Management Science

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

Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies

Turgay Ayer, Oguzhan Alagoz, Natasha K. Stout, Elizabeth S. Burnside

To cite this article:
Turgay Ayer, Oguzhan Alagoz, Natasha K. Stout, Elizabeth S. Burnside (2016) Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies. Management Science 62(5):1339-1362. http://dx.doi.org/10.1287/mnsc.2015.2180

MANAGEMENT SCIENCE
Vol. 62, No. 5, May 2016, pp. 1339–1362
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2015.2180

Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies

Turgay Ayer
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, ayer@isye.gatech.edu

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

Natasha K. Stout
Department of Population Medicine, Harvard Medical School/Harvard Pilgrim Health Care Institute, Boston, Massachusetts 02215, natasha_stout@hms.harvard.edu

Elizabeth S. Burnside
Department of Radiology, University of Wisconsin–Madison, Madison, Wisconsin 53706, eburnside@uwhealth.org

Most major health institutions recommend women to undergo repeat mammography screening for early diagnosis of breast cancer, the leading cause of cancer deaths among women worldwide. Although the proportion of women who ever had a mammogram is increasing, a majority of women do not get repeat mammograms. This paper analyzes the role of heterogeneity in women's adherence on optimal mammography screening recommendations. We develop a dynamic modeling framework that considers imperfect and heterogeneous adherence to screening recommendations. We carefully calibrate our model and solve it using real data. Unlike the existing breast cancer screening guidelines, our results suggest that adherence and heterogeneity in women's adherence behaviors should be explicitly considered. In particular, we find that when screening strategies are optimized assuming average adherence for everyone, the effect on patients with already low adherence would be relatively small, but patients with high adherence would be adversely affected. Considering imperfect and heterogeneous adherence in the population, our model suggests (1) given the current low adherence rates, an aggressive screening policy such as annual screening between the ages of 40 and 79 should be promoted to the general population; (2) screening strategies may be adjusted in clinical practice based on women's adherence, and screening intervals can be extended to two years for women with a history of high adherence; and (3) if the screening patterns change in the long run and most regular screeners adopt the most recent U.S. Preventive Services Task Force guidelines, then improving overall mammography adherence in society becomes more critical.

Keywords: adherence behavior; health policy; incomplete information; dynamic programming; breast cancer; mammography screening
History: Received July 17, 2012; accepted January 16, 2015, by Rakesh Sarin, decision analysis. Published online in *Articles in Advance* October 22, 2015.

1. Introduction

Mammography is the most effective screening modality to detect breast cancer, the most common female cancer, at earlier and more treatable stages (Kerlikowske et al. 1995). Repeat mammography can identify breast abnormalities an average of one to three years before a woman can feel a lump during a self-examination (CDC 1995) and may reduce breast cancer mortality by 20%–30% (Kerlikowske et al. 1995).

However, there is no consensus on whom to screen and how often, and the topic often sparks great controversy and receives extensive media coverage (see, e.g., Parker-Pope 2009, Szabo 2009, Dellorto 2009). There are at least four different mammography screening guidelines in the United States alone. For example, whereas the American Cancer Society (ACS) recommends annual mammography screening for women over age 40, recent guidelines by the U.S. Preventive Services Task Force (USPSTF), a government-appointed independent expert panel, recommend biennial screening for women between the ages of 50 and 74. Guidelines from other developed countries further differ from those in the United States (Table 1).

Existing mammography screening guidelines are not only controversial, but they also do not encompass several important dynamics of mammography screening, such as differences in individual breast cancer risk factors and adherence to screening recommendations. In this study, we account for both of these factors and

Table 1 **Recommended Mammography Screening Policies by Various Institutions in the United States (Qaseem et al. 2007, USPSTF 2009) and Other Countries with Organized Population-Based Cancer Screening Programs (Klabunde and Ballard-Barbash 2007, Shapiro et al. 1998)**

Institution/Country	Age groups to screen	Screening intervals in years	
		Age 40–49	Age 50+
American Cancer Society	40+	1	1
National Cancer Institute	40+	1–2	1–2
U.S. Preventive Services Task Force	50–74	—	2
American College of Preventive Medicine	50+	—	1–2
Canada, Italy, Japan	50–69	—	2
France, Netherlands	50–74	—	2
Sweden	40–74	1.5	2
United Kingdom	50–70	—	3

focus on the latter, which is defined by the World Health Organization (WHO) as "the extent to which the patient follows medical instructions" and identified as "the primary reason for suboptimal clinical benefit" (WHO 2003, p. 3). In developed countries, adherence to medical therapies in general is estimated to be around 50% and it is much lower in developing countries (O'Donohue and Levensky 2006). In breast cancer screening context, the scenario is similar; less than half of eligible U.S. women are estimated to undergo annual mammograms at least two years in a row, and less than 10% undergo annual mammograms over a period of nine to 10 years (Partin et al. 2005).

Recent literature in psychology and behavioral sciences shows that there exists significant heterogeneity among women in terms of adherence to cancer screening recommendations (O'Neill et al. 2008, Flynn et al. 2011). On the other hand, existing mammography screening guidelines simply assume that all women at the same age group would follow the same screening recommendations regardless of the differences in their adherence behaviors (Qaseem et al. 2007, USPSTF 2009). Behavioral sciences literature reveals that this is a false assumption that might lead to severe errors, including waste of healthcare resources, reduction in life expectancy, and an increased mortality of breast cancer (O'Donohue and Levensky 2006). A recent study estimates that each year 4,775 breast cancer deaths could be prevented if all eligible American women adhered to the screening recommendations (Sabatino et al. 2008).

Although adherence has not received much attention in operations research/management science (OR/MS) literature, there is extensive literature on the subject in medical and behavioral sciences. This literature can be broadly categorized as (a) adherence enhancement studies, (b) studies that focus on adherence-influencing factors, and (c) studies that focus on measurement of adherence. Although none of these areas are the focus of this paper, we find it noteworthy to summarize this literature.

A majority of the literature focuses on adherence enhancement methods. Health promotion checklists, mailed reminders, telephone calls, physician counseling, and administrative changes have been shown to enhance adherence in particular settings (O'Donohue and Levensky 2006). Nevertheless, despite the significant improvements in adherence-enhancing interventions in mammography screening, only a minority of patients meet adherence goals (Ornstein et al. 1993).

A significant body of the literature considers identifying the reasons for poor adherence (Christensen 2004). Several behavioral models are utilized (see, e.g., the Health Belief Model, Theory of Reasoned Action, and Psychological Reactance Theory in Christensen 2004) and numerous empirical studies are conducted to understand the factors that influence patient adherence. A wide range of factors have been associated with adherence behavior, including physician recommendation, patient's socioeconomic status (such as education level), psychological characteristics (such as expectations and perceptions), healthcare team and system-related factors (such as service quality), and social environment factors (such as social network size) (O'Donohue and Levensky 2006). Among these factors, physician recommendation is consistently reported as the single best predictor of mammography adherence (Mayne and Earp 2003, Yanovitzky and Blitz 2000, Friedman et al. 1995, Lerman et al. 1990).

Lastly, some literature is devoted to measurement of adherence, which is critical for tailoring clinical recommendations based on patients' readiness to adhere (WHO 2003). Recent research in behavioral sciences has revealed that there is no single standard strategy that fits all patients; therefore, tailoring interventions based on patients' measured (estimated) adherence behaviors is strongly encouraged (WHO 2003, O'Donohue and Levensky 2006). In preventive intervention programs such as cancer screening, the extent of patients' adherence can be directly observed based on their prior appointment attendance (Shumaker et al. 2009, Sohl and Moyer 2007).

An ideal cancer screening strategy would segment patients and tailor cancer screening recommendations based on individuals' adherence behaviors and their personal breast cancer risks (WHO 2003, IOM 2005), which is the focus of this paper. We propose an analytical framework for simultaneously incorporating dynamic adherence behaviors of patients and their breast cancer risks into mammography screening decisions. Throughout the decision process, we update a patient's estimated adherence level and her individual breast cancer risk based on the prior history of adherence behavior and observed test outcomes

(such as a negative or a false-positive mammogram), respectively. We use real data in our numerical analysis and shed light on the controversial mammography screening guidelines by providing concrete screening recommendations at the clinical and policy level. To our knowledge, this is the first analytical study that considers adherence and heterogeneity in adherence to a cancer screening strategy, and aims to optimize a cancer screening schedule accordingly.

The rest of this paper is organized as follows. Section 2 provides the background material on breast cancer epidemiology and mammography screening and reviews the relevant literature on breast cancer screening and adherence. In §3, we develop a mathematical model for this problem. In §4, we represent the optimality equations of the mathematical model in an alternative form and based on this new representation, we propose a solution algorithm. In §5, we describe our data sources and parameter estimations. In §6, we present the results of our numerical experiments. Finally, we discuss our findings and conclude in §7.

2. Background

2.1. Breast Cancer Epidemiology and Trade-offs in Mammography Screening

Breast cancer is a malignant tumor that develops from cells of the breast. There are two main types of breast cancer: *in situ* and *invasive*. If the cancer cells remain confined to lobules (milk-producing glands) or ducts (tiny tubes that carry the milk), they are called in situ breast cancers; whereas if the cancer cells spread beyond the duct lining, they are called invasive breast cancers. Prognosis and survival rate vary significantly depending on breast cancer type; therefore, differentiation of these two types is critical in the management of the disease.

A key trade-off in breast cancer screening is between the benefits of early cancer detection and harms of mammography screening, including pain, psychological distress, and especially risks associated with false-positive mammograms. False-positive mammograms cause anxiety, depression, morbidity, time loss, and often lead to unnecessary invasive diagnostic follow-up exams (such as biopsy) (Brewer et al. 2007). The risk of experiencing a false-positive mammogram is not small (about 11% per mammogram) and increases as screening becomes more frequent. For every 1,000 women who undergo 10 mammograms, more than half are estimated to have at least one false positive and nearly 200 of them are estimated to undergo at least one unnecessary biopsy (Elmore et al. 1998).

Aggressiveness of breast cancer is significantly higher in younger women (Jayasinghe et al. 2005), which motivates some guidelines to recommend more frequent screening for younger women. On the other hand, incidence and mortality of breast cancer are significantly higher among older women compared to younger women (Jemal et al. 2007a). Furthermore, younger women are likely to experience more false-positive mammograms than older women, as the accuracy of mammography increases with age (Kerlikowske et al. 2000). Motivated by these facts, other guidelines recommend more frequent screening for older women (Qaseem et al. 2007). Because of these trade-offs, breast cancer screening guidelines aim to balance the benefits and harms of mammography screening. Considering screening adherence and heterogeneity in women's adherence behaviors further complicates the problem.

2.2. Relevant Literature on Breast Cancer Screening and Adherence

Several modeling studies have investigated breast cancer screening in both medical and OR/MS literature. These studies can be broadly categorized as simulation or analytical models.

Most of the studies that appear in the medical literature are cost-effectiveness analyses based on simulation modeling (see Knudsen et al. 2007 for a review). Most simulation models assume 100% participation to the screening, and a few evaluate the cost-effectiveness at various fixed participation levels (Mandelblatt et al. 2003, Stout et al. 2006). Among these simulation studies, research by the Cancer Intervention and Surveillance Modeling Network (CISNET), a National Cancer Institute (NCI)-sponsored consortium that focuses on modeling cancer screening and control, is particularly noteworthy because the CISNET models use NCI-provided common data sets and provide evidence for the actual screening guidelines (see http://www.cisnet.cancer.gov).

There are also several analytical studies that consider breast cancer screening problems. However, none of these studies consider the role of adherence, but instead they assume full adherence to screening recommendations. Furthermore, most analytical studies make several limiting assumptions such as existence of an error-free test; age independency of disease progression, mortality rates, and test outcomes; that an in situ tumor is not harmful; or a self-detection never happens. The relevant analytical models published in the OR/MS literature since the early 1990s are summarized below.

Ozekici and Pliska (1991) model an infinite-horizon delayed Markov decision process to minimize the total expected cost of screening, false positives, treatments, and deaths. They assume stationary disease aggression and do not consider age dependency of screening test accuracies. Parmigiani (1993) develops a continuous-time non-Markovian stochastic model for disease progression and calculates the optimal policy that trades off between the costs of inspection and undetected failure, assuming that there exists an

error-free test. Baker (1998) builds a stochastic model similar to that of Parmigiani (1993) to evaluate costs of screening and life years lost because of breast cancer. Baker (1998) assumes in situ tumors are not harmful, and thus can be ignored.

Zelen (1993) presents a continuous time stochastic model to study different screening schedules, assuming that all parameters except incidence are independent of age. Ivy (2002) analyzes the competing objectives of multiple decision makers (patients and payers) in breast cancer screening and treatment. Maillart et al. (2008) use a partially observable Markov chain to evaluate alternative screening scenarios with respect to lifetime mortality risk using sample path enumeration. Lastly, unlike other analytical models that focus on routine mammography schedules, Ayer et al. (2012) consider individualizing mammography screening decisions based on a woman's breast cancer risk. However, similar to other analytical models, Ayer et al. (2012) also do not consider heterogeneity in women's adherence behaviors and assume 100% adherence to screening recommendations.

Although we are not aware of any analytical studies that focus on the design of a breast cancer prevention strategy while accounting for imperfect adherence, some studies consider the effects of patient adherence in other domains. For example, a few studies analyze the effect of imperfect adherence to human immunodeficiency virus (HIV) therapy and aims to find a minimal adherence level to avoid resistance to treatment (see, e.g., Miron et al. 2010, Smith 2006). A related study on cervical cancer screening considers how much adherence can drop in vaccinated women for human papillomavirus (HPV) before the incidence of cervical cancer increases too much (Bauch et al. 2010). Krakovska and Wahl (2007) show that drug regimens should be adjusted for poor adherence for optimal treatment benefit for HIV. Lastly, Mason et al. (2011) consider a similar problem in the context of diabetes, where they aim to optimize statin treatment decisions for diabetes patients when future adherence is uncertain.

3. Model Formulation

We formulate this problem as a discrete-time, finite-horizon partially observable Markov decision process (POMDP) that simultaneously considers uncertainties regarding a patient's adherence behavior and breast cancer risk. Similar to the actual breast cancer screening guidelines by the USPSTF (USPSTF 2009), the objective is to maximize patients' total expected quality-adjusted life years (QALYs), which account for both the quality and the quantity of life years (Drummond et al. 2005). The decision maker is the physician (he) who makes recommendations to the patient (she). The physician is risk neutral and acts in the best interest of the patient. The physician's broad aim is to develop a breast cancer screening schedule that is most beneficial to the patient, considering her prior appointment attendance behavior (i.e., adherence), existing breast cancer risk factors, and history of prior screening outcomes. Until the optimal age to initiate mammography screening, the physician recommends the patient to *wait* (W). Once screening starts, at any decision epoch t, depending on the patient's breast cancer risk and her adherence behavior, the physician could recommend an *annual* ($M(1)$), *biennial* ($M(2)$), or *triennial* ($M(3)$) mammogram. That is, he might recommend an immediate mammogram and another at time $t+1$, $t+2$, or $t+3$. However, because of imperfect adherence, the patient may not follow the physician's recommendations (i.e., not adhere) and might return at a later time or might not return at all (i.e., no show). The patient may also show up before a scheduled mammogram in the case of a self-detected lump. That is, the time interval between successive decision epochs is variable and "violations" (i.e., not following the physician's recommendations) might occur either at the decision epochs in which the next screening is scheduled or when a self-detection occurs. For example, as illustrated in Figure 1, when a patient is recommended to undergo biennial mammography screening and does not have any self-detection, she undergoes mammography now (at time t), waits next year, but may or may not show up (adhere or not adhere) for the next mammography exam two years later (at time $t+2$). On the other hand, she might also show up earlier than two years if she experiences a self-detected lump.

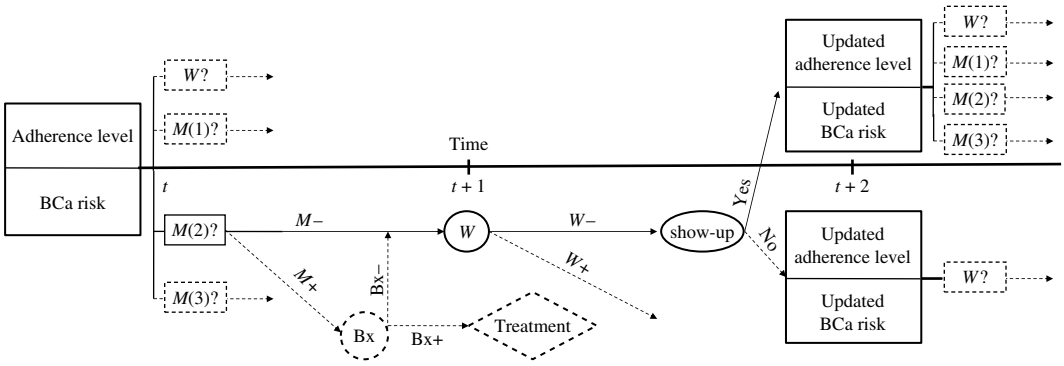
Before presenting the optimality equations, we first introduce the notation used in the model.

- a: Action available to the physician (i.e., screening recommendation), $a \in A = \{W, M(1), M(2), M(3)\}$, where W denotes a *wait*, i.e., defer the mammography decision for one year, $M(1)$ denotes *annual mammography recommendation*, $M(2)$ denotes *biennial mammography recommendation*, and $M(3)$ denotes *triennial mammography recommendation*. Without loss of generality, we define $\bar{M} = \{M(1), M(2), M(3)\}$ and assume that the time interval between two successive mammography recommendations is at most three years, because this is the maximum interval among the actual mammography screening guidelines (Qaseem et al. 2007).

Following a positive mammography result, we assume that the patient undergoes a perfect diagnostic test such as biopsy, which is a reasonable assumption since the accuracy of biopsy is reported to be close to 100% (Verkooijen 2002). If the biopsy reveals that the patient has cancer, then the patient starts treatment and the screening decision process ends (Figure 1).

- t: Decision epoch, $t = 0, 1, 2, \ldots, T$; $T < \infty$, where t denotes the number of years above the age 40 (e.g.,

Figure 1 An Illustrative Figure Showing a Possible Scenario (Path) When the Selected Action Is $M(2)$, and $o = W-$



Notes. Solid line represents the realized scenario and dashed lines represent other possible scenarios. BCa, breast cancer; Bx, biopsy; Bx+, positive biopsy; Bx−, negative biopsy.

$t = 0$ represents age 40, $t = 1$ represents age 41, etc.). Depending on the action selected at time t, test outcomes, and patient's adherence behavior, the next decision can be made at time $t+1$, $t+2$, $t+3$, or at a later time (if the patient is absent for over three years). We assume that screening decisions start at age 40 because this is the earliest age to start mammography screening among the actual screening guidelines. Similarly, we assume that the decisions end at age 100 because none of the actual screening guidelines recommend mammography screening over this age (Klabunde and Ballard-Barbash 2007).

- s: Core health state, $s \in \{0, 1, 2, 3\}$, where s denotes the true health state of the patient at time t. In particular, 0 denotes a cancer-free patient, 1 denotes a patient with in situ cancer, 2 denotes a patient with invasive cancer, and 3 denotes death. We remark that cancer stages are not fully observable, but only partially observable via imperfect test outcomes such as mammography or a self-detection.
- Θ_a: Observation space, which includes observations seen on taking action a. Upon undergoing a mammography exam, either a positive mammography result ($M+$) or a negative mammography result ($M-$) could be observed. On the other hand, even if the patient does not receive any mammograms at certain decision epochs, observations can still be seen in the form of a self-detection (i.e., finding of a breast lump denoted by $W+$) or no self-detection ($W-$).
- $\mathcal{O}_t^a(o \mid s)$: Observation probability, which represents the probability of seeing observation o at time t, given that action a is selected and core health state is s at time t. We define the observation probability matrix as $\mathbb{O}_t^a = [\mathcal{O}_t^a(o \mid s)]$. In breast cancer screening, observations are imperfect and provide only partial information about the core health state of a patient. When a patient is cancer-free, with a certain probability (specificity) the test result will be negative; and when the patient has cancer, with a certain probability (sensitivity) the test result will be positive. If we let $spec_t(M)$ and $spec_t(W)$ denote the specificity of mammography and self-detection at time t, respectively, and $sens_t(s, M)$ and $sens_t(s, W)$ denote the sensitivity of mammography and self-detection for core health state s at time t, respectively, then

$$\mathbb{O}_t^a = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cc} Test- & Test+ \end{array} \\ \begin{pmatrix} spec_t(a) & 1 - spec_t(a) \\ 1 - sens_t(1, a) & sens_t(1, a) \\ 1 - sens_t(2, a) & sens_t(2, a) \\ 1 & 0 \end{pmatrix} \end{array},$$

where $Test-$ represents $W-$ or $M-$ and $Test+$ represents $W+$ or $M+$.

- $\mathcal{P}_t^a(s' \mid s)$: Core health state transition probability, i.e., the probability that the patient will be in state $s' \in S$ at time $t+1$, given that she was in state s and action a was selected at time t.
- $b_t \in B_t(S)$: Health belief state, which is an element of the space of all probability distributions over the core health states, $B_t(S)$. The sth component of a health belief state b_t, denoted by $b_t(s)$, represents the probability of being in core health state s. Intuitively, b_t represents the physician's belief about the patient's health status. For example, if $b_t = [97\%\ 1\%\ 2\%\ 0\%]$, then the physician's belief about the patient's in situ and invasive cancer risks are 1% and 2%, respectively, and the remaining 97% and 0% represent cancer-free and death probabilities, respectively.

The initial belief state is estimated based on a patient's existing breast cancer risk factors (see §5.2). As the system evolves (i.e., the woman ages), the core health states and the observations seen (e.g., $M-$ or $M+$) may change, and these may affect the health belief state b_t. Therefore, at each decision epoch, the health belief state is updated based on the actions taken and observations seen. The s'th component of the updated health belief state, $b_{t+1}(s') = \Pr_{t+1}(s' \mid b_t, a, o)$, represents the probability of being in core health state s' at time

$t+1$, given that the health belief state was b_t, action a was selected, and observation o was seen at time t. Although the updated health belief state b_{t+1} depends on b_t, a, and o, we drop these dependencies to simplify the notation. The updated health belief state b_{t+1} is computed by

$$b_{t+1}(s') = \Pr_{t+1}(s' \mid b_t, a, o) = \begin{cases} \dfrac{\sum_{s\in S} b_t(s)\mathcal{O}_t^a(o \mid s)\mathcal{P}_t^a(s' \mid s)}{\sum_{s\in S} b_t(s)\mathcal{O}_t^a(o \mid s)} \\ \quad \text{if } a = W,\ o \in \Theta_W \text{ or } a \in \bar{M},\ o = M-, \\ \mathcal{P}_t^M(s' \mid 0) \quad \text{if } a \in \bar{M},\ o = M+. \end{cases} \tag{1}$$

Note that if the patient has cancer and is detected by a mammogram (i.e., a true positive mammogram), there is no update to the health belief state as she starts treatment and the decision process ends (Figure 1).

- $\mathcal{O}_t^a(o \mid b_t)$: Probability of seeing an observation o at time t when action a is selected and health belief state is b_t at time t, which is computed by $\mathcal{O}_t^a(o \mid b_t) = \sum_{s\in S} b_t(s)\mathcal{O}_t^a(o \mid s)$.

Similar to health states, the physician does not have full information about a patient's adherence behavior. Instead, he has a belief about the patient's adherence level, which is updated through the patient's observed adherence behavior measured by prior appointment attendance. The parameters related to the patient's adherence behavior are as follows.

- h: Core adherence state, i.e., the type of patient regarding her adherence behavior. In particular, $h \in H = \{R, I\}$, where R denotes a regular screener and I denotes an irregular screener. We adopt the CISNET models' definitions for regular and irregular screeners. Regular screeners are very likely to adhere to the screening recommendations and irregular screeners are very likely to skip them. On the other hand, with small probabilities, regular screeners may fail to adhere and irregular screeners may still adhere to future appointments. More detailed descriptions of regular and irregular screeners can be found in §5.
- z: Patient's observed appointment attendance behavior, $z \in \{\varsigma, \varsigma'\}$, where ς denotes that the patient "shows up" at her scheduled mammography appointment and ς' denotes a "no show."
- $\mathcal{Z}_{t+k}^a(z \mid h)$: Adherence rate, i.e., probability of observing appointment attendance behavior $z \in \{\varsigma, \varsigma'\}$ at time $t+k$, given that action a is selected and the core adherence state is h at time t. We remark that k denotes the duration between the recommended subsequent mammograms and the function $\mathcal{Z}_{t+k}^a(z \mid h)$ captures the relationship between adherence and duration between subsequent mammography exams.
- $\mathcal{L}_{t+k}(h' \mid h)$: Adherence state transition probability, i.e., the probability that the patient will be of type $h' \in H$ at time $t+k$, given that she was of type h at time t. We remark that there is strong clinical evidence that adherence behavior may change over time (see, e.g., Gierisch et al. 2010, Watson-Johnson et al. 2011). For example, a regular screener at time t may no longer continue to be a regular screener at time $t+k$, which is captured by the adherence state transition probability.
- $\pi_t \in \Pi_t(H)$: Adherence belief state, which is an element of the space of all probability distributions over the core adherence states, $\Pi_t(H)$. The hth component of adherence belief state π_t, denoted by $\pi_t(h)$, represents the probability of being in core adherence state h at time t. Intuitively, π_t denotes the physician's belief about the patient's adherence behavior (i.e., regular or irregular screener). For example, if $\pi_t = [90\%\ 10\%]$, then the physician believes this patient is a regular screener with 90% probability.

Note that adherence belief state is a sufficient statistic to capture the entire history of past mammography attendance behavior, and similar to a health belief state, an adherence belief state is updated based on the action and the observed patient behavior. The h'th component of the updated adherence belief state, denoted by $\pi_{t+k}(h') = \Pr_{t+k}(h' \mid \pi_t, a, z)$, is the probability of being in core adherence state h' at time $t+k$, given that the adherence belief state was π_t, action a was selected at time t, and the patient's observed behavior was z at time $t+k$. The updated adherence belief state π_{t+k} depends on π_t, a, and z, but we drop these dependencies to simplify the notation. The updated adherence belief state is computed by

$$\pi_{t+k}(h') = \Pr_{t+k}(h' \mid \pi_t, a, z) = \begin{cases} \dfrac{\sum_{h\in H} \pi_t(h)\mathcal{L}_{t+k}(h' \mid h)\mathcal{Z}_{t+k}^a(z \mid h')}{\sum_{h,h'\in H} \pi_t(h)\mathcal{L}_{t+k}(h' \mid h)\mathcal{Z}_{t+k}^a(z \mid h')} \\ \qquad \text{if } k = f, \\ \pi_t(h') \quad \text{if } k < f, \end{cases} \tag{2}$$

where f represents the length of the interval time (in years) between subsequent actions, i.e., $f = 1$ if $a \in \{W, M(1)\}$, $f = 2$ if $a = M(2)$, and $f = 3$ if $a = M(3)$. That is, updates are made at the end of the recommended screening interval. For example, if $a = M(3)$ (i.e., $f = 3$), then the update to the adherence belief state will be made at the end of the third year ($k = 3$) based on whether the patient shows up or not.

- $\mathcal{Z}_{t+k}^a(z \mid \pi_t)$: Probability of observing patient behavior $z \in \{\varsigma, \varsigma'\}$ at time $t+k$, given that action a is selected and adherence belief state is π_t at time t, which is computed by $\mathcal{Z}_{t+k}^a(z \mid \pi_t) = \sum_{h,h'\in H} \pi_t(h)\mathcal{L}_{t+k}(h' \mid h) \cdot \mathcal{Z}_{t+k}^a(z \mid h')$. In our base case analysis, we assume that a patient's adherence behavior depends only on the patient's true adherence state and the physician's recommendation, but not the breast cancer risk. In §6.6, we relax this assumption to illustrate the potential impacts

of any new discoveries about the relationship between known breast cancer risk factors and adherence behavior on optimal screening policies.

Lastly, we define the rewards used in our model, which incorporate life expectancy and harms associated with a mammogram:

- $r_t(s, a, o)$: Expected QALYs between time t and $t+1$ when the patient's true health state is s, action a is selected, and observation o is seen.
- $r_t(s, a)$: Expected QALYs between time t and $t+1$ when the patient's true health state is s and action a is selected, which is computed by $r_t(s, a) = \sum_{o \in O} \mathcal{O}_t^a(o \mid s) \cdot r_t(s, a, o)$.
- $r_t(b_t, a)$: Expected QALYs between time t and $t+1$ when the health belief state is b_t and action a is selected, which is computed by $r_t(b_t, a) = \sum_{s \in S} b_t(s) r_t(s, a)$.
- $R_t(s)$: Total expected post-cancer QALYs accrued at time t when the patient is in one of the cancer states ($s \in \{1, 2\}$), detected by biopsy following a positive mammography result ($M+$), and cancer treatment has been initiated.
- $r_T(s)$: Total expected remaining QALYs at time T given that the patient is alive at time T and her true health state is s.

3.1. Optimality Equations

Let $V_t^*(b_t, \pi_t)$ be the maximum total expected QALYs that a patient can attain at time t when the current state is (b_t, π_t). Similarly, let $V_t^a(b_t, \pi_t)$ be the maximum total expected QALYs that a patient can attain at time t when the recommended action is a and the current state is (b_t, π_t). Then,

$$V_t^*(b_t, \pi_t) = \max\{V_t^{M^*}(b_t, \pi_t), V_t^W(b_t, \pi_t)\}, \quad t = 0, \ldots, T-3, \tag{3}$$

where,

$$M^* = \arg\max_{a \in \bar{M}} \{V_t^{M(1)}(b_t, \pi_t), V_t^{M(2)}(b_t, \pi_t), V_t^{M(3)}(b_t, \pi_t)\}. \tag{4}$$

If a patient with health belief state b_t and adherence belief state π_t is recommended *annual mammography screening* ($M(1)$) at time t, the immediate mammography result may either be negative ($M-$) or positive ($M+$). If the immediate mammogram is negative (with probability $\mathcal{O}_t^{M(1)}(M- \mid b_t)$) or false positive (with probability $b_t(0)\mathcal{O}_t^{M(1)}(M+ \mid 0)$), the patient is scheduled to undergo another mammography exam at time $t+1$, for which she either shows up (with probability $\mathcal{X}_{t+1}^{M(1)}(s \mid \pi_t)$) or does not show up (with probability $\mathcal{X}_{t+1}^{M(1)}(s' \mid \pi_t)$). If the patient shows up, she is recommended the best available mammography action (i.e., M^*) at time $t+1$. Otherwise, the action at time $t+1$ is a *wait* (W) by default. On the other hand, if the immediate mammogram at time t is true positive (i.e., the patient has cancer), the patient starts treatment, accrues a one-time $\sum_{s=1}^{2} b_t(s)\mathcal{O}_t^{M(1)}(M+ \mid s)R_t(s)$ QALYs, and leaves the process. That is,

$$\begin{aligned} V_t^{M(1)}(b_t, \pi_t) &= r_t(b_t, M(1)) \\ &+ \mathcal{O}_t^{M(1)}(M- \mid b_t)(\mathcal{X}_{t+1}^{M(1)}(s \mid \pi_t) V_{t+1}^{M^*}(b_{t+1}, \pi_{t+1}) \\ &+ \mathcal{X}_{t+1}^{M(1)}(s' \mid \pi_t) V_{t+1}^{W}(b_{t+1}, \pi_{t+1})) \\ &+ b_t(0)\mathcal{O}_t^{M(1)}(M+ \mid 0)(\mathcal{X}_{t+1}^{M(1)}(s \mid \pi_t) V_{t+1}^{M^*}(b_{t+1}, \pi_{t+1}) \\ &+ \mathcal{X}_{t+1}^{M(1)}(s' \mid \pi_t) V_{t+1}^{W}(b_{t+1}, \pi_{t+1})) \\ &+ \sum_{s=1}^{2} b_t(s)\mathcal{O}_t^{M(1)}(M+ \mid s)R_t(s), \\ & \qquad t = 0, \ldots, T-1. \end{aligned} \tag{5}$$

The *wait* (W) action models both the optimal age to initiate mammography screening and a no-show. In particular, until the optimal age to initiate screening, the *wait* action corresponds to deferring mammography screening decisions. Once the screening starts, the *wait* action represents a no-show.

For the belief (b_t, π_t), a patient taking the *wait* action at time t (either on a recommendation or because of a no-show) may either have no self-detection ($W-$) or have a self-detection ($W+$) between time t and $t+1$. If she has no self-detection ($W-$), she may either show up (with probability $\mathcal{X}_{t+1}^{W}(s \mid \pi_t)$) and be recommended the optimal action at time $t+1$, or does not show up (with probability $\mathcal{X}_{t+1}^{W}(s' \mid \pi_t)$) and follows action W at time $t+1$. On the other hand, if she has a self-detection between t and $t+1$, we assume she will show up with certainty and undergo mammography at time $t+1$, regardless of her cancer risk and adherence level. Note that in this case there will be no update to the adherence belief state (i.e., $\pi_{t+1} = \pi_t$) because showing up after a self-detection does not reflect the patient's adherence to a screening recommendation. That is,

$$\begin{aligned} V_t^{W}(b_t, \pi_t) &= r_t(b_t, W) \\ &+ \mathcal{O}_t^{W}(W- \mid b_t)(\mathcal{X}_{t+1}^{W}(s \mid \pi_t) V_{t+1}^{*}(b_{t+1}, \pi_{t+1}) \\ &+ \mathcal{X}_{t+1}^{W}(s' \mid \pi_t) V_{t+1}^{W}(b_{t+1}, \pi_{t+1})) \\ &+ \mathcal{O}_t^{W}(W+ \mid b_t) V_{t+1}^{M^*}(b_{t+1}, \pi_t), \\ & \qquad t = 0, \ldots, T-1. \end{aligned} \tag{6}$$

Similar to the definition of $V_t^{M(1)}(b_t, \pi_t)$, we define $V_t^{M(2)}(b_t, \pi_t)$ and $V_t^{M(3)}(b_t, \pi_t)$ as follows:

$$\begin{aligned} V_t^{M(i)}(b_t, \pi_t) &= r_t(b_t, M(i)) + \mathcal{O}_t^{M(i)}(M- \mid b_t) V_{t+1}^{W(i-1)}(b_{t+1}, \pi_{t+1}) \\ &+ b_t(0)\mathcal{O}_t^{M(i)}(M+ \mid 0) V_{t+1}^{W(i-1)}(b_{t+1}, \pi_{t+1}) \end{aligned}$$

$$+\sum_{s=1}^{2} b_t(s)\mathcal{O}_t^{M(i)}(M+\mid s)R_t(s),$$

$$t=0,\ldots,T-i, \text{ for } i=2,3, \quad \text{where,} \qquad (7)$$

$$V_t^{W(1)}(b_t,\pi_t) = r_t(b_t,W)+\mathcal{O}_t^{W}(W-\mid b_t)\left(\mathcal{X}_{t+1}^{W}(s\mid\pi_t)V_{t+1}^{M^*}(b_{t+1},\pi_{t+1}) + \mathcal{X}_{t+1}^{W}(s'\mid\pi_t)V_{t+1}^{W}(b_{t+1},\pi_{t+1})\right)+\mathcal{O}_t^{W}(W+\mid b_t) \cdot\left(V_{t+1}^{M^*}(b_{t+1},\pi_t)\right), \quad t=0,\ldots,T-1, \quad \text{and,} \qquad (8)$$

$$V_t^{W(2)}(b_t,\pi_t) = r_t(b_t,W)+O_t^{W}(W-\mid b_t)V_{t+1}^{W(1)}(b_{t+1},\pi_{t+1}) + O_t^{W}(W+\mid b_t)V_{t+1}^{M^*}(b_{t+1},\pi_t), \quad t=0,\ldots,T-2. \qquad (9)$$

Note that we introduce $V_t^{W(1)}(b_t,\pi_t)$ and $V_t^{W(2)}(b_t,\pi_t)$ to simplify the notation, which represent the maximum total expected QALYs that a patient can attain at time t when the current belief is (b_t,π_t), the patient does not have a mammogram at time t, and is recommended to *return* 1 *year later to undergo mammography* and *return* 2 *years later to undergo mammography*, respectively. Note also that the patient may have a self-detection at any time during the interval between screenings and may show up and undergo a mammogram before the next scheduled screening exam.

Lastly, we add boundary conditions for $t\in\{T-2, T-1, T\}$ as follows:

$$V_t^*(b_t,\pi_t) = \begin{cases} \max\{V_t^{M(2)}(b_t,\pi_t), V_t^{M(1)}(b_t,\pi_t), V_t^{W}(b_t,\pi_t)\} & \text{if } t=T-2,\\ \max\{V_t^{M(1)}(b_t,\pi_t), V_t^{W}(b_t,\pi_t)\} & \text{if } t=T-1,\\ V_t^{a}(b_t,\pi_t)=\sum_{s\in S} b_t(s)r_t(s) \quad \forall a\in A, \pi_t\in\Pi_t(H) & \text{if } t=T. \end{cases}$$

4. An Alternative Representation for the Optimality Equations and the Optimal Solution Algorithm

A POMDP is a continuous state space Markov decision process (MDP). Because there are infinitely many states, unlike in MDPs, classical backward induction algorithms do not work to solve a POMDP model. A pioneering finding in the POMDP research is due to Smallwood and Sondik (1973), who showed that the value function of a POMDP is piecewise linear and convex (pwlc). Based on this finding, several solution algorithms are developed for conventional POMDP models (Cassandra and Kaelbling 1998).

Our proposed POMDP model is different than the conventional POMDP models (see, e.g., Smallwood and Sondik 1973, Monahan 1982, Cassandra and Kaelbling 1998) in three ways. First, our POMDP model has two belief states with observations seen at different decision epochs. Furthermore, the outcomes for different observations may lead to different pathways (including an absorbing state following a stopping condition). As a result, combining and factoring these two belief states into a single one is not feasible. Note also that such a factored representation (a probability distribution over product space of the health and adherence state spaces) is clinically not meaningful. Second, the order of events and stopping rules are different. In conventional POMDP models, first an action is taken, then a state transition may occur, and afterward an observation is seen. Whereas this ordering is preserved for the adherence belief state, it is not the case for the health belief state. For a health belief state, an action is taken, then an observation is seen (e.g., a normal mammogram), and afterward transition to a new state occurs. Furthermore, in the case of a true-positive mammography observation, the decision ends before a transition occurs (i.e., a conditional optimal stopping problem). Third, the structures of the future value functions, V_{t+1}, are different. In conventional POMDP models, future value functions are defined by the optimal action at that decision epoch. On the other hand, because of the structure of our problem, future value functions are sometimes defined by a specific action, or a restricted set of actions (see Equations (5)–(9)).

In the remainder of this section, we show that the value function of our POMDP model preserves piecewise linear convexity and hence can be solved optimally. Then, we present an alternative equivalent representation for the optimality equations based on this finding. Finally, using this new representation, we present a modified version of Monahan's algorithm (Monahan 1982) with Eagle's reduction (Eagle 1984) to optimally solve this nonconventional POMDP. Proofs of the analytical results can be found in the online appendix (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2015.2180).

Proposition 1. *Value functions* $V_t^{W}(b_t,\pi_t)$,$V_t^{M(1)}(b_t,\pi_t)$,$V_t^{M(2)}(b_t,\pi_t)$, *and* $V_t^{M(3)}(b_t,\pi_t)$ *are all pwlc in the product space* $B_t(S)\times\Pi_t(H)$ *for all* $t\le T$, *and hence can be expressed as the maximum of a finite number of linear functions. That is,*

$$V_t^{a}(b_t,\pi_t) = \max_k\left\{\sum_{s\in S} b_t(s)\sum_{h\in H}\pi_t(h)\alpha_t^{k,a}(s,h)\right\}, \quad a\in A \qquad (10)$$

for some $\{\alpha_t^{0,a},\alpha_t^{1,a},\ldots\mid a\in A\}$, *where the components* $\alpha_t^{i,a}=[\alpha_t^{i,a}(s,h)]_{s\in S,\,h\in H}$ *are called the* α*-vectors generated by action* a.

THEOREM 1. *The optimal value function $V_t^*(b_t, \pi_t)$ is pwlc in the product space $B_t(S) \times \Pi_t(H)$ for all $t \le T$, and hence can be expressed as*

$$V_t^*(b_t, \pi_t) = \max_{a \in A} \left\{ \max_k \left\{ \sum_{s \in S} b_t(s) \sum_{h \in H} \pi_t(h) \alpha_t^{k,a}(s,h) \right\} \right\} \quad (11)$$

for some $\{\alpha_t^{0,a}, \alpha_t^{1,a}, \ldots \mid a \in A\}$.

Next, we describe the optimal solution algorithm for our POMDP model, which works as follows: for each decision epoch t, we compute all possible α-vectors generated by each action a. We refer to this finite set of α-vectors generated by action a as $\Lambda_t^a = \{\alpha_t^{k,a} \mid a \in A\}$. Then, for each action a, we eliminate $\alpha_t^{k,a} \in \Lambda_t^a$ if it is dominated at all possible belief and adherence states by any other $\alpha_t^{l,a} \in \Lambda_t^a$. We refer to this reduced set of α-vectors generated by action a as $\bar{\Lambda}_t^a = \{\alpha_t^{l,a} \mid a \in A\}$. Note that, the Monahan algorithm (Monahan 1982) keeps only the α-vectors that generate the optimal value function at each decision epoch. However, because of the structure of optimality equations given in (3)–(9), where the future value functions may be specified by a particular (and not necessarily optimal) action a, we need to keep track of all α-vectors generated by action a that are not dominated by any other α-vector generated by the same action a.

The elimination procedure works in two steps. In the first step, for each action a, we eliminate $\alpha_t^{k,a} \in \Lambda_t^a$ if its components are completely dominated by any other α-vector generated by the same action a. That is, if $\exists\, r$ such that $\alpha_t^{r,a}(s,h) \ge \alpha_{t,a}^{k}(s,h)\ \forall\, s \in S$ and $h \in H$, then we eliminate $\alpha_t^{k,a}$ from Λ_t^a. In the second step, we eliminate the remaining α-vectors in Λ_t^a that are dominated for some belief and adherence states but are not completely dominated. We eliminate such α-vectors at each decision epoch by solving a similar linear program (LP) to that of the Monahan (1982) LP as follows:

$$\begin{aligned} \max\ & \sigma \\ \text{s.t.}\ & \sum_{s \in S,\, h \in H} x_t(s,h)(\alpha_t^{k,a}(s,h) - \alpha_t^{r,a}(s,h)) - \sigma \ge 0 \\ & \qquad\qquad \forall\, \alpha_t^{r,a} \in \Lambda_t^a, \qquad (12) \\ & \sum_{s \in S,\, h \in H} x_t(s,h) = 1 \\ & x_t(s,h) \ge 0 \quad \forall\, s \in S,\, h \in H, \end{aligned}$$

where x_t and σ are the decision variables. We solve this LP for each $\alpha_t^{k,a} \in \Lambda_t^a$ and if the LP yields a solution $\sigma < 0$, then $\alpha_t^{k,a}$ is dominated and hence can be eliminated from the possible set of α-vectors generated by action a. Otherwise, we add $\alpha_t^{k,a}$ to $\bar{\Lambda}_t^a$, which is carried to decision epoch $t-1$.

5. Parameter Estimation

5.1. Estimating Input Parameters

Our data sources used in parameter estimation are listed in Table 2. Our primary source of data for parameter estimation comes from a validated computer simulation model developed as part of NCI's CISNET program. This model, which we refer to as the breast cancer simulation (BCS), is a detailed simulation model of breast cancer epidemiology that can replicate breast cancer natural history, detection, treatment, and mortality rates in the U.S. population. Our BCS model is calibrated to the breast cancer statistics reported in NCI's Surveillance, Epidemiology, and End Results (SEER) program and the National Center for Health Statistics, and cross-validated against the Wisconsin State Cancer Reporting System. Along with other CISNET breast cancer simulation models, our BCS model has been used to investigate population-based mammography screening strategies (Mandelblatt et al. 2009) and provided evidence for some of the actual breast cancer screening policies (USPSTF 2009, Mandelblatt et al. 2009). Detailed descriptions of model design and validation for the BCS model have been described elsewhere (http://www.cisnet.cancer.gov).

We use our BCS model in estimating age-specific core health state transition probabilities representing the natural history of breast cancer and life expectancy at each decision epoch. In our base case analysis, we use the following values for the disutilities associated with mammography: (a) 0.5 days for a negative mammogram (Mandelblatt et al. 1992), (b) two weeks for a true positive mammogram (Velanovich 1995), and (c) four weeks for a false-positive mammogram, because the literature reports that the disutility for a false-positive mammogram should be higher than that of a true positive mammogram (Earle et al. 2000). We do not consider any disutility associated with radiation exposure, as with the current technology, the radiation

Table 2 Input Data Sources for Parameter Estimation

Parameter	Data source
Core state transition probabilities	BCS
Intermediate rewards	BCS
Post-cancer life expectancy	SEER
Adherence rates	BCS and Cronin et al. (2005)
Adherence distribution	PRISM trial
Age distribution	U.S. Census Bureau (2010)
Risk distribution	BCSC Risk Calculator (2013)
Sensitivity and specificity of mammography	Kerlikowske et al. (2000)
Sensitivity and specificity of CBE	Barton et al. (1999)
Sensitivity and specificity of BSE	Baxter (2001)
CBE proportion in the population	Elmore et al. (2005)
BSE proportion in the population	Messina et al. (2004)

Notes. BCS, breast cancer simulation; SEER, surveillance, epidemiology, and end results; CBE, clinical breast exam; BSE, breast self-exam.

risk due to mammography is minimal (ACS 2012). We do sensitivity analyses on these parameters and report the corresponding results in §6.6. To estimate the postcancer life expectancies, we use age-specific mortality rates for patients under cancer treatment based on SEER data (Jemal et al. 2009), the details of which are available in Arias (2006).

We estimate the age distribution for the female U.S. population over age 40 using U.S. Census data (U.S. Census Bureau 2010), adherence distribution using the results of the Personally Relevant Information about Screening Mammography (PRISM) trial (Gierisch et al. 2010), and risk distribution based on the findings of a recent study that estimated distribution of breast cancer risk based on the Gail model (BCSC Risk Calculator 2013). We estimate patients' adherence and nonadherence rates as well as changes in adherence behavior using CISNET and NCI-reported data. To estimate these rates, we use the CISNET definitions of annual, biennial, and irregular screeners. This model classifies women as annual screeners if their mean time between two mammograms is less than 1.5 years; biennial screeners if it is between 1.5 years to 2.5 years; and irregular screeners if it is greater than 2.5 years. In our model, we consolidate annual and biennial screeners as regular screeners. To estimate the transition probability for adherence states, we count the number of transitions among different adherence states, and then transform these counts into transition probabilities by normalizing each count based on the row sum. Detailed descriptions of the CISNET model for mammography dissemination and usage patterns in the United States can be found elsewhere (Cronin et al. 2005, 2006, CISNET 2010). We estimate adherence rates for annual, biennial, and irregular screeners based on the observed mammography dissemination and usage patterns in the United States, which were informed by data from the National Health Interview Survey[1] and the Breast Cancer Surveillance Consortium (BCSC).[2]

We estimate the sensitivity and specificity of mammography and self-detection from the published studies in the literature. In particular, we estimate age-specific sensitivity and specificity values for mammography using Kerlikowske et al. (2000). When estimating the sensitivity and specificity of a self-detection, we consider both a breast self-exam (BSE) and a clinical breast exam (CBE). In particular, we first estimate the sensitivity and specificity of BSE and CBE using Baxter (2001) and Barton et al. (1999), respectively. We then estimate the CBE and BSE proportions in the population from Elmore et al. (2005) and Messina et al. (2004) and calculate the weighted average of CBE and BSE sensitivity and specificity values using these proportions to estimate the sensitivity and specificity of self-detection (see Table 3).

Table 3 Sensitivity and Specificity Values of Tests Used in the Model

(a) Mammograpy

Age group	Sensitivity (%)	Specificity (%)
40–49	72.2	88.9
50–54	72.2	89.3
55–59	81	89.3
60–69	81	89.7
70+	86.2	89.7

(b) BSE, CBE, and self-detection

Test	Sensitivity (%)	Specificity (%)
BSE	26	90
CBE	54	94
Self-detection	44	92

Note. BSE, breast self-exam; CBE, clinical breast exam.

5.2. Estimating Adherence and Breast Cancer Risk

Estimated adherence (i.e., probability that a patient is a regular screener) and breast cancer risk are not used as direct inputs to the model, but are necessary for calculating the initial adherence and health belief states. These initial belief states are then used to operationalize the policies presented in the model results. For this purpose, we need to estimate these parameters only for the age at which the decision process starts. For subsequent ages, these values are updated based on the test outcomes (e.g., $M-$ or $M+$) and a patient's observed adherence behaviors using (1) and (2).

Mammography adherence depends on several factors such as economic status, education level, existence of health insurance, and lifestyle. There already exist several statistical models that identified adherence-influencing factors, and the initial adherence can be estimated using one of such models (e.g., Rahman et al. 2005, Armstrong et al. 2004, Gierisch et al. 2010, Lerman et al. 1990). Indeed, a recent meta-analysis combined and synthesized the findings of 195 such models and data from more than 4.5 million women (Schueler et al. 2008). Table 4 summarizes strong predictors of mammography utilization based on the findings of this meta-analysis.

To estimate a patient's initial breast cancer risk, we can use one of several existing breast cancer risk estimation models (see, e.g., Reinier et al. 2007, Claus et al. 2001, Trentham-Dietz et al. 2000, Gail et al. 1989). Among these risk estimation models, the most widely used one is the Gail model, which is a validated model that estimates breast cancer risk based on current age, age at menarche, age at first live birth, number of first-degree relatives with breast cancer, number of previous biopsies, and presence of atypical findings

[1] http://www.cdc.gov/nchs/nhis.htm (accessed September 29, 2015).
[2] http://breastscreening.cancer.gov/ (accessed September 29, 2015).

Table 4 Strong Predictors of Mammography Utilization

Factor	No. of women	No. of studies	Multivariate OR	(95% CI)
Financial				
No insurance	217,381	40	0.47	(0.39, 0.57)
Public vs. private insurance	81,039	12	0.74	(0.63, 0.86)
Low income/money concerns	3,515,336	43	0.74	(0.67, 0.82)
Educational				
Low education level	359,212	52	0.78	(0.73, 0.83)
Access				
No physician recommendation	122,975	23	0.16	(0.08, 0.33)
No physician visit within year	3,316,763	17	0.34	(0.25, 0.47)
No primary care physician/usual source of care	171,635	33	0.41	(0.32, 0.53)
Rural	69,910	9	0.75	(0.63, 0.90)
Use of past screening				
Clinical breast examination	82,252	10	9.15	(3.49, 23.98)
Pap test	49,964	13	3.45	(2.12, 5.62)
Breast disease				
No personal history of benign breast disease	57,773	13	0.51	(0.42, 0.62)
No family history of breast cancer	206,930	21	0.69	(0.61, 0.78)
Breast health beliefs				
Mammogram harmful	39,809	6	0.54	(0.43, 0.67)
Mammogram only needed if symptoms	44,458	5	0.56	(0.43, 0.72)
Modesty/embarrassment	18,628	3	0.55	(0.39, 0.76)
Fatalism	18,614	7	0.61	(0.40, 0.92)
Breast cancer knowledge				
Poor knowledge of screening	25,482	11	0.46	(0.35, 0.60)
Acculturation				
Recent immigrant	15,186	6	0.54	(0.37, 0.79)
Lifestyle				
Smoker	147,256	12	0.69	(0.60, 0.80)
Not married	204,813	33	0.79	(0.68, 0.90)
Drinker	40,822	6	1.3	(1.09, 1.54)

Note. OR, odds ratio.
Source. Adapted from Schueler et al. (2008).

on a biopsy. To estimate in situ breast cancer risk, Gail et al. (1999) proposed using the incidence ratio between in situ and invasive cancers. The Gail model has been modified several times to allow applicability to various populations. The latest version of the Gail model is available on the NCI's website[3] and is used around 20,000 to 30,000 times each month to quantify an individual woman's risk of breast cancer in daily clinical practice (Elmore and Fletcher 2006).

6. Numerical Results

In this section, we present our numerical results. We used an Intel Xeon 2.67 GHz processor with 12 GB RAM for our computational experiments and solved our model using the algorithm presented in §4. The computation time for solving the base case problem using our solution algorithm was 153 hours. Although this is a long time, we remark that we need to solve these POMDPs only once to generate the policy graphs, and actual decisions can be made by utilizing these policy graphs.

The remainder of this section is organized as follows. Section 6.1 discusses the optimal screening recommendations from a clinical decision maker's perspective. Section 6.1.1 illustrates the dynamic updates for adherence and breast cancer risk and discusses how the optimal screening recommendations may change based on a patient's observed adherence behavior and prior test outcomes. Section 6.2 discusses screening recommendations from a health policy perspective and compares the performance of the optimal screening policy to several actual screening policies. Sections 6.3 and 6.4 quantify the value of considering adherence and the role of heterogeneity in adherence, respectively. Section 6.5 estimates the potential impact of improving adherence to current guidelines. Lastly, §6.6 presents extended model results and sensitivity analysis. Throughout §6, we define a high-risk woman as someone who has a family history of breast cancer. Also, we assume that a patient with a detected cancer is treated and kept under surveillance with annual mammograms till age 85, regardless of the policy adopted.

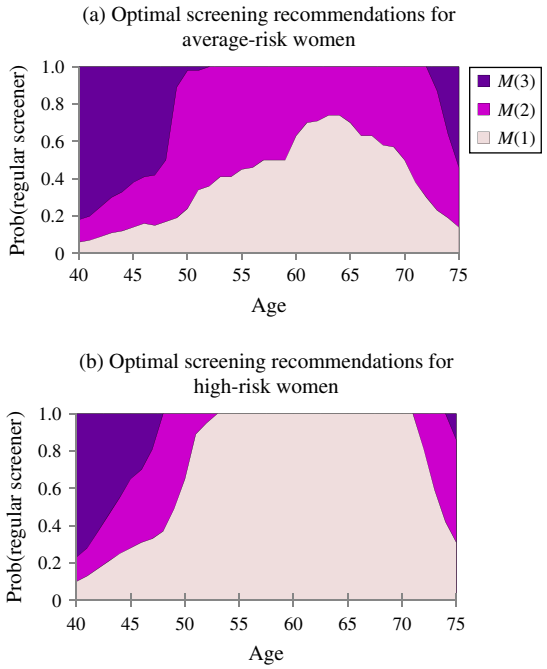
6.1. Optimal Screening Recommendations: A Clinical Decision Maker's Perspective

Figure 2 demonstrates the effects of adherence on the optimal recommendations for average (Figure 2(a)) and high-risk women (Figure 2(b)) at different ages. Note that, in Figure 2, the y-axis represents the probability of being a regular screener, which is a measure for the adherence behavior of women. That is, women with a high (low) likelihood of being regular screeners represent patients who are highly likely (unlikely) to adhere to screening recommendations. The optimal screening recommendations can be interpreted as follows: for a 40-year-old average-risk woman who has an 80% likelihood of being a regular screener, triennial screening (M(3)) recommendation is optimal (Figure 2(a)).

These results shed light on the controversial breast cancer screening guidelines. For example, should annual or biennial mammography screening be recommended? Is triennial screening ever optimal? Our findings suggest that there is not a "one-size-fits-all" strategy, and the optimal screening recommendations may change based on women's adherence behaviors and risk levels. For example, whereas triennial screening may be optimal for younger or older women who are highly likely to adhere, annual or biennial screening could be recommended for women aged 50–70, depending on their adherence levels. On the other

[3] http://www.cancer.gov/bcrisktool/ (last accessed September 29, 2015).

Figure 2 **(Color online) Optimal Screening Recommendations as a Function of Varying Adherence Rates for Different Age Groups**



Note. The y-axis represents the probability of being a regular screener, which is a measure for the adherence behavior of women.

hand, triennial screening between 50–70, as recommended in the United Kingdom, is never optimal for the U.S. population. Furthermore, regardless of women's adherence behaviors, our findings suggest the most frequent screening should be recommended between the ages of 50 and 70, an age group for which breast cancer is known to be more aggressive compared to older ages, and cancer incidence and mammography accuracy are higher compared to younger ages (Jemal et al. 2007b, Kerlikowske et al. 2000).

Figure 3 shows the optimal screening recommendations as a function of in situ and invasive breast cancer risks for women with different adherence behaviors at various ages. We define the threshold in situ and invasive breast cancer risks for action a as the probabilities over which action a or more aggressive screening recommendations are optimal. The optimal risk-based screening policy can be interpreted as follows: for a 40-year-old woman who is estimated to be a 70% regular screener, the threshold in situ and invasive breast cancer risks, respectively, are 2.4% and 1.9% for annual screening, 2% and 1.7% for biennial screening, and 0.1% and 0.1% for triennial screening (Figure 3(a)).

Tables 5 and 6 present the corresponding QALYs, expected number of mammograms, and false positives for average and high-risk women with various adherence levels, respectively, and compare them to the existing policies. For average-risk women, when adherence level is high (e.g., 70% or 100%), biennial screening strategies result in higher QALYs than annual screening strategies. On the other hand, when adherence level is low (e.g., 40%), annual screening strategies perform better (Table 5). This implies that for average-risk women who are less (more) likely to adhere, more (less) aggressive screening strategies are more appropriate.

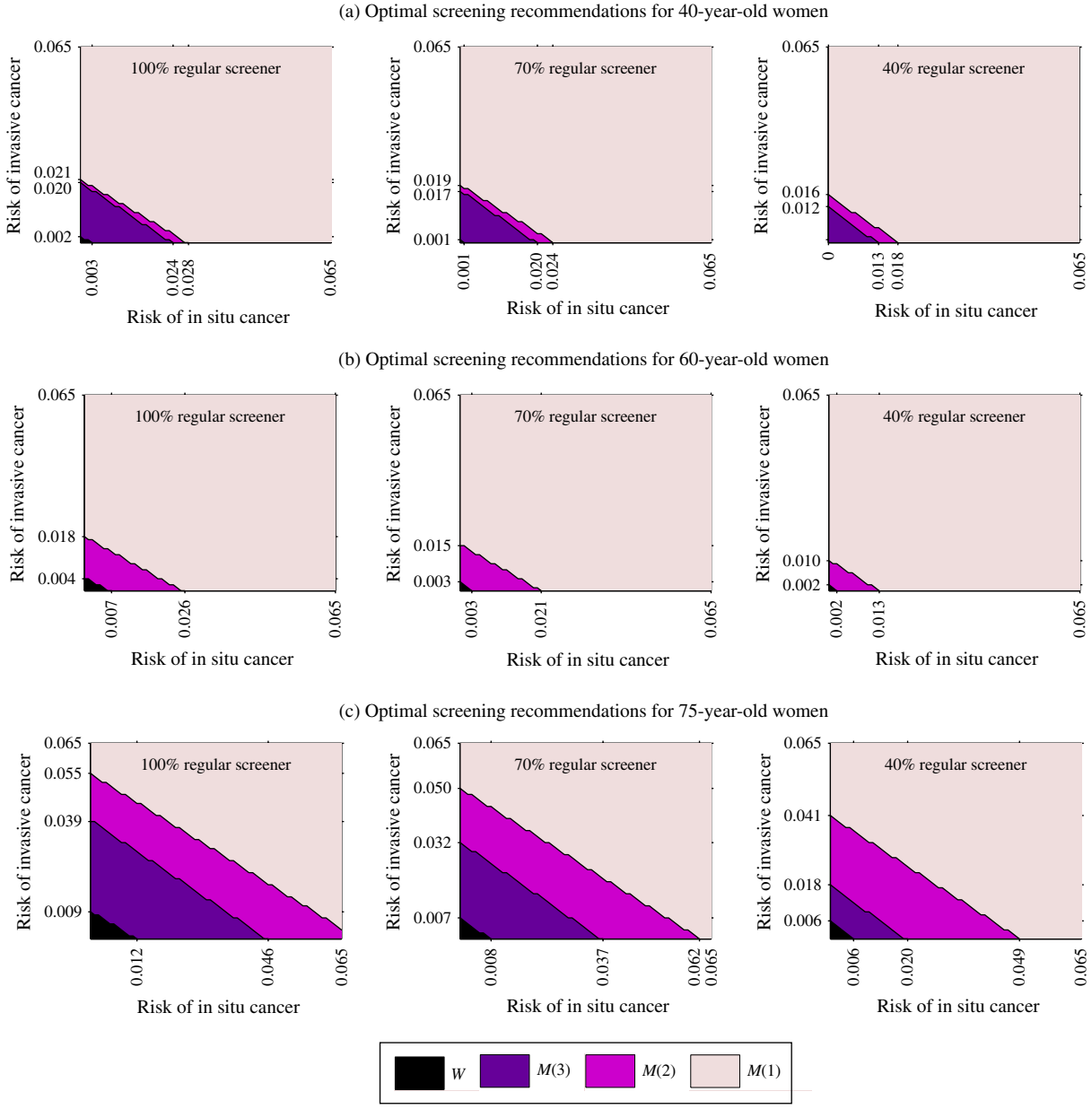
For women at increased risk of breast cancer, contrary to the average-risk women, annual screening strategies result in higher QALY gains than biennial screening strategies (Table 6). Furthermore, at any adherence rate, starting screening at age 40 results in greater QALY gains compared to starting at age 50 (QALY gains almost double). These results imply that starting screening early and having aggressive screening between the ages of 40 and 49 are especially critical for high-risk women.

6.1.1. Case Examples. Based on a patient's adherence behavior (measured by existing adherence-influencing factors and prior appointment attendance history) and her breast cancer risk (measured by existing breast cancer risk factors and prior test outcomes), the optimal screening recommendations may change significantly. In the following case examples, we illustrate the effects of initially estimated adherence level and appointment attendance history on optimal screening recommendations. In both of these cases (summarized in Table 7), we assume that the outcomes of the screening mammograms are negative and no self-detection occurs.

Case 1a. We consider a patient whose breast cancer risk characteristics are listed in row 1 of Table 7. We estimate her current risks of in situ and invasive cancers as 0.2% and 0.4%, respectively, using the Gail model described in §5.2. With respect to adherence-influencing factors, we do not know anything about this woman other than the fact that she has no family or personal history of breast cancer; therefore, we assume that she is a woman with average adherence and estimate her likelihood of adherence as 46%. Hence, we specify her initial health belief state as $b = [99.4\%\ 0.2\%\ 0.4\%\ 0\%]$ and her adherence belief state as $\pi = [46\%\ 54\%]$. Figure 4(a) shows the optimal screening recommendations for this patient. She undergoes mammography at age 40 and is recommended to undergo another one two years later. However, she misses her appointment at age 42 and does not show up until age 45. As a result, her estimated likelihood of being a regular screener decreases to 34% and her in situ and invasive cancer risks are updated as 0.3% and 0.2% at age 45, when she is recommended an annual screening.

Case 1b. We consider another patient who has the same breast cancer risk factors as the patient in Case 1a (row 2 of Table 7), but for this woman, we also know that she has a history of clinical breast exam. Based on the Schueler model (Schueler et al. 2008) described in §5.2, we estimate her initial adherence as 89%.

Figure 3 (Color online) Optimal Screening Recommendations as a Function of the Risk of In Situ and Invasive Cancers for Women with Different Ages and Screening Behaviors



Note. W, wait; $M(3)$, triennial screening; $M(2)$, biennial screening; $M(1)$, annual screening.

Figure 4(b) shows the optimal screening recommendations for this patient. She undergoes a mammography at age 40 and is recommended to undergo another one three years later. If she shows up at age 43, her likelihood of being a regular screener is updated to 94%. In addition, as a result of the negative mammogram, her invasive cancer risk decreases to 0.2%. Given these belief states (i.e., $\pi = [95\%\ 5\%]$ and $b = [99.6\%\ 0.2\%\ 0.2\%\ 0\%]$), the optimal policy again prescribes triennial screening at age 43. If she continues to follow the physician's recommendations and attends her future appointments, triennial screening is recommended until age 52 and biennial screening thereafter.

6.2. Optimal Screening Policy: A Policy Maker's Perspective

Although screening recommendations can be tailored to a patient's risk level and adherence behavior at the clinical level, a policy maker may still prefer to promote a routine screening policy for the general population, as a routine screening policy might be easier to operationalize at the policy level. Therefore, in

Table 5 Expected QALYs, Number of Mammograms, and Number of False Positives for 40-Year-Old Average-Risk Women Under Different Screening Strategies

	Expected number of mammograms (per 1,000 women)			Expected number of false positives (per 1,000 women)			Expected QALYs gained (vs. no screening) (per 1,000 women)		
Strategy	100% adh.	70% adh.	40% adh.	100% adh.	70% adh.	40% adh.	100% adh.	70% adh.	40% adh.
Annual screening									
40–74 y	35,644	25,825	15,878	3,338	2,337	1,335	583	618	599
40–79 y	41,168	29,790	18,281	3,873	2,711	1,549	599	643	630
50–74 y	26,632	19,485	12,209	2,421	1,695	968	540	556	523
50–79 y	32,157	23,450	14,612	2,956	2,069	1,182	556	581	555
Biennial screening									
40–74 y	19,318	14,286	9,044	1,666	1,166	666	660	620	511
40–79 y	22,215	16,405	10,375	1,933	1,353	773	688	651	541
50–74 y	15,265	11,420	7,365	1,258	881	503	592	549	445
50–79 y	18,163	13,539	8,696	1,525	1,067	610	620	579	475
Triennial screening									
50–70 y	10,313	7,799	5,044	779	545	311	555	484	358
Optimal screening	19,452	15,767	11,452	1,647	1,296	882	701	691	638

Notes. adh., Adherence. Expected number of mammograms represents the attended number of mammograms, not the recommended ones.

Table 6 Expected QALYs, Number of Mammograms, and Number of False Positives for 40-Year-Old High-Risk Women Under Different Screening Strategies

	Expected number of mammograms (per 1,000 women)			Expected number of false positives (per 1,000 women)			Expected QALYs gained (vs. no screening) (per 1,000 women)		
Strategy	100% adh.	70% adh.	40% adh.	100% adh.	70% adh.	40% adh.	100% adh.	70% adh.	40% adh.
Annual screening									
40–74 y	36,806	27,656	18,326	3,097	2,168	1,238	1,796	1,765	1,610
40–79 y	41,931	31,334	20,556	3,594	2,515	1,437	1,811	1,788	1,639
50–74 y	27,706	21,033	14,179	2,246	1,572	898	912	903	824
50–79 y	32,831	24,711	16,409	2,743	1,920	1,097	928	926	853
Biennial screening									
40–74 y	21,565	16,815	11,748	1,546	1,082	618	1,712	1,572	1,280
40–79 y	24,253	18,781	12,983	1,793	1,255	717	1,738	1,601	1,307
50–74 y	17,146	13,496	9,526	1,167	817	467	953	872	701
50–79 y	19,835	15,462	10,763	1,415	990	566	979	900	728
Triennial screening									
50–70 y	12,539	10,082	7,198	723	506	289	914	793	587
Optimal screening	22,533	18,144	13,982	1,605	1,203	818	1,900	1,810	1,645

Notes. adh., Adherence. Expected number of mammograms represents the attended number of mammograms, not the recommended ones.

this section, we focus on a policy maker's perspective, and compare the optimal screening policy to the routine screening policies listed in Table 1 across a range of adherence and risk levels. In this regard, the optimal screening policy provides an upper bound for the routine screening policies in terms of the QALYs gained per year (compared to no screening). Other performance measures include expected number of mammograms per year and expected number of false positives per year imposed by each screening scenario. We evaluate policies for the entire population of women over age 40 based on the current adherence rates, population dynamics, and risk distribution.

Table 7 Case Examples Illustrating the Effects of Adherence Level and Personal Risk Factors on Screening Policies

Case	Age	Race	FH	PH	Prior Bx	Age at menstruation	Age at first birth	Prob(reg. screener) (%)
1a	40	Caucasian	No	No	No	11	30	46
1b	40	Caucasian	No	No	No	11	30	89

Note. FH, family history of breast cancer; PH, personal history of breast cancer; Prior Bx, prior biopsy; Prob(reg. screener), likelihood of being a regular screener.

Table 8 presents our results comparing numerous screening strategies, and Figure 5(a) presents the societal trade-offs based on the results presented in Table 8. Our results show that the total number of mammograms performed each year in the United States would be between 33.29 and 36.16 million, when an annual screen-

Figure 4 (Color online) Effect of Appointment Attendance on the Optimal Screening Recommendations

(a) Optimal screening recommendations for the patient presented in Case 1a

Age	Belief states	Show-up?	Action
40	b = [99.4% 0.2% 0.4% 0%] π = [46% 54%]	Yes	$M(2)$
42	b = [99.7% 0.2% 0.1% 0%] π = [20% 80%]	No	W
43	b = [99.6% 0.2% 0.2% 0%] π = [16% 84%]	No	W
44	b = [99.5% 0.3% 0.2% 0%] π = [7% 93%]	No	W
45	b = [99.5% 0.3% 0.2% 0%] π = [34% 66%]	Yes	$M(1)$

(b) Optimal screening recommendations for the patient presented in Case 1b

Age	Belief states	Show-up?	Action
40	b = [99.4% 0.2% 0.4% 0%] π = [89% 11%]	Yes	$M(3)$
43	b = [99.6% 0.2% 0.2% 0%] π = [94% 6%]	Yes	$M(3)$
46	b = [99.5% 0.3% 0.2% 0%] π = [97% 3%]	Yes	$M(3)$
49	b = [99.4% 0.4% 0.2% 0%] π = [98% 2%]	Yes	$M(3)$
52	b = [99.4% 0.4% 0.2% 0%] π = [99% 1%]	Yes	$M(2)$

ing policy is in place. Given that most regular screeners follow the ACS guidelines (Howard and Adams 2012), our estimations for the annual screening policies and corresponding total number of mammograms and false positives reflect the current screening patterns in the United States, which is in line with the findings of previous research (Wainer 2011, Lewis et al. 2006).

From a policy maker's perspective, when cost considerations are set aside (which is the current practice in mammography screening policies), promoting annual screening (such as annual screening between the ages of 40 and 74 or 79) for the general population appears to be more appropriate (and perform close to the optimal policy) than promoting a less aggressive screening policy, such as biennial screening. This is because more than half of the U.S. women do not adhere to screening recommendations and biennial screening is optimal only for patients who are not at high risk and are highly likely to adhere. Therefore, based on current adherence data, our findings indicate that, for the general population, the ACS guidelines may result in higher QALYs as compared to the recent USPSTF recommendations that promote biennial screening between the ages of 40 and 74.

Recent research has shown that the 2009 revision to the USPSTF breast cancer screening guidelines did not have a short-term impact on screening patterns and most regular screeners still follow the ACS guidelines (Howard and Adams 2012), which lead to higher QALYs than the USPSTF guidelines under the current

Table 8 Expected Annual QALYs, Number of Mammograms, and Number of False Positives Under Different Screening Strategies with Current Adherence Rates

Strategy	Expected number of mammograms/year (in millions)	Expected number of false positives/year (in millions)	Expected QALYs gained/year (vs. no screening)
Annual screening			
40–74 y	33.297	2.876	413,522
40–79 y	37.411	3.297	464,471
50–74 y	31.853	2.730	402,531
50–79 y	36.168	3.172	453,480
Biennial screening			
40–74 y	19.289	1.509	361,212
40–79 y	21.340	1.716	410,800
50–74 y	18.640	1.445	350,923
50–79 y	20.781	1.661	400,511
Triennial screening			
50–70 y	14.946	1.097	268,031
Optimal screening	31.588	2.784	478,439

Note. Expected number of mammograms represents/year the attended number of mammograms, not the recommended ones.

Table 9 Expected Annual Additional QALY Savings When Adherence Is Considered

Adherence rate (%)	Additional expected QALYs gained/year (vs. no screening) (in thousands)		QALY savings due to the consideration of adherence/year[a] (in thousands) (C = B − A)
	Optimal policy from perfect adherence model (A)	Optimal policy from imperfect adherence model (B)	
20	231.07	381.60	150.53 (65%)
30	306.03	440.36	134.33 (44%)
40	363.71	469.22	105.51 (29%)
46[b]	391.93	478.44	86.51 (22%)
50	408.56	482.57	74.02 (18%)
60	443.69	490.08	46.39 (10%)
70	471.33	498.16	26.83 (6%)
80	493.09	506.49	13.40 (3%)
90	510.15	514.99	4.84 (1%)

[a]Numbers in parentheses show the percentage increase when compared to the results from the perfect adherence model (column A).
[b]Represents current adherence rate.

adherence level of the overall population. However, in the long run, if most regular screeners adopt the USPSTF guidelines without any change in adherence rates, then we estimate that this change in screening patterns would result in about 62.5 to 113.5 thousand QALY reduction per year.

6.3. Value of Considering Adherence

In this section, we analyze the benefits of considering adherence in a mammography screening schedule. For this purpose, we have done the following analysis. First, we found the optimal policy when adherence is assumed to be 100%, which resulted in 523.42 thousand QALY savings per year over no screening. Then, we simulated this policy allowing imperfect adherence (we allowed adherence to range from 10% to 90%), and calculated the corresponding expected QALYs (column A in Table 9). Next, we calculated the optimal policy and corresponding expected QALYs at these adherence rates using the model that considered imperfect adherence (column B in Table 9). Lastly, to estimate the value of considering adherence, we calculated the difference in QALYs between the optimal policy when imperfect adherence is considered, and simulated optimal policy with imperfect adherence (column C in Table 9, which is equal to B–A).

As presented in Table 9, when adherence is explicitly considered, annual QALY savings due to considering adherence range from 4.84 thousand (when adherence = 90%) to 150.53 thousand (when adherence = 20%) (see column C). When adherence is equal to 46%, which represents the current overall population adherence, then annual QALY savings due to considering adherence is equal to 86.51 thousand (22% increase). Noting that the maximum annual QALY savings for the optimal screening policy over no screening could be at most 523.42 thousand (when adherence = 100%), we believe such a savings is remarkable.

6.4. Quantifying the Role of Heterogeneity in Adherence

To quantify the role of heterogeneity, we calculated the outcomes under the assumption that adherence is homogeneous in the population and is equal to 46% (average adherence in the population). Our results show that when average adherence is assumed for everyone, the effect on patients with already low adherence would be relatively small, but patients with high adherence would be adversely affected (see Table 10). This is because screening policies optimized with respect to average adherence lead to unnecessarily aggressive strategies for women with already high adherence. In particular, patients with adherence rates of 60% and over will have substantial QALY reduction with significantly more mammograms and false positives. For example, for every 1,000 women with 80% adherence rate, following a policy optimized based on average

Table 10 Expected Additional QALY Savings When Heterogeneity in Adherence Is Considered

Adherence rate (%)	QALYs gained (vs. no screening) (per 1,000 women)			Expected number of mammograms (per 1,000 women)			Expected number of false positives (per 1,000 women)		
	Optimal policy assuming 46% adherence for everyone	Optimal policy considering actual adherence	QALY difference	Optimal policy assuming 46% adherence for everyone	Optimal policy considering actual adherence	Number of mammograms difference	Optimal policy assuming 46% adherence for everyone	Optimal policy considering actual adherence	Number of false positives difference
20	518	528	10	10,305	11,453	1,148	775	882	107
30	594	603	9	14,358	15,768	1,410	1,162	1,297	135
40	630	638	8	18,282	19,738	1,456	1,550	1,690	140
50	646	653	7	22,146	22,302	156	1,937	1,947	10
60	649	663	14	25,978	21,349	−4,629	2,324	1,848	−476
70	643	672	29	29,791	20,530	−9,261	2,712	1,763	−949
80	632	682	50	33,591	20,333	−13,258	3,099	1,742	−1,357
90	617	691	74	37,383	20,740	−16,643	3,487	1,780	−1,707

Table 11 Expected QALYs, Number of Mammograms, and Number of False Positives per Year Under Different Screening Strategies with Various Adherence Rates

	Expected number of mammograms/year (in millions)				Expected number of false positives/year (in millions)				Expected QALYs gained/year (vs. no screening) (in thousands)			
Strategy	Current	60% adh.	70% adh.	80% adh.	Current	60% adh.	70% adh.	80% adh.	Current	60% adh.	70% adh.	80% adh.
Annual screening												
40–74 y	33.30	42.12	48.35	54.56	2.88	3.75	4.38	5.00	413.52	424.66	424.69	420.37
40–79 y	37.41	47.43	54.54	61.61	3.30	4.30	5.02	5.73	464.47	471.12	467.04	458.17
50–74 y	31.85	40.25	46.17	52.07	2.73	3.56	4.16	4.75	402.53	414.95	416.09	412.98
50–79 y	36.17	45.82	52.66	59.46	3.17	4.14	4.83	5.52	453.48	461.41	458.44	450.78
Biennial screening												
40–74 y	19.29	24.24	27.70	31.10	1.51	1.97	2.30	2.62	361.21	402.32	423.14	438.73
40–79 y	21.34	26.85	30.71	34.52	1.72	2.24	2.61	2.98	410.80	453.85	474.86	489.98
50–74 y	18.64	23.41	26.72	29.99	1.45	1.88	2.20	2.51	350.92	391.75	412.60	428.37
50–79 y	20.78	26.13	29.87	33.56	1.66	2.17	2.53	2.89	400.51	443.28	464.32	479.61
Triennial screening												
50–70 y	14.95	18.86	21.57	24.23	1.10	1.43	1.67	1.91	268.03	313.75	340.25	362.63
Optimal screening	31.59	31.86	31.31	30.74	2.78	2.80	2.75	2.69	478.44	490.08	498.00	506.49

Notes. adh., Adherence. Expected number of mammograms represents the attended number of mammograms, not the recommended ones.

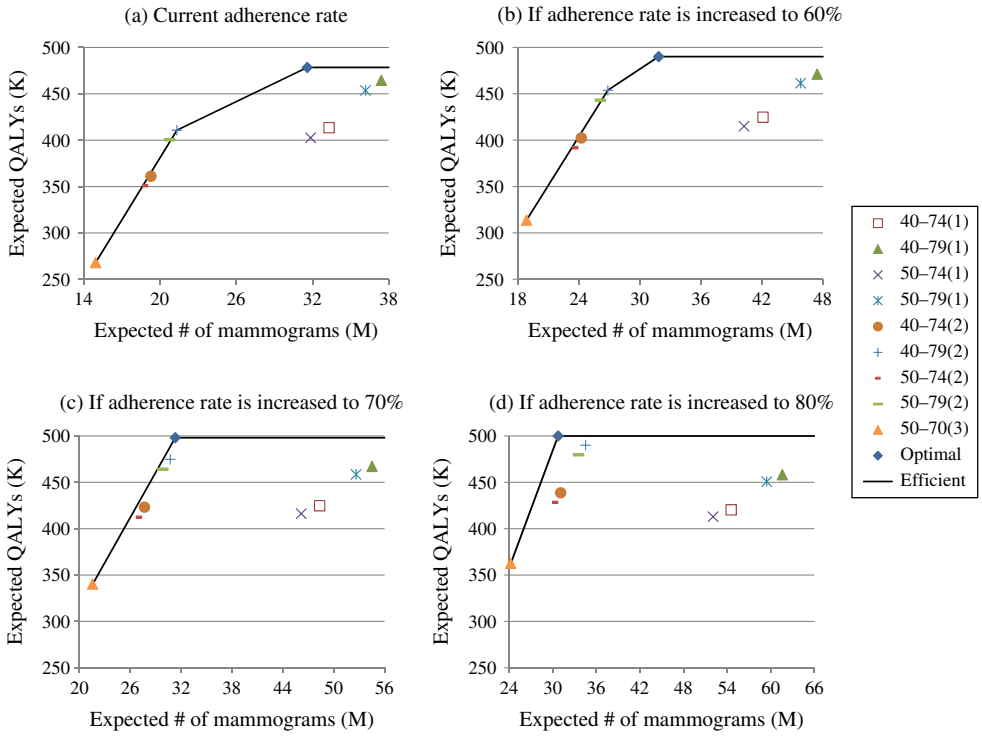
adherence will lead to 50 less QALYs, 13,258 additional (unnecessary) mammograms and 1,357 more false positives.

6.5. Potential Impact of Improving Adherence to Current Guidelines

Health services research has shown that simple reminder systems can significantly improve adherence (O'Donohue and Levensky 2006). In this section, we analyze the impact of improving adherence to current guidelines. In particular, we consider the impact of improving overall adherence rates to 60%, 70%, and 80% on annual number of mammograms/year, false positives/year, and QALYs gained/year.

Table 11 presents the overall impact of improving adherence in the context of different screening guidelines, and Figure 5 presents the changes in the societal trade-offs as a function of improving adherence rates.

Figure 5 (Color online) Societal Trade-off Representation of the Screening Policies Reported in Table 11



Our results show that, if the screening patterns remain the same over the long run (i.e., most regular screeners follow the ACS guidelines, annual screening between age 40–74), improving overall adherence rate from 46% to 60% would result in an increase of 11.14 thousand QALYs per year. As overall adherence increases further, QALY gains start to decrease. Any improvement on adherence rates between 60%–70% would translate into nonsignificant gains in QALY savings and any improvement beyond 70% adherence would lead to a reduction in QALY gains. Although this finding may appear counterintuitive, there is an intuitive explanation. This is because when adherence rates are very high, annual screening may be "too aggressive" for the general population, the majority of which consists of average-risk women. As a result, at such high adherence levels, such aggressive policies would result in too many "unjustified" false positives and unnecessary biopsies without any significant increase in detection rates, which in turn would lead to a reduction in QALY gains.

On the other hand, if the screening patterns change in the long run and most regular screeners adopt the USPSTF guidelines (biennial screening between the ages of 50 and 74), then improving overall mammography adherence becomes more critical. For example, for the USPSTF guidelines, improving overall adherence from 46% to 60%, 70%, and 80% would result in about 40.83, 61.68, and 77.45 thousand QALY increase per year, respectively. Given that prior guideline changes had long-term impacts on screening patterns, it is likely to observe a similar trend as a result of the recent guideline change over the long run (Howard and Adams 2012). If this happens, then improving adherence to mammography would become much more critical than it was in the past.

6.6. Extended Results and Sensitivity Analysis

In this section, we present the results for the case where we relax the assumption that adherence does not depend on breast cancer risk and history of false positives. We also assess the sensitivity of our results against relative changes in disutilities associated with mammography.

At present, there is a limited knowledge base and conflicting literature on the effects of a patient's risk on adherence behavior (see, e.g., Gierisch et al. 2010, Lipkus et al. 1996, Calvocoressi et al. 2004, Rahman et al. 2003, Rakowski et al. 2006). Therefore, in our base case analysis, we assume that the existence of known breast cancer risk factors does not affect adherence behavior. On the other hand, we acknowledge that, as more evidence comes to light, future research may lead to different findings on the effect of breast cancer risk on adherence behavior. In this regard, from a modeling standpoint, our framework is general enough and capable of incorporating the effects of many risk factors on adherence behavior.

Breast cancer risk factors can be categorized as (1) risk factors that usually do not change after age 40 (e.g., ethnicity, a known family history, age at menarche, age at first live birth), and (2) dynamic risk factors that may change after age 40 (e.g., body-mass index). Most risk factors are static and do not change with age, hence they do not dynamically interact with adherence behavior. Therefore, any known effects of such static risk factors on adherence could be considered by adjusting adherence parameters in our model. For example, if a known family history is discovered to increase adherence rate in the future, our adherence rate parameters could be adjusted accordingly to capture this affect. To illustrate this idea, we reran our model to analyze how our results would change if a known family history is discovered to increase adherence rates by 10% (see Figure 6). As illustrated in Figure 6, if a known family history increases the adherence rates by 10%, then this increased adherence leads to a slightly less aggressive optimal screening strategy. However, as presented in Figure 6, the results are not very sensitive against such changes.

Similarly, because of the limited knowledge base, in our base case analysis we assume that a history of false-positive mammograms does not affect adherence

Figure 6 **(Color online) Optimal Policies for Women with a Family History of Breast Cancer Under Different Modeling Assumptions**

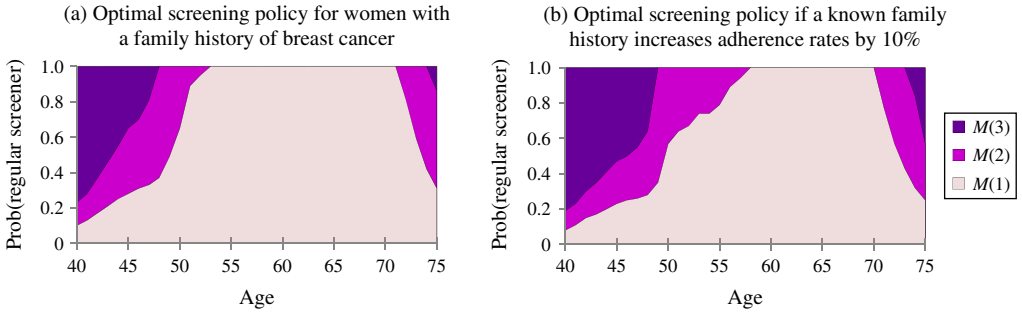
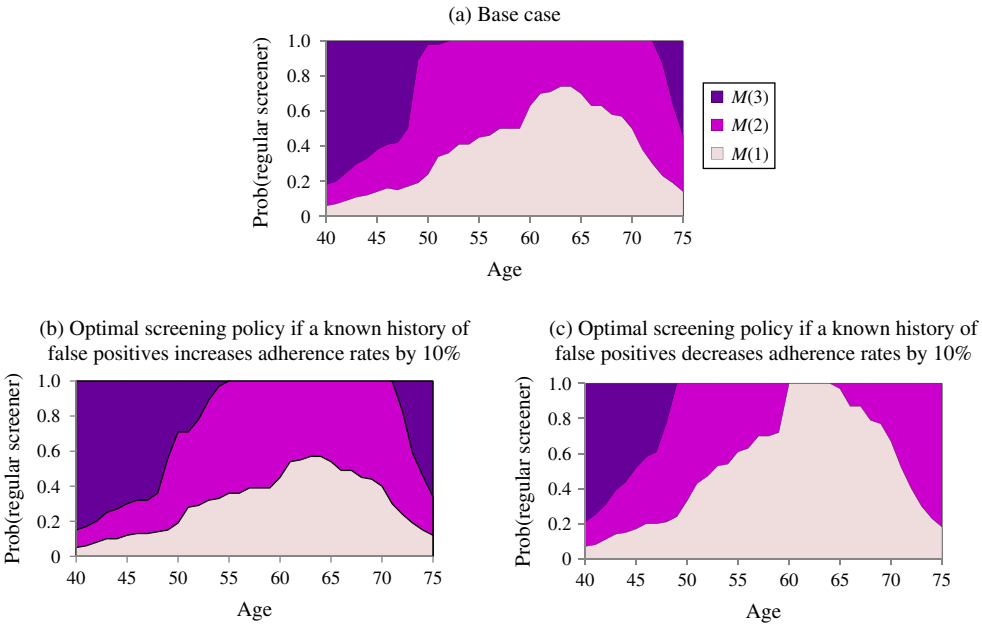


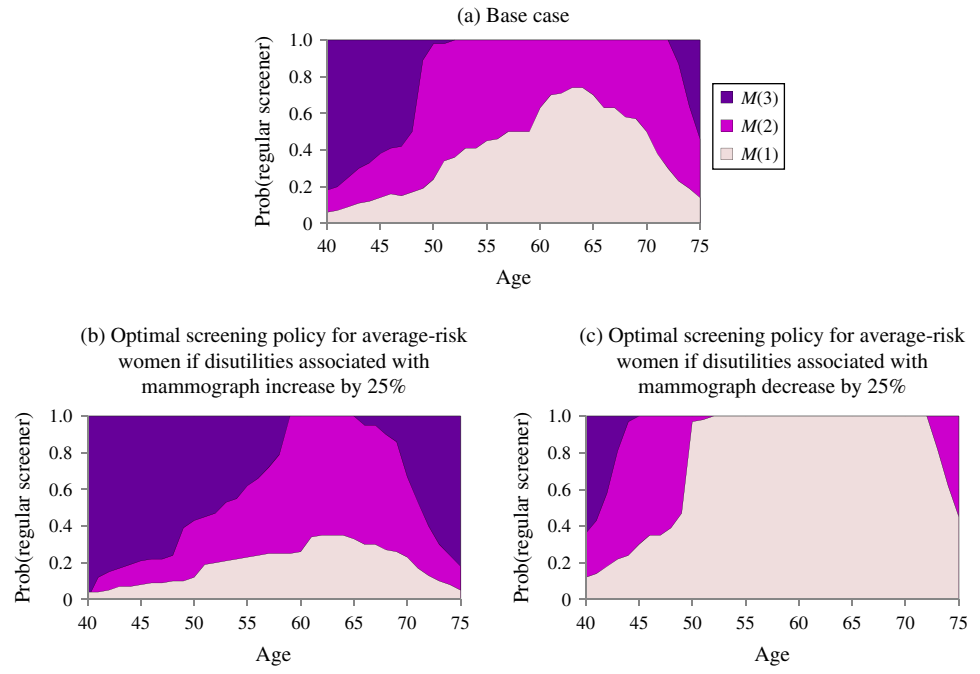
Figure 7 **(Color online) Optimal Policies for Average-Risk Women with a History of False Positives Under Different Modeling Assumptions**



behavior. To assess if and how our results would change if a known history of false positives changed the adherence behavior, we rerun our model for the cases where a known history of false positives were assumed to increase or decrease adherence rates by 10% (see Figure 7). As shown in Figure 7, whereas an increase in adherence rates would not lead to significant changes in optimal policies, a decrease would lead to significantly more aggressive policies especially for women between the ages of 50 and 70.

Regarding the sensitivity of our results with respect to the disutilities associated with mammography, we observe that the results are sensitive to relative changes in these parameters Figure 8), which is in line with the findings of previous studies (Stout et al. 2006, Mandelblatt et al. 2009). In particular, relative increases

Figure 8 **(Color online) Sensitivity of the Optimal Screening Policies Against Varying Disutilities**



(decreases) in the disutilities associated with mammography lead to higher (lower) threshold probabilities for being a regular screener, below which more aggressive screening is optimal (see Figure 8).

7. Discussion and Conclusion

Although adherence is identified as the primary reason for suboptimal clinical benefit (WHO 2003), the existing breast cancer screening guidelines do not explicitly consider the possibility of nonadherence to the recommended screening policies. In this study, we present an analytical framework to study the role of mammography adherence and heterogeneity in women's adherence behaviors on breast cancer screening policies.

Our research has important policy implications. First, given that the existing breast cancer screening studies mostly ignored the effect of imperfect adherence and resulted in conflicting guidelines, our results have the potential to shed light on the trade-offs inherent in different breast cancer screening policies. In particular, we find that, contrary to recent USPSTF recommendations, which advocate biennial screening, promoting more aggressive screening policies for the general public, such as annual screening as recommended by the ACS, is more beneficial, because more than half of U.S. women do not adhere to screening recommendations. On the other hand, when we assume 100% adherence, our results show that biennial screening is better than annual screening for the general population. In fact, the studies that provided the basis for the USPSTF policy updates assumed 100% adherence to breast screening recommendations, which could be one of the main reasons why such conclusions were reached (Mandelblatt et al. 2009).

Second, our results show that although promoting annual screening is better for the general public and beneficial for overall society, women with already high adherence would be adversely affected by such aggressive strategies. In particular, we find that biennial screening is better for such women, and annual screening would result in too many unnecessary mammograms and false positives, resulting in a substantial reduction in QALYs. This suggests that heterogeneity in adherence should also be considered when operationalizing breast cancer screening policies. In that regard, our results suggest that screening intervals should be adjusted and extended to two years (as recommended by the USPSTF) for women with a history of high adherence in clinical practice. We remark that although guidelines exist for the general population, similar adjustments are already being made in clinical practice in an ad-hoc manner, depending on the specific cases and practicing physician. With the improved technology, optimal screening recommendations tailored to women's adherence and risk levels might be operationalized at the clinical level via the use of a decision support system that estimates and keeps track of the patient's risk level and adherence behavior in the future. Given that the existing breast cancer screening guidelines also emphasize the need for personalizing cancer screening policies based on women's needs (Smith et al. 2009), we believe our framework is a promising step in this direction.

Third, although the recent revision to the USPSTF guidelines did not have a short-term impact on screening patterns and most regular screeners still follow annual screening policies, prior experience has shown that patients tend to adopt to policy changes over the long run (Howard and Adams 2012). In that regard, if the screening patterns change in the long run and most regular screeners adopt the new USPSTF guidelines as expected, then improving overall mammography adherence becomes more critical, which otherwise results in many undetected cancers and a significant reduction in QALYs.

Fourth, we find that irrespective of women's adherence levels and breast cancer risks, the most aggressive screening should be performed between the ages of 50 and 70, rather than younger (40–50) or older (70+) age groups. This is because breast cancer risk increases as women age, but cancer progression is more aggressive in younger women. These two factors balance each other, which results in more aggressive screening strategies in this intermediate age category.

Although our finding that women who are likely to adhere should receive a less frequent screening recommendation is more intuitive, the finding that women who are less likely to adhere should receive a more frequent screening recommendation may appear somewhat contradictory, since these women are already known to show low participation. However, we remark that adherence behavior is strongly correlated with physician recommendations (Sohl and Moyer 2007) and the cumulative adherence rate within a specified period of time (such as in three years) is likely to increase as the frequency of screening recommendations during this period increases. For example, a patient with a low likelihood of adherence may be more likely to come in three years after three successive annual mammography recommendations compared to a patient who is recommended a single mammography during this three-year period. To increase the likelihood of mammography attendance of women who have a weak appointment attendance history, frequent screening recommendations might be delivered through mailed reminders or telephone calls, which have been shown to be effective at increasing mammography attendance (Taplin et al. 2000).

Our model can also be used to assess the potential benefits of adherence-enhancement interventions. As noted in §1, a majority of the adherence literature is

devoted to the adherence-enhancement studies. However, there is limited research on cost-effectiveness of adherence-enhancement interventions (Saywell et al. 2003). Although our model is developed for breast cancer screening, it provides a general framework for improving screening policies for other diseases that may be sensitive to the cohorts' adherence behaviors and risk levels.

Our study has some limitations. First, in our numerical analysis, we assume that adherence to personalized dynamic screening schedules would be the same as that of a routine screening strategy with a constant interval. At present, there is a limited knowledge base on the effects of personalized dynamic screening schedules on adherence, which is an active area of research (Yuan et al. 2012, Schousboe et al. 2011, Zhu et al. 2012). Personalized screening schedules could be operationalized through the means of personalized automated decision support systems, and currently randomized controlled trials (RCTs) are being conducted on this area for prostate cancer, and they are planned to be extended to breast cancer over time (Yuan et al. 2012). The hypothesis is that individualized screening schedules will provide each patient a well-defined screening schedule, reduce patients' confusion due to variations in guidelines, and hence improve overall adherence (Yuan et al. 2012, Zapka et al. 2011). Note that, if this hypothesis is true, the benefit of such a personalized screening schedule would be even higher. Our modeling framework might provide more insight on this as more evidence comes to light. Second, although our modeling framework can capture the relationship between static adherence-influencing factors and their effects on adherence behavior, it does not explicitly consider the interaction between the dynamic risk factors that may change over time (such as insurance status, marriage status, etc.). Unfortunately, any reliable data to dynamically capture the effects of such factors do not exist. One possible approach to ameliorate the effects of this limitation would be to reset the estimated adherence at the time of change (e.g., reestimate adherence using the statistical model when insurance status changes). In addition, there is a paucity of data reflecting the effects of a patient's risk on adherence behavior. Although a few small-scale studies reported some (and often conflicting) associations between breast cancer risk and adherence behavior (see Gierisch et al. 2010, Calvocoressi et al. 2004, Lipkus et al. 1996, Rahman et al. 2003), existing nationwide studies did not find any association between repeat mammography behavior and risk of breast cancer (Rakowski et al. 2006). As more evidence comes to light, future research may lead to different findings on the effect of breast cancer risk on adherence behavior. Third, we assume that a history of false-positive mammograms does not affect adherence behavior. There is also a limited knowledge base in this area, and the existing literature reports conflicting results on the effects of false-positive mammograms on adherence behavior. In particular, whereas some studies reported that false positives interfere with subsequent screening (Álamo-Junquera et al. 2012, Chiarelli et al. 2003, Hofvind et al. 2003), several other studies reported that false positives did not have a negative effect on subsequent screening (Burman et al. 1999, Lerman et al. 1991, O'Sullivan et al. 2001, Pisano et al. 1998, Lampic et al. 2003, Pinckney et al. 2003). In addition, a recent study showed that the effect of a false-positive mammogram on adherence to subsequent screening was conditional on the receipt of a physician's recommendation for mammography. In particular, this study showed that among women who had false positives, those who said their physicians did not advise them to have mammograms were less likely to adhere (compared with those who had normal results), whereas those who reported that their physicians advised them to undergo screening had similar adherence rates with those who had normal results. That is, physician advice for screening was found to dominate the influence of false positives. Note that as more evidence comes to light about the effects of a history of false-positive mammograms on future adherence, any such effect can be captured in our model by recomputing the policies for such women (see §6.6). Lastly, in our modeling framework, there is a single decision maker (the physician), and the patients' adherence behavior is modeled as a stochastic event in response to the physician's recommendations. An alternative to a single decision-maker model is a multiagent game-theoretic model, where the physician and patient are the model agents. However, to our knowledge, there does not exist any medical data (and clinical evidence) to capture and reflect the dynamics of such a game. Further, game-theoretic models have not received much credit in modeling doctor-patient relationship, as physicians are often altruistic, and patients often trust physicians.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2015.2180.

Acknowledgments

The authors thank Rakesh Sarin, an anonymous associate editor, and three anonymous referees. This research was supported by the National Institutes of Health [Grants R01LM010921 and R01CA165229].

References

ACS (2012) Can breast cancer be found early? American Cancer Society, Atlanta. Accessed September 3, 2013, http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-detection.

Álamo-Junquera D, Murta-Nascimento C, Macià F, Baré M, Galcerán J, Ascunce N, Zubizarreta R, et al. (2012) Effect of false-positive results on reattendance at breast cancer screening programmes in Spain. *Eur. J. Public Health* 22(3):404–408.

Arias E (2006) United States life tables, 2004. *National Vital Statist. Reports* 54(14):1–40.

Armstrong K, Long JA, Shea JA (2004) Measuring adherence to mammography screening recommendations among low-income women. *Preventive Medicine* 38(6):754–760.

Ayer T, Alagoz O, Stout NK (2012) A POMDP approach to personalize mammography screening decisions. *Oper. Res.* 60(1): 1017–1018.

Baker RD (1998) Use of a mathematical model to evaluate breast cancer screening policy. *Health Care Management Sci.* 1(2): 103–113.

Barton MB, Harris R, Fletcher SW (1999) The rational clinical examination. Does this patient have breast cancer? The screening clinical breast examination: Should it be done? How? *J. Amer. Medical Assoc.* 282(13):1270–1280.

Bauch CT, Li M, Chapman G, Galvani AP (2010) Adherence to cervical screening in the era of human papillomavirus vaccination: How low is too low? *Lancet Infectious Diseases* 10(2): 133–137.

Baxter N (2001) Should women be routinely taught breast self-examination to screen for breast cancer? *Canadian Medical Assoc. J.* 164(13):1837–1845.

BCSC (Breast Cancer Surveillance Consortium) Risk Calculator (2013) Distribution of 5-year breast cancer risk in the breast cancer surveillance consortium (BCSC) estimated with the BCSC risk calculator. *J. National Cancer Inst.* 105(1):5.

Brewer NT, Salz T, Lillie SE (2007) Systematic review: The long-term effects of false-positive mammograms. *Ann. Internal Medicine* 146(7):502–510.

Burman ML, Taplin SH, Herta DF, Elmore JG (1999) Effect of false-positive mammograms on interval breast cancer screening in a health maintenance organization. *Ann. Internal Medicine* 131(1):1–6.

Calvocoressi L, Kasl SV, Lee CH, Stolar M, Claus EB, Jones BA (2004) A prospective study of perceived susceptibility to breast cancer and nonadherence to mammography screening guidelines in African American and white women ages 40 to 79 years. *Cancer Epidemiology Biomarkers Prevention* 13(12): 2096–2105.

Cassandra AR, Kaelbling LP (1998) Exact and approximate algorithms for partially observable Markov decision processes. Doctoral thesis, Brown University, Providence, RI.

CDC (Centers for Disease Control and Prevention) (1995) The national breast and cervical cancer early detection program at-a-glance. U.S. Department of Health and Human Services Centers for Disease Control and Prevention, Atlanta. Last accessed September 29, 2015, http://www.cdc.gov/cancer/nbccedp/about.htm.

Chiarelli AM, Moravan V, Halapy E, Majpruz V, Mai V, Tatla RK, (2003) False-positive result and reattendance in the Ontario breast screening program. *J. Medical Screening* 10(3):129–133.

Christensen AJ (2004) *Patient Adherence to Medical Treatment Regimens: Bridging the Gap Between Behavioral Science and Biomedicine* (Yale Univeristy Press, New Haven, CT).

CISNET (Cancer Intervention and Surveillance Modeling Network) (2010) Input parameter generator interfaces. National Cancer Institute, Bethesda, MD. Accessed September 2, 2010, http://cisnet.cancer.gov/interfaces.

Claus EB, Stowe M, Carter D (2001) Breast carcinoma in situ: Risk factors and screening patterns. *J. National Cancer Inst.* 93(23): 1811–1817.

Cronin KA, Mariotto AB, Clarke LD, Feuer EJ (2006) Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *JNCI Cancer Spectrum* 2006(36):26–29.

Cronin KA, Yu B, Krapcho M, Miglioretti DL, Fay MP, Izmirlian G, Ballard-Barbash R, Geller BM, Feuer EJ (2005) Modeling the dissemination of mammography in the United States. *Cancer Causes Control* 16(6):701–712.

Dellorto D (2009) Task force opposes routine mammograms for women age 40-49. *CNN* (November 16), http://www.cnn.com/2009/HEALTH/11/16/mammography.recommendation.changes/index.html?eref=onion.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005) *Methods for the Economic Evaluation of Health Care Programmes* (Oxford University Press, New York).

Eagle JN (1984) The optimal search for a moving target when the search path is constrained. *Oper. Res.* 32(5):1107–1115.

Earle CC, Chapman RH, Baker CS, Bell CM, Stone PW, Sandberg EA, Neumann PJ (2000) Systematic overview of cost-utility assessments in oncology. *J. Clinical Oncology* 18(18):3302–3317.

Elmore JG, Fletcher SW (2006) The risk of cancer risk prediction: "What is my risk of getting breast cancer?" *J. National Cancer Inst.* 98(23):1673–1675.

Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW (1998) Ten-year risk of false positive screening mammograms and clinical breast examinations. *New England J. Medicine* 338(16): 1089–1096.

Elmore JG, Reisch LM, Barton MB, Barlow WE, Rolnick S, Harris EL, Herrinton LJ, et al. (2005) Efficacy of breast cancer screening in the community according to risk level. *J. National Cancer Inst.* 97(14):1035–1043.

Flynn PM, Betancourt H, Ormseth SR (2011) Culture, emotion, and cancer screening: An integrative framework for investigating health behavior. *Ann. Behavioral Medicine* 42(1):79–90.

Friedman LC, Woodruff A, Lane M, Weinberg AD, Cooper HP, Webb JA (1995) Breast cancer screening behaviors and intentions among asymptomatic women 50 years of age and older. *Amer. J. Preventive Medicine* 11(4):218–223.

Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. National Cancer Inst.* 81(24):1879–1886.

Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, Vogel V (1999) Weighing the risks and benefits of Tamoxifen treatment for preventing breast cancer. *J. National Cancer Inst.* 91(21):1829–1846.

Gierisch JM, Earp JA, Brewer NT, Rimer BK (2010) Longitudinal predictors of nonadherence to maintenance of mammography. *Cancer Epidemiology Biomarkers Prevention* 19(4):1103–1111.

Hofvind SS-H, Wang H, Thoresen S (2003) The Norwegian breast cancer screening program: Re-attendance related to the women's experiences, intentions and previous screening result. *Cancer Causes Control* 14(4):391–398.

Howard DH, Adams EK (2012) Mammography rates after the 2009 U.S. Preventive Services Task Force breast cancer screening recommendation. *Preventive Medicine* 55(5):485–487.

IOM (Institute of Medicine) (2005) *Saving Women's Lives: Strategies for Improving Breast Cancer Detection and Diagnosis* (National Academy Press, Washington, DC).

Ivy JS (2002) A maintenance model for breast cancer detection and treatment. Working paper.

Jayasinghe UW, Taylor R, Boyages J (2005) Is age at diagnosis an independent prognostic factor for survival following breast cancer? *ANZ J. Surgery* 75(9):762–767.

Jemal A, Ward E, Thun M (2007b) Recent trends in breast cancer incidence rates by age and tumor characteristics among U.S. women. *Breast Cancer Res.* 9(3):1–6.

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ (2009) Cancer statistics, 2009. *Cancer J. Clinicians* 59(4):225–249.

Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ (2007a) Cancer statistics, 2007. *Cancer J. Clinicians* 57(1):43–66.

Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL (1995) Efficacy of screening mammography. A meta-analysis. *J. Amer. Medical Assoc.* 273(2):149–154.

Kerlikowske K, Carney PA, Geller B, Mandelson MT, Taplin SH, Malvin K, Ernster V, et al. (2000) Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann. Internal Medicine* 133(11): 855–863.

Klabunde CN, Ballard-Barbash R (2007) Evaluating population-based screening mammography programs internationally. *Seminars Breast Disease* 10(2):102–107.

Knudsen AB, McMahon PM, Gazelle GS (2007) Use of modeling to evaluate the cost-effectiveness of cancer screening programs. *J. Clinical Oncology* 25(2):203–208.

Krakovska O, Wahl LM (2007) Optimal drug treatment regimens for HIV depend on adherence. *J. Theoret. Biol.* 246(3):499–509.

Lampic C, Thurfjell E, Sjödén P-O (2003) The influence of a false-positive mammogram on a woman's subsequent behaviour for detecting breast cancer. *Eur. J. Cancer* 39(12):1730–1737.

Lerman C, Rimer B, Trock B, Balshem A, Engstrom PF (1990) Factors associated with repeat adherence to breast cancer screening. *Preventive Medicine* 19(3):279–290.

Lerman C, Track B, Rimer BK, Boyce A, Jepson C, Engstrom PF (1991) Psychological and behavioral implications of abnormal mammograms. *Ann. Internal Medicine* 114(8):657–661.

Lewis RS, Sunshine JH, Bhargavan M (2006) A portrait of breast imaging specialists and of the interpretation of mammography in the United States. *Amer. J. Roentgenology* 187(5):W456–W468.

Lipkus IM, Rimer BK, Strigo TS (1996) Relationships among objective and subjective risk for breast cancer and mammography stages of change. *Cancer Epidemiology Biomarkers Prevention* 5(12): 1005–1011.

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

Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu AL, Atkins D, Klein J, Helfand M (2003) The cost-effectiveness of screening mammography beyond age 65 years: A systematic review for the U.S. Preventive Services Task Force. *Ann. Internal Medicine* 139(10):835–842.

Mandelblatt JS, Wheat ME, Monane M, Moshief RD, Hollenberg JP, Tang J (1992) Breast cancer screening for elderly women with and without comorbid conditions: A decision analysis model. *Ann. Internal Medicine* 116(9):722–730.

Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning JH, Draisma G, Huang H, et al. (2009) Effects of mammography screening under different screening schedules: Model estimates of potential benefits and harms. *Ann. Internal Medicine* 151(10): 738–747.

Mason JE, England DA, Denton BT, Smith SA, Kurt M, Shah ND (2011) Optimizing statin treatment decisions for diabetes patients in the presence of uncertain future adherence. *Medical Decision Making* 32(1):154–166.

Mayne L, Earp JA (2003) Initial and repeat mammography screening: Different behaviors/different predictors. *J. Rural Health* 19(1): 63–71.

Messina CR, Lane DS, Glanz K, West DS, Taylor V, Frishman W, Powell L (2004) Relationship of social support and social burden to repeated breast cancer screening in the women's health initiative. *Health Psych.* 23(6):582–594.

Miron RE, Smith RJ, (2010) Modelling imperfect adherence to HIV induction therapy. *BMC Infectious Diseases* 10(1):6–22.

Monahan GE (1982) A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Sci.* 28(1):1–16.

O'Donohue WT, Levensky ER (2006) *Promoting Treatment Adherence: A Practical Handbook for Health Care Providers* (Sage Publications, New York).

O'Neill SC, Bowling JM, Brewer NT, Lipkus IM, Skinner CS, Strigo TS, Rimer BK (2008) Intentions to maintain adherence to mammography. *J. Womens Health* 17(7):1133–1141.

Ornstein SM, Musham C, Reid A, Jenkins RG, Zemp LD, Garr DR (1993) Barriers to adherence to preventive services reminder letters: The patient's perspective. *J. Family Practice* 36(2): 195–200.

O'Sullivan I, Sutton S, Dixon S, Perry N (2001) False positive results do not have a negative effect on reattendance for subsequent breast screening. *J. Medical Screening* 8(3):145–148.

Ozekici S, Pliska SR (1991) Optimal scheduling of inspections: A delayed Markov model with false positives and negatives. *Oper. Res.* 39(2):261–273.

Parker-Pope T (2009) Benefits and risks of cancer screening are not always clear, experts say. *New York Times* (October 21), http://www.nytimes.com/2009/10/22/health/22screen.html?_r=0.

Parmigiani G (1993) On optimal screening ages. *J. Amer. Statist. Assoc.* 88(422):622–628.

Partin MR, Slater JS, Caplan L (2005) Randomized controlled trial of a repeat mammography intervention: Effect of adherence definitions on results. *Preventive Medicine* 41(3–4):734–740.

Pinckney RG, Geller BM, Burman M, Littenberg B (2003) Effect of false-positive mammograms on return for subsequent screening mammography. *Amer. J. Medicine* 114(2):120–125.

Pisano ED, Earp J, Schell M, Vokaty K, Denham A (1998) Screening behavior of women after a false-positive mammogram. *Radiology* 208(1):245–249.

Qaseem A, Snow V, Sherif K, Aronson M, Weiss KB, Owens DK, Clinical Efficacy Assessment Subcommittee of the American College of Physicians (2007) Screening mammography for women 40 to 49 years of age: A clinical practice guideline from the American College of Physicians. *Ann. Internal Medicine* 146(7):511–515.

Rahman SMM, Dignan MB, Shelton BJ (2003) Factors influencing adherence to guidelines for screening mammography among women aged 40 years and older. *Ethnicity Disease* 13(4): 477–484.

Rahman SMM, Dignan MB, Shelton BJ (2005) A theory-based model for predicting adherence to guidelines for screening mammography among women age 40 and older. *Internat. J. Cancer Prevention* 2(3):169–179.

Rakowski W, Meissner H, Vernon SW, Breen N, Rimer B, Clark MA (2006) Correlates of repeat and recent mammography for women ages 45 to 75 in the 2002 to 2003 health information national trends survey (hints 2003). *Cancer Epidemiology Biomarkers Prevention* 15(11):2093–2101.

Reinier K, Vacek P, Geller B (2007) Risk factors for breast carcinoma in situ versus invasive breast cancer in a prospective study of pre- and post-menopausal women. *Breast Cancer Res. Treatment* 103(3):343–348.

Sabatino SA, Habarta N, Baron RC, Coates RJ, Rimer BK, Kerner J, Coughlin SS, Kalra GP, Chattopadhyay S (2008) Interventions to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers systematic reviews of provider assessment and feedback and provider incentives. *Amer. J. Preventive Medicine* 35(1): S67–S74.

Saywell RM, Champion VL, Zollinger TW, Maraj M, Skinner CS, Zoppi KA, Muegge CM (2003) The cost effectiveness of 5 interventions to increase mammography adherence in a managed care population. *Amer. J. Managed Care* 9(1):33–44.

Schousboe JT, Kerlikowske K, Loh A, Cummings SR (2011) Personalizing mammography by breast density and other risk factors for breast cancer: Analysis of health benefits and cost-effectiveness. *Ann. Internal Medicine* 155(1):10–20.

Schueler KM, Chu PW, Smith-Bindman R (2008) Factors associated with mammography utilization: A systematic quantitative review of the literature. *J. Womens Health* 17(9):1477–1498.

Shapiro S, Coleman EA, Broeders M, Codd M, de Koning H, Fracheboud J, Moss S, et al. (1998) Breast cancer screening programmes in 22 countries: Current policies, administration and guidelines. *Internat. J. Epidemiology* 27(5):735–742.

Shumaker SA, Ockene JK, Riekert KA (2009) *The Handbook of Health Behavior Change* (Springer, New York).

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

Smith RA, Cokkinides V, Brawley OW (2009) Cancer screening in the United States, 2009: A review of current American cancer society guidelines and issues in cancer screening. *Cancer J. Clinicians* 59(1):27–41.

Smith RJ (2006) Adherence to antiretroviral HIV drugs: How many doses can you miss before resistance emerges? *Proc. Royal Soc. B: Biol. Sci.* 273(1586):617–624.

Sohl SJ, Moyer A (2007) Tailored interventions to promote mammography screening: A meta-analytic review. *Preventive Medicine* 45(4):252–261.

Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG (2006) Retrospective cost-effectiveness analysis of screening mammography. *J. National Cancer Inst.* 98(11): 774–782.

Szabo L (2009) Report: Mammograms may not be needed until age 50. *USA Today* (November 17), http://usatoday30.usatoday.com/news/health/2009-11-17-1Amammogram17_ST_N.htm.

Taplin SH, Barlow WE, Ludman E, MacLehos R, Meyer DM, Seger D, Herta D, Chin C, Curry S (2000) Testing reminder and motivational telephone calls to increase screening mammography: A randomized study. *J. National Cancer Inst.* 92(3):233–242.

Trentham-Dietz A, Newcomb PA, Storer BE, Remington PL (2000) Risk factors for carcinoma in situ of the breast. *Cancer Epidemiology Biomarkers Prevention* 9(7):697–703.

U.S. Census Bureau (2010) Census of population and housing. Last accessed September 29, 2015, http://www.census.gov/prod/www/decennial.html.

USPSTF (2009) Clinical guidelines: Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann. Internal Medicine* 151(10):716–726.

Velanovich V (1995) Immediate biopsy versus observation for abnormal findings on mammograms: An analysis of potential outcomes and costs. *Amer. J. Surgery* 170(4): 327–332.

Verkooijen HM (2002) Diagnostic accuracy of stereotactic large-core needle biopsy for nonpalpable breast disease: Results of a multicenter prospective study with 95% surgical confirmation. *Internat. J. Cancer* 99(6):853–859.

Wainer H (2011) How should we screen for breast cancer? *Significance* 8(1):28–30.

Watson-Johnson LC, DeGroff A, Steele CB, Revels M, Smith JL, Justen E, Barron-Simpson R, Sanders L, Richardson LC (2011) Mammography adherence: A qualitative study. *J. Womens Health* 20(12):1887–1894.

WHO (World Health Organization) (2003) *Adherence to Long-Term Therapies: Evidence for Action* (World Health Organization, Geneva).

Yanovitzky I, Blitz CL (2000) Effect of media coverage and physician advice on utilization of breast cancer screening by women 40 years and older. *J. Health Comm.* 5(2):117–134.

Yuan MJ, Hébert ET, Johnson RK, Long J, Vandewater EA, Vickers AJ (2012) A personalized automated messaging system to improve adherence to prostate cancer screening: Research protocol. *JMIR Res. Protocols* 1(2):e20. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3626152/.

Zapka JM, Klabunde CN, Arora NK, Yuan G, Smith JL, Kobrin SC (2011) Physicians' colorectal cancer screening discussion and recommendation patterns. *Cancer Epidemiology Biomarkers Prevention* 20(3):509–521.

Zelen M (1993) Optimal scheduling of examinations for the early detection of disease. *Biometrika* 80(2):279–293.

Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schröder FH, Vickers AJ (2012) Risk-based prostate cancer screening. *Eur. Urology* 61(4):652–661.