the polyp onset probabilities from the same study (Loeve et al. 2004). We estimate the probability of having CRC at age 50 for low-risk patients as 0.002, which is the limited prevalence of CRC (prevalence within one year of age 50) from the SEER database (SEER 2012).

Figure 6 illustrates how path enumeration derives the optimal preventative colonoscopy screening policy for a 50-year-old low-risk male. The belief states in this figure show only the probabilities of being in the $LR0$, $LR1$, and $LR2$ states for the brevity. The belief state for CRC risk = 0.002 and polyp risk = 0.08 falls into the light-colored area of the graph for 50-year-old low-risk males. Therefore, the optimal policy recommends a colonoscopy at age 50 for this patient. Then, assuming $o_0 = T-$, we update the belief state using Equation (3). The new belief state at age 51 indicates a smaller CRC and adenomatous polyp risk (0.0003 and 0.0267) because colonoscopy is very accurate and having a $T-$ colonoscopy result is a significant indicator for having no colorectal lesion at age 50. The optimal action for this updated belief state is DN.

Assuming that no self-diagnosis takes place ($o_t = T-$; $1 \le t \le 4$), the belief state of this patient passes the threshold at age 55 and moves to the light-colored area, which suggests undergoing colonoscopy. Therefore, Figure 6 shows that it is optimal for a 50-year-old low-risk male patient to undergo colonoscopy, and if the screening test result is $T-$, he should undergo another colonoscopy after five years unless a self-diagnosis occurs. In contrast, the guidelines suggest that he should undergo the next colonoscopy after 10 years. Because we consider screening decisions only at and after age 50, this result only indicates that if an asymptomatic 50-year-old patient has not been screened yet, he should undergo colonoscopy in this year. However, this may not be the optimal time for initial colonoscopy screening if the earlier ages are considered. Because of data limitations, finding the optimal time of initial colonoscopy screening is not within the scope of this paper.

Age and gender affect optimal screening policies significantly. For example, Figure 7 illustrates the optimal colonoscopy screening policies after a polyp diagnosis ($P+$) for 60-, 70-, and 79-year-old low-risk patients who become high-risk after the polypectomy. Note that in this and the following figure, the belief states for the initial ages show only the probabilities of being

Figure 6 Optimal Colonoscopy Screening Policy Generation

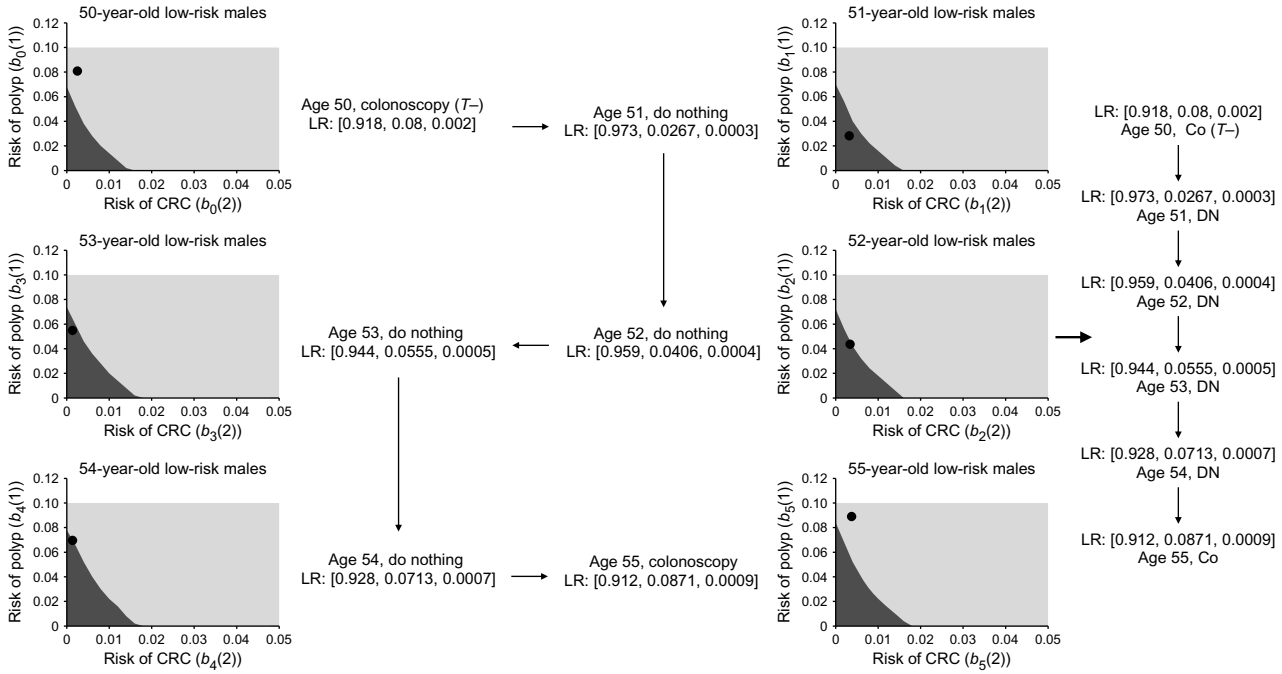
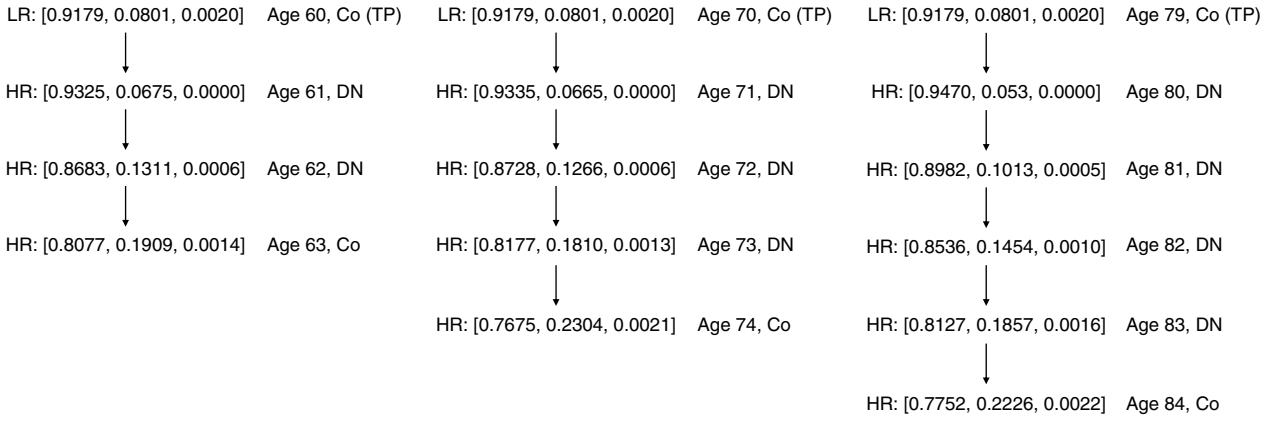


Figure 7 Effect of Age on Colonoscopy Screening Policies After Polypectomy for Males



in the $LR0$, $LR1$, and $LR2$, although they show the probabilities of being in the $HR0$, $HR1$, and $HR2$ states for the other ages. Current guidelines suggest a follow-up colonoscopy in three years, regardless of age. The optimal screening policies suggest the same policy for 60-year-old males. However, the optimal follow-up interval is four and five years for 70- and 79-year-old males, respectively, because the benefit of follow-up decreases as patients age because of the increasing mortality risk. Furthermore, Figure 8 compares the optimal colonoscopy policies for 76-year-old low-risk males and females after a polypectomy. The optimal follow-up interval is four years for males and five years for females. This difference is a result of low-risk and high-risk females being slightly less prone to CRC than males (SEER 2012).

The optimal screening policies obtained from our POMDP model also determine the stopping age for screening. Most guidelines do not specify a stopping age, except the recent USPSTF guidelines recommend stopping CRC screening at age 75 for low-risk patients only (USPSTF 2008), based on the simulation model of Zauber et al. (2008), which evaluates four different termination scenarios. Unlike Zauber et al. (2008), our model evaluates all possible stopping ages for colonoscopy screening. In addition, our POMDP approach determines the optimal stopping age as a function of the patient's current risk. We determine the age to stop screening as follows: If a patient with belief b_t undergoes colonoscopy with a $T-$ outcome and his future belief states always fall into the do nothing area, then he should stop screening until a self-diagnosis of CRC (SD) takes place. The following definition presents the conditions under which optimal screening decisions recommend terminating colonoscopy screening at age $50+t$.

Definition 1. Let b_n, $N > n > t$ be a sequence of belief states where $b_n = b^{DN}_{b_{n-1},\,T-}$ for $n > t$ and $b_t = b$. Then a $50+t$-year-old patient with belief state b should undergo the last colonoscopy and then stop screening, if $a^*_t(b) = Co$ and $a^*_n(b_n) = DN$ for $n > t$.

Figure 8 Colonoscopy Screening Policies After Polypectomy and Effect of Gender

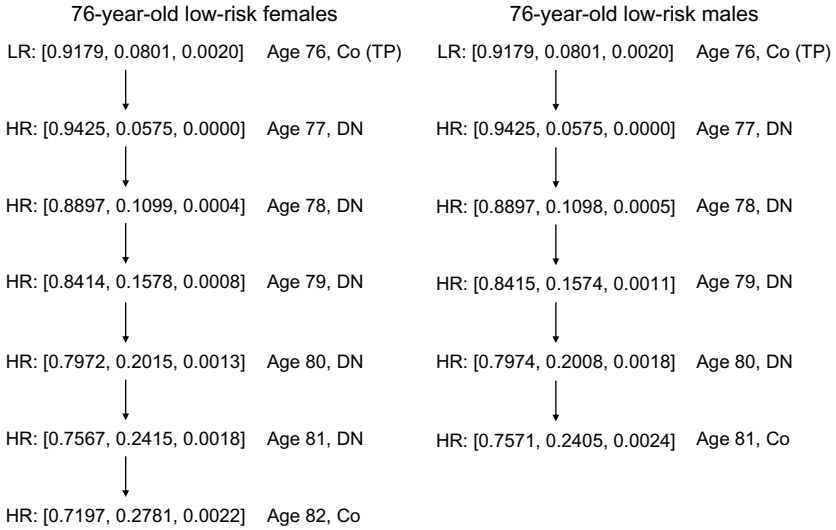
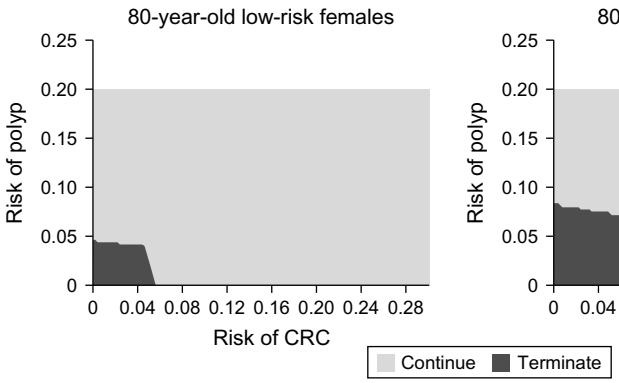


Figure 9 Decisions on Colonoscopy Screening Termination

(a) Colonoscopy screening termination decisions for 80-year-old patients



(b) Age to stop colonoscopy screening

	Males	Females
Low-risk patients	78	80
High-risk patients	83	83
Post-CRC patients	86	91

This definition guarantees that once screening is terminated, all later ages also satisfy the termination criterion. Figure 9(a) illustrates the screening termination decisions for 80-year-old low-risk patients, where the dark-colored area refers to the CRC and adenomatous polyp risk combinations for which CRC screening should be terminated. Note that the dark area is smaller for low-risk females, which means that low-risk females should terminate CRC screening later than males. This may seem counterintuitive, as the optimal colonoscopy decisions imply less frequent colonoscopy screening for low-risk females than males (see Figure 5). However, this is because elderly females generally live longer and may benefit more from a future screening than males. Moreover, Figure 9(b) illustrates the ages to stop screening for the patients who follow the optimal colonoscopy screening policies and always have a $T-$ screening result after age 50. We find that females, especially in post-CRC level, should stop screening later than males, and patients in higher risk levels should stop colonoscopy screening later than those in lower risk levels.

6.3. Comparison of the Optimal Policies and Guidelines

We compare the performance of our proposed policies to that of the guidelines of AGA (Table 1) and no screening using the patient-level simulation model that is described in §F.4 of Online Appendix F. We perform 10 million replications for each gender, risk level, and policy type. We set the stopping age for colonoscopy screening as 75 for low-risk patients, as suggested by USPSTF (2008), and as 85 for both high-risk and post-CRC patients, which is the current practice at the gastroenterology center of our institution's hospital.

We use six performance measures: (1) expected TQALYs; (2) expected number of colonoscopies per patient; (3) lifetime CRC risk, referring to the cumulative probability of having CRC between ages 50 and 100; (4) CRC mortality, referring to the probability of mortality from CRC between ages 50 and 100; (5) total costs, referring to the expected CRC-related costs, including those of screening and treatment; and (6) screening interval, which refers to the average time (in years) between two consecutive colonoscopy screenings. The percentage improvement refers to percentage reduction for CRC risk and CRC mortality criteria, whereas it refers to percentage increase for others.

Table 6 illustrates the results for 50-year-old patients. Both the guidelines and the optimal policies reduce the lifetime CRC risk and mortality (up to 95%) and are associated with higher TQALYs. Note that the improvements are higher for post-CRC patients than for others. The optimal policies outperform the guidelines in TQALYs, CRC risk, and CRC mortality. Compared to the guidelines, the optimal policies are associated with up to 86% and 89% less lifetime CRC risk and mortality, respectively. Moreover, a patient is expected to have up to 3.2% more TQALYs by following the optimal policies rather than the guidelines. Although the improvements in the TQALY criterion may appear low for low- and high-risk patients, considering that there were 4.5 million 50-year-old U.S. residents in 2008, these improvements translate into thousands of life-year savings in the United States.

Performance improvements of the optimal policies over guidelines are similar in males and females. However, the performance improvements in post-CRC females are significantly higher than those in any other patient group, including post-CRC males. This is because the CRC progression is significantly more aggressive in post-CRC females than it is for any other patient group (Erenay et al. 2011). In addition, there is evidence that the CRC progression in post-CRC patients may be faster than it is reported in the literature (Erenay et al. 2011). This implies that the current guidelines may underemphasize the CRC surveillance in their recommendations.

Also note that, for each gender and risk level, the optimal policies recommend more frequent screening than the guidelines, which is feasible in terms of

Table 6 Comparison of Guidelines and the Optimal Policies for 50-Year-Old Patients

	No screening	Guidelines	Optimal policies	Improvement over no screening (%)	Improvement over guidelines (%)
Males					
Low-risk patients					
TQALYs	28.272	28.502	28.547	0.97	0.16
No. of colonoscopies	0	3.72	6.96	—	87
Lifetime CRC risk	7.59%	3.41%	2.01%	74	41
CRC mortality	3.12%	1.24%	0.67%	79	46
Total cost	$26,635	$25,949	$27,319	2.57	5.28
Screening interval	—	8.45	4.12	—	—
High-risk patients					
TQALYs	27.634	28.348	28.403	2.78	0.19
No. of colonoscopies	0	6.43	10.58	—	64
Lifetime CRC risk	14.72%	3.93%	2.59%	82	34
CRC mortality	5.73%	1.17%	0.75%	87	35
Total cost	$34,440	$29,513	$31,397	−8.83	6.39
Screening interval	—	4.50	2.54	—	—
Post-CRC patients					
TQALYs	21.375	22.088	22.288	4.27	0.90
No. of colonoscopies	0	4.54	12.46	—	175
Lifetime CRC risk	29.78%	12.47%	3.79%	87	70
CRC mortality	13.12%	4.90%	1.13%	91	77
Total cost	$135,551	$130,875	$132,919	−1.94	1.56
Screening interval	—	4.16	1.77	—	—
Females					
Low-risk patients					
TQALYs	32.121	32.353	32.396	0.85	0.13
No. of colonoscopies	0	4.00	7.05	—	76
Lifetime CRC risk	6.96%	3.16%	1.95%	72	38
CRC mortality	3.09%	1.24%	0.65%	79	48
Total cost	$25,689	$25,338	$26,627	3.65	5.09
Screening interval	—	8.37	4.52	—	—
High-risk patients					
TQALYs	31.496	32.191	32.239	2.36	0.15
No. of colonoscopies	0	7.00	10.18	—	45
Lifetime CRC risk	13.13%	3.56%	2.66%	80	25
CRC mortality	5.55%	1.16%	0.76%	86	34
Total cost	$33,204	$29,198	$30,589	−7.88	4.76
Screening interval	—	4.50	2.97	—	—
Post-CRC patients					
TQALYs	24.067	25.237	26.038	8.19	3.18
No. of colonoscopies	0	5.36	25.99	—	384
Lifetime CRC risk	42.62%	26.00%	3.63%	91	86
CRC mortality	20.26%	10.41%	1.10%	95	89
Total cost	$159,055	$156,468	$161,843	1.75	3.44
Screening interval	—	4.30	1.02	—	—

screening capacity because many clinicians already recommend colonoscopy screening as frequently as the optimal policies. For example, the average screening interval of the optimal policy for 50-year-old low-risk males is 4.12 years, and Klabunde et al. (2009) report that approximately 35% of the clinicians recommend colonoscopy every 3–5 years for low-risk patients. The increasing number of colonoscopy screenings increases screening costs and decreases CRC treatment costs by preventing more CRC cases. Therefore, the reduction in treatment costs compensates a proportion of the screening costs, which gives an edge to the optimal policies to be cost-effective. For example, the cost-effectiveness ratios based on the guidelines for 50-year-old low-risk, high-risk, and post-CRC males are $42,283, $49,436, and $14,917 per QALY, respectively, which are all less than the acceptable cost-effectiveness ration thresholds of $50,000 or $100,000 per QALY.

Several input parameters involve uncertainty because there is limited research on them or they depend on patient preferences. Therefore, we conduct sensitivity analyses on all cause disutilities, colonoscopy disutilies, CRC treatment disutilities, probability of complications, colonoscopy sensitivity, compliance rate, and discount rate. We perform (1) robustness analyses, which measure the performance of the base-case optimal policy

with different parameter settings; and (2) sensitivity analyses, which measure the effect of these parameters on the performance of the optimal policies. The details of the sensitivity analysis are provided in Online Appendix I. The results of the sensitivity analyses illustrate that the optimal policies are robust to the changes in several parameters, including colonoscopy sensitivity and CRC treatment disutility, and somewhat sensitive to the others. Colonoscopy disutility appears to be the most sensitive input parameter. In contrast, the optimal policies, which are more aggressive than the guidelines, generally outperform the guidelines except for the cases with high colonoscopy disutility. These results imply that the practices of the clinicians who advocate more aggressive CRC screening might be justifiable.

Deviating from the guidelines was discouraged, as it may unnecessarily increase disutility and complication risk (Mysliwiec et al. 2004). However, Table 6 shows that improvements in health benefits from more frequent screening than the guidelines may compensate the negative effects of increasing disutility and complication risk. To verify this observation, we compare the performance of the guidelines with the screening policies that are recommended by the clinicians to low-risk and high-risk patients as reported in Klabunde et al. (2009) and Mysliwiec et al. (2004). We observe that the conclusions we draw above are correct for both risk levels. For example, practitioners recommend undergoing a colonoscopy every 3–10 years for the low-risk patients. Among these fixed-interval screening policies, which we refer to as *simple policies*, we find that screenings every 4, 5, and 6 years perform best with respect to the TQALYs, with around 0.11%–0.12% TQALY improvements over guidelines.

We also measure the performances of these simple policies for different colonoscopy disutility scenarios for low-risk patients. We observe several things: (1) As the disutilities decrease, the performance gap between the optimal and simple policies increases and becomes significant. For example, the percentage improvements in TQALYs are 0.16% and 0.12% for the optimal policy and best simple policies with base-case disutility values, whereas the performance improvements are 0.30% versus 0.22% in 0.5× base-case disutility scenario, respectively. (2) The simple policy with the best TQALY performance varies as the disutility values change; e.g., screenings every 3, 4–5, and 6 years have the best performance for 0.5× , 1×, and 1.5× base-case disutility scenarios, respectively. We conduct similar experiments for high-risk patients and observe similar results with higher optimality gaps.

7. Conclusions

We develop a POMDP model and determine the optimal colonoscopy screening policies for CRC prevention and surveillance. The optimal annual colonoscopy decisions are of control-limit type based on the risk of having CRC and adenomatous polyp for all age, gender, and risk-level categories that appeal to the clinicians who are already familiar with making decisions based on the estimated risk values and thresholds. Some of estimated parameters of the model, e.g., the disutility values, may not represent the preferences of all patients. Therefore, we suggest the readers focus on the insights obtained by our study, which are summarized as follows. Although the current guidelines are effective, as they provide a significant portion of the possible QALY benefits compared to no screening, there is some room for improvement in colonoscopy screening practice to increase TQALYs and decrease CRC risk and CRC mortality. Clinicians may develop better screening policies by incorporating factors including (but not limited to) age, gender, and personal history into the screening decisions, i.e., more frequent screening, decreasing the screening frequency as patients age, recommending different screening intervals for males and females, readjusting the screening termination ages, and recommending later screening termination for females. The incorporation of these factors and individual risk of CRC and adenomatous polyp into screening decisions is also a significant step toward personalizing CRC screening (Fletcher 2008). The POMDP framework that we propose in this work may be useful for the endeavors in this direction.

As illustrated in §6, age, gender, and risk level affect optimal colonoscopy screening policies. For example, the optimal policies suggest slightly less frequent CRC screening for low- and high-risk females and more frequent screening for post-CRC females than males in the corresponding risk levels. Moreover, the optimal policies suggest that females stop CRC screening later than males. Current guidelines and the literature have not recognized these results as asymptomatic females and males having similar CRC incidence and survival (SEER 2012). However, females live longer than males and lose more QALYs once they develop CRC; thus, they can benefit from a colonoscopy at late ages. In addition, current guidelines do not consider the high MCRC incidence in females (Erenay et al. 2011); therefore, they ignore gender effect on CRC surveillance.

We also observe that the optimal screening policies generally suggest more frequent colonoscopy screening than the current guidelines. Although the majority of the clinicians follow the guidelines, several studies report a significant number (43%) of clinicians suggest more frequent screening than the guidelines (Mysliwiec et al. 2004, Klabunde et al. 2009). The practice of some of these clinicians is in line with our findings; i.e., colonoscopy every 4–6 years performs acceptably well compared to the optimal policies. However, this

insight should be considered cautiously for two reasons. (1) Screening more frequently may improve the performance of CRC screening and surveillance if the colonoscopy disutility is low or moderate relative to our base case disutility estimations, whereas it may lead to limited additional QALY benefit for a patient with high disutility (see Online Appendix I). (2) Some studies criticize the clinicians who suggest more frequent screening for unnecessarily increasing the risk of complications, screening disutility, and expenditures (Mysliwiec et al. 2004, Goodwin et al. 2011). We find that this criticism may not be justifiable in terms of increased risk of complications and disutility because the increase in the expected TQALYs compensates the corresponding increase in the expected disutility and risk of complications in most cases. However, it is important to note that our disutility value estimates may not be accurate and we do not consider costs in our optimization model.

Furthermore, we observe that simple periodic colonoscopy screening policies such as screening every 4, 5, or 6 years perform reasonably well compared to the optimal policies for low-risk patients when only age, gender, and personal history are considered. We conduct a similar analysis for high-risk patients and observe that triennial colonoscopy screening performs close to the optimal policy. Therefore, clinicians who prefer simplicity may recommend a periodic screening policy with limited loss in health outcomes. However, if more risk factors such as body mass index, histology of removed polyps, etc. are incorporated into the POMDP model, the performance of the optimal policies may become significantly better than that of the simple policies. In that case, simple policies may be enhanced by considering the insights gained from the optimal policies in the current work, such as varying screening intervals with age and gender and gender-based screening termination. Readers should also note that the choice of the best simple policy depends on patient type and associated disutility values because the performance gap between the optimal and simple policies increases as the aggressiveness of CRC progression increases and the disutility values decrease.

Some studies and recent developments comply with the insights provided by the proposed model. For instance, the potential of a more aggressive colonoscopy screening policy has already been realized in some countries; e.g., Austrian health organizations are now recommending that asymptomatic (low-risk) patients undergo colonoscopy every 7–10 years instead of every 10 years (Haidinger et al. 2008). In addition, a recent clinical simulation analysis reports that undergoing more aggressive colonoscopy screening (e.g., every 7 years for low-risk patients) can be more rewarding, as it is associated with a better cost-effectiveness ratio (Wilschut et al. 2011). Furthermore, recent clinical studies pointed out that factors such as gender and age might be considered in CRC screening (Ferlitsch et al. 2011) and, again recently, a few simulation analyses are conducted to test the feasibility of incorporating such factors into the screening decisions (Wilschut et al. 2011, Lansdorp-Vogelaar et al. 2009b). These developments indicate that the current CRC literature has already started to show a move toward more sophisticated and personalized CRC screening.

Limitations of this study are due to the data. Our model considers the polyp and cancer states regardless of the lesion size and location. It would not be practical to add these factors to our model because of the lack of appropriate data and potential issues with the interpretation of our results. Even if we use only two polyp sizes, three CRC stages, and two lesion locations, the optimal policies would be a function of a 10-dimensional belief state instead of a three-dimensional one. Considering only one adenomatous polyp and CRC state enables us to conveniently translate our results to clinicians, as they can easily understand the optimal colonoscopy decisions based on CRC and polyp risk. We anticipate that our single-size polyp and single-stage CRC state definitions do not significantly affect our model's performance. For instance, for verification purposes, we extended our POMDP model to incorporate two CRC stages in the state definition. Although this extension improved the results slightly, the improvements were not high and the trends observed in the original POMDP model were not affected. Similarly, we restrict our analysis to the screening decisions after age 50 because of the data limitations. Therefore, the proposed optimal policies are valid under the assumption of initiating the screening decisions at age 50. As this assumption complies with the current practice, our results are comparable with the guidelines, thus, the provided insights are valid.

Many other issues are related to the optimal colonoscopy screening and surveillance that are yet to be explored. Although we mainly focus on maximizing health outcomes (TQALYs), cost is an important issue to analyze, as colonoscopy costs vary significantly (Tangka et al. 2013). We implicitly consider costs associated with CRC; however, incorporating costs explicitly (representing the society's perspective) into screening decisions may lead to a different set of optimal policies. In addition, several new screening methods such as CT colonograhpy and new generation fecal tests (e.g., fecal immunochemical testing (FIT)) are proposed as alternatives for colonoscopy. For example, biennial FIT screening has been shown to provide comparable performance with one-time colonoscopy in CRC detection (Quintero et al. 2012). However, colonoscopy is associated with better adenoma detection, leading to better CRC prevention (Quintero et al. 2012) and these alternative methods lead to a higher number

of diagnostic follow-up colonoscopies (Zauber et al. 2008). Consequently, a recent survey illustrates that a large portion of clinicians reported an increase in their colonoscopy order rate (Zapka et al. 2012). Therefore, as suggested by Bretthauer and Kalager (2012), analyzing how colonoscopy and other screening methods can be used together would be a good future research direction.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2014.0484.

Acknowledgments

The authors thank Stephen Graves, an anonymous associate editor, and anonymous referees for their suggestions and insights, which improved this paper. The authors thank FICO for providing a license for FICO XPRESS Optimizer to solve the LP problems to determine the undominated alpha-vectors in the POMDP solution procedure. This research was supported by the National Science Foundation [Grant CMMI-0844423].

References

Alagoz O, Ayer T, Erenay FS (2011) Operations research models for cancer screening. Cochran J, Cox LA Jr, Keskinocak P, Kharoufeh JP, Smith JC, eds. *Wiley Encyclopedia of Operations Research and Management Science* (John Wiley & Sons, New York).

Arias E (2007) United States life tables, 2004. *Natl. Vital Statist. Reports* 56(9):1–40.

Ayer T, Alagoz O, Stout NK (2012) OR forum: A POMDP approach to personalize mammography screening decisions. *Oper. Res.* 60(5):1019–1034.

Biondo S, Parés D, Frago R, Martí-Ragué J, Kreisler E, De Oca J, Jaurrieta E (2004) Large bowel obstruction: Predictive factors for postoperative mortality. *Diseases Colon Rectum* 47(11):1889–1897.

Bond JH (2000) Polyp guideline: Diagnosis, treatment, and surveillance for patients with colorectal polyps. *Amer. J. Gastroenterology* 95(11):3053–3063.

Bretthauer M, Kalager M (2012) Colonoscopy as a triage screening test. *New England J. Medicine* 366(8):759–760.

Cassandra AR (1994) Optimal policies for partially observable Markov decision processes. Technical report, Brown University, Providence, RI.

Erenay FS (2010) Optimal screening policies for colorectal cancer prevention and surveillance. Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada.

Erenay FS, Alagoz O, Banerjee R, Cima RR (2011) Estimating the unknown parameters of the natural history of metachronous colorectal cancer using discrete-event simulation. *Medical Decision Making* 31(4):611–624.

Ferlitsch M, Reinhart K, Pramhas S, Wiener C, Gal O, Bannert C, Hassler M, et al. (2011) Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. *J. Amer. Medical Assoc.* 306(12):1352–1358.

Fletcher RH (2008) Personalized screening for colorectal cancer. *Medical Care* 46(9):S5–S9.

Frazier AL, Colditz GA, Fuchs CS, Kuntz KM (2000) Cost-effectiveness of screening for colorectal cancer in the general population. *J. Amer. Medical Assoc.* 284(15):1954–1961.

Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR, Neugut AI (2003) Risk of perforation after colonoscopy and sigmoidoscopy: A population-based study. *J. Natl. Cancer Inst.* 95(3):230–236.

Goodwin JS, Singh A, Reddy N, Riall TS, Kuo YF (2011) Overuse of screening colonoscopy in the Medicare population. *Arch. Internal Medicine* 171(15):1335–1343.

Haidinger G, Waldhoer T, Vutuc C (2008) Self-reported colonoscopy screening in Austria. *Eur. J. Cancer Prevention* 17(4):354–357.

Klabunde CN, Lanier D, Nadel MR, McLeod C, Yuan G, Vernon SW (2009) Colorectal cancer screening by primary care physicians: Recommendations and practices, 2006–2007. *Amer. J. Preventive Medicine* 37(1):8–16.

Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Habbema JDF (2009a) At what costs will screening with CT colonography be competitive? A cost-effectiveness approach. *Internat. J. Cancer* 124(5):1161–1168.

Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Winawer SJ, Habbema JDF (2009b) Individualizing colonoscopy screening by sex and race. *Gastrointestinal Endoscopy* 70(1):96–108.

Leshno M, Halpern Z, Arber N (2003) Cost-effectiveness of colorectal cancer screening in the average risk population. *Health Care Management Sci.* 6(3):165–174.

Levin TR, Zhao W, Conell C, Seeff LC, Manninen DL, Shapiro JA, Schulman J (2006) Complications of colonoscopy in an integrated health care delivery system. *Ann. Internal Medicine* 145(12):880–886.

Loeve F, Boer R, Zauber AG, van Ballegooijen M, van Oortmarssen GJ, Winawer SJ, Habbema JDF (2004) National polyp study data: Evidence for regression of adenomas. *Internat. J. Cancer* 111(4):633–639.

Longo WE, Virgo KS, Johnson FE, Oprian CA, Vernava AM, Wade TP, Phelan MA, Henderson WG, Daley J, Khuri SF (2000) Risk factors for morbidity and mortality after colectomy for colon cancer. *Diseases of the Colon and Rectum* 43(1):83–91.

Maheshwari S, Patel T, Patel P (2008) Screening for colorectal cancer in elderly persons: Who should we screen and when can we stop? *J. Aging Health* 20(1):126–139.

Maillart LM, Ivy JS, Ransom S, Diehl K (2008) Assessing dynamic breast cancer screening policies. *Oper. Res.* 56(6):1411–1427.

Mysliwiec PA, Brown ML, Klabunde CN, Ransohoff DF (2004) Are physicians doing too much colonoscopy? A national survey of colorectal surveillance after polypectomy. *Ann. Internal Medicine* 141(4):264–271.

Ness RM, Holmes AM, Klein R, Dittus R (2000) Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. *Amer. J. Gastroenterology* 95(7):1800–1811.

Ohlsson B, Pålsson B (2003) Follow-up after colorectal cancer surgery. *Acta Oncologica* 42(8):816–826.

Parmigiani G, Skates S, Zelen M (2002) Modeling and optimization in early detection programs with a single exam. *Biometrics* 58(1):30–36.

Pfister DG, Benson AB III, Somerfield MR (2004) Surveillance strategies after curative treatment of colorectal cancer. *New England J. Medicine* 350(23):2375–2382.

Pierskalla WP, Brailer DJ (1994) Applications of operations research in health care delivery. Pollock SM, Rothkopf MH, Barnett A, eds. *Handbooks in Operations Research and Management Science*, Operations Research in the Public Sector, Vol. 6 (Elsevier Science, Amsterdam), 469–505.

Quintero E, Castells A, Bujanda L, Cubiella J, Salas D, Lanas Á, Andreu M, Carballo F, Morillas JD, Hernández C, et al. (2012) Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. *New England J. Medicine* 366(8):697–706.

Regueiro CR (2005) AGA future trends committee report: Colorectal cancer: A qualitative review of emerging screening and diagnostic technologies. *Gastroenterology* 129(3):1083–1103.

Roberts S, Wang L, Klein R, Ness R, Dittus R (2007) Development of a simulation model of colorectal cancer. *ACM Trans. Modeling Comput. Simulation* 18(1):1–30.

Schoen RE, Cooper GS, Scott LD (2006) Debate: Should screening colonoscopy be performed on an 88-yr-old healthy patient? *Amer. J. Gastroenterology* 101(8):1713–1715.

SEER (2012) Surveillance, Epidemiology and End Results Program. Accessed April 4, 2012, http://seer.cancer.gov.

Severens JL, Milne RJ (2004) Discounting health outcomes in economic evaluation: The ongoing debate. *Value in Health* 7(4):397–401.

Siegel R, Ward E, Brawley O, Jemal A (2011) Cancer statistics, 2011. *CA-Cancer J. Clin.* 61(4):212–236.

Smallwood RD, Sondik EJ (1973) The optimal control of partially observable Markov processes over a finite horizon. *Oper. Res.* 21(5):1071–1088.

Syngal S, Weeks JC, Schrag D, Garber JE, Kuntz KM (1998) Benefits of colonoscopic surveillance and prophylactic colectomy in patients with hereditary nonpolyposis colorectal cancer mutations. *Ann. Internal Medicine* 129(10):787–796.

Tangka FKL, Subramanian S, Beebe MC, Hoover S, Royalty J, Seeff LC (2013) Clinical costs of colorectal cancer screening in 5 federally funded demonstration programs. *Cancer* 119(S15):2863–2869.

USPSTF (2008) Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann. Internal Medicine* 149(9):627–638.

Vijan S, Inadomi J, Hayward RA, Hofer TP, Fendrick AM (2004) Projections of demand and capacity for colonoscopy related to increasing rates of colorectal cancer screening in the United States. *Alimentary Pharmacology Therapeutics* 20(5):507–515.

Vijan S, Hwang I, Inadomi J, Wong RKH, Choi JR, Napierkowski J, Koff JM, Pickhardt PJ (2007) The cost-effectiveness of CT colonography in screening for colorectal neoplasia. *Amer. J. Gastroenterology* 102(2):380–390.

Wilschut JA, Steyerberg EW, Van Leerdam ME, Lansdorp-Vogelaar I, Habbema JDF, Van Ballegooijen M (2011) How much colonoscopy screening should be recommended to individuals with various degrees of family history of colorectal cancer? *Cancer* 117(18): 4166–4174.

Winawer S, Fletcher R, Rex D, Bond J, Burt R, Ferrucci J, Ganiats T, Levin T, Woolf S, Johnson D, et al. (2003) Colorectal cancer screening and surveillance: Clinical guidelines and rationale-update based on new evidence. *Gastroenterology* 124(2):544–560.

Winawer SJ (2007) Colorectal cancer screening. *Best Practice Research Clinical Gastroenterology* 21(6):1031–1048.

Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, Mulrow CD, Woolf SH, et al. (1997) Colorectal cancer screening: Clinical guidelines and rationale. *Gastroenterology* 112(2):594–642.

Yun HR, Lee LJ, Park JH, Cho YK, Cho YB, Lee WY, Kim HC, Chun HK, Yun SH (2008) Local recurrence after curative resection in patients with colon and rectal cancers. *Internat. J. Colorectal Disease* 23(11):1081–1087.

Zapka J, Klabunde CN, Taplin S, Yuan G, Ransohoff D, Kobrin S (2012) Screening colonoscopy in the US: Attitudes and practices of primary care physicians. *J. General Internal Medicine* 27(9): 1150–1158.

Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM (2008) Evaluating test strategies for colorectal cancer screening: A decision analysis for the US Preventive Services Task Force. *Ann. Internal Medicine* 149(9): 659–669.

Zhang J, Denton BT, Balasubramanian H, Shah ND, Inman BA (2012a) Optimization of prostate biopsy referral decisions. *Manufacturing Service Oper. Management* 14(4):529–547.

Zhang J, Denton BT, Balasubramanian H, Shah ND, Inman BA (2012b) Optimization of PSA screening policies: A comparison of the patient and societal perspectives. *Medical Decision Making* 32(2):337–349.