7-2015

Preventive Maintenance Decision Modeling in Health and Service Systems

Mahboubeh Madadi
University of Arkansas, Fayetteville

Preventive Maintenance Decision Modeling in Health and Service Systems

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Industrial Engineering

by

Mahboubeh Madadi
Shahid Beheshti University
Bachelor of Science in Applied Mathematics, 2005
University of Tehran
Master of Science in Industrial Engineering, 2009

July 2015
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

_____
Dr. Shengfan Zhang
Dissertation Director

_____
Dr. C. Richard Cassady
Committee Member

_____
Dr. Edward A. Pohl
Committee Member

_____
Dr. Lisa M. Maillart
Committee Member

Abstract

This dissertation focuses on the preventive maintenance decision modeling in healthcare and service systems. In the first part of this dissertation, some issues in preventive health decisions for breast cancer are addressed, and in the second part, the required characteristics for preventive maintenance of an unreliable queuing system are derived.

Adherence to cancer screening is the first issue that is addressed in this dissertation. Women's adherence or compliance with mammography screening remained low in the recent years. In this dissertation, we first develop a design-based logistic regression model to quantify the probability of adherence to screening schedules based on women's characteristics. In Chapter 3, we develop a randomized finite-horizon partially observable Markov chain to evaluate and compare different mammography screening strategies for women with different adherence behaviors in terms of quality adjusted life years (QALYs) and lifetime breast cancer mortality risk. The results imply that for the general population, the American Cancer Society (ACS) policy is an efficient frontier policy. In Chapter 4, the problem of overdiagnosis in cancer screening is addressed. Overdiagnosis is a side effect of screening and is defined as the diagnosis of a disease that will never cause symptoms or death during a patient's lifetime. We develop a mathematical framework to quantify the lifetime overdiagnosis and mortality risk for different screening policies, and derive the (near) optimal policies with minimum overdiagnosis risk.

In the second part, we consider an unreliable queuing system with servers stored in a shared stack. In such a system, servers have heterogeneous transient usage since servers on the top of the stack are more likely to be used. We develop a continuous-time Markov chain model to

derive the utilization and usage time of servers in the system. These quantities are critical for the decision maker for deriving a maintenance policy.

Acknowledgments

I would like to express my gratitude to my advisor, Dr. Shengfan Zhang, for her tremendous supports, guidance, patience, and for providing me with an excellent atmosphere for doing research from the very first day of my doctoral education. My research would not have been possible without her supports. I also would like to thank Dr. Richard Cassady for providing me with the idea of expanding my dissertation to the area of reliability and maintenance and for his invaluable guidance. I would like to thank Dr. Lisa Maillart and Dr. Edward Pohl for their vital contribution and guidance to this dissertation. I would like to thank all my committee members for every opportunity they have given me to fulfill my goals.

I would like to thank the faculty and staff of the Department of Industrial Engineering at the University of Arkansas for their help and support in many ways. I would also like to thank all my fellow graduate students and friends for the laughter they shared with me and for making this journey an enjoyable experience for me.

I want to say thank you from my heart to my family. I cannot express how thankful I am for all their supports and love. This dissertation is dedicated to my parents, Fatemeh and Mohammad, for their unconditional love.

Contents

List of Figures

List of Tables

List of Papers

Chapter 2: Madadi, M., Zhang, S., Yeary, K. H. K., & Henderson, L. M. (2014). Analyzing factors associated with women's attitudes and behaviors toward screening mammography using design-based logistic regression. *Breast cancer research and treatment*, *144*(1), 193-204.

Chapter 3: Madadi, M., Zhang, S., & Henderson, L. M. (2015). Evaluation of breast cancer mammography screening policies considering adherence behavior. *European Journal of Operational Research*, In Press.

1 Introduction: Preventive Maintenance in Healthcare and Service Systems

"Prevention is better than treatment." This proverb has inspired much research in a wide range of areas such as health, service and manufacturing systems.

For health systems, preventive healthcare, or preventive medicine, focuses on disease prevention and health maintenance. Preventive healthcare includes early diagnosis of disease, counseling, and other interventions to avert a health problem. The goal of preventive healthcare is to prevent health problems or identify and treat them early before their effect on the patient's life is irreversible. Chronic diseases, such as heart diseases and cancers, are the causes for 7 in every 10 deaths among Americans each year, and they account for 75% of the nation's health spending [1]. These chronic diseases can be largely preventable through appropriate preventive care like screening tests. Screening tests can advance the time of disease detection and detect the disease at stages when it is more curable and the treatment works more effectively. For example, mammography screening can detect breast cancer 1.7 years before a woman can feel a lump in her breast [2], and ultimately it can reduce breast cancer mortality rate by 15% [3].

Preventive maintenance in manufacturing and service systems, on the other hand, focuses on the care and maintenance of equipment and facilities to keep them in a satisfactory operating condition and to prevent them from sudden failure. The goal of preventive maintenance is to avoid and/or mitigate the consequences of failure of equipment. Preventive maintenance can save overall cost by decreasing equipment downtime, increasing life expectancy and efficiency of the equipment, enhancing customers' satisfaction in service-oriented systems, etc.

The focus of this dissertation is on the preventive care modeling in two areas: health systems and service-oriented systems. The former addresses the problem of preventive healthcare decisions

for chronic diseases with a focus on breast cancer, while the latter part provides insight on preventive maintenance of a service system with unreliable servers. Specifically, in the first part, we focus on the evaluation and optimization of preventive healthcare decisions which includes analysis and estimation of individual adherence behavior to preventive care, as well as the evaluation and optimization of cancer screening policies in terms of different patient outcomes. The following breast cancer screening issues are addressed: (1) characterizing the factors associated with patient's adherence to mammography screening recommendations and estimating the probability of patient adherence to a prescribed screening test, (2) evaluating and comparing various screening mammography policies while incorporating heterogeneity in women's adherence behaviors, and (3) minimizing overdiagnosis in cancer screening under various sources of uncertainty such as disease progression and efficacy of mammography tests. In the second part of this dissertation, an unreliable multi-server queuing system with stacked servers is studied. In such a system, the equipment/servers are used in a non-uniform fashion, i.e., servers on the top of the stack are more likely to be used in a finite time horizon than servers at the bottom. This implies that some servers deteriorate faster than the others. For such systems, an estimate of the transient utilization and usage time of servers can provide insight on deterioration patterns and preventive maintenance decisions.

A more detailed outline of this dissertation is presented below.

Chapter 2 addresses the problem of adherence to preventive mammography screening since the efficacy of mammography screening policies is strongly associated with women's compliance with the recommendations. We developed a statistical model that characterizes significant factors associated with women's adherence to mammography screening recommendations. The model is then used to estimate the probability that a woman complies with her physician's

recommendation of undergoing a screening mammogram. The second chapter is the cornerstone for the next chapter since the results of this chapter serve as inputs for Chapter 3.

In Chapter 3, we develop a randomized discrete-time finite-horizon partially observable Markov chain model to investigate the effect of imperfect adherence on the patients' health outcomes for a wide range of screening mammography policies. The policies are evaluated in terms of the expected remaining quality adjusted life years (QALYs) and breast cancer lifetime mortality risk. Three possibilities for the cancer detection are considered in the proposed model: mammography screening detection, self-detection and interval cancer which represents fast growing cancers that show symptoms within 12 months following a negative mammogram. Various uncertainties such as breast cancer natural history, unreliability of mammography screening, interval cancer detection, and heterogeneity in patients' adherence behaviors are incorporated into the proposed model.

Chapter 4 addresses the overdiagnosis issue in cancer screening. Overdiagnosis is the major disadvantage of cancer screening and is defined as the diagnosis of screen-detected cancers that would not have presented clinically in a woman's lifetime in the absence of screening [4]. In this chapter, a probabilistic model is developed to evaluate the overdiagnosis and cancer mortality risks of a screening schedule. Furthermore, we propose an optimization model to minimize the risk of overdignosis of preventive screening schedules while maintaining the patient's lifetime breast cancer mortality risk at a predefined level. We incorporate uncertainty in disease onset, cancer sojourn time, and competing causes of death.

Chapter 5 focuses on a service-oriented environment. In this chapter, an unreliable queueing system with n identical servers that are stored in a stack is studied. In such a queuing system,

customers may find it more convenient to select the server that is on the top of the stack. Similarly, customers finished with a server may find it more convenient to place the server back to the top of the stack. As a result, servers that begin at or near the top of the stack are used more frequently than the servers that are placed lower in the stack. In such a system, a server's utilization is stochastically dependent on its initial location in the stack and the usage of the other servers. In Chapter 5, we develop a continuous-time Markov chain model to estimate transient utilization and usage time of servers in such a system. This study provides estimate on the deterioration behavior of servers which is critical information for the systems maintenance decisions.

References

1. Centers for Disease Control and Prevention, http://www.cdc.gov/healthcommunication/toolstemplates/entertainmented/tips/preventivehealth.html, accessed October 2014.

2. Center for Disease Control and Prevention, The National Breast and Cervical Cancer Early Detection Program, Available at: http://www.cdc.gov/cancer/nbccedp. Accessed March 2013

3. Gøtzsche P.C., Nielsen M. (2009). Screening for breast cancer with mammography. Cochrane Database Syst Rev (4):CD001877

4. Heywang-Köbrunner, S. H., Hacker, A., & Sedlacek, S. (2011). Advantages and disadvantages of mammography screening. Breast Care, 6(3), 199-207.

2 Analyzing Factors Associated with Women's Attitudes and Behaviors Toward Screening Mammography Using Design-Based Logistic Regression

2.1 Introduction

Breast cancer is the most common cancer worldwide in women contributing more than 25% of the total number of cancer cases diagnosed in 2012 [1]. It is also the most common non-skin cancer affecting women in the United States. According to the American Cancer Society, about one in eight US women will develop breast cancer in her lifetime. In 2015, an estimated 231,840 invasive breast cancer cases will be diagnosed, and about 40,290 will die from breast cancer [2].

Mammography plays an important role in detection of breast cancer when the cancer is in its early stage. It is known to be an effective method and the current standard practice for early diagnosis of breast cancer. On average, mammography can detect breast cancer 1.7 years before a woman can feel a lump in her breast [3]. In addition, it can reduce breast cancer mortality rate by a reasonable estimate of 15% [4].

There are varying recommendations from different agencies for screening mammography in the U.S. The two most referred guidelines are from the American Cancer Society (ACS) and the U.S. Preventive Services Task Force (USPTF). The ACS recommends starting annual mammography at age 40 but recommends no ending age [5]. In November 2009 the USPSTF issued new breast cancer screening guidelines, which recommend biennial screening for women ages 50 to 74. The USPSTF also recommended that women ages 40 to 49 should not undergo screening mammography unless they are in a high-risk group [6].

The impact of mammography on women's outcomes is strongly associated with women's compliance with the mammography recommendations. Examination of a clinic-based group of 216 women with a strong family history of breast or ovarian cancer showed that 50%, 83%, 69%, and 53% of the women at age groups of 30–39, 40–49, 50–64, and 65 years and older, respectively, had only one mammogram between 1995 and 1999 [7]. More recently, a study by the Centers for Disease Control and Prevention (CDC) revealed that, in 2010 about 67.1% of women aged 40 and older had one mammogram between 2008 and 2010 [8].

In this study, we aim to (1) identify significant factors (socio-demographic, health-related, behavioral attributes and knowledge mammography) associated with women's adherence to mammography screening, and (2) study the attitudes toward mammography in non-adherent women. We assume that women's compliance with screening mammography is strongly correlated with their intentions to get a mammogram (i.e., their expectations about their next mammogram or their thoughts about getting a mammogram [9]). Based on this assumption, if a woman has a positive attitude toward screening mammograms and her plan for the next mammogram is within the next one or two years, then she is considered to be adherent.

2.2 Methods

2.2.1 Data Source

We used the 2003 Health Information National Trends Survey (HINTS) data for this study [10]. HINTS, developed by the National Cancer Institute (NCI), is a nationally representative telephone survey of 6,369 adults aged 18 and older. The 2003 data is currently the only HINTS data that provides information on women's attitudes and perceptions toward mammography. HINTS 2005 and 2007 data do not include questions regarding mammography. HINTS 2012

only reports the time of a most recent mammogram and whether a doctor has informed about a mammogram; however, no data is collected on women's intentions or attitudes toward mammograms, which is the focus of this study.

2.2.2 Study Design

There are two stages for this study. In the first stage, we aim to identify significant factors associated with women's adherence to mammography screenings. To be consistent with current mammography guidelines and practices, we focus on women older than 40. We perform the analyses separately for two age groups: women younger than 65, and women 65 and older. In contrast to earlier studies, we measure adherence based on women's intentions and attitudes toward mammography.

In the second stage, we focus only on women older than 42 with poor mammography history, i.e. women who have never had mammography or their most recent mammography was more than two years ago and they did not say "never heard of it" when asked why they had not had a mammogram. Age 42 is used as the lower boundary for these analyses because an every-other-year interval for mammography requires enough time for women over age 40 to have a mammogram. We do not separate women younger and older than 65 in order to achieve an adequate sample size. We examine the differences between those who think about getting a mammogram and other women in this population.

2.2.3 Predicting Factors

We categorized factors that may be associated with women's adherence to mammography screening into four groups: *Socio-demographic characteristics* including age, race, marital status, employment, education, income, health coverage, and living area based on Census Division;

Health-related characteristics including body mass index (BMI), average number of visits to a healthcare provider a year, family history of cancer, psychological distress composite score, being advised to have a mammogram, and having had a Pap smear; *Behavioral characteristics* including trusting cancer information from doctors, looking for cancer information, perception of chance of getting breast cancer, eating habits (consumption of fruits and vegetables), having exercised in the last month, and lifetime number of smoked cigarettes; *Knowledge of breast cancer/mammography* including knowing the age at which mammography should begin and the frequency of receiving mammograms.

2.2.4 Statistical Analysis

HINTS data includes a set of 50 replicate weights, which were generated according to jackknife variance estimation [10]. Data were weighted to produce overall and stratified estimates that represent the U.S. population. We incorporate weights in the analyses for the population study. To handle missing values for the predicting factors, we use the *naive nearest neighbor hot deck* method, which takes advantage of the similarity between the observations to impute the missing values. Each missing value is replaced with the observed response from the nearest observation in the data set [11]. The distance between two observations (x and y) is

$$d(x,y) = \sqrt{\sum_{i=1}^{p} \delta_i(x_i, y_i)}, \qquad 2\text{-}1$$

Where p is the number of the variables in the data set and $\delta_i(x_i, y_i)$ is 1 if $x_i \neq y_i$ and 0 if $x_i = y_i$.

We defined *adherence to mammography (γ)* separately for women aged 40 and 41 years versus those aged 42 and older, as a women aged 40 or 41 without any mammography history can still be considered adherent if she plans to get a mammogram in the near future (2 years or 1 year

respectively). We defined a woman older than 42 as adherent if the interval between her most recent self-reported mammogram and the next expected mammogram was less than two years. Otherwise she was considered non-adherent. We categorized those who answered the question "Do you have a plan to get mammography" with "undecided" or "not planned" as non-adherent. For this response variable, we built separate models and compared results between those less than 65 years and those age 65 and older.

For the outcome of *thinking about getting a mammogram in non-adherent women (τ),* we narrow our attention to women who never had a mammogram or their most recent mammogram was more than two years ago. The response variable is one if the answer to the question "Have you thought about getting a mammogram" is yes, or zero otherwise.

We examined bivariate associations between all potential independent variables to select significant covariates for inclusion in subsequent models. We used Rao-Scott corrections to chi-squared tests to incorporate the survey design (50 jackknife weights) of the HINTS data.

The standard logistic regression assumes that the data come from a simple random sample (i.e., independent and identically distributed (iid)). Thus, the standard logistic regression may be inappropriate for analyzing sample survey data due to clustering and stratification methods used in the survey designs [12]. Since the HINTS data are collected based on complicated multistage sampling designs, the classic methods for simple random sampling data may be inefficient for analyzing the HINTS data.

Suppose that in the complex survey data collection, the population is separated into *k* strata ($k = 1, 2, \dots, K$), with *j* primary sampling units (PSUs, $j = 1, 2, \dots, M_k$) in each stratum and *i* elements $i = 1,2, \dots, N_{kj}$ in the kj^{th}PSU. Suppose the number of observations is: n =

$\sum_{k=1}^{K}\sum_{j=1}^{m_k} n_{kj}$ where n_{kj} is the number of elements from m_k primary sampling unit from stratum k. The pseudo-maximum likelihood function for the "design-based" logistic regression model is

$$\sum_{k=1}^{K}\sum_{j=1}^{m_k}\sum_{i=1}^{n_{kj}}\left[w_{kji}\times y_{kji}\right]\times\ln\left[\pi\left(x_{kji}\right)\right]+\left[w_{kji}\times(1-y_{kji})\right]\times\ln\left[1-\pi\left(x_{kji}\right)\right], \qquad 2\text{-}2$$

where w_{kji} is the known sampling weight for the kji^{th} observation [13].

Design-based multiple logistic regression models [12-13] were developed to identify significant factors associated with (1) adherence to mammography and (2) thinking about getting a screening mammogram in non-adherent women. The design based multivariate logistic regression model for the logistic regression is

$$\gamma(or\ \tau)=ln\left(\frac{\pi(x)}{1-\pi(x)}\right)=\alpha+\beta\times SD+\delta*HR+\gamma*BR+\theta*KBC\&M, \qquad 2\text{-}3$$

where SD, HR, BR and KBC&M are the vectors of socio-demographic, health related, behavior related and knowledge related features and $\alpha,\beta,\delta,\gamma,and\ \theta$ are the coefficients vectors of parameter estimates.

We used the Wald tests [13] to determine the statistical significance of each coefficient in the model. Classification tables [13] are used to evaluate the performance of the logistic regression models. We considered a predicted value of 0.5 as the cut off point for our classification analyses. Suppose $\widehat{\gamma_1}$ and $\widehat{\gamma_2}$ represent the predicted values from the regression models for adherence in women younger than 65 and women 65 and older, respectively. We consider a woman younger than 65 (65 and older) to be adherent if her probability to be adherent, $\widehat{\gamma_1}$ ($\widehat{\gamma_2}$), is greater than 0.5 and non-adherent otherwise. The same strategy is implemented for the outcome "thinking about

getting a mammogram" (τ) in non-adherent women. The observed values for the dependent outcomes and the predicted values are then cross-classified to evaluate the predictive accuracy of the models. All analyses are performed at the significance level of 5%.

2.3 Results

The distributions of all predicting factors for the two response variables are presented in Table 2-1. There are a total of 2,370 women included for the *adherence to mammography* response variable and 451 women included for the *thinking about getting a mammogram* response variable.

2.3.1 Comparing mammography adherence among women in the two age groups

The logistic regression results (Table 2-2) suggest that in the socio-demographic characteristics category, women who are single, separated, divorced or widowed are less likely to adhere to screening mammography compared to married women in both age groups (OR=0.685 and OR=0.499 for women younger than 65 and women aged 65 and older, respectively). As income increases, adherence to screening mammography also increases. Specifically among women younger than 65 the difference between women with income more than $75,000 and women whose income is less than $25,000 is significant (OR=2.526 for income greater than $75,000 versus income less than $25,000). Women younger than 65 who have health insurance are also more likely to adhere to screening mammography (OR= 2.959).

Table 2-1 Distribution of socio-demographic, health related, behavioral, and knowledge related characteristics

Independent Variables		Women's intentions and attitudes toward mammography				Thinking about mammogram in non-adherent women	
		Age < 65 (N=1612)		Age ≥ 65 (N=758)			
		$\gamma_1^* = 0$ (N=441)	$\gamma_1 = 1$ (N=1171)	$\gamma_2^{**} = 0$ (N=265)	$\gamma_2 = 1$ (N=493)	$\tau^\dagger = 0$ (N=199)	$\tau = 1$ (N=252)
Socio-demographic characteristics							
Race	Black	64 (14.51)	154 (13.15)	17 (6.41)	43 (8.72)	12 (6.03)	40 (15.87)
	White	284 (64.40)	914 (78.05)	199 (75.09)	420 (85.19)	14 (7.03)	16 (6.35)
	Other	36 (8.16)	53 (4.53)	11 (4.15)	10 (2.03)	160 (80.40)	183 (72.62)
	Missing	57 (12.95)	50 (4.27)	38 (14.34)	20 (4.06)	13 (6.53)	13 (5.16)
Marital status	Married	200 (45.35)	726 (62.00)	45 (16.98)	197 (39.96)	56 (28.14)	115 (45.63)
	Single/Divorced/widowed	202 (45.80)	429 (36.63)	192 (72.45)	291 (59.03)	137 (68.84)	135 (53.57)
	Missing	39 (8.84)	16 (1.36)	28 (10.56)	5 (1.01)	6 (3.01)	2 (0.79)
Employment	Employed	257 (58.28)	775 (66.18)	23 (8.68)	62 (12.58)	53 (26.63)	145 (57.54)
	Unemployed	144 (32.65)	383 (32.71)	214 (80.75)	424 (86.00)	141 (70.85)	102 (40.48)
	Missing	40 (9.07)	13 (1.11)	28 (10.56)	7 (1.42)	5 (2.51)	5 (1.98)
Education	Some high school or less	65 (14.74)	94 (8.03)	68 (25.66)	73 (14.81)	45(22.61)	39 (15.48)
	High school Graduate	134 (30.38)	334 (28.52)	84 (32.70)	177 (35.90)	68 (34.17)	127 (50.40)
	Some college/College graduate	204 (46.26)	723 (61.74)	86 (32.45)	237 (48.07)	81 (40.70)	84 (33.33)
	Missing	38 (8.62)	114 (9.73)	27 (10.19)	6 (1.22)	5 (2.51)	2 (0.79)
Income	≤ $25,000	146 (33.11)	249 (21.26)	147 (55.47)	204 (41.38)	106 (53.27)	100 (39.68)
	>$25,000 and ≤$75,000	161 (36.51)	543 (46.37)	50 (18.87)	178 (36.11)	50 (25.13)	94 (37.30)
	>$75,000	47 (10.66)	291 (24.85)	5 (1.89)	22 (4.06)	10 (5.02)	31 (12.30)
	Missing	87 (19.73)	337 (28.78)	66 (23.77)	89 (18.05)	33 (16.58)	27 (10.71)
Insurance	No	107 (24.26)	79 (6.74)	4 (1.5)	2 (0.41)	31 (15.58)	46 (18.25)
	Yes	296 (67.12)	1071 (91.46)	232 (87.54)	487 (98.78)	161 (80.90)	204 (80.95)
	Missing	38 (8.62)	21 (1.79)	29 (10.94)	4 (0.81)	7 (3.52)	2 (0.79)
Living area based on Census Division	East North Central	65 (14.74)	201 (17.16)	45 (16.98)	78 (15.82)	37 (18.59)	30 (11.90)
	East South Central	28 (6.35)	67 (5.72)	30 (11.32)	29 (5.88)	17 (8.54)	15 (5.95)
	Middle Atlantic	66 (14.97)	164 (14.00)	25 (9.43)	73 (14.81)	26 (13.06)	38 (15.08)
	Mountain	44 (9.97)	67 (5.72)	20 (7.54)	37 (7.50)	16 (8.04)	25 (9.92)
	New England	21 (4.76)	76 (6.49)	3 (1.13)	23 (4.66)	7 (3.52)	8 (3.17)
	Pacific	51 (11.56)	171 (14.60)	36 (13.58)	55 (11.16)	27 (13.57)	31 (12.30)
	South Atlantic	88 (19.95)	224 (19.13)	45 (16.98)	113 (22.92)	37 (18.59)	57 (22.62)
	West North Central	32 (7.26)	90 (7.69)	30 (11.32)	37 (7.50)	19 (9.55)	21 (8.33)
	West South Central	46 (10.43)	111 (9.49)	31 (11.7)	48 (9.74)	13 (6.53)	27 (10.71)

Table 2-1 Distribution of socio-demographic, health related, behavioral, and knowledge related characteristics

Independent Variables		Women's intentions and attitudes toward mammography				Thinking about mammogram in non-adherent women	
		Age < 65 (N=1612)		Age ≥ 65 (N=758)			
		$\gamma_1^* = 0$ (N=441)	$\gamma_1 = 1$ (N=1171)	$\gamma_2^{**} = 0$ (N=265)	$\gamma_2 = 1$ (N=493)	$\tau^\dagger = 0$ (N=199)	$\tau = 1$ (N=252)
	Missing	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Health-related characteristics							
Body mass index (BMI)	Underweight	8 (1.81)	18 (1.54)	10 (3.77)	8 (1.63)	9 (4.52)	5 (1.98)
	Normal	148 (33.56)	475 (40.56)	96 (36.23)	184 (37.32)	86 (43.22)	84 (33.33)
	Overweight	112 (25.40)	321 (27.41)	68 (25.66)	190 (38.54)	51 (25.63)	70 (27.78)
	Obese	114 (25.85)	306 (26.13)	50 (18.87)	97 (1.97)	36 (18.09)	78 (30.95)
	Missing	59 (13.37)	51 (4.35)	41 (15.47)	14 (2.84)	17 (8.54)	15 (5.95)
	Less than twice a year	85 (19.27)	197 (16.82)	29 (10.94)	60 (12.17)	86 (43.22)	107 (42.46)
	2 to 4 times a year	135 (30.61)	519 (44.32)	120 (45.28)	214 (43.41)	59 (29.65)	88 (34.92)
	5 or more a year	93 (21.09)	405 (34.59)	74 (27.92)	201 (40.77)	49 (24.62)	56 (22.22)
	Missing	128 (29.02)	50 (4.23)	42 (15.85)	18 (3.65)	5 (2.51)	1 (0.40)
Family history of cancer	No	145 (32.88)	324 (27.67)	97 (36.60)	148 (30.02)	75 (37.69)	77 (30.56)
	Yes	294 (66.67)	844 (72.07)	166 (62.64)	342 (69.37)	123 (61.81)	175 (69.44)
	Missing	2 (0.45)	3 (0.26)	2 (0.75)	3 (0.61)	1 (0.50)	0 (0)
Psychological distress	No	355 (80.50)	1084 (92.57)	221 (83.40)	471 (95.54)	179 (89.95)	226 (89.68)
	Yes	46 (10.43)	69 (5.89)	11 (4.15)	11 (2.23)	10 (5.02)	24 (9.52)
	Missing	40 (9.07)	15 (1.54)	33 (12.45)	11 (2.23)	10 (5.02)	2 (0.79)
Have been advised to have a mammography	No	133 (30.16)	111 (9.48)	120 (45.28)	49 (9.94)	111 (55.78)	96 (38.09)
	Yes	155 (35.15)	1012 (86.42)	92 (34.72)	429 (87.02)	33 (16.58)	94 (37.30)
	Missing	153 (34.69)	48 (4.09)	53 (20)	15 (3.04)	55 (27.64)	62 (24.60)
Having had Pap smear	No	8 (1.81)	6 (0.51)	26 (9.81)	10 (2.03)	23 (11.56)	8 (3.17)
	Yes	383 (86.84)	1127 (96.24)	209 (78.87)	472 (95.74)	165 (82.91)	234 (92.86)
	Missing	50 (11.34)	38 (3.24)	30 (11.32)	11 (2.23)	11 (5.53)	10 (3.97)
Behavioral characteristics							
Trust cancer info from doctors	Not at all	10 (2.27)	4 (0.34)	9 (3.34)	4 (0.81)	9 (4.52)	4 (1.59)
	A little	35 (7.94)	46 (3.93)	15 (5.66)	20 (4.06)	19 (9.55)	14 (5.56)
	Some	162 (36.73)	392 (33.48)	89 (33.58)	131 (26.57)	73 (36.68)	88 (34.92)
	A lot	224 (50.79)	716 (61.14)	148 (55.85)	330 (66.94)	95 (47.74)	143 (56.75)
	Missing	20 (4.53)	13 (1.11)	4 (1.51)	8 (1.62)	3 (1.51)	3 (1.19)
Looking for cancer information	No	217 (49.21)	443 (37.83)	182 (68.68)	280 (56.79)	144 (72.36)	123 (48.81)
	Yes	224 (50.79)	726 (62.00)	82 (30.94)	210 (42.60)	54 (27.14)	129 (51.19)
	Missing	0 (0)	2 (0.17)	1 (0.38)	3 (0.61)	1 (0.50)	0 (0)

Table 2-1 Distribution of socio-demographic, health related, behavioral, and knowledge related characteristics

Independent Variables		Women's intentions and attitudes toward mammography				Thinking about mammogram in non-adherent women	
		Age < 65 (N=1612)		Age ≥ 65 (N=758)			
		$\gamma_1^* = 0$ (N=441)	$\gamma_1 = 1$ (N=1171)	$\gamma_2^{**} = 0$ (N=265)	$\gamma_2 = 1$ (N=493)	$\tau^{\dagger} = 0$ (N=199)	$\tau = 1$ (N=252)
Perception of chance of getting breast cancer	Low	209 (47.39)	541 (46.20)	157 (59.24)	265 (53.75)	134 (67.34)	130 (51.59)
	Moderate	129 (29.25)	414 (35.35)	58 (21.89)	156 (31.64)	43 (21.61)	82 (32.54)
	High	49 (11.11)	184 (15.71)	11 (4.15)	44 (8.92)	6 (3.01)	31 (12.30)
	Missing	54 (12.24)	32 (2.73)	39 (14.72)	28 (5.68)	16 (8.04)	9 (3.57)
Fruit and vegetable consumption	Per day	193 (43.76)	615 (52.52)	137 (51.70)	306 (62.07)	32 (16.08)	43 (17.06)
	Per week	113 (25.62)	333 (28.44)	59 (22.26)	110 (22.31)	52 (26.13)	74 (29.36)
	Per Month	76 (17.23)	199 (16.92)	37 (13.96)	70 (14.20)	99 (49.75)	122 (48.41)
	Missing	59 (13.38)	24 (2.05)	32 (12.07)	7 (1.42)	16 (8.04)	13 (5.16)
Lifetime number of smoked cigarettes	<100	187 (42.4)	642 (54.82)	152 (57.36)	283 (57.40)	112 (56.28)	120 (47.62)
	≥ 100	221 (50.11)	526 (44.92)	87 (32.83)	207 (41.99)	86 (43.22)	131 (51.98)
	Missing	33 (7.48)	3 (0.26)	29 (9.81)	3 (0.61)	1 (0.50)	1 (0.40)
Exercise in the past month	No	137 (31.06)	840 (71.73)	133 (50.19)	334 (67.75)	77 (38.69)	90 (35.71)
	Yes	267 (60.54)	321 (27.41)	105 (39.62)	154 (31.24)	116 (58.69)	159 (63.09)
	Missing	37 (8.39)	10 (0.85)	27 (10.19)	5 (1.01)	6 (3.01)	3 (1.19)
Knowledge of Breast Cancer/Mammography							
Age at which women should start getting mammography	40-50	227 (51.47)	755 (64.47)	87 (32.83)	280 (56.79)	89 (44.72)	132 (52.38)
	Others	140 (31.75)	371 (31.68)	83 (31.32)	144 (29.21)	58 (29.15)	95 (37.70)
	Missing	74 (16.78)	45 (3.84)	95 (35.84)	69 (13.99)	52 (26.13)	25 (9.92)
Frequency of getting mammograms	Every 1 to 2 years	257 (58.28)	909 (77.63)	136 (51.32)	384 (77.89)	106 (53.27)	159 (63.09)
	Others	118 (26.76)	249 (21.26)	68 (25.66)	97 (19.67)	57 (28.64)	79 (31.35)
	Missing	66 (14.97)	13 (1.11)	61 (23.02)	12 (2.43)	36 (18.09)	14 (5.55)

For the health-related characteristics, as the number of visits to a health care provider increases, women are more likely to adhere to screening mammography (OR=2.358 for women with more than 5 visits to doctor per year versus women with less than two visits per year). Being advised to have mammogram is the strongest factor in this category OR= 5.298 and 10.711 for women younger than 65 and women 65 and older, respectively compared to women who are not advised. Women who have had a prior Pap smear are more adherent than those who have not (OR=3.203 and OR=3.809 for women younger than 65 and women 65 and older, respectively).

For the behavioral characteristics, women who trust the cancer information from their doctor "a little", "some" and "a lot" are 6.727, 13.194, and 14.824 times more likely to adhere compared to women who do not trust their doctor at all. Women who have an increased risk perception of breast cancer are more adherent. For women younger than 65, the corresponding ORs for women with "moderate" and "high" risk perception are 1.519 and 1.541, respectively. The corresponding values for women older than 65 are 1.645 and 3.036, respectively. Women younger than 65, who have smoked at least 100 cigarettes in their lifetimes are less likely to be adherent (OR=0.687).

For the knowledge related class of factors, women whose answer to the appropriate interval between two subsequent mammograms is other than every one or two years are less likely to adhere. The corresponding odds ratios are 0.812 and 0.636 for women younger than 65 and women 65 and older, respectively.

Table 2-2 Results for design-based multiple logistic regression on mammography adherence comparing women younger than 65 and women aged 65 and older

Independent Variables		Age < 65		Age ≥ 65	
		Odds ratio	P-value	Odds ratio	P-value
Socio-demographic characteristics					
Race	Black	-		-	
	White	0.874	0.606	1.329	0.536
	Other	0.548	0.077	0.425	0.148
Marital status	Married	-		-	
	Single/Divorced/Separated	0.685	0.025	0.499	0.012
Education	Some high school or less	-		-	
	High school Graduate	1.750	0.089	1.460	0.310
	Some college/ College graduate	1.457	0.231	1.507	0.320
Income	≤ $25,000	-		-	
	>$25,000 and ≤$75,000	1.324	0.186	1.341	0.044
	>$75,000	2.526	0.004	1.992	0.609
Insurance	No	-		-	
	Yes	2.959	<0.0001	2.771	0.130
Health-related characteristics					
Number of visits to health provider a year	Less than twice a year	-		-	
	2 to 4 times a year	1.618	0.019	1.016	0.969
	5 or more a year	2.358	0.001	1.610	0.313
Being advised to have a mammography	No	-		-	
	Yes	5.298	<0.0001	10.711	<0.0001
Having had Pap smear	No	-		-	
	Yes	3.203	0.0004	3.809	0.001
Behavioral characteristics					
Trust cancer info from doctors	Not at all	-		NA	
	A little	6.727	0.052		
	Some	13.194	0.007		
	A lot	14.824	0.004		
Looking for cancer information	No	-		-	
	Yes	1.134	0.457	0.883	0.484
Perception of chance of getting breast cancer	Low	-		-	
	Moderate	1.519	0.011	1.645	0.054
	High	1.541	0.079	3.036	0.043
Lifetime number of smoked cigarettes	<100	-		NA	
	≥ 100	0.687	0.047		
Exercise in the past month	Yes	NA		1.134	0.623
	No			-	
Knowledge of Breast Cancer/Mammography					
Age at which women should start getting mammography	40-50	-		-	
	Others	0.921	0.605	0.692	0.101
Frequency of getting mammograms	Every 1 to 2 years	-		-	
	Others	0.812	0.043	0.636	0.048

"-" represents the reference group.

2.3.2 Thinking about Getting a Mammogram in Non-Adherent Women

In the bivariate analysis, age, marital status, income, and insurance are among the variables, which proved to be significant in the socio-demographic class of factors. In the health related class of factors, being advised to have a mammography, having had Pap smear are statistically significant in bivariate analyses. Looking for cancer information, risk perception of breast cancer, lifetime number of smoked cigarettes, and exercise in the past month are significant factors in the bivariate analyses in the behavioral characteristics class.

However, in the multiple logistic regression analyses (Table 2-3), the only significant factor is age. As women's age increased, they reported thinking about getting mammograms less often. Women ages 60 to 69 years, and older than 70 years are 0.377 and 0.112 times less likely to think about getting a mammogram compared to women aged 42 to 49 years.

Table 2-3 Results for design-based multiple logistic regression on attitudes toward mammography in non-adherent women

Independent Variables		Non-Adherent Women (N=510)	
		Odds ratio (OR)	P-value
Socio-demographic characteristics			
Age	42-49	-	
	55-59	0.784	0.572
	60-69	0.377	0.015
	70+	0.112	<0.0001
Marital status	Married	-	
	Single/Divorced/Separated	0.965	0.917
Income	≤ $25,000	-	
	>$25,000 and ≤$75,000	1.579	0.193
	>$75,000	1.738	0.363
Insurance	No	-	
	Yes	1.277	0.496
Health-related characteristics			
Being advised to have a mammography	No	-	
	Yes	1.789	0.083
Having had Pap smear	No	-	
	Yes	1.156	0.727
Behavioral characteristics			
Looking for cancer information	No	-	
	Yes	1.676	0.081
Perception of chance of getting breast cancer	Low	-	
	Moderate	1.412	0.207
	High	2.049	0.337
Lifetime number of smoked cigarettes	<100	-	
	≥ 100	0.817	0.541
Exercise in the past month	No	-	
	Yes	1.061	0.848
Knowledge of Breast Cancer/Mammography			
Age at which women should start getting mammography	40-50	-	
	Others	1.102	0.737
Frequency of getting mammograms	Every 1 to 2 years	-	
	Others	0.967	0.925

2.3.3 Evaluation of Regression Models

Table 2-4 and 2-5 present the classification tables for the first response variable and the second response variable, respectively. Sensitivity, specificity and accuracy of the proposed logistic regression models are presented in the tables. Note that the numbers presented in the tables are weighted according to the replication jackknife weights in the data to be representative of the national population. For example, the number 4,727,850 in Table 2-4 is the number of non-adherent women younger than 65 who are correctly classified as non-adherent by the design based logistic regression.

Table 2-4 Classification table of logistic regression model on mammography adherence*

			Actual Data		
			Non-Adherent ($\gamma_1(\gamma_2)$=0)	Adherent ($\gamma_1(\gamma_2) = 1$)	
Logistic Regression Model Result	Younger than 65	$\hat{\gamma}_1 < 0.5$	4,727,850	7,816,121	Accuracy=77.7
		$\hat{\gamma}_1 \geq 0.5$	2,234,154	30,291,566	
	Measures		Specificity=67.91	Sensitivity=79.49	
	65 and older	$\hat{\gamma}_2 < 0.5$	3,914,902	2,164,589	Accuracy=80.50
		$\hat{\gamma}_2 \geq 0.5$	1,086,976	9,505,231	
	Measures		Specificity=78.27	Sensitivity=81.45	

* $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are the adherence likelihood calculated from the logistic regression models for women younger than 65 and women 65 and older, respectively.

Table 2-5 Classification table of logistic regression model on attitudes toward mammography in non-adherent women

		Actual Data		
		Do not think about getting a mammogram (τ=0)	Think about getting a mammogram (τ=1)	
Logistic Regression Model Result	$\hat{\tau} < 0.5$	3,427,551	2,071,371	Accuracy=74.43
	$\hat{\tau} \geq 0.5$	1,290,733	6,360,130	
	Measure	Specificity=72.64	Sensitivity=75.43	

* $\hat{\tau}$ is the likelihood of being concerned about getting mammogram in non-adherent women calculated from the logistic regression model.

As the results show, for the adherence to mammography response variable the proposed design-based logistic regression models correctly classify 77.7% of women younger than 65 and 80.5% of women 65 and older. For the thinking about getting a mammogram response variable 73.43% of women are correctly classified.

2.4 Conclusion

In this study, we evaluated the association between women's adherence to mammography screening and four classes of factors including various socio-demographic, health-related, behavioral characteristics and knowledge of breast cancer/mammography factors using the 2003 HINTS data.

In the literature, definitions of adherence to mammography screening vary widely across studies. Zapka et al. [14] and Wu et al. [15] defined adherence based on the ACS guideline, "had a mammogram in the last year". Maxwell et al. [16], Murabito et al. [17], Tejeda et al. [18], Schonberg et al., [19] and Vyas, et al. [20], considered "had a mammogram in the past two years" as a measure of adherence, which is consistent with the USPSTF guideline. Wu et al. [15], Maxwell et al. [16], Tejeda et al. [18] and Meissner et al. [21] considered "ever had a mammogram" to measure women's decisions in mammography participation. However, this measure cannot be interpreted as the adherence to mammography guidelines since it does not incorporate a specific time interval. Other time intervals are also considered in the literature. For example, Harrison et al. [22] defined "had any mammogram in the preceding 5-year period" as their response variable. Many studies defined adherence based on "ever had" or "the most recent" mammogram rather than considering whether the women returned for repeat mammography or not. To bridge this gap, Carney et al. [23] defined the response variable as "returning for a screening mammogram within 24 months of the initial exam" and identified factors that had

significant effect, such as health insurance coverage, first degree relatives with breast cancer, and knowledge about breast cancer. In other studies, Allen et al. [24. 25] explored the relationship between mammography use and social network characteristics while considering “receipt of at least two mammograms, the most recent of which was within the past 2 years with a maximum interval of 2 years between screenings” as the adherence measure. Rakowski et al. [26] considered “two exams on schedule, based on an every-other-year interval” in their study. A systematic literature review of the studies on the mammography use and the associated factors published from 1988 to 2004 was completed by Schueler et al. [27]. In summary, they found that the strongest predictors of mammography use were past screening behavior (clinical breast examination and Pap test), having access to a physician, and having a physician-recommend mammography. These results are in line with our findings as the design based logistic regression result show that “having been advised to get mammogram”, “have had Pap test” are the two most important factors. We also found that “perception of chance of getting breast cancer”, “income”, “insurance” and “knowledge about mammography” are significant factors.

Most of these studies used data at the local, regional or state level, making national generalizations difficult. In addition, these studies did not consider if a woman underwent a mammogram because she intended to do so, or for a diagnostic purpose (e.g. after the woman had felt a lump in her breast), or because her doctor prescribed the mammogram. In addition, for the studies that used longitudinal data (sequential mammograms), it is difficult to keep track of women who died, moved, etc. This can potentially cause bias in the data analysis, and as a result, these measures may not truly represent women’s adherence to mammography guidelines.

This study differs from prior studies [14-27] in that adherence is defined based on women’s intention and plans of obtaining mammograms rather than their mammography history behavior.

Results from our work will likely helps decision makers to differentiate between women who are concerned about mammography and have regular plans for getting mammograms from those who are not concerned. The results of this analysis can help policy makers identify non-adherent populations. We also examine the relationship between aforementioned four classes of factors with women's attitudes toward getting mammograms for women with poor mammography history. By characterizing the factors associated with attitudes toward mammography in the non-adherent population, policy makers can differentiate between the women who are concerned about receiving a mammogram but may have some barriers versus those who are not concerned at all.

As for the limitations of our study, it is based on self-reported data and may yield biased results if a woman did not report the exact time of her last mammogram. We also used the 2003 HINTS data set since it is the only HINTS data that provides information on women's intention toward mammography.

Adherence to screening mammography is lower among minorities [28]. Based on the 2010 National Health Interview Survey, 69.4% of American Indian/Alaska native women and 64.1% of Asian women had a mammogram between 2008 and 2010, while participation rates among white and black women were about 72.8% and 73.2%, respectively [29]. Our results also imply that the percent of women intending to have regular screening mammography is lower in races other than white and black. However, we do not have enough evidence to conclude that these differences are statistically significant. The results reveal that for both age groups, married women are more likely to comply with the mammography guidelines. For women younger than 65, financial aspects such as insurance and income are also significant factors the policy makers should consider for improving adherence.

Physicians' involvement in achieving mammography adherence is important. However, many physicians do not make referrals at the recommended intervals, even though they may endorse the guidelines [30]. Our results show that prior advice to have a mammogram by the woman' physician is the most significant factor in health related class of factors. This suggests that to enhance women's adherence, doctors should recommend that their patients have regular screening mammograms. Mammography referrals are more frequent for women who have access to the health care system (i.e. women with regular physician and health insurance), but less frequent among vulnerable women, (i.e., older women with lower educational attainment or lower annual family income [31]). It would be worth investigating the factors associated with low mammogram referral rates by physicians. Having history of Pap smear was found to be significant for both age groups. This is in line with a previous study by Augustson et al. [32], in which they characterize the association between clinical breast exam (CBE), Pap smear, fecal occult blood testing (FOBT) adherence, and mammography adherence. Therefore, low compliance with other cancer screenings can help identify women in need of additional interventions to improve mammography adherence. Although the number of visits to a health provider is significant for women younger than 65, this factor is not significant for older women. This is interesting as older women, regardless of being adherent or not, are expected to visit their doctors more frequently than younger women because of aging associated diseases.

In the behavioral characteristics class, "lifetime number of smoked cigarettes" is a significant factor for younger women but not for older women. Trusting doctor's information about cancer is proved to be significant for younger women, however, we do not have enough evidence to say that it is also significant for older women. Previous studies differ about whether fear motivated or inhibits precautionary behaviors. McCaul et al. tested different predictors in the context of fear

of breast cancer and breast cancer screening and showed that greater fear was related to higher levels of screening intentions and behaviors [33]. Our data analysis confirms this finding and shows that risk perception of breast cancer is a significant predictor on intention of getting regular mammograms for both age groups.

For the last class of covariates, as expected, knowledge about frequency of getting mammograms is significant. Thus, increasing women's knowledge about breast cancer and mammography recommendations is also likely to significantly influence women's adherence.

For the second response variable (thinking about getting a mammogram in non-adherent women), we found little difference between women who are and are not concerned about getting mammograms. For this response variable, younger women are more likely to think about obtaining a mammogram. Women who have been advised to get a mammogram before and those who are looking for cancer information may be more concerned about getting mammograms, although these findings are not statistically significant.

This study enables decision makers to identify barriers women may face to obtain a screening mammogram and also allows for the identification of characteristics of non-adherent women who are not concerned about getting a mammogram. Our findings suggest the most significant factors influencing women's compliance are being advised to have a mammogram and insurance. The former has been applied on a systematic level in some European countries (e.g. the UK, Sweden, Norway), which have organized population-based screening programs where women are invited through a personal letter for a free mammogram every two or three years [34]. A recent analysis of the European Network for Information on Cancer (EUNICE) data for years 2005, 2006 and/or 2007 (10 national and 16 regional programs for women aged 50-69 in 18

European countries) reported that in 13 of the 26 programs the participation rate exceeded the European Union benchmark of 70%, and 9 programs achieved a participation level >75% [35]. According to the CDC, 66.6% and 67.1% of U.S. women 40 years and over received a mammogram in the past two years based on data from 2005 and 2008, respectively [8]. Although a direct comparison of screening mammography participation rates between European countries and the U.S. is difficult due to difference in recommendations (i.e. age at which to begin screening, interval between mammograms), our results support the benefit of future policies that systematically advise and provide resources to women to receive a mammogram, as is already done in other countries. Several studies also identified significant factors associated with attendance in population-based mammography programs in European countries (the UK, France and Sweden) [36-41]. Some of these factors such as income, visiting healthcare providers, having had Pap smear and perceived risk of breast cancer are also found significant in our analysis. As in most European countries mammography is free, none of these studies reported insurance as an influencing factor for getting a mammogram. However, as our results show, lack of insurance is still a barrier in the U.S., especially for women younger than 65.

In summary, sending reminders and having insurance were shown to be the most influential factors in screening mammography. Other significant factors include marital status, number of visits to health provider a year, having had Pap smear, risk perception of breast cancer, and knowledge of breast cancer/mammography. These findings can help in designing programs aimed at improving screening rates and provides policy makers with data to allow for interventions to remove these barriers and make the targeted population more concerned about getting breast cancer screenings.

References

1. World Cancer Research Fund International,

   http://www.wcrf.org/cancer_statistics/world_cancer_statistics.php, accessed March 2014.

2. American Cancer Society, http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-key-statistics, accessed January, 2015.

3. Center for Disease Control and Prevention, The National Breast and Cervical Cancer Early Detection Program, Available at: http://www.cdc.gov/cancer/nbccedp. Accessed March 201

4. Gøtzsche P.C., Nielsen M. (2009). Screening for breast cancer with mammography. Cochrane Database Syst Rev (4):CD001877

5. American Cancer Society web site. Available at: http://www.cancer.org/healthy/findcancerearly/cancerscreeningguidelines/american-cancer-society-guidelines-for-the-early-detection-of-cancer. Accessed May 25, 2013.

6. The US Preventive Services Task Forces. Available at: http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm. Accessed May 25, 2013

7. Isaacs C, Peshkin BN, Schwartz M, DeMarco TA, Main D, Lerman C (2002) Breast and ovarian cancer screening practices in healthy women with a strong family history of breast or ovarian cancer. Breast Cancer Research and Treatment 71 (2): 103-112

8. Center for Disease Control and Prevention. Available at: http://www.cdc.gov/nchs/fastats/mamogram.htm. Accessed June 27, 2013

9. Becker MH, Maiman LA. (1975) Socio-behavioral determinants of compliance with health and medical care recommendations. Med Care 13:10–24

10. Nelson D, Kreps G, Hesse B, Robert C, Gordon W, Neeraj A, Barbara R, K. Vish V, Neil W, Sara A. (2004) The Health Information National Trends Survey (HINTS): development, design, and dissemination. Journal of Health Communication 9:443- 460

11. Andridg RR, Little RJA (2010) A Review of Hot Deck Imputation for Survey Non-response. International Statistical Review 78(1): 40–64

12. Roberts G, Rao JNK, Kumar S. (1987) Logistic regression analysis of sample survey data. Biometrika 74(1): 1-12

13. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression, 2nd edn. John Wiley. &. Sons., NewYork

14. Zapka JG, Hosmer D, Costanza ME (1992) Changes in Mammography Use: Economic, Need, and Service Factors. American Journal of Public Health 82(10):1345-1351
15. Wu H, Zhu K, Jatoi I, Shah M, Shriver CD, Potter J. (2007) Factors associated with the incompliance with mammogram screening among individuals with a family history of breast cancer or ovarian cancer, Breast Cancer Research and Treatment 101(3): 317-324

16. Maxwell CJ, Bancej CM, Snider J (2001) Predictors of mammography use among Canadian women aged 50–69: findings from the 1996/97 National Population Health Survey. Canadian Medical Association Journal 164 (3):329-334

17. Murabito JM, Evans JC, Larson MG, Kreger BE, Splansky GL, Freund KM, Mark A. M, Wilson PW (2001) Family Breast Cancer History and Mammography Framingham Offspring Study. American Journal of Epidemiology 154(10):916-923

18. Tejeda S., Thompson B., Coronado GD, Martin DP, Heagerty PJ (2009) Predisposing and Enabling Factors Associated with Mammography Use Among Hispanic and Non-Hispanic White Women Living in a Rural Area. Journal of Rural Health 25(1):85-92

19. Schonberg MA, McCarthy EP, York M, Davis RB, Marcantonio ER (2007) Factors influencing elderly women's mammography screening decisions: implications for counseling. BMC Geriatrics 7(1),26

20. Vyas A, Madhavan S, LeMasters T, et al (2011) Factors Influencing Adherence to Mammography Screening Guidelines in Appalachian Women Participating in a Mobile Mammography Program. Journal of Community Health 37 (3):632-646

21. Meissner HI, Breen N, Taubman ML, Vernon SW, Graubard BI (2007) Which women aren't getting mammograms and why? (United States). Cancer Causes Control 18(1):61–70

22. Harrison RV, Janz NK, Wolfe RA, Tedeschi PJ, Huang X, McMahon LF (2003) 5-Year Mammography Rates and Associated Factors for Older Women. Cancer 97(5):1147-1155

23. Carney PA, Harwood BG, Weiss JE, et al. (2002) Factors Associated with Interval Adherence to Mammography Screening in a Population-Based Sample of New Hampshire Women. Cancer 95(2): 219-227

24. Allen JD, Stoddard AM, Sorensen G. (2008) Do Social Network Characteristics Predict Mammography Screening Practices? Health Education & Behavior 35(6):763-776

25. Allen JD, Sorensen G, Stoddard AM (1999) The Relationship between Social Network Characteristics and Breast Cancer Screening Practices among Employed Women. Ann Behav Med 21(3):193-200

26. Rakowski W, Meissner H, Vernon SW, et al (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 Epidemiol Biomarkers Prev 15(11):2093-2101

27. Schueler KM, Chu PW, Smith-Bindman R (2008) Factors Associated with Mammography Utilization: A Systematic Quantitative Review of the Literature. Journal of Women's Health 17 (9):1477-1498

28. Phillips CE, Rothstein JD, Beaver K, et al (2011) Patient navigation to increase mammography screening among inner city women. Journal of general internal medicine 26(2): 123-129

29. Center for Disease Control and Prevention, National Health Interview Survey, United States, 2010. Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6103a1.htm. Accessed July 3, 2013

30. Blechman EA, Brownell KD (1998) Behavioural medicine and women: a comprehensive handbook. Guilford Press, New York

31. O'Malley MS, Earp JA, Hawley ST, et al. (2001) The association of race/ethnicity, socioeconomic status, and physician recommendation for mammography: who gets the message about breast cancer screening? American Journal of Public Health 91(1): 49

32. Augustson EM, Vadaparampil ST, Paltoo DN, et al. (2003) Association between CBE, FOBT, and Pap smear adherence and mammography adherence among older low-income women. Preventive medicine 36(6): 734-739

33. McCaul KD, Reid PA, Rathge RW, et al. (1996) Does concern about breast cancer inhibit or promote breast cancer screening? Basic and Applied Social Psychology 18(2):183-194

34. Perry N, Broeders M, deWolf C, et al. (2006) European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis. [Printed in Belgium]. European Communities 2006. http://ec.europa.eu/health/ph_projects/2002/cancer/fp_cancer_2002_ext_guid_01.pdf. Accessed December 6, 2013.

35. Giordano L, von Karsa L, Tomatis M, et al. 2012. Mammographic screening programmes in Europe: organization, coverage and participation. Journal of Medical Screening 19(suppl 1): 72-82.

36. Duport, N. 2012. Characteristics of women using organized or opportunistic breast cancer screening in France. Analysis of the 2006 French Health, Health Care and Insurance Survey.Revue d'epidemiologie et de sante publique 60(6): 421-430.

37. Zackrisson S, Lindström M, Moghaddassi M, et al. 2007. Social predictors of non-attendance in an urban mammographic screening programme: a multilevel analysis. Scandinavian journal of public health 35(5): 548-554.

38. Lagerlund M, Maxwell AE, Bastani R. 2002. Sociodemographic predictors of non-attendance at invitational mammography screening–a population-based register study (Sweden). Cancer Causes & Control 13(1): 73-82.
39. Lagerlund M, Widmark C, Lambe M, et al. 2001. Rationales for attending or not attending mammography screening-a focus group study among women in Sweden. European journal of cancer prevention 10(5): 429-442.

40. Aro AR, de Koning HJ, Absetz P, et al. 1999. Psychosocial predictors of first attendance for organised mammography screening. Journal of Medical Screening 6(2): 82-88.

41. Lidbrink, E., Frisell, J., Brandberg, Y., Rosendahl, I., & Rutqvist, L. E. (1995). Nonattendance in the Stockholm mammography screening trial: relative mortality and reasons for nonattendance. Breast cancer research and treatment, 35(3), 267-275.

3 Evaluation of breast cancer mammography screening policies considering adherence behavior

3.1 Introduction

Mammography screening recommendations have been the subject of significant debate in recent years. As discussed in Chapter 2, there are varying screening guidelines from different organizations about when to start and end mammograms and how frequently a woman should undergo mammography screenings. For example, the American Cancer Society (ACS), the department of Health and Human Services (HHS), the American Medical Association (AMA), and the American College of Radiology (ACR) recommend screening mammography every year, beginning at age 40. In 2009, the U.S. Preventive Services Task Force (USPSTF) issued revised screening mammography guidelines, which resulted in a significant controversy. According to the USPSTF guidelines, screening mammograms should be done every two years between age 50 and 75 for women at average risk of breast cancer [1].

Although mammography is known to be one of the most effective methods of detecting breast cancer, there are many concerns on its adverse effects on women's quality of life. Unreliability of mammography, i.e., false negative and false positive results and overdiagnosis are among the deficiencies of mammography [2]. A prospective cohort study of 7 mammography registries of the Breast Cancer Surveillance Consortium (BCSC) showed that the adjusted sensitivity increased with age from 68.6% in women ages 40 to 44 to 83.3% in women ages 80 to 89 [3]. Similarly, the specificity rate increased from 88.2% for women aged 40 through 44 years to 93.4% in women older than 75 [4]. Moreover, about one-third of all aggressive cancers are diagnosed in the interval between successive annual mammograms [5]. These cancers are known

as interval cancers and are defined as cancers detected within 12 months after a negative mammogram [6]. Some of the interval cancers are present at the time of mammography screening (false-negatives) while others grow rapidly in the interval between a mammogram and detection [7]. Generally, interval cancers grow rapidly and are frequently diagnosed at advanced stages [8]. In addition, studies have shown that receiving mammograms increases a woman's chance of developing breast cancer due to exposure to radiation [2, 9-10]. Each radiation-absorbed dose (rad) of exposure increases breast cancer risk by 1 percent [2]. Moreover, Bleyer et al. also found that up to a third of diagnosed breast cancers are overdiagnosed cases and do not need treatment [11].

Various studies have been conducted to evaluate and compare different screening policies or to identify optimal policies. Kirch et al. [12] developed an inspection strategy for the detection of an age-dependent disease with the objective of minimizing detection delay. To illustrate their methodology, they developed optimal schedules for breast cancer examinations. Shwartz [13] developed a mathematical model to evaluate life expectancy gain associated with different screening policies under different assumptions about the rate of disease progression, the characteristics of the screening technique, and the frequency of screenings. Parmigiani [14] presented a continuous time non-Markovian stochastic process model of disease progression to analyze which age groups and what part of the population should undergo breast cancer screening, while minimizing the total expected loss or risk including financial costs, side effects, wasted time, stress due to false-negative test results, etc. Zelen [15] formulated a screening scheduling problem to maximize a weighted utility function in a continuous time setting. In Zelen's model all the parameters were assumed to be stationary except for the incidence rate. Baker [16] used a mathematical parametric model to assess different screening policies in terms

of minimizing the cost of cancer plus the cost of carrying out any screening. Baker's model was developed based on a small data set used to estimate the model parameters, and compares a small set of cost optimal policies under different sets of constraints. Hanin et al. [17] developed a model to study screening strategies based on tumor size at detection. In their model they combined several processes such as tumor latency, tumor growth and tumor detection to estimate the tumor size at detection. As a measure of the screening efficiency, they proposed the difference between the expected tumor sizes at detection with and without screening, when both spontaneous and screening-based detections are in place. They concluded that in the case of exponential tumor growth, the optimal screening schedules with a fixed number of exams are uniform or very close to uniform (same interval between subsequent screening tests). Maillart et al. [18] formulated a partially observable Markov chain model to evaluate a wide range of dynamic mammography screening policies as well as current practices. They compared different policies in terms of the resulting lifetime breast cancer mortality risk and the expected number of mammograms women should undergo for each policy and generated a frontier of efficient policies. In their formulation, they considered age-based dynamics of breast cancer (i.e., increasing incidence, decreasing aggression). They also incorporated the imperfect nature of the screening outcomes and dynamics of test result accuracy (increasing sensitivity and specificity rates with age). Ayer et al. [19] developed a finite-horizon partially observable Markov decision process (POMDP) model to determine the optimal personalized mammography screening strategy based on personal risk characteristics of women such as their prior screening history. Ahern et al. [20] developed a mathematical model to optimize cancer screening schedules, taking into account the trade-off between the benefits and costs of screenings in their proposed utility function. They considered two different optimization frameworks: optimize the number of

screening examinations with equal screening intervals between exams but without a pre-fixed total cost; and optimize the ages at which screening should be given for a fixed total cost and prove that the optimal solution exists under each of the two frameworks. In another study, Hanin et al. [21] developed a model to construct optimal schedules of cancer screening that maximize the probability that by the time of primary tumor detection, it has not yet metastasized. They applied their model to mammographic screening for breast cancer. Their model includes all main stages of cancer progression: cancer latency, primary tumor growth, its detection and metastatic spread.

However, none of the above mentioned studies have taken patient behavior into consideration. They all assume that patients adhere to the guidelines perfectly and undergo the prescribed screening mammograms. This is also true for the current screening recommendations which are based on the assumption of 100% adherence to the guidelines. However, as discussed in Chapter 2, not all women have the same attitudes toward breast cancer screening. There are a limited number of studies taking individual adherence behavior into consideration. Brailsford et al. [22] used a three-phase discrete event simulation to model breast cancer and screening policies incorporating women's behavioral factors in their model. They assigned behavioral attributes to each simulated woman to control her compliance with the prescribed mammograms in their model. They compared a limited number of screening policies, including the current UK policy, in terms of the number of screen detected cancers, and life yeas saved. In another study, Ayer et al. [23] analyzed the role of behavioral heterogeneity in women's adherence on optimal mammography screening recommendations.

In this research, incorporating women's adherence behavior to mammography recommendations, a mathematical framework is proposed to evaluate and compare various screening policies in

terms of QALYs and the lifetime mortality risk of breast cancer. Our study is different from the two studies discussed above in the approach of incorporating patient adherence to the screening tests. In contrast to earlier studies, we allow uncertainty in a patient's adherence probabilities in the model. Adherence is also a function of the length of the interval between two subsequent screenings. Moreover, assuming a patient's adherence behavior is related to and can be estimated from her perception/planning toward mammography, not only can we evaluate and compare screening policies for a specific patient knowing her personal characteristics, but we can identify efficient policies for the general population as well. In addition, we incorporate in our model the possibility of interval cancer detection, and the increased risk of developing breast cancer due to exposure to X-ray radiation during a mammography test.

The remainder of this chapter is organized as follows. In Section 3.2, the proposed model is presented. In Section 3.3, model inputs and parameters estimation, and computational results are presented. Finally, Section 3.4 summarizes the findings and discusses future work.

3.2 Model Formulation

A randomized discrete-time finite-horizon partially observable Markov chain model is developed to evaluate various mammography screening policies in terms of the expected QALYs and lifetime breast cancer mortality risk. This problem is formulated as a partially observable Markov chain because the women's true health states are not outwardly observable due to the imperfect nature of mammography screening tests.

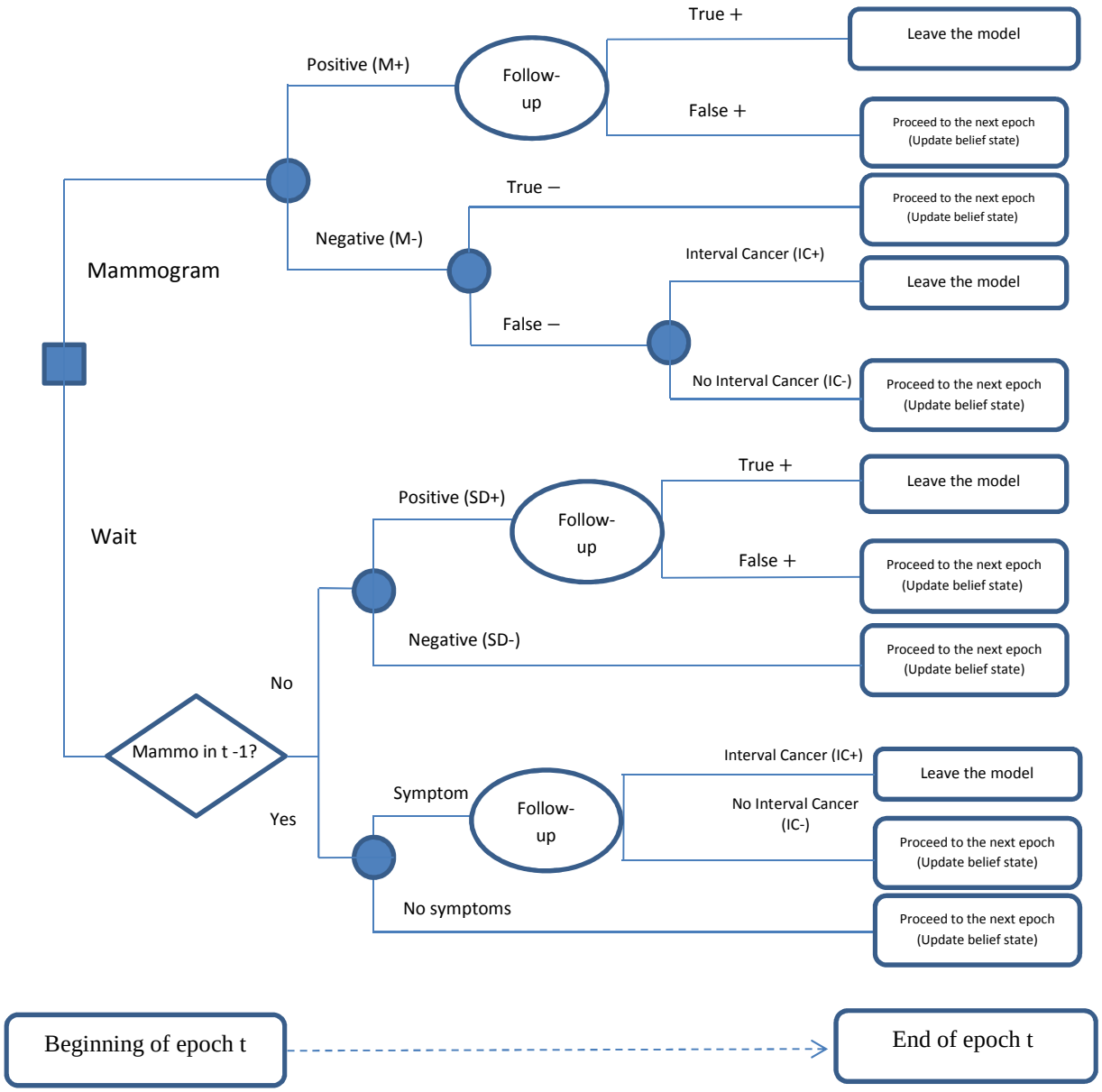


Figure 3-1 One period sample path of the breast cancer detection process

The model takes into account two methods of detection: screening mammography and self-detection (SD), i.e., breast self-examination. Interval cancer detection is considered separately from self-detection, depending on the time interval between the detection and the last

mammogram. The detailed mammography screening process is presented in Figure 1 and is described as follows.

At epoch *t* if a mammogram is recommended, a woman may undergo the recommended mammogram with probability $q(t)$ or skip it with probability $1 - q(t)$. If she undergoes the prescribed mammogram, she may receive a positive or negative result. After a positive result, since mammography is not perfect and has a relatively low sensitivity (the probability of receiving true positive result), a breast biopsy test is usually conducted to check the suspicious area found on a mammogram and see whether the result is true positive or not. Biopsy is assumed to be perfect since it has a relatively high sensitivity and specificity (probability of receiving true negative) rates [24]. If the biopsy result is negative suggesting a false positive mammogram, the patient proceeds to the next epoch. However, if the biopsy confirms the mammographic findings, cancer is detected and the patient leaves the model. We do not model the treatment explicitly. Instead, we assume that when the patient is diagnosed with breast cancer, regardless of the method of detection, she receives a terminal (lump sum) reward and leaves the model. After a negative result, the woman proceeds to the next epoch. However, it is possible that after a negative result the cancer becomes symptomatic and be detected (interval cancer detection in this case).

If the woman skips the recommended mammogram or the recommended action at time *t* is to wait, she may or may not develop some symptoms suggesting there is cancer present. If she develops symptoms, depending on the action in the previous epoch we have either a self-detection (if the previous action is a wait), or an interval cancer (if the previous action is a mammogram). Note that the interval between two decision epochs in the process is six month.

Therefore, if the action in a previous epoch was a mammogram and symptoms occur after the current epoch, we have an interval cancer and not a self-detection. We assume that when the woman feels a lump in her breast (symptom) she would go for a mammogram, and if the result of the mammogram is positive, she would have a biopsy test to confirm that the cancer is present. If both the mammography and biopsy tests are positive, similar to the previous case the patient leaves the model. However, if the follow-up tests (i.e., mammogram or biopsy) are negative, the patient proceeds to the next epoch.

For consistency, we refer to the woman as the "patient," irrespective of her health condition.

Below is the list of notation used in the problem formulation.

t Decision epochs, $t = 0, \dots, T$. We consider the earliest and latest age a policy may recommend a patient undergo a mammogram to be 40 and 100, respectively. The start age of 40 is considered in order to be consistent with the current mammography screening policies (e.g. ACS). In addition, we assume the process ends at age 100 because of the negligible probability of surviving after this age according to the U.S. life tables [25]. Mammography decisions are made every six months. Patients enter the process at age 40 ($t = 0$) and, if cancer is not detected, stay in the process up to age 100 ($t = T = 120$), regardless of the policy under study.

s_t Underlying health state a patient occupies at epoch t and $s_t \in \mathrm{S} = \{0, 1, 2, 3, 4\}$. The five states considered in the model are cancer free (state 0), early breast cancer (state 1), advanced breast cancer (state 2), death from breast cancer (state 3), and death from other causes (state 4). We define cancer states similarly to Maillart et al.

[18]. States 0 through 2 are partially observable due to the imperfect nature of mammography tests. However, states 3 and 4 are fully observable.

a_t Action prescribed at time *t*, and $a_t \in A = \{\mathrm{W}, \mathrm{M}\}$. W and M represent "Wait" and "Mammogram", respectively.

π_t Belief state distribution, where $\pi_t = [\pi_t(0), \pi_t(1), \pi_t(2)]$ and represents the occupancy distribution of partially observable states at which the patient is believed to be, i.e., $\pi_t(s)$ represents the probability that a patient is believed to be in state *s*.

$\mathrm{P}_t(\acute{s}|s, a)$ Underlying transition probability, i.e., the probability of being in state $\acute{s}$ at time $t + 1$, when the patient is in state *s* and takes action *a* at time *t*. Note that the transition probability is a function of action a since undergoing a mammogram increases the patient's chance of developing radiation-induced breast cancer.

o^{a_t} Observation when a patient takes action a_t. If $a_t = \mathrm{M}$, the observation space is $O^{\mathrm{M}} = \{\mathrm{M}+, \mathrm{M}-\}$, where $\mathrm{M}+$ and, $\mathrm{M}-$ represent a positive mammogram, and a negative mammogram, respectively. In the case of receiving a negative mammogram ($\mathrm{M}-$), the patient may show symptoms within one year, which results in interval cancer detection. In this case the observation space is $O^{\mathrm{IC}} = \{\mathrm{IC}+, \mathrm{IC}-\}$, where $\mathrm{IC}+$ and $\mathrm{IC}-$ represent interval cancer and no interval cancer, respectively. If $a_t = \mathrm{W}$ and the patient did not have a mammogram test within the past year, then $O^{\mathrm{W}} = \{\mathrm{SD}+, \mathrm{SD}-\}$, where $\mathrm{SD}+$ and $\mathrm{SD}-$ represent self-detection and no self-detection, respectively.

$Q_t(\mathrm{o}|\mathrm{s}, a)$ Observation probabilities, i.e., the probability of observing *o* at time *t* when the patient is in state *s* and action *a* is taken at time *t*. Note that for the case that observation is $\mathrm{IC}+$ or $\mathrm{IC}-$, for simplicity, we use the notation $Q_t(\mathrm{o}|\mathrm{s})$ since

	interval cancers can happen both after a wait and a mammogram action. However, in "wait" cases we should have had a mammogram in the previous epoch (at time $t-1$ or six month prior to time t).
$r_t(s\|a,o)$	Immediate reward that a woman receives when she is at state *s*, takes action *a* and receives observation *o* at time *t*.
$R_{1,t}(s)$	The total expected reward a patient receives when she is in one of the cancer states (*s*= 1 or 2) and her cancer is detected through screening mammography.
$R_{2,t}(s)$	The total expected reward when a patient is in cancer state *s* in the case of interval cancer or when the cancer is diagnosed via self-detection. Different total expected reward values are considered for screen-detected and symptomatic cancers since previous studies have shown a systematic survival benefit for screen-detected cancers comparing to symptomatic cancers [26]. However, the expected reward for interval cancers and self-detected cancers is assumed to be the same as it has been shown that there is no statistical difference in survival between interval cancers and clinical cancers [27].
$\rho_{1,t}(s)$	Probability of eventually dying from breast cancer when the cancer is diagnosed through screening mammography and the patient is in state *s* (=1 or 2) at time *t*.
$\rho_{2,t}(s)$	Probability of eventually dying from breast cancer in the case of interval cancer or when the cancer is diagnosed through self-detection.
γ_t	Probability of dying from causes other than breast cancer. Note that γ_t is independent of the state the patient is in and $\gamma_t = P_t(4 \mid s,a)$.
$I(t)$	Indicator function representing whether the previous action at epoch $t-1$ was a mammogram ($I(t)=1$) or not ($I(t)=0$).

d Mammography screening policy or a sequence of prescribed actions d=$(d_0, d_1, \dots, d_T)$ where d_t represents the action at time *t* under policy *d*.

D_t^{M} Disutility of mammography at time *t*.

D_t^{B} Disutility of biopsy at time *t*.

$q(t)$ Probability that a woman undergoes the prescribed mammogram at epoch *t*. *q(t)* varies in the interval of $[q_1(t)\ \ q_2(t)]$, where $q_1(t)$ and $q_2(t)$ are bounds of the confidence interval of the adherence probabilities and can be estimated based on the woman's various characteristics. Estimation for $q_1(t)$ and $q_2(t)$ is discussed in Section 3.1.

The state transition diagram of the underlying Markov chain representing the disease natural history is presented in Figure 3-2. Note that cancer death only occurs when the cancer progresses to an advanced stage [18].

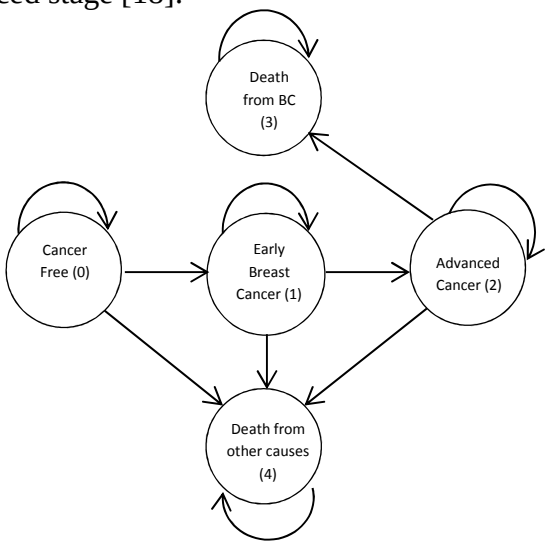


Figure 3-2 Transitions between different health states

The policies are evaluated and compared in terms of two outcome measures: total expected quality adjusted life years (QALYs), and lifetime mortality risk of breast cancer.

The value functions of policy *d* in our model follows a randomized Markov model in which the actions are chosen under uncertainty, i.e., at age *t* the woman may take action M (mammogram) with probability $q(t)$ or choose to skip a mammogram (W) with probability $1-q(t)$. Given that $q_1(t)$ and $q_2(t)$ are the lower and upper confidence bounds for $q(t)$, the value function of policy *d* at time *t* is defined as below:

$$V_t^d(\pi) = \begin{cases} \dfrac{1}{q_2(t)-q_1(t)} \int_{q_1(t)}^{q_2(t)} [q(t)\cdot V_t^{\mathrm{M}}(\pi) + (1-q(t))\cdot V_t^{\mathrm{W}}(\pi)]dq(t) & \text{if } d_t = \mathrm{M} \\ V_t^{\mathrm{W}}(\pi) & \text{if } d_t = \mathrm{W} \end{cases} \qquad 3\text{-}1$$

$$= \begin{cases} \dfrac{q_1(t)+q_2(t)}{2}\left(V_t^{\mathrm{M}}(\pi) - V_t^{\mathrm{W}}(\pi)\right) + V_t^{\mathrm{W}}(\pi) & \text{if } d_t = \mathrm{M}, \\ V_t^{\mathrm{W}}(\pi) & \text{if } d_t = \mathrm{W}, \end{cases}$$

where $V_t^a(\pi)$ is the value function at time *t* when the action *a* is taken and the occupancy distribution is π. Note that although π is a function of time as introduced in the notation section, we drop this time index in all equations for the clarity of the notation. Equation 3-1 calculates the average value of function $q(t)\cdot V_t^{\mathrm{M}}(\pi) + (1-q(t))\cdot V_t^{\mathrm{W}}(\pi)$ over the interval $[q_1(t), q_2(t)]$.

3.2.1 Total Expected Quality Adjusted Life Years (QALYs)

Suppose $\varphi_t^a(\pi)$ represents the value function ($V_t^a(\pi)$ in Equation 3-1) for our first outcome measure: the total expected quality adjusted life years (QALYs) the patient can attain when the current belief state distribution is π and action *a* is taken. If at time *t* the screening policy recommends a mammogram, then

$$\varphi_t^M(\pi) = \pi(0)\cdot Q_t(\mathrm{M}-\mid 0,\mathrm{M})\left[r_t(0,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-) + \sum_{s'=0}^{4} P_t(s'\mid 0,\mathrm{M})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-))\right]$$
$$+ \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{M}-\mid s,\mathrm{M})\left\{\begin{array}{l} Q_t(\mathrm{IC}+\mid s)\cdot R_{2,t}(s) + Q_t(\mathrm{IC}-\mid s) \\ \cdot\left[r_t(s,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-) + \sum_{s'=0}^{4} P_t(s'\mid s,\mathrm{M})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-))\right]\end{array}\right\} \qquad 3\text{-}2$$
$$+\pi(0)\cdot Q_t(\mathrm{M}+\mid 0,\mathrm{M})\left[r_t(0,\mathrm{M},\mathrm{M}+) + \sum_{s'=0}^{4} P_t(s'\mid 0,\mathrm{M})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}+))\right]$$
$$+ \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{M}+\mid s,\mathrm{M}).R_{1,t}(s),$$

where $\vartheta(\pi,a,o)$ is the updated occupancy distribution at time t+1 when the current occupancy distribution (at time *t*) is π and action *a* is chosen and the result is observation *o*. $\vartheta(\pi,a,o)$ is calculated using Equations 3-3a and 3-3b.

$$\xi_{(\pi,a,o)}(s) = \begin{cases} \dfrac{\pi(s)\cdot Q_t(M-\mid s,M)\cdot Q_t(IC-\mid s)}{\sum_{j=0}^{2}\pi(j)\cdot Q_t(M-\mid j,M)\cdot Q_t(IC-\mid j)} & \text{if } o_t = \mathrm{M}-\&\ \mathrm{IC}- \\ \dfrac{\pi(s)\cdot Q_t(SD-\mid s,W)}{\sum_{j=0}^{2}\pi(j)\cdot Q_t(SD-\mid j,W)} & \text{if } a_{t-1}=\mathrm{W},\ a_t=\mathrm{W}, o_t=\mathrm{SD}- \\ \dfrac{\pi(s)\cdot Q_t(IC-\mid s)}{\sum_{j=0}^{2}\pi(j)\cdot Q_t(IC-\mid j)} & \text{if } a_{t-1}=\mathrm{M},\ a_t=\mathrm{W},\ o_t=\mathrm{IC}- \\ 1 & \text{if } s_t=0,\ o_t=\mathrm{M}+\text{ or SD}+ \end{cases} \qquad 3\text{-}3a$$

and,

$$\vartheta_{(\pi,a,o)}(s') = \sum_{s=0}^{2}\xi_{(\pi,a,o)}(s)\cdot P_t(s'\mid s,a). \qquad 3\text{-}3b$$

Note that $\xi_{(\pi,a,o)}(s)$ is the updated occupancy upon observing o at the current epoch and $\vartheta_{(\pi,a,o)}(s')$ is the updated occupancy at the beginning of the next epoch. Time indices are dropped

for clarity. The first line in Equation 3-3a represents the case when the patient undergoes a prescribed mammogram, receives a negative result and does not develop any symptoms within six months of the mammogram test. The second and third lines both represent the case when the recommended action is to wait in the current epoch. However, the second line is when the action at epoch $t-1$ is a wait and the third line is when the recommended action at $t-1$ is a mammogram. These two cases need to be considered separately since the third case represents an interval cancer rather than self-detection because of the negative results in the previous mammogram at epoch $t-1$ (within 12 month of a negative mammogram). The fourth line represents a false positive result. In such cases the follow-up tests reveal that the patient is healthy. In the case when the result is true positive (M+, SD+ and IC+ and $s=1$ or 2), the woman receives a lump sum reward and leaves the model. Therefore, no occupancy distribution update is needed.

The logic of Equation 3-2 is as follows. When a patient undergoes a mammogram, she either receives a positive or a negative result. If the patient is in the cancer free state and she receives a negative result, an immediate reward is calculated and the patient will transit to time $t+1$. If the patient is in cancer states but she receives a negative result, then she may develop some symptoms with probability $Q_t(\text{IC+}\,|\,s)$ before the next epoch, which results in cancer detection. In this case, interval cancer is identified and the patient receives the lump sum reward and leaves the model. We assume that interval cancer can only occur in cancer states since even the fastest growing cancers cannot grow from a single cell to a symptomatic size within six months [28], which is the interval between two subsequent epochs in our model. If the patient does not develop any symptoms, she stays in the model and transits to time $t+1$. If the woman is in the cancer free state and she receives a positive result, her true health state will be determined via a

follow-up biopsy and she remains in the model. However, if she is in any of the two cancer states and receives a positive result, she will receive a lump sum reward and the process is terminated.

If the screening policy does not recommend a mammogram at time *t*, then the expected QALYs is calculated as

$$
\begin{aligned}
\varphi_{\mathrm{t}}^{\mathrm{W}}(\pi) = {} & \pi(0)\cdot Q_t(\mathrm{SD}-\mid 0,\mathrm{W})\left[r_t(0,\mathrm{W},\mathrm{SD}-) + \sum_{s'=0}^{4} P_t(s'|0,\mathrm{W})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}-)) \right] \\
& + (1-I(t))\cdot \left\{ \begin{array}{l} \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{SD}-\mid s,\mathrm{W}) \left[r_t(s,\mathrm{W},\mathrm{SD}-) + \sum_{s'=0}^{4} P_t(s'\mid s,\mathrm{W})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}-)) \right] \\ + \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{SD}+\mid s,\mathrm{W})\cdot R_{2,t}(s) \end{array} \right\} \\
& + \pi(0)\cdot Q_t(\mathrm{SD}+\mid 0,\mathrm{W})\left[r_t(0,\mathrm{W},\mathrm{SD}+) + \sum_{s'=0}^{4} P_t(s'\mid 0,\mathrm{W})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}+)) \right] \\
& + I(t)\cdot \left\{ \begin{array}{l} \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{IC}-\mid s) \left[r_t(s,\mathrm{W},\mathrm{IC}-) + \sum_{s'=0}^{4} P_t(s'\mid s,\mathrm{W})\cdot \varphi_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{IC}-)) \right] \\ + \sum_{s=1,2} \pi(s)\cdot Q_t(\mathrm{IC}+\mid s)\cdot R_{2,t}(s) \end{array} \right\}.
\end{aligned}
\qquad \text{3-4}
$$

The logic of Equation 3-4 is as follows. When the action at time *t* is to wait ($a_t = \mathrm{W}$), as discussed above, two separate cases need to be considered: 1) the prescribed action at time $t-1$ is to wait, and 2) the prescribed action at time $t-1$ is a mammogram. In Case (1), if the woman waits or does not get a mammogram at time *t*, there are two possible outcomes: whether symptoms are developed or not during the time interval between epoch *t* and $t+1$. In the case when no symptom develops, the patient receives an immediate reward $r_t(\mathrm{s},\mathrm{W},\mathrm{SD}-)$, remains in the model and her belief state distribution is updated using Equation 3-3a and 3-3b. However, if some symptoms are developed (with probability $Q_t(\mathrm{SD}+\mid s,\mathrm{W})$) and the patient is in the cancer free state, a follow-up test reveals her true health state, then her occupancy distribution is

updated and she remains in the model. If she is in a cancer state *s* (s=1, 2) and some symptoms are shown, the follow-up tests will detect the presence of breast cancer. In this case, the patient receives a lump sum reward $R_{2,t}(s)$ and leaves the model. In Case (2), given that the patient had a negative mammogram at epoch $t-1$, she may develop some symptoms (with probability $Q_t(\text{IC}+\mid s)$) that results in interval cancer detection in which case she receives a lump sum reward $R_{2,t}(s)$ and leaves the model. It is also possible that no symptom occurs, which means the patient receives an immediate reward $r_t(s, \text{W}, \text{IC}-)$, her belief state gets updated and she proceeds to the next epoch.

Assuming $r_t(s, W, SD-)$ is determined, the immediate rewards in Equation 3-2 and 3-4 are calculated using the following relationship:

$$r_t(s,a,o)=\begin{cases} r_t(s,\text{W},\text{SD}-)-D_t^{\text{M}} & \text{if } a_t=\text{M} \text{ and } o_t=\text{M}- \text{ and IC}- \\ r_t(s,\text{W},\text{SD}-)-D_t^{\text{M}}-D_t^{\text{B}} & \text{if } \text{s}_t=0,\ a_t=\text{M} \text{ and } o_t=\text{M}+ \\ r_t(s,\text{W},\text{SD}-) & \text{if } a_t=\text{W} \text{ and } o_t=\text{IC}- \text{ or SD}- \\ r_t(s,\text{W},\text{SD}-)-D_t^{\text{M}}-Q_t(\text{M}+\mid s,\text{M})\cdot D_t^{\text{B}} & \text{if } \text{s}_t=0,\ a_t=\text{W} \text{ and } o_t=\text{SD}+. \end{cases} \quad 3\text{-}5$$

3.2.2 Lifetime Breast Cancer Mortality Risk

Suppose $\omega_t^a(\pi)$ represents the value function ($V_t^a(\pi)$) for the second outcome measure: the probability that the patient eventually dies from breast cancer when the current belief state is π, and action *a* is taken at time *t*. If the screening policy recommends a mammogram at time *t*, then

$$
\begin{aligned}
\omega_t^{\mathrm{M}}(\pi) = {} & \pi(0)\cdot Q_t(\mathrm{M}-\mid 0,\mathrm{M})\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-)) \\
& +\pi(1)\cdot Q_t(\mathrm{M}-\mid 1,\mathrm{M})[Q_t(\mathrm{IC}+\mid 1)\cdot\rho_{2,t}(1) \\
& +Q_t(\mathrm{IC}-\mid 1)\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-))] \\
& +\pi(2)\cdot Q_t(\mathrm{M}-\mid 2,\mathrm{M})\{Q_t(\mathrm{IC}+\mid 2)\cdot\rho_{2,t}(2) \\
& +Q_t(\mathrm{IC}-\mid 2)\cdot[P_t(3\mid 2,\mathrm{M})+(1-P_t(3\mid 2,\mathrm{M})-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-))]\} \\
& +\pi(0)\cdot Q_t(\mathrm{M}+\mid 0,\mathrm{M})\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}+)) \\
& +\sum_{s=1,2}\pi(s)\cdot Q_t(\mathrm{M}+\mid s,\mathrm{M})\cdot\rho_{1,t}(s).
\end{aligned}
\qquad 3\text{-}6
$$

If the patient does not get the prescribed mammogram at time *t* or the policy recommends the patient to wait, then her lifetime breast cancer mortality risk is

$$
\begin{aligned}
\omega_t^{\mathrm{W}}(\pi) = {} & \pi(0)\cdot Q_t(\mathrm{SD}-\mid 0,\mathrm{W})\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}-)) \\
& +(1-I(t))\cdot\left\{\begin{array}{l}\pi(1)\cdot Q_t(\mathrm{SD}-\mid 1,\mathrm{W})\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}-)) \\ +\pi(2)\cdot Q_t(\mathrm{SD}-\mid 2,\mathrm{W})[(1-P_t(3\mid 2,\mathrm{W})-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}-))+P(3\mid 2,\mathrm{W})] \\ +\sum_{s=1,2}\pi(s)\cdot Q_t(\mathrm{SD}+\mid s,\mathrm{W})\cdot\rho_{2,t}(s)\end{array}\right\} \\
& +\pi(0)\cdot Q_t(\mathrm{SD}+\mid 0,\mathrm{W})\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{SD}+)) \\
& +I(t)\cdot\left\{\begin{array}{l}\pi(1)\cdot Q_t(\mathrm{IC}-\mid 1)\cdot(1-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{IC}-)) \\ +\pi(2)\cdot Q_t(\mathrm{IC}-\mid 2)\cdot[P_t(3\mid 2,\mathrm{W})+(1-P_t(3\mid 2,\mathrm{W})-\gamma_t)\cdot\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{W},\mathrm{IC}-))] \\ +\sum_{s=1,2}\pi(s)\cdot Q_t(\mathrm{IC}+\mid s)\cdot\rho_{2,t}(s)\end{array}\right\},
\end{aligned}
\qquad 3\text{-}7
$$

where $\vartheta(\pi,a,o)$ is the updated occupancy distribution as presented in Equations 3-3a and 3-3b.

The logic behind the Equations 3-6 and 3-7 is as follows. If a patient undergoes the mammogram (Equation 3-6), she either receives a negative result or a positive one. If the mammography result is negative (with probability $Q_t(\mathrm{M}-\mid s,\mathrm{M})$) and the woman is in the cancer free state (*s*=0), she may survive to time $t+1$ with probability $1-\gamma_t$, and her risk of eventually dying from breast cancer is calculated by $\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-))$. If she is in the early stage cancer state (*s*=1) or

advanced cancer state (s=2), after receiving a negative result (false negative) she may develop some symptoms with probability $Q_t(\mathrm{IC}+\mid s)$, which results in her cancer detection by follow-up tests (mammogram and biopsy if needed). In this case, her lifetime breast cancer mortality risk would be $\rho_{2,t}(s)$ and she leaves the model. However, if she is in state $s = 1$ (or $s = 2$), and she does not develop any symptoms after a false negative result (with probability $Q_t(\mathrm{IC}-\mid s)$), she survives to time $t+1$ with probability $1-\gamma_t$ (or $1-P_t(3\mid 2,\mathrm{M},\mathrm{M}-)-\gamma_t$) and her lifetime breast cancer mortality risk at time $t+1$ is calculated by $\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}-\&\mathrm{IC}-))$. Note that if the patient is in the advanced cancer state ($s = 2$), she may die from breast cancer in the current period with probability $P_t(3\mid 2,\mathrm{M},\mathrm{M}-)$. If the mammography result is positive and the woman is in the cancer free state, her true health state will be determined via a follow-up biopsy and she survives to time $t+1$ with probability $1-\gamma_t$ and her value function is determined through $\omega_{t+1}^{d_{t+1}}(\vartheta(\pi,\mathrm{M},\mathrm{M}+))$. However, if she is in one of the cancer states, her probability of eventually dying from breast cancer would be $\rho_{1,t}(s)$.

In Equation 3-7, similar to Equation 3-4, two cases need to be distinguished, based on the action at time $t-1$, to account for the possibility of interval cancer. If the action at epoch $t-1$ is to wait then two outcomes in the current period are possible: developing symptoms ($Q_t(\mathrm{SD}+\mid s,\mathrm{W})$) and no symptoms ($Q_t(\mathrm{SD}-\mid s,\mathrm{W})$). If the patient does not develop any symptom when she is in the cancer free or early cancer state, she may survive to the next time period with probability $1-\gamma_t$. Then her lifetime breast cancer mortality risk is calculated using the recursive function at time $t+1$. However, if she is in the advanced cancer state, she will either die from breast cancer

before the next period with probability $P_t(3 \mid 2, \mathrm{W}, \mathrm{SD}-)$ or survive to the next period with probability $1 - P_t(3 \mid 2, \mathrm{W}, \mathrm{SD}-) - \gamma_t$ and her probability of eventually dying from breast cancer is calculated through $\omega_{t+1}^{d_{t+1}}(\vartheta(\pi, \mathrm{W}, \mathrm{SD}-))$. The patient might develop some symptoms with probability $Q_t(\mathrm{SD}+ \mid s, \mathrm{W})$ when she is in state *s*. If the patient is in the cancer free state and develops some symptoms, a follow-up mammogram (or biopsy) reveals her true health status, so her belief state occupancy distribution is updated, and the probability of her eventually dying from breast cancer is calculated by $\omega_{t+1}^{d_{t+1}}(\vartheta(\pi, \mathrm{W}, \mathrm{SD}+))$. If the patient is in cancer states (*s*=1 or 2) and develops symptoms, her lifetime breast cancer mortality risk would be $\rho_{2,t}(s)$. In this case she leaves the model. In the second case when the previous action was a mammogram, the patient may develop interval cancer in the current epoch or not. Similar to the first part, the corresponding lifetime breast cancer mortality risks for this case is calculated.

3.3 Numerical Studies

A wide range of routine screening policies and two-phase policies (with changing screening intervals) are evaluated in this section. Policies are defined as a quintuple (starting age, first screening interval, switching age, second screening interval, stopping age). For example, policy (40,1,50,2,80) represents a policy that recommends women start getting mammograms at age 40 and undergo the screening test every year up to age 50 and then undergo mammogram every 2 years up to age 80. Another example for routine screening policy would be (50,1,n,n,100) that recommends women undergo annual screenings from age 50 to 100. Table 3-1 provides polices considered in the analysis. The earliest screening start age considered is 40 since it is the earliest age among the current screening guidelines. We also include the USPSTF policy and no mammogram policy in our analysis. In total we evaluate 362 policies. Note that the numerical

examples presented in this section do not include six-month policies (i.e., screening every six month) because the data source for estimating patient adherence discussed in Section 3.1 does not include six-month policies. However, the six-month decision epoch in our model formulation allows better representation of disease natural progression. The readers can evaluate more policies of interest if they have data for estimating patient adherence behaviors. Different adherence cases are also presented as numerical examples in this section.

Table 3-1 Screening policies considered in the numerical analysis*

Policy Parameter	Range Evaluated
Start age	40,45,50,55
First screening interval	1,2,3
Switch age	50,60,70
Second screening interval	1,2,3
Stop age	80,85,90,95,100

**Two special policies "No mammography" and "USPSTF" policies are also included in the numerical analysis but are not shown in the table.*

3.3.1 Model Inputs

As discussed in Chapter 2, a patient's level of adherence can be approximated based on her personal attributes such as socio-demographic, health, behavioral and knowledge related characteristics [29-31]. Three different adherence cases are evaluated in this study: a specific woman, the general population, and a perfect adherence case. Adherence probabilities for these three cases are extracted using the 2003 Health Information National Trend Survey (HINTS) data and the developed design-based logistic regression in Chapter 2. The three adherence cases considered in this study are:

1) A white unemployed woman who is not married and does not have a family history of breast cancer and believes she has a low chance of developing breast cancer. This case represents very low adherence.

2) The general population. This case represents the average adherence rate of the U.S. population derived from the HINTS data, which is a nationally-representative stratified survey. For this case we adjust our logistic regression model for patients' ages.
3) Perfect adherence for which a patient adheres to a prescribed mammogram test with certainty.

Table 3-2 shows the confidence intervals of adherence probabilities for the general population (Case 2). These results are consistent with the CDC report for 2-year interval between 2008 and 2010 [32]. As expected, adherence probabilities are increasing in age until late 60's and then start to decrease. This decrease can be related to morbidity and aging associated diseases in older ages. Adherence probabilities for Case 1 are significantly lower than those of Case 2. For 1-year screening interval, the highest adherence probabilities appear in ages 65-69 (confidence interval (0.0581, 0.2226)), with lowest in ages 40-44 (confidence interval (0.0240, 0.0906)). Adherence is only slightly better for 2-year and 3-year screening intervals. Summary of these numbers is omitted for brevity.

Table 3-2 Confidence interval of adherence probabilities for the general population (Case 2)

Age	1-year Interval	2-year Interval	3-year Interval
40-44	(0.4018, 0.5007)	(0.5383, 0.6362)	(0.5539, 0.6512)
45-49	(0.4667, 0.5680)	(0.6746, 0.7656)	(0.6914, 0.7385)
50-54	(0.5440, 0.6490)	(0.7078, 0.7998)	(0.7204, 0.8109)
55-59	(0.5968, 0.7067)	(0.7466, 0.8398)	(0.7579, 0.8493)
60-64	(0.5723, 0.6945)	(0.6863, 0.7972)	(0.7043, 0.8127)
65-69	(0.6033, 0.7302)	(0.6982, 0.8138)	(0.6983, 0.8138)
70-74	(0.5629, 0.6972)	(0.6365, 0.7638)	(0.6365, 0.7638)
>75	(0.4006, 0.5012)	(0.5256, 0.6256)	(0.5391, 0.6386)

Initial belief states (risks of early and invasive cancers) for women aged 40 are estimated using the Gail model [33]. The Gail model estimates the invasive breast cancer risk for an individual based on some risk factors such as age, age at menarche, age at first live birth, number of first-

degree relatives with breast cancer, etc. To estimate the in situ cancer risk, Gail et al. [34] proposed using the incidence ratio between in situ and invasive cancers. For example, for average population, the initial risk of early and advanced cancer at age 40 is 0.3% and 0.6%, respectively. For the policies with start age greater than 40, the initial belief states are updated using the belief states for women at age 40, and Equations 3-3a and 3-3b.

Table 3-3 presents the data source for our numerical analyses. To specify the post-cancer life expectancies we use age-specific mortality rates for patients under cancer treatment from the SEER data [35] based on the method described in Arias et al. [25]. To adjust the life expectancies for symptomatic cancers, we use the relative hazard rates presented in Wishart et al. [26].

Table 3-3 Input data sources for model parameter estimation

Parameter	Reference
Transition probabilities	Maillart et al. [18], and Epstein et al. [2]
BSE sensitivity and specificity	Baxter [38]
Mammography sensitivity and specificity	Maillart et al. [18]
Lump sum rewards for screen detected cancers	SEER [35], and Arias et al. [25]
Lump sum rewards for symptomatic cancers	SEER [35], Arias et al. [25], and Wishart et al.[26]
Immediate rewards	Sonnenberg and Beck [36], and Stout et al. [37]
Initial risk of early and advanced breast cancer	Gail model [33], and Gail et al. [34]
Interval cancer rate	Croteau et al. [7]
Lifetime breast cancer mortality risk under treatment for screen detected cancers	Schairer et al. [39]
Lifetime breast cancer mortality risk under treatment for symptomatic cancers	Schairer et al. [39], and Wishart et al. [26]
Adherence Probabilities	Madadi et al. [29], and HINTS [40]

The expected immediate reward for the case the woman waits and no symptom is developed, i.e., $r_t(s, \mathrm{W}, \mathrm{SD}-)$ is determined using the half-cycle correction method [36] which assigns 0.5 (six months) life years if the patient is alive in the current decision epoch and 0.25 life years if the

patient dies in the current decision epoch. The immediate rewards for the other cases are calculated using Equation 3-5 and the numbers reported in Stout et al. [37].

3.3.2 Results

3.3.3 Analyses on QALYs

Table 3-4 presents the top five policies for the three different adherence cases in terms of QALYs. It also presents the associated lifetime breast cancer mortality risk for each policy.

Table 3-4 The five best policies for different adherence cases in terms of QALYs

Policy	QALYs	Lifetime BC mortality risk
Case 1- A Specific Woman		
(40,1,n,n,95)	37.4439	0.0263
(40,1,n,n,100) *	37.4439	0.0263
(40,1,n,n,90)	37.4429	0.0264
(40,1,n,n,85)	37.4405	0.0267
(40,1,n,n,80)	37.4371	0.0276
Case 2- General Population		
(40,1,n,n,95)	37.5307	0.0256
(40,1,n,n,100)	37.5307	0.0256
(40,1,n,n,90)	37.5306	0.0256
(40,1,n,n,85)	37.5268	0.0258
(45,1,n,n,95)	37.5240	0.0264
Case 3- Perfect Adherence		
(40,1,70,2,90)	37.5252	0.0262
(40,1,70,2,95)	37.5252	0.0262
(40,1,70,2,100)	37.5249	0.0262
(40,1,60,2,90)	37.5205	0.0275
(40,1,60,2,95)	37.5205	0.0274

**Policy in bold refers to the ACS policy.*

The best policies for both Case 1 and the general population include the ACS policy. However, Case 1 has lower QALYs and higher breast cancer mortality risk comparing to the general population due to lower adherence probabilities. The best policies for the perfectly adherent case, however, does not include the ACS policy and prescribe less frequent mammograms (2 year

interval) at older ages. For very frequent policies (e.g., the ACS), the imperfect adherence cases benefit more. Note that the USPSTF policy is not among the best policies for any of the adherence cases presented above.

Evaluation of the impact of adherence behavior on the outcomes of a specific screening policy shows that higher adherent women always have lower breast cancer mortality risk. However, this does not apply to the QALYs; for some policies Case 2 with imperfect adherence behavior have higher QALYs than perfect adherent women (Case 3). Let Δ_{ij} denote the percent of QALYs a woman in the adherence Case i $(= 1,2)$ loses due to her imperfect adherence to the prescribed screening policy j, i.e.,

$$\Delta_{ij} = \frac{QALYs_{3j} - QALYs_{ij}}{QALYs_{3j}}, \qquad i = 1,2 \tag{3-8}$$

where $QALYs_{3j}$ is the QALYs of policy j for Case 3, a perfectly adherent woman. Figure 3-3 shows 29 policies that have negative losses, implying that women with average adherence as the general population (Case 2) benefit more than perfectly adherent women. Note that for all the policies considered in this study, Case 1 always has the lowest QALYs and thus positive values of Δ_{ij} for all policies.

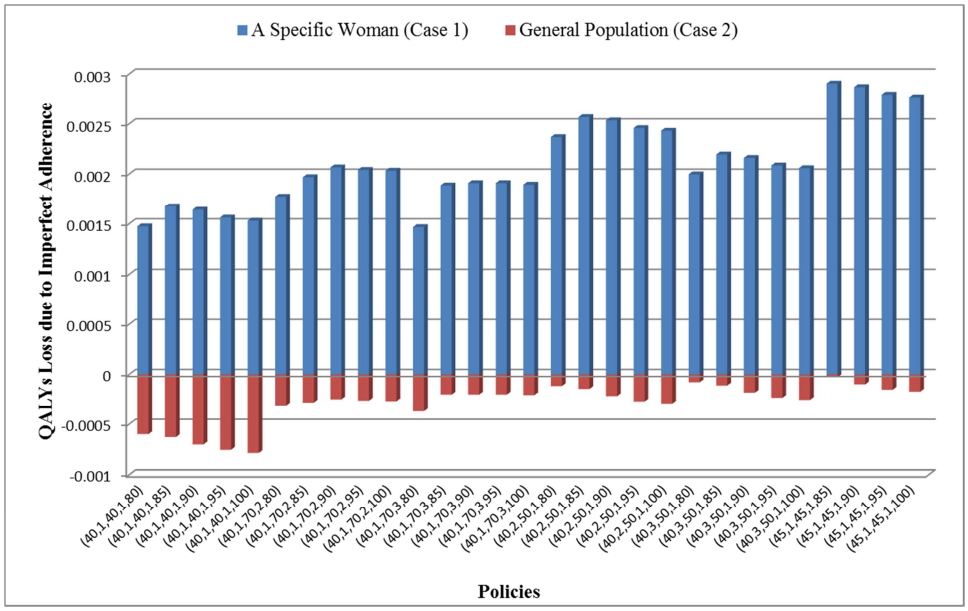


Figure 3-3 QALYs loss due to imperfect adherence

3.3.4 Analyses on Lifetime Breast Cancer Mortality Risk

The five best policies in terms of lifetime breast cancer mortality risk, interestingly, are the same. These policies are presented in Table 3-5. Note that the calculated breast cancer mortality presented here are very close to the CDC estimate of 1 in 37 women (~ 0.027) [41]. Unlike the first objective function (QALYs), the breast cancer lifetime mortality risk has a straightforward relation to the expected number of prescribed mammograms and the expected number of actual mammograms the woman undergoes. In other words, policies with more frequent mammograms decrease the risk of dying from breast cancer. The results show that among the adherence cases presented here, Case 3 (perfect adherent women) benefits the most and has the lowest risk of dying from breast cancer. This implies that, although exposure to mammography X-ray may increase women's risk of developing breast cancer, on average it is not strong enough to affect the benefits of mammography examination in deceasing breast cancer lifetime mortality risk.

This is in line with a previous study on estimating the risk of radiation-induced breast cancer following exposure of the breast to ionizing radiation occurring during mammography tests [9].

Table 3-5 The five best policies for different adherence cases in terms of lifetime breast cancer mortality risk

Policy	Lifetime breast cancer mortality risk	QALYs
Case 1- A Specific Woman		
(40,1,n,n,100)	0.0263	37.4437
(40,1,n,n,95)	0.0263	37.4439
(40,1,n,n,90)	0.0264	37.4429
(40,1,n,n,85)	0.0267	37.4405
(40,2,50,1,100)	0.0273	37.4193
Case 2- General Population		
(40,1,n,n,100)	0.0256	37.5304
(40,1,n,n,95)	0.0256	37.5307
(40,1,n,n,90)	0.0256	37.5306
(40,1,n,n,85)	0.0258	37.5268
(40,2,50,1,100)	0.0273	37.5214
Case 3- Perfect Adherence		
(40,1,n,n,100)	0.0242	37.5013
(40,1,n,n,95)	0.0242	37.5025
(40,1,n,n,90)	0.0242	37.5044
(40,1,n,n,85)	0.0244	37.5033
(45,1,n,n,100)	0.0249	37.5172

3.3.5 Efficient Frontier Policies and Policies in General Practice

Table 3-6 presents the results for policies in general practice and no mammography policy for the three adherence cases.

Table 3-6 QALYs and, mortality risk for screening policies in general practice

Policy	Case 1- A Specific Woman		Case 2-General Population		Case 3- Perfect Adherence	
	QALYs	**Mortality**	**QALYs**	**Mortality**	**QALYs**	**Mortality**
ACS*	37.4437	0.0263	37.5304	0.0256	37.5013	0.0242
USPSTF	37.2772	0.0362	37.3890	0.0348	37.4255	0.0339
Every 2 Years*	37.3585	0.0313	37.4756	0.0299	37.5107	0.0289
Every 3 Years*	37.3066	0.0347	37.4091	0.0334	37.4514	0.0327
No mammography	37.0142	0.0527	37.0142	0.0527	37.0142	0.0527

* Screening start age and end age for this policy is 40 and 100, respectively.

As the results in sections 3.2.1 and 3.2.2 suggest, there is a tradeoff between the two objective functions. Figure 3-4 provides all policies (excluding *"no mammogram"* policy) along with the lists of efficient policies for each adherence cases. The efficient policies, ACS and USPSTF policies are highlighted in Figure 3-4. The screening starting age for all efficient policies (except for the perfect adherence case) is 40. It is in contrary to the earlier study [42] in which the author argues that the benefits for women younger than 50 are less certain because of the lower incidence of the disease. However, since disease progression rate is higher and the cancer sojourn time is shorter in women younger than 50, benefits of getting screened before age 50 is justifiable [43].

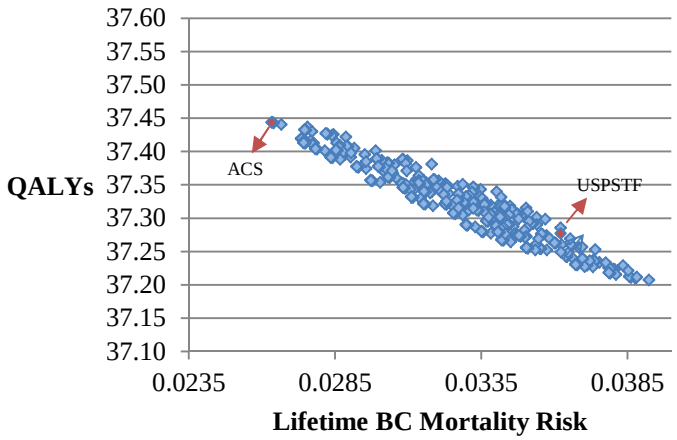
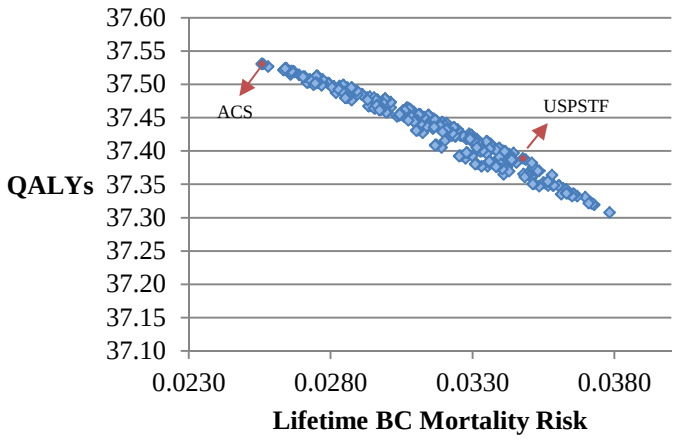
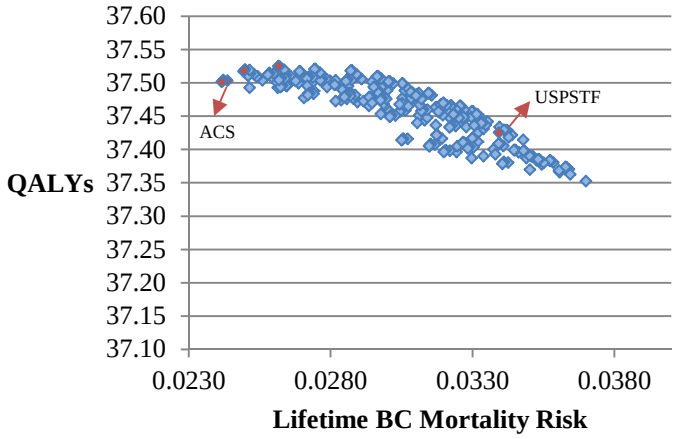
Policies	QALYs	BC Mortality Risk	
Case 1- A Specific Woman			
(40,1,n,n,95) **(40,1,n,n,100)**	37.4439 **37.4439**	0.0263 **0.0263**	
Case 2- General Population			
(40,1,n,n,95) **(40,1,n,n,100)**	37.5307 **37.5307**	0.0256 **0.0256**	
Case 3- Perfect Adherence			
(40,1,n,n,90) (45,1,n,n,90) (45,1,n,n,95) (40,1,70,2,90) (40,1,70,2,95)	37.5044 37.5202 37.5183 37.5252 37.5252	0.0242 0.0250 0.0249 0.0262 0.0262	

Figure 3-4 Efficient frontier policies for the three adherence cases

The current ACS policy (listed in bold in Figure 3-4) is only efficient for Case 1 and Case 2. For perfect adherence case, efficient frontier policies are similar to the ACS with one year interval between subsequent tests but smaller ending age. Note that, the USPSTF policy and other policies in general practice (for which the associated QALYs and breast cancer lifetime mortality risk are presented in Table 3-6) are not among the efficient policies for any of the adherence groups.

It can be seen that there are more efficient policies for the perfect adherence case compared to the other two cases. This occurs because of the tradeoff between the two objective functions. For the perfect adherence case, the first objective (QALYs) requires less frequent mammograms due to disutility of mammogram tests and perfect compliance of this case with the prescribed policies. On the other hand, the second objective (lifetime breast cancer mortality risk) requires more frequent mammography tests to have a lower risk of dying from breast cancer. Therefore, for this case the efficient policies are more spread out. For the other two adherence cases, since women are more likely to skip a prescribed mammogram, the first objective requires more frequent mammograms (every year) to compensate for the skipped mammograms. Thus, we have fewer efficient policies for these cases.

The tradeoff plots presented in Figure 3-4 are more spread for the perfect adherence case and gets denser and less spread as the adherence level decreases. This happens because as the adherence level decrease, policies converge to “no mammogram” policy with associated QALYs and lifetime breast cancer mortality risk of 37.0142 and 0.0527, respectively.

3.4 Conclusion

Current screening mammography guidelines assume that women's compliance with the recommendations is perfect, i.e., women undergo their mammograms as prescribed by their physicians/health providers. However, this is not the case in reality. Women skip mammograms for different reasons. This study investigates the effect of imperfect adherence for a wide range of screening mammography policies. Two different method of breast cancer detection are considered: mammography examination and self-detection. We also incorporate the possibility of detecting breast cancer within 12 months after a normal mammography screening (interval cancer), and the risk of developing radiation-induced breast cancer during mammography examination. A randomized partially observable Markov chain is proposed to evaluate and compare various screening policies and current practices in terms of two important health outcomes: quality adjusted life years (QALYs), and the lifetime mortality risk of breast cancer while incorporating women's adherence behavior to mammography guidelines.

Three different adherence cases are considered, including perfect adherence and two imperfect adherence cases. The adherence probabilities for imperfect adherence cases are estimated using a woman's characteristics (e.g., age, race, marital status, perception of breast cancer risk, etc.). Our results indicate that depending on the policy under study, there is one adherence group that benefits the most in terms of QALYs. As the policies get more intense (more frequent policies), the general population case obtains higher QALYs comparing to the perfect adherence case. However, Case 1 (a specific woman), which has a very low adherence level, always has lower QALYs than the other two cases. In terms of lifetime breast cancer mortality risk, higher adherence cases experience lower mortality risk for all policies. A set of efficient frontier

policies are extracted for different adherence cases. The ACS policy was among the efficient frontier policies for the two imperfect cases, but not for the perfect adherence case. Comparing efficient frontier policies for different adherence cases indicates that on average women with higher level of adherence experience higher QALYs, lower lifetime breast cancer mortality risk and longer interval between two subsequent prescribed mammograms.

The outcomes of this chapter along with the results of Chapter 2 can help physicians/health providers to tailor screening mammography recommendations based on their patient's estimated adherence likelihood. In other words, based on the patient characteristics and estimated adherence level, physicians can decide if they should shorten or lengthen the interval between two subsequent mammograms.

The direction for our future research is to incorporate the risk behavior of the decision maker and explore how different risk attitudes (risk averse, risk seeking) along with adherence behavior may affect a policy's efficiency. In this study, our purpose is to evaluate the impact of adherence on patient outcomes. Another possible direction is to optimally find a screening policy based on the patient estimated adherence behavior. It seems that screening adherence may be a function of a woman's screening history; i.e., a woman with better screening history are more likely to adhere to the current screening recommendation. Although our study considers the screening history in estimating adherence probabilities based on a survey data, it does not incorporate the dynamic feature of this factor. In addition, adherence strongly depends on a patient's experience in her previous screenings (e.g., false positives). We do not include these factors in our study due to lack of data with regard to these characteristics.

References

1. The US Preventive Services Task Forces. Available at: http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm. Accessed May 25, 2013

2. Epstein S.S., Bertell R., Seaman B. (2001). Dangers and unreliability of mammography: breast examination is a safe, effective, and practical alternative. Int J Health Serv.,31(3), 605-15

3. Breast Cancer Surveillance Consortium, http://breastscreening.cancer.gov/data/performance/screening/2009/perf_age.html, accessed November 30, 2013

4. Carney, P. A., Miglioretti, D. L., Yankaskas, B. C., Kerlikowske, K., Rosenberg, R., Rutter, C. M., ... & Ballard-Barbash, R. (2003). Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of internal medicine*, *138*(3), 168-175

5. Epstein S.S., Steinman D., LeVert S. (1998). The breast cancer prevention program. Ed. 2., Macmillan, new York

6. Burhenne, H. J., Burhenne, L. W., Goldberg, F., Hislop, T. G., Worth, A. J., Rebbeck, P. M., & Kan, L. (1994). Interval breast cancers in the Screening Mammography Program of British Columbia: analysis and classification. *AJR. American journal of roentgenology*, *162*(5), 1067-1071.

7. Hébert-Croteau N., Théberge I., Langlois A., Major D., Brisson J. (2005). Interval cancer in women following a normal initial mammogram in the Quebec breast cancer screening program (PQDCS)) in 1998-2000, Direction Systèmes De Soins Et Services

8. Caumo, F., Vecchiato, F., Strabbioli, M., Zorzi, M., Baracco, S., & Ciatto, S. (2010). Interval cancers in breast cancer screening: comparison of stage and biological characteristics with screen-detected cancers or incident cancers in the absence of screening. Tumori, 96(2), 198-201

9. Yaffe, M. J., & Mainprize, J. G. (2011). Risk of Radiation-induced Breast Cancer from Mammographic Screening 1. Radiology, 258(1), 98-105.

10. Gøtzsche, P. C., & Nielsen, M. (2009). Screening for breast cancer with mammography. Cochrane Database Syst Rev, 4(1).

11. Bleyer, A., & Welch, H. G. (2012). Effect of three decades of screening mammography on breast-cancer incidence. New England Journal of Medicine,367(21), 1998-2005

12. Kirch, R.L.A., Klein M. (1974). Surveillance schedules for medical examinations. Management Science 20(10), 1403–1409

13. Shwartz, M. (1978). A mathematical model used to analyze breast cancer screening strategies. Operation Research 26(6), 937–955

14. Parmigiani, G., (1993). On optimal screening ages, Journal of American Statistics Association, 88(422), 622–628

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

16. Baker, R., (1998). Use of a mathematical model to evaluate breast cancer screening policy, Health Care Management Science 1, 103–113

17. Hanin, L. G., Tsodikov, A. D., & Yakovlev, A. Y. (2001). Optimal schedules of cancer surveillance and tumor size at detection. Mathematical and computer modelling, 33(12), 1419-1430.

18. Maillart L. M., Ivy J. S., Ransom S., Diehl K., (2008). Assessing Dynamic Breast Cancer Screening Policies. Operations Research 56(6),1411–1427

19. Ayer T., Alagoz O., Stout N. K., (2011). A POMDP Approach to Personalize Mammography Screening Decisions. Operations Research, Articles in Advance, 1–16

20. Ahern C. H., Cheng Y., Shen Y., (2011). Risk-specific optimal cancer screening schedules: an application to breast cancer early detection. Stat Biosci. 3(2), 169–186

21. Hanin, L., & Pavlova, L. (2013). Optimal screening schedules for prevention of metastatic cancer. Statistics in medicine, 32(2), 206-219.

22. Brailsford S. C., Harper P. R. Sykes J., (2012). Incorporating Human Behavior in Simulation Models of Screening for Breast Cancer. European Journal of Operational Research 219, 491-507

23. T. Ayer, O. Alagoz, N.K. Stout, E.S. Burnside. Heterogeneity in Women's Adherence and Its Role on Optimal Breast Cancer Screening Policies. under review at Management Science

24. Parker, S. H., F. Burbank, R. J. Jackman, C. J. Aucreman, G. Cardenosa,T. M. Cink, J. L. Coscia. (1994). Percutaneous large-core breast biopsy: A multi-institutional study. Radiology 193(2), 359–364

25. Arias, E., Heron, M., & Tejada-Vera, B. (2012). National vital statistics reports. United States Centers for Disease Control and Prevention

26. Wishart, G. C., Greenberg, D. C., Britton, P. D., Chou, P., Brown, C. H., Purushotham, A. D., & Duffy, S. W. (2008). Screen-detected vs symptomatic breast cancer: is improved survival due to stage migration alone?. *British journal of cancer*, *98*(11), 1741-1744.

27. Frisell, J., Von Rosen, A., Wiege, M., Nilsson, B., & Goldman, S. (1992). Interval cancer and survival in a randomized breast cancer screening trial in Stockholm. *Breast cancer research and treatment*, *24*(1), 11-16.

28. Chen HH, Duffy SW, Tabar L: A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. Statistician 45:307-317, 1996

29. Madadi M., Zhang S., Yeary K. H. K., Henderson L. (2013). Analyzing Factors Associated with Women's Attitudes and Behaviors toward Screening Mammography Using Design-Based Logistic Regression. Breast Cancer Research and Treatment, 144(1), 193-204.

30. Murabito JM, Evans JC, Larson MG, Kreger BE, Splansky GL, Freund KM, Mark A. M, Wilson PW (2001). Family Breast Cancer History and Mammography Framingham Offspring Study. American Journal of Epidemiology 154(10), 916-923

31. Schonberg MA, McCarthy EP, York M, Davis RB, Marcantonio ER (2007) Factors influencing elderly women's mammography screening decisions: implications for counseling. BMC Geriatrics 7(1), 26

32. National Cancer Institute, http://www.cancer.gov/bcrisktool/, accessed December 2014.

33. Gail, M. H., J. P. Costantino, J. Bryant, R. Croyle, L. Freedman, K. Heizlsouer, V. Vogel. (1999). Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. J. Natl. Cancer Inst. 91(21) 1829–1846.

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

35. Sonnenberg, F. A., J. R. Beck. (1993). Markov models in medical decision making: A practical guide. Medical Decision Making 13(4), 322–338.

36. Stout, N. K., M. A. Rosenberg, A. Trentham-Dietz, M. A. Smith, S. M. Robinson, D. G. Fryback. (2006). Retrospective cost effectiveness analysis of screening mammography. J. Natl. Cancer Inst. 98(11), 774–782

37. Baxter, N. (2001). Should women be routinely taught breast selfexamination to screen for breast cancer? Canadian Medical Assoc. J. 164(13), 1837–1845.

38. Schairer, C., P. Mink, L. Carroll, S. Devesa. (2004). Probabilities of death from breast cancer and other causes among female breast cancer patients. 96(17), 1311–1321

39. Nelson D, Kreps G, Hesse B, et al. The Health Information National Trends Survey (HINTS): development, design, and dissemination. Journal of Health Communication 2004; 9:443- 460.

40. Center for Disease Control and Prevention, http://www.cancer.org/cancer/cancerbasics/ lifetime-probability-of-developing-or-dying-from-cancer, Accessed January, 2015.

41. Moss, S. (2004). Should women under 50 be screened for breast cancer? British J. Cancer 91, 413–417

42. Shen Y., Zelen M., (2001). Screening Sensitivity and Sojourn Time From Breast Cancer Early Detection Clinical Trials: Mammograms and Physical Examinations. Journal of Clinical Oncology 19, 3490-3499

4 Minimizing Overdiagnosis in Breast Cancer Screening

4.1 Introduction

In Chapter 3, the potential benefits and risks associated with screening mammography are discussed. Unreliability of screening examinations, increased risk of radiation-induced cancer, and overdiagnosis are among the potential risks of screening. Overdiagnosis is known to be the most important disadvantage of cancer screening [1] as it can adversely affect people's lives. Overdiagnosis is defined as the diagnosis of screen-detected cancers that would not have presented clinically in a woman's lifetime in the absence of screening [2]. Overdiagnosis causes physical, psychosocial and economic harms by unnecessary labeling of patient with a lifelong diagnosis as well as unneeded treatments and surveillance [3].

There are two possible explanations for overdiagnosis: 1) The cancer never progresses (or, in fact, regresses), or 2) the cancer progresses slowly enough that the patient dies from a competing cause before the cancer becomes symptomatic [3]. In other words, overdiagnosis occurs when "very slow" growing cancers (more precisely, at a slow enough pace that individuals die from something else before the cancer ever causes symptoms) are detected. The second explanation incorporates the interaction of three factors: the tumor size at detection, its growth rate, and the patient's competing risks for mortality. Thus, even a rapidly growing cancer may still represent overdiagnosis if detected when it is very small or in a patient with limited life expectancy [4].

Note that overdiagnosis is different from false positive. Overdiagnosis occurs when a disease is diagnosed *correctly*, but the diagnosis is *irrelevant* suggesting that the treatment for the disease is not needed. However, false positive is an initial test result that suggests the presence of a disease, but it is later proved (with additional testing) that the disease is not present.

As it is not possible to distinguish between lethal and harmless cancers, all detected cancers are treated. Overdiagnosis and overtreatment are therefore inevitable [5]. However, tailoring screening strategies can control the probability of overdiagnosis, and thus decrease the cost associated with overdiagnosis and overtreatment.

Quantifying overdiagnosis, however, is challenging because it is impossible to distinguish between an overdiagnosed cancer and one that will become clinical at the time of diagnosis. There are various studies that quantified overdiagnosis resulting from cancer screenings. The magnitude of overdiagnosis estimates varies widely from one study to another. For example, according to Welch and Black [2], 25% of mammographically detected breast cancers are overdiagnosed. However, based on Jørgensen & Gøtzsche [4], one in three breast cancers detected in publicly organized mammography screening programs is overdiagnosed. Biesheuvel et al. [6] provided a very rough estimate of overdiagnosis, ranging from around 3 to 50% or more. In an earlier study in 2008, Duffy et al. [7] estimated the breast cancer overdiagnosis risk to be 39%. However, in a later study [8] after a prolonged follow-up of a screening program in England and Wales, Duffy and Parmar reported an overdiagnosis risk of 7–8% for a biannual screening schedule. They found this estimate to be considerably more plausible than their own previous estimate of 39%. Carter et al. [9] conducted a systematic review on overdiagnosis studies, and identified the methods that have been used for measuring overdiagnosis of cancer. They analyzed the advantages and disadvantages of each method in providing reliable estimates of overdiagnosis.

Most of the studies mentioned above are cohort studies or randomized controlled trial follow-up studies which estimate the overdiagnosis risk based on the changes in breast cancer incidence

following the introduction of screening programs in a population setting [4-8]. However, the downside of these studies is the substantial time and resource requirements due to the need for follow-up observations over a long period of time in order to get a reliable overdiagnosis estimate. Another approach in estimating overdiagnosis is through mathematical modeling. There are a few relevant studies in the literature that propose a mathematical framework for estimating overdiagnosis risk [10-12]. Davidov et al. [10] developed a mathematical model to evaluate the probability of overdiagnosis for cancer screening. They applied their model to hypothetical early detection programs for prostate cancer. Gunsoy et al. [11] developed a Markov simulation model for the evaluation of mammography screening in a cohort of British women born in 1935-40. They evaluated the impact of the screening frequency, starting and ending ages on breast cancer mortality reduction and overdiagnosis. Seigneurin et al. [12] developed a stochastic simulation model and an approximate Bayesian computation approach to quantify overdiagnosis.

In this chapter, we propose a mathematical framework to estimate breast cancer lifetime overdiagnosis and mortality risks for different screening policies. In addition, we propose an optimization model to minimize lifetime risk of overdignosis of breast cancer while maintaining a patient's lifetime breast cancer mortality risk at a predefined level. Our model provides a more flexible framework for modeling different uncertainty sources such as cancer sojourn time. Cancer sojourn time is defined as a time interval between the onset of a detectable preclinical cancer and the point when the cancer progresses to the clinical stage, causing signs and/or symptoms [13]. In addition, different from previous studies, we seek for the optimal screening policy that yields the minimum lifetime overdiagnosis risk.

The remainder of this chapter is organized as follows. In Section 4.2, the proposed model is presented. Section 4.3 presents the data sources and parameters estimation for our numerical studies. In Section 4.4, some mammography screening policies are evaluated and (near) optimal policies are obtained. Finally, Section 4.5 summarizes the findings and concludes this chapter.

4.2 Proposed Model

In the proposed model, we incorporate uncertainty in the onset of detectable preclinical cancer, variation in the cancer sojourn time, and competing causes of death. Our goal is to find the optimal breast cancer screening policies that minimize the overdiagnosis risk which is defined as the probability of detecting a cancer that would never develop symptoms or cause death during the patient's lifetime in the absence of screening. The model also maintains the lifetime breast cancer mortality risk at a predefined threshold. Following is the list of notation used in the problem formulation.

m number of screening examinations in the prescribed policy

τ screening schedule,$\tau = \{\tau_1, \dots, \tau_m\}$,where τ_i's are the decision variables representing the age at which a mammogram should be prescribed. We denote the beginning and ending of the decision period with τ_0 and τ_{m+1}, respectively. Note that τ_0 and τ_{m+1} are not decision variables and just represent the decision horizon.

T random variable representing the patient's age at the onset of the detectable preclinical cancer

S random variable representing the cancer sojourn time

S_j random variable representing the remaining sojourn time (also known as forward

	recurrence time) measured from age τ_j
Ξ	patient health state space, $\Xi = \{CF, SC, CC\}$, where CF, SC, CC represent a cancer-free individual, a patient with screen detected breast cancer, and a patient with clinical (symptomatic) breast cancer, respectively.
R^{ξ}	random variable representing the life years of an individual in health state $\xi \in \Xi$
R_j^{ξ}	random variable representing the remaining life years of a patient in health state $\xi \in \Xi$ at age τ_j
$f(.)$	probability density function of T
$g_i(.)$	probability density function of the cancer sojourn time when the cancer preclinical onset is in the i^{th} screening interval $[\tau_{i-1}, \tau_i)$
$g_{j\|i}(.)$	probability density function of the forward recurrence time measured from age τ_j given that the cancer preclinical onset (T) is in the i^{th} screening interval $[\tau_{i-1}, \tau_i)$
$h_{\xi\|x}(.)$	conditional probability density function of R_j^{ξ}, the remaining life years of an individual in health state ξ given that she has survived to age x
$H_{\xi\|x}(.)$	conditional cumulative distribution function of R_j^{ξ}, the remaining life years of an individual in health state ξ given that she has survived to age x
α_j	sensitivity of mammography (probability of detecting a cancer when it is present) at age τ_j

4.2.1 Overdiagnosis

The objective of the model is to minimize the probability of breast cancer overdiagnosis, i.e., the probability of detecting a cancer through screening that would never develop symptoms or causes death during a patient's lifetime. Figure 4-1 shows the case when overdiagnosis occurs.

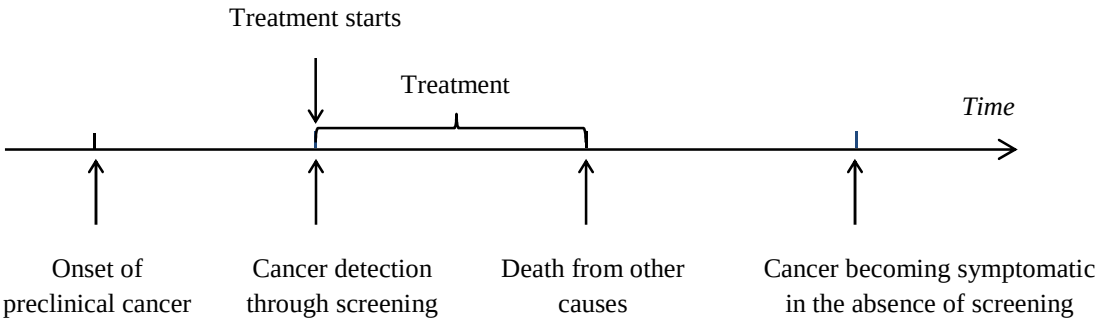


Figure 4-1 Representation of overdiagnosis in cancer screening

In the case of overdiagnosis, the breast cancer is detected through a screening examination and treatment starts upon cancer detection. However, if the breast cancer was not detected through screening, the patient would have died from a competing cause before the breast cancer grows to the clinical stage. In other words, the cancer grows slowly enough that the patient would die from a cause other than breast cancer before the cancer became symptomatic. Therefore, the probability of overdiagnosis is equivalent to the conditional probability that at the time of cancer detection (τ_j), the remaining cancer sojourn time (S_j) is greater than the remaining life years of a cancer-free individual (R_j^{CF}), given that the cancer is diagnosed through screening. Suppose D is the event that the cancer is detected through screening examinations. Then the probability of overdiagnosis is

$$Pr\left(S_j > R_j^{CF} \middle| D\right). \qquad 4\text{-}1$$

Let $t \in [\tau_{i-1}, \tau_i), i = 1, 2, \dots, m$, be the onset of the detectable preclinical breast cancer. Given that the cancer is diagnosed at the j^{th}screening (at age τ_j, $j = i, i+1, \dots, m$), the conditional probability of overdiagnosis is

$$\begin{aligned} \omega_{ij}(t) &= \int_0^{\infty} \int_0^{s} g_{j|i}(s,t) h_{CF|j}(r) dr ds \\ &= \int_0^{\infty} g_{j|i}(s,t) H_{CF|j}(s) ds, \end{aligned} \qquad 4\text{-}2$$

which calculates the probability that the remaining sojourn time is greater than the remaining life years of a cancer-free individual at age τ_j. Note that the remaining life years of a cancer-free patient is independent of the cancer onset and remaining sojourn time. The upper bound of the remaining sojourn time is infinity ($+\infty$) implying that it is possible that the cancer would never advance to the clinical stage or produce symptoms. In addition, the conditional probability density function of the cancer sojourn time ($g_i(.)$) and the conditional probability density function of the forward recurrence time (the remaining sojourn time, $g_{j|i}(.)$) are related through the following equation.

$$g_{j|i}(s,t) = Pr\left(S = \tau_j - t + s \middle| S > \tau_j - t\right) = \frac{g_i\left(\tau_j - t + s\right)}{\bar{G}_i\left(\tau_j - t\right)}, \qquad \forall s > 0 \qquad 4\text{-}3$$

where $\bar{G}_i(.)$ is the survival function of the cancer sojourn time when the cancer onset is at $t \in [\tau_{i-1}, \tau_i)$.

In addition, diagnosing the cancer at age τ_j through screening implies that the cancer has not become symptomatic yet. Let $\gamma_j(t)$ be the probability that a cancer with onset $t \in [\tau_{i-1}, \tau_i)$ has

not developed to a clinical stage at age τ_j yet (i.e., the cancer sojourn time is greater than $\tau_j - t$, Figure 4-2), given that the cancer sojourn time is greater than $\tau_{j-1} - t$, i.e.,

$$\gamma_j(t) = \frac{Pr(S > \tau_j - t)}{Pr(S > \tau_{j-1} - t)} = \frac{\bar{G}_i(\tau_j - t)}{\bar{G}_i(\tau_{j-1} - t)}, \qquad j > i. \tag{4-4}$$

Note in the case that $j = i$, $\gamma_j(t)$ reduces to $Pr(S > \tau_i - t)$.

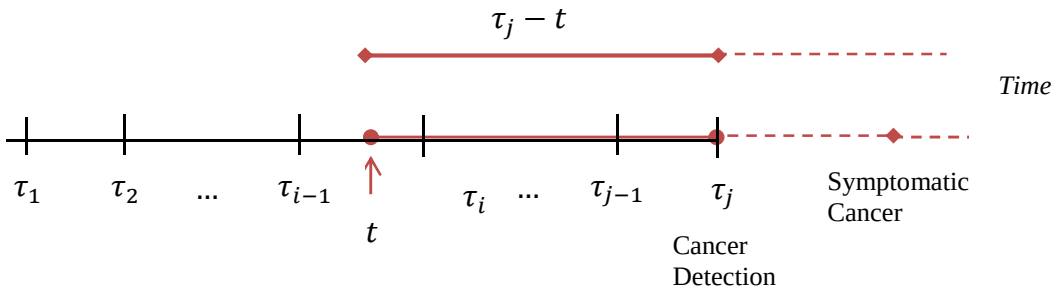


Figure 4-2 Representation of a hypothesized screening policy, cancer onset and detection

Let D_{ij} be the event that the cancer is diagnosed at the j^{th} screening when the onset of the detectable preclinical cancer is in the i^{th} screening interval. This implies that the cancer was missed in all previous screenings $(i^{th}, (i+1)^{th}, \ldots, (j-1)^{th})$, and then detected at the j^{th} examination which occur with probability

$$Pr(D_{ij}) = \begin{cases} \alpha_j & j = i, \\ \alpha_j \prod_{l=i}^{j-1} (1 - \alpha_l) & j > i. \end{cases} \tag{4-5}$$

Moreover, assume that D_i is the event that the cancer is diagnosed through screening given that its preclinical onset is in the i^{th} screening interval. Then, $D_i = \cup_{j=i}^{m} D_{ij}$, and $D'_{ij}s$ are mutually

exclusive events. Therefore, the probability that a cancer with preclinical onset in the i^{th} interval is diagnosed through screening is

$$Pr(D_i) = \sum_{j=i}^{m} Pr(D_{ij}). \quad 4\text{-}6$$

Hence, following is the probability of overdiagnosis given that the cancer onset is at $t \in [\tau_{i-1}, \tau_i)$

$$\omega_i(t) = \sum_{j=i}^{m} \omega_{ij}(t)\, \gamma_j(t)\, \frac{Pr(D_{ij})}{Pr(D_i)}. \quad 4\text{-}7$$

In addition, given that the patient develops breast cancer in her lifetime, the probability that the cancer onset is in the screening interval $[\tau_{i-1}, \tau_i)$ is $\frac{\int_{\tau_{i-1}}^{\tau_i} f(t)dt}{K}$, where $K$is the lifetime probability of developing breast cancer and $K = \int_0^{\tau_{m+1}} f(t)dt$. Therefore, since cancer onset at different intervals are mutually exclusive, the probability of overdiagnosis for screening schedule $\tau = \{\tau_1, \tau_2, \dots, \tau_m\}$ is

$$\Omega_\tau = \frac{1}{K} \sum_{i=1}^{m} \int_{\tau_{i-1}}^{\tau_i} f(t)\omega_i(t)dt \quad 4\text{-}8$$

$$= \frac{1}{K} \sum_{i=1}^{m} \sum_{j=i}^{m} \frac{Pr(D_{ij})}{Pr(D_i)} \int_{\tau_{i-1}}^{\tau_i} f(t)\, \gamma_j(t) \int_0^{\infty} g_{ij}(s,t) H_{CF|j}(s)\, ds dt\,.$$

4.2.2 Lifetime Breast Cancer Mortality Risk

To calculate the lifetime breast cancer mortality risk, two cases need to be considered: i) when the cancer is diagnosed through screening examination, and ii) when the cancer progresses to a

clinical stage and becomes symptomatic in the interval between two prescribed screening tests. These two distinct cases are considered because symptomatic cancers, as opposed to screening detected cancers, are more advanced and may lead to higher breast cancer mortality risk [14].

Suppose that the cancer preclinical onset is in the i^{th} prescribed screening interval, i.e. $T = t \in [\tau_{i-1}, \tau_i)$. The lifetime cancer mortality risks for these two cases are discussed below.

Consider the first case when the cancer is diagnosed at age τ_j or the j^{th} screening test ($j \geq i$). In this case the probability that the patient dies from breast cancer is

$$v_{j,1} = Pr\left(R_j^{SC} < R_j^{CF}\right) = \int_0^{\infty} h_{CF|j}(r) H_{SC|j}(r)\, dr. \qquad 4\text{-}9$$

Given that the cancer is detected at the j^{th} screening test, the patient should survive to age τ_j. The probability that the patient does not die from other causes before age τ_j given that she has survived to age τ_{j-1} is

$$\frac{P\left(R^{CF} > \tau_j\right)}{P\left(R^{CF} > \tau_{j-1}\right)}. \qquad 4\text{-}10$$

Detecting the cancer at the j^{th} screening test implies that the cancer sojourn time is greater than $\tau_j - t$. This is also suggesting the cancer does not develop any symptoms up to age τ_{j-1}. The associated probability of this event as presented in Equation 4-4 is $\gamma_j(t) = \frac{\bar{G}_i(\tau_j - t)}{\bar{G}_i(\tau_{j-1} - t)}$.

In addition, the probability that the cancer is diagnosed at the j^{th}screening is presented in Equation 4-5. Therefore, using the total probability rule the unconditional probability that a screen detected cancer causes death is

$$\sum_{j=i}^{m} v_{j,1} \frac{P(R^{CF} > \tau_j)}{P(R^{CF} > \tau_{j-1})} \gamma_j(t) Pr(D_{ij}). \quad 4\text{-}11$$

Since cancer onset at different intervals are mutually exclusive, using the total probability rule, the lifetime breast cancer mortality risk for the first case is

$$\theta_{\tau,1} = \sum_{i=1}^{m} \sum_{j=i}^{m} Pr(D_{ij}) \frac{P(R^{CF} > \tau_j)}{P(R^{CF} > \tau_{j-1})} v_{j,1} \int_{\tau_{i-1}}^{\tau_i} f(t) \gamma_j(t) dt. \quad 4\text{-}12$$

The second case considers a situation in which the cancer becomes symptomatic. Assume that the cancer becomes symptomatic at age $u \in (\tau_{j-1}, \tau_j), j = i+1, \ldots, m$. Given that the cancer has not been detected up to age τ_{j-1}, the probability that the cancer becomes symptomatic at age u in the interval between τ_{j-1} and τ_j is

$$\gamma_2(t,u) = \frac{g_i(u)}{\bar{G}_i(\tau_{j-1} - t)}. \quad 4\text{-}13$$

In this case, the conditional breast cancer lifetime mortality risk at age u is

$$v_{j,2}(u) = \int_0^{\infty} h_{CF|u}(r) \cdot H_{CC|u}(r)\, dr. \quad 4\text{-}14$$

Similar to the first case, the probability that the patient does not die from a competing cause given that she has survived to age τ_{j-1} (when she received a false negative screening result) is

$$\frac{P(R^{CF} > u)}{P(R^{CF} > \tau_{j-1})}. \quad 4\text{-}15$$

Therefore, the associated cancer mortality risk when the cancer onset is in the interval $[\tau_{i-1}, \tau_i)$, and the cancer becomes symptomatic at age u in the interval $(\tau_{j-1}, \tau_j),\ j = i+1, \ldots m$ is

$$\prod_{l=i}^{j-1} (1 - \alpha_l) \int_{\tau_{j-1}}^{\tau_j} \gamma_2(t,u) \frac{P(R^{CF} > u)}{P(R^{CF} > \tau_{j-1})} v_{j,2}(u)\, du, \quad 4\text{-}16$$

which calculates the probability that the cancer does not get detected through previous screenings ($\prod_{l=i}^{j-1}(1-\alpha_l)$) and the probability that the individual dies from breast cancer when the cancer becomes symptomatic in the interval (τ_{j-1}, τ_j). Summing over all screening intervals after the cancer onset interval, the cancer mortality risk is

$$\Psi_1(t) = \sum_{j=i+1}^{m+1} \prod_{l=i}^{j-1}(1-\alpha_l) \int_{\tau_{j-1}}^{\tau_j} \gamma_2(t,u) \frac{P(R^{CF} > u)}{P\left(R^{CF} > \tau_{j-1}\right)} \nu_{j,2}(u)\, du. \qquad 4\text{-}17$$

In addition, it is possible that the cancer becomes symptomatic before the first scheduled screening in the interval (t, τ_i). In such a case, the following is the associated cancer mortality risk, which has a similar logic as Equation 4-17.

$$\Psi_2(t) = \int_{t}^{\tau_i} g_i(u) \frac{P(R^{CF} > u)}{P\left(R^{CF} > \tau_{j-1}\right)} \nu_{j,2}(u)\, du \qquad 4\text{-}18$$

Equation 4-18 calculates the probability that the cancer becomes symptomatic at age u ($g_i(u)$), the patient does not die from a competing cause in the interval (τ_{j-1}, u), i.e., $\frac{P(R^{CF}>u)}{P(R^{CF}>\tau_{j-1})}$, and she eventually dies from cancer ($\nu_{j,2}(u)$).

Similar to the first case, since cancer onset at different intervals are mutually exclusive, the lifetime cancer mortality risk of screening schedule τ for a symptomatic cancer is

$$\theta_{\tau,2} = \sum_{i=1}^{m} \int_{\tau_{i-1}}^{\tau_i} f(t)[\Psi_1(t) + \Psi_2(t)]\, dt. \qquad 4\text{-}19$$

Considering both cases of cancer detection, the lifetime cancer mortality risk for schedule τ is

$$\Theta_\tau = \theta_{\tau,1} + \theta_{\tau,2}. \quad \text{4-20}$$

Optimization Model

The proposed optimization model can then be written as

$$\text{Min } \Omega_\tau$$
$$\text{Min } m$$
s.t.
$$\Theta_\tau < \epsilon \quad \text{4-21}$$
$$\tau_j < \tau_{j+1} \quad \forall j \quad \text{4-22}$$
$$\tau_j \in \{\tau_0, \tau_0 + \delta, \tau_0 + 2\delta \dots, \tau_{m+1} - \delta\} \; \forall j, \quad \text{4-23}$$

where constraint 4-21 ensures that the lifetime mortality risk is less than a predefined threshold; constraints 4-22 and 4-23 determine the order of screening decisions and age range in which the screening tests should be prescribed, respectively. τ_0 and τ_{m+1} in equation 4-23 represent starting and ending age of a screening schedule. In addition, δ in constraint 4-23 represents the minimum interval between two subsequent screening tests.

4.3 Model Inputs

Table 4-1 presents the summary of data sources for the different inputs (parameters and probability distributions) incorporated in the model.

Table 4-1 Data sources for the model input estimations

Description	Reference
Distribution of onset of preclinical detectable breast cancer	Parmigiani and Skates [15]
Parameters of preclinical sojourn time distribution (exponential case)	Shen and Zelen [16] Tabar et al. [17]
Parameters of preclinical sojourn time distribution (lognormal case)	Peer et al. [21]
Probability density of remaining life years of a cancer-free individual	The US life table for females (Center for Disease Control and Prevention) [22] and Surveillance, Epidemiology and End Results (SEER) [23]
Age and stage specific probability distribution of death from breast cancer	Schairer et al. [24] and Zhang [25]
Stage distribution of screen detected breast cancers	Bleyer and Welch [26]
Stage distribution of clinically detected breast cancers	Plevritis et al. [27]
Age specific mammography sensitivity	Kerlikowske et al. [28]

The probability distribution of breast cancer onset is the most challenging distribution in this study to estimate. Parmigiani and Skates [15] developed a convolution model to estimate the age of disease onset distribution based on the natural history of disease and the data available on disease incidence rate, cancer sojourn time, competing causes of death, etc. They then used singular value decomposition method to solve their developed model numerically. We applied their method to estimate breast cancer preclinical onset age distribution in this study. Figure 4-3 shows the extracted probability density for the preclinical breast cancer onset.

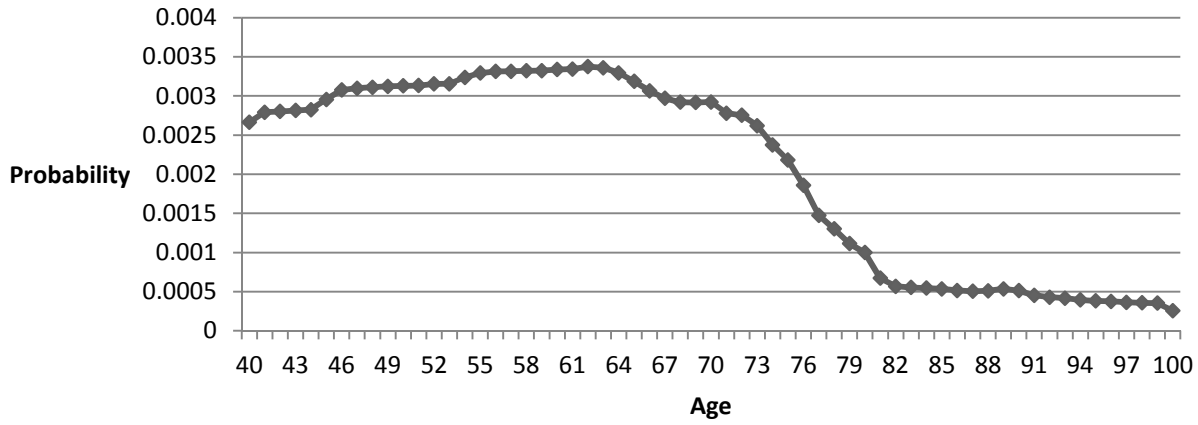


Figure 4-3 Estimated distribution of the age of onset of detectable breast cancer

There are various studies in the literature estimating the distribution of breast cancer sojourn time. Some of these studies used exponential specification in estimating the sojourn time distributions [16-19]. However, the assumption of exponential duration in the preclinical stage has some limitations. The first concern is the implausible assumption of mode at zero, which corresponds to an instant transition from preclinical to clinical stage, and the fast decaying tail, which does not adequately account for slow growing tumors [15]. The second limitation is due to the memoryless property of the exponential distribution, which implies that the sojourn time and remaining sojourn time upon cancer detection through screening have the same distribution. In other words, the hazard function of sojourn time is constant, implying that as time passes the instantaneous probability that the cancer develops to a clinical stage is constant. However, it would be more plausible that the instantaneous probability of developing to a clinical stage increases with time. A second distribution proposed for modeling cancer sojourn time is lognormal distribution. Spratt et al. [20] postulated a lognormal distribution for sojourn time based on the growth patterns of breast tumors. In this study, we examine both exponential and lognormal sojourn time distributions.

For the exponential case, the age-specific rate parameters are adopted from the mean sojourn time provided in Shen and Zelen [16], and Tabar et al. [17]. Figure 4-4 presents the estimated exponential breast cancer sojourn time distributions for the different age groups: 40-49, 50-59, and 60 and older. It also presents the corresponding hazard function for different age groups.

For the case with lognormal sojourn time, the age-specific distribution parameters are estimated using the median and upper 95% quantile matching Peer et al. [21]. The estimated lognormal breast cancer sojourn time distributions for the three age groups, along with the corresponding hazard functions are presented in Figure 4-5. As the results show, the hazard functions for this case are increasing for most parts. However, for the first and second age groups, the hazard rate starts decreasing very slowly for after about three years and five years, respectively. This suggests that if no symptom appears by a certain amount of time after the cancer onset, then the instantaneous probability that the cancer actually becomes symptomatic starts to decrease (because the tumor stops growing or it regresses).

The probability distribution of the remaining life years for a cancer-free individual is extracted from the 2008 US life table for females [22] and the Surveillance, Epidemiology and End Results (SEER) data [23]. SEER data is used to exclude the probability of death from breast cancer in the life table and adjust the probabilities for women without breast cancer.

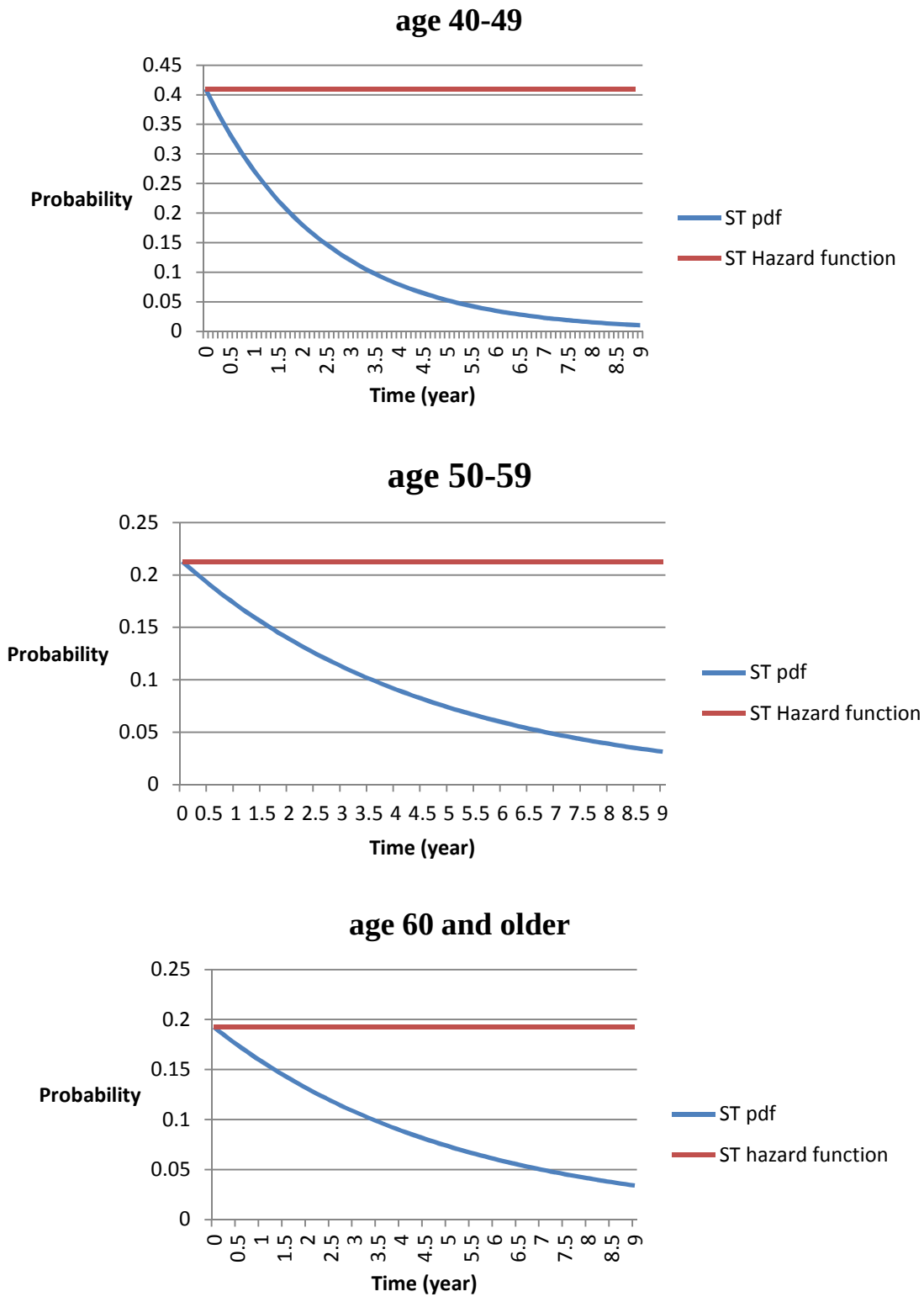


Figure 4-4 Estimated exponential breast cancer sojourn time distribution for different age groups

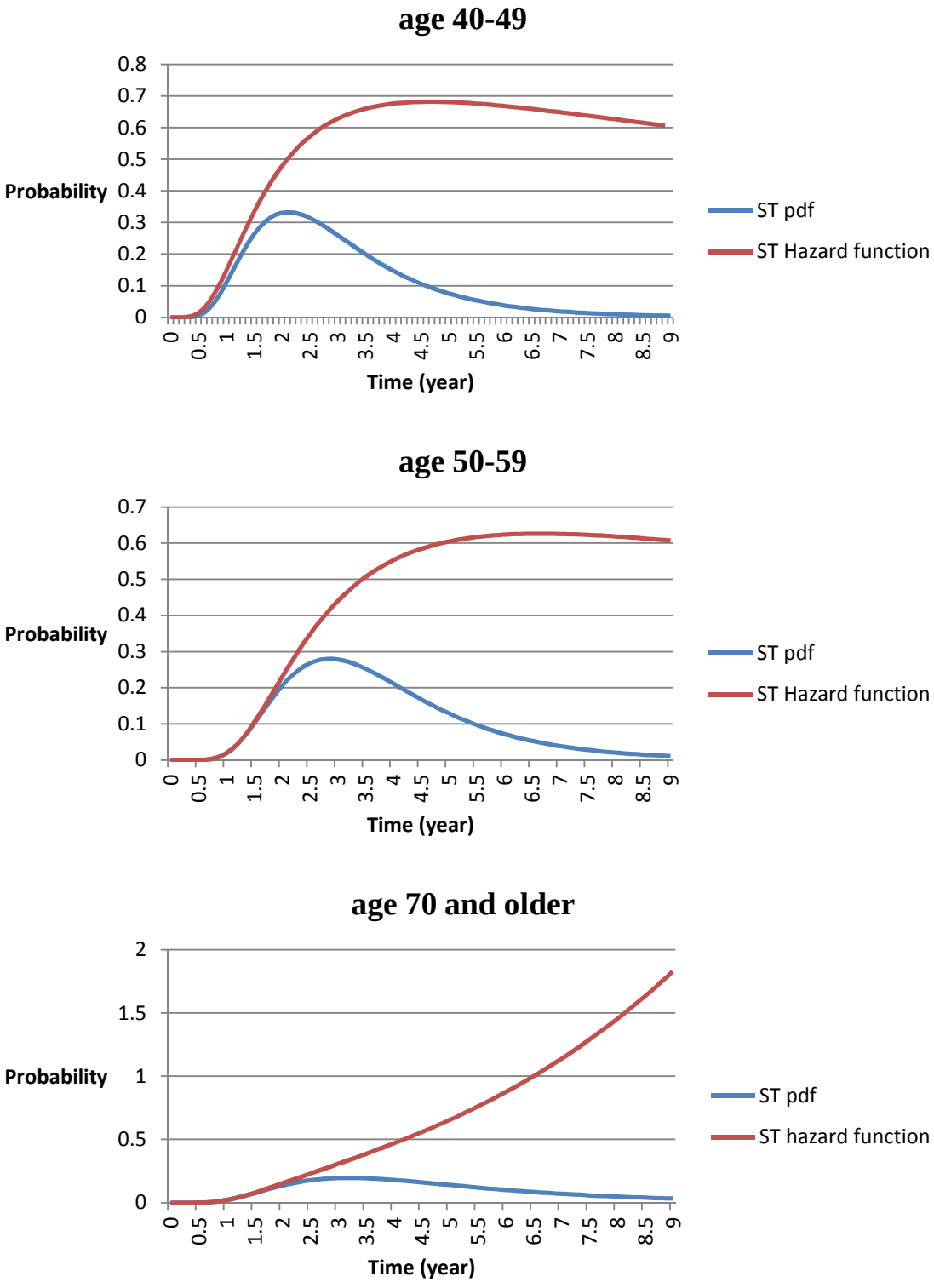


Figure 4-5 Estimated lognormal breast cancer sojourn time distribution for different age groups

We estimate the probability density of remaining life years for the two cases of cancer detection (screen detected and symptomatic breast cancer) using the data available in Schairer et al. [24] and Zhang [25]. We adopt the stage distributions of screen detected and symptomatic breast cancer from Bleyer and Welch [26] and Plevritis et al. [27], respectively, and adjust the probability density of remaining life years for the two cases.

The mortality probabilities in [22, 24-25] are presented in tabular format, and therefore we used piecewise-constant density function to model the probability density functions. Let p_i^{ξ} be the probability that an individual of age i in health state ξ dies in the age interval $(i, i+1]$. Therefore,

$$Pr\left(R^{\xi} \leq k\right) = 1 - \prod_{i \leq k} (1 - p_i^{\xi}) \qquad 4\text{-}24$$

and,

$$h_{j|\xi}(r) = \frac{Pr\left(R^{\xi} \leq k+1\right) - Pr\left(R^{\xi} \leq k\right)}{Pr(R^{\xi} \geq j)}, \qquad k < r \leq k+1, \qquad k > j. \qquad 4\text{-}25$$

Lastly, age specific mammography screening sensitivities are extracted from Kerlikowske et al. [28].

4.4 Computational Results

Disease onset ad conditional remaining life year functions do not have a general closed form and are only available in tabular format. This makes analytical calculation of integration very challenging or even impossible. Therefore, in this study we exploit Monte Carlo integration method to calculate the integrations in the proposed model. Monte Carlo integration methods are

sampling methods to calculate complicated integrations, based on the central limit theorem and the law of large numbers. Please refer to Robert and Casella [29] for more details.

In addition, due to the complexity of the model, solving the optimization model through analytical approaches is not feasible. Therefore, we apply a combination of bisection method and Particle Swarm Optimization (PSO) method to find the (near) optimal solution of the proposed model.

The results presented in this section have two parts. In the first part we evaluate different screening policies, including the ACS and the USPSTF policies, in terms of the breast cancer overdiagnosis and mortality risks. The second part presents the (near) optimal policies obtained using the bisection and PSO algorithm.

4.4.1 Policies Evaluation

In this section we present the overdiagnosis and lifetime mortality risks of different policies for two reasons: (1) to validate our model by comparing the obtained risk values with numbers reported in the literature, and (2) to evaluate these polices and compare their performance in terms of breast cancer overdiagnosis and mortality risks. We evaluate different static as well as dynamic screening policies with two and three different screening intervals between subsequent tests. Table 4-2 presents the policies evaluated in this chapter. Including the USPSTF policy, in total we evaluated 241 policies. We present these policies using the same structure as introduced in Chapter 3. For a policy with three different screening intervals, we use a similar structure as well. For example, policy (40,1,50,2,60,3,80) recommends women get annual screening tests between age 40 to 50, then switch to biennial screenings up to age 60, and finally undergo screenings every three years up to age 80.

Table 4-2 Screening policies considered in the numerical analysis

policy	Starting age	1st interval	1st switch age	2nd interval	2nd switch age	3rd interval	Ending age
Static	40,50	1,2,3	-	-	-	-	80,90,100
Dynamic with 2 screening intervals	40,50	1,2,3	50,60,70	1,2,3	-	-	80,90,100
Dynamic with 3 screening intervals	40,50	1,2,3	50,60,70	1,2,3	60,70,80	1,2,3	80,90,100

4.4.2 Exponential Sojourn Time

Figure 4-6 presents the breast cancer overdiagnosis and mortality risks of the policies considered in this chapter along with the efficient frontier policies (highlighted) for the case with exponential sojourn time.

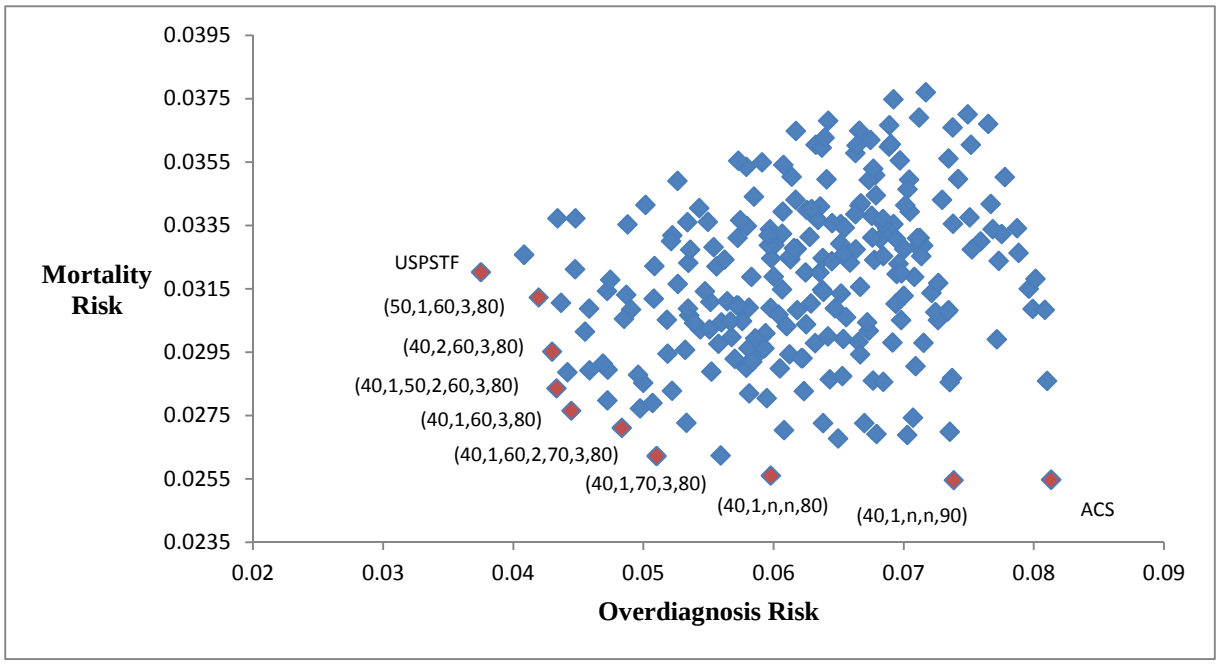


Figure 4-6 Overdiagnosis and mortality risks of the screening policies in Table 4-2 under exponential sojourn time

The overdiagnosis risk estimates for biannual screening schedules are very close to the values reported in the literature: Duffy et al.'s estimate of 7%-8% [8], Zackrisson et al.'s estimate of 10%

[30], De Gelder estimate of 8.9‑15.2% [31] and 7.2% [32]. To the best of our knowledge there is no overdiagnosis reported in the literature for annual screening policies. Therefore, we did not have a reference to compare our results with. The breast cancer mortality risks are also comparable with the mortality risks estimated in Chapter 3 (reported in Table 3-6). Note that, the underlying sojourn time distribution for both results in Table 3-6 and Figure 4-6 is exponential.

The results in Figure 4-6 show that both the ACS and the USPSTF policies are efficient. In fact, the USPSTF policy and the ACS policy have the lowest overdiagnosis and mortality risk among the evaluated policies, respectively. Most of the efficient frontier policies (except for the ACS and (40,1,n,n,90)) recommend women stop screening at or before age 80. This is because after age 80, the probability of death from competing causes significantly increases which results in high overdiagnosis risk. In addition, most efficient policies (except for the USPSTF and (50,1,60,3,80)) recommend women start screenings at age 40. This maintains both the overdiagnosis and mortality risks at a lower level. In addition, in terms of the distribution of screening tests in the efficient frontier policies, more frequent tests are recommended in the age interval 50 to 60 when the breast cancer onset is more likely.

4.4.3 Lognormal Sojourn Time

The corresponding overdiagnosis and mortality risks for the case with lognormal sojourn time with efficient frontier policies highlighted are presented in Figure 4-7. The overdiagnosis risks for this case are very close to the exponential case, and therefore comparable with the reported values in the literature [8, 30-32]. Also, the mortality risks are very close to the mortality risk estimations reported in Table 3-6 in Chapter 3.

Similar to the case with exponential sojourn time, the ACS policy with the highest number of screening tests has the highest overdiagnosis and lowest lifetime breast cancer mortality risk and the USPSTF, on the other hand has the lowest overdiagnosis risk. In addition, the recommended stopping age in most efficient policies is 80 to maintain a low overdiagnosis risk. Similar to the exponential sojourn time case, more screening tests are distributed between age 50 to 60 comparing to the other age intervals.

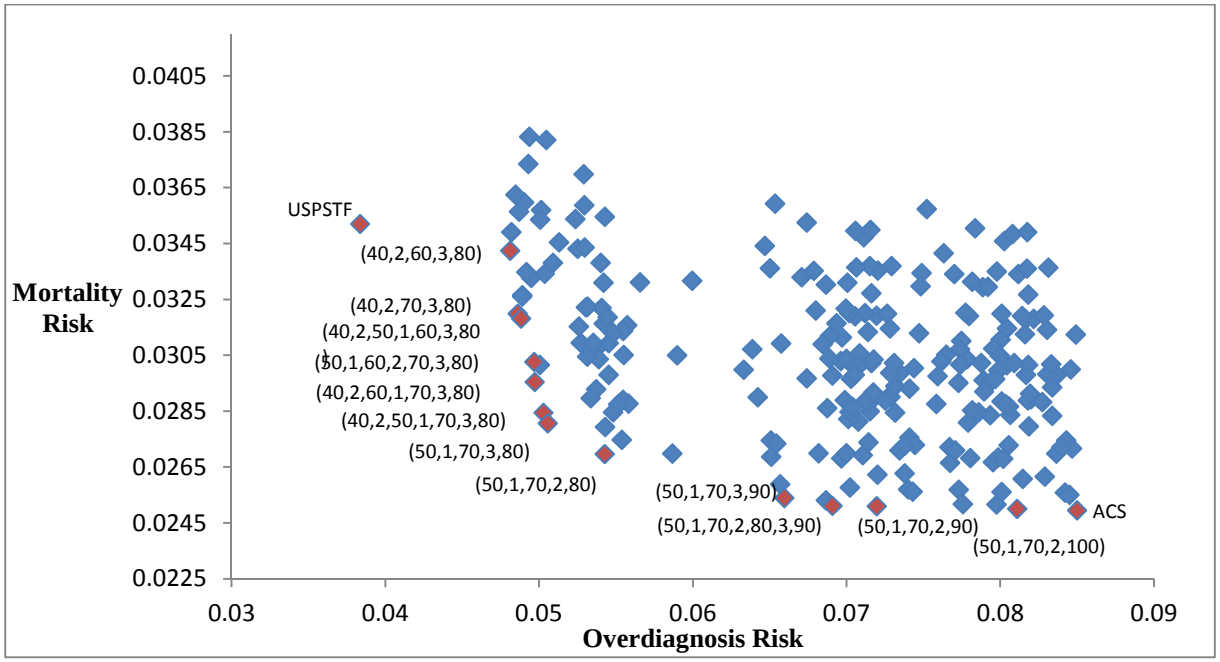


Figure 4-7 Overdiagnosis and mortality risks of the screening policies in Table 4-2 under lognormal sojourn time

4.5 Policy Optimization

In this section, the algorithm to obtain the (near) optimal polices and the extracted (near) optimal policies are presented. A combination of bisection and binary Particle Swarm Optimization (PSO) is applied to obtain near optimal policies.

PSO is originally attributed to Kennedy, Eberhart [33] and is a biologically-inspired algorithm motivated by a social analogy, such as flocking, herding, and schooling behavior in animal populations. The idea behind PSO is that the individuals that are part of a society hold an opinion that is part of a "belief space" and shared by every possible individual. Individuals may modify this "opinion state" based on the knowledge of the environment and the individual's previous history of states (its memory). For more details on PSO, please refer to [33].

The solutions or particles (policies) are encoded as a binary string so that a one (zero) represents a screening (no screening) at the corresponding age. There are two stages for the optimization process. In the first stage, we determine the minimum number of screening examinations (m^*) needed to guarantee a mortality risk lower than the predefined threshold, ϵ. We use the bisection method and PSO to determine the optimum number of screening tests. The number of tests is determined based on the bisection method. For each fixed number of tests, the policy with minimum mortality risk is obtained using PSO. Note that, it is critical for our algorithm to preserve the number of screening tests when generating a new particle. Therefore, we develop a variant of PSO in which the number of swaps between 0's and 1's in a particle (solution) is the same (Please refer to Figure 4-8).

In the second stage, the (near) optimal policy with the following objective function is extracted using a similar PSO algorithm.

$$G = \Omega_\tau + M * (\Theta_\tau - \epsilon) \qquad 4\text{-}27$$

In Equation 4-27, M is the penalty for violating constraint 4-21 in the proposed model. The pseudocode of the bisection and particle swarm optimization is presented in Figure 4-8. Note that we used the (near) optimal policy found in the first stage as an input for the second stage.

We assume that the threshold for the mortality risk (ϵ) is 0.0270 (the lifetime breast cancer mortality risk reported by the American Cancer Society: 1 in 37 [34]). We also assume that $\tau_0 = 40$, $\tau_{m+1} = 100$, and $\delta = 1$ year.

4.5.1 Parameter Tuning

The selection of PSO parameters influences the performance of the heuristic. Thus, we conduct a formal investigation through design of experiments (DOE) to select the best combination. Following is the list of parameters considered in the experimental design.

1) *Swarm size*: the swarm size influences performance of PSO. Too few particles may cause the algorithm to get trapped in local optima, while too many particles slow down the algorithm. There is no exact rule in the literature for selection of swarm size. But generally, when dimension of the problem increases, the swarm size should also be increased [36].
2) *Cognitive acceleration coefficient* (c_1): this represents the weighting of the stochastic acceleration terms that pull particles towards P_{best}, i.e., the best solution found by a specific particle. If the value of this constant is too high, the particles move abruptly and the risk of getting trapped in local optima increases. Conversely, if the value is too low, the particles move too slowly, and computational effort increases significantly [35].
3) *Social acceleration coefficient* (c_2): this represents the weighting of the stochastic acceleration terms that pull particles towards g_{best}, i.e., the best solution found. Similar to the cognitive acceleration coefficient, too high or too low value for this parameter can cause getting stuck to local optima or non-convergence of the algorithm, respectively [36].

Step 1: Bisection method and Particle Swarm Optimization (PSO_1) **to determine** $\boldsymbol{m}^*$

procedure $BiSec(l, u)$

$m \leftarrow \frac{l+u}{2}$

While $(u - l \geq 1)$

If $(PSO_1\ (m) <\ \epsilon)$ **then** $u \leftarrow m$

If $(PSO_1(m)\ \geq \epsilon)$ **then** $\ l \leftarrow m$

EndWhile

end procedure

procedure $PSO_1(m)$

initialization;

repeat

for i = 1 to number of individuals **do**

if $G(ind_i) < G(Best - ind_i)$ **then** ◁ $G()$ evaluates the mortality risk (Eq. 4-21)

$Best - ind_i \leftarrow ind_i$

end if

if $G(ind_i) < G(Best)$ **then**

$Best \leftarrow ind_i$

end if

end for

for i = 1 to number of individuals **do**

for d = 1 to number of dimensions **do**

$v_{id}(t) = v_{id}(t-1) + c_1\varphi_1(LBestPol_i - Pol_i) + c_2\varphi_2(GBestPol_i - Pol_i)$

$a(t) = v_{id}(t) - v_{id}(t-1)$

sort $a(t)$ according to a descending order

randomly choose an index (i_1) from the Max range of $a(t)$

$ind_i(i_1) \leftarrow 1$

randomly choose an index (i_2) from the Min range of $a(t)$ and

$ind_i(i_2) \leftarrow 0$

end for

end for

until stopping criteria

end procedure

Step 2: Particle Swarm optimization $(\boldsymbol{PSO_2})$ **with** $\boldsymbol{m} = \boldsymbol{m}^*$ **to determine the (near) optimal policy with the minimum overdiagnosis and a mortality risk less than** ϵ

$PSO_2(m^*, GBest)$ is similar to PSO_1 with $G = \Omega_\tau + M * (\Theta_\tau - \epsilon)$

Figure 4-8 Pseudocode of the bisection and Particle Swarm Optimization methods

In addition, the relative value of these cognitive and social acceleration coefficients is critical and affects the algorithm's performance. When the value of the cognitive acceleration coefficient (c_1) increases, it enhances particles' attraction towards P_{best} and decreases their attraction towards g_{best}, and vice versa. In the literature, the setting of $c_1 = c_2 = 2$ has been proposed as a generally acceptable setting for most problems, and is widely used in practical applications of PSO [35].

Three levels of each factors considered in the experimental design analysis are shown in Table 4-3. A 3^3 factorial design is conducted. Response variables for each combination are calculated as the average of objective function values over all best values found by particles in the swarm. The algorithm performance is not sensitive to the swarm size, and does better with lower value for cognitive acceleration coefficient comparing to social acceleration coefficient.

Table 4-3 Factorial design results

Parameter	**Levels**	**Final Selection**
Swarm size	30, 40, 50	30
Cognitive acceleration coefficient (c_1)	1, 2, 3	1
Social acceleration coefficient (c_2)	1, 2, 3	2

4.5.2 Optimization Results

Table 4-4 presents the (near) optimal policies for the cases with exponential and lognormal sojourn time.

For the case with the exponential sojourn time, the (near) optimal policy includes 20 examinations ($m^* = 20$). The corresponding breast cancer overdiagnosis and mortality risks for this policy are 0.0326 and 0.0269, respectively. Note that most screenings are distributed in early ages (40 to 65) and there are very few tests in the interval between 66 and 74, and no tests after

age 74. This is due to the fact that at older ages, sojourn time is longer, and there are more aging-related mortality causes. Therefore, from both the overdiagnosis and mortality risk perspectives, it is plausible to have fewer tests at older ages. Increasing the number of prescribed screenings yields a policy with slightly higher overdiagnosis and lower mortality risks.

The (near) optimal policy for the case with lognormal sojourn time includes 19 screening tests ($m^* = 19$). This policy yields overdiagnosis and mortality risks of 0.0398 and 0.0269, respectively, which are very close to the risks of the (near) optimal policy with exponential sojourn time. Similar to the exponential case, most screening tests are prescribed at early ages (between age 40 and 65) and there is no prescribed examination after age 81.

Note that in both exponential and lognormal sojourn time models, the (near) optimal policies have the same number of tests as or fewer tests than the "every three years" policy, and yet has a lower mortality risk and evidently lower overdiagnosis risk. Therefore, the results suggest that, in-practice policies with evenly distributed screening tests throughout the patient life are not efficient.

Comparing the results for the exponential and lognormal cases implies that the (near) optimal policies follow a similar structure in both cases: more tests at younger ages and fewer tests as a woman ages. However, the (near) optimal policy in the case with lognormal distribution prescribes screenings till older ages (age 81). This is due to the increasing probability of instantaneous symptomatic cancer with time when the sojourn time is modeled as lognormal, as opposed to the constant probability of instantaneous symptomatic cancer when sojourn time is exponential.

Table 4-4 (Near) optimal policies found for both exponential and lognormal sojourn time

Age	Exponential Sojourn Time	Lognormal Sojourn Time
	$\boldsymbol{m^* = 20}$	$\boldsymbol{m^* = 19}$
40	0	1
41	1	0
42	1	1
43	1	0
44	1	1
45	1	0
46	1	1
47	1	0
48	1	1
49	0	1
50	1	1
51	0	0
52	1	0
53	0	1
54	1	0
55	1	1
56	1	0
57	0	1
58	1	1
59	0	0
60	1	0
61	0	1
62	1	0
63	0	1
64	1	0
65	0	1
66	1	0
67	0	0
68	0	1
69	1	0
70	0	0
71	0	1
72	0	0
73	0	0
74	1	1
75	0	0
76	0	0
77	0	1
78	0	0
79	0	0
80	0	0
81	0	1
Mortality Risk	0.0269	0.0269
Overdiagnosis Risk	0.0326	0.0398

4.6 Conclusion

Ideally, screening interventions aim to detect diseases that will ultimately cause harms, and the purpose of screening intervention is to advance the detection time, when the disease is in its early stages and is more likely to be cured. However, there is always risk of overdiagnosis and overtreatment when detecting a disease in its early stages. Overdiagnosis of a disease is defined as the diagnosis of an asymptomatic disease having no signs or symptoms, which would have never become symptomatic during an individual's remaining lifetime. The fundamental idea of overdiagnosis is that screening intervention detects the disease at an earlier age than it would have been diagnosed under usual care.

In this chapter we derive the equations for the probability of breast cancer overdiagnosis and mortality risks. Although applied to breast cancer, the proposed model can be generalized to calculate risks for any other types of cancer. We evaluate the risks associated with in-practice and alternative mammography screening policies recommended with different health agencies. The overdiagnosis and mortality risks for several in-practice policies including the ACS and the USPSTF policies are derived. Our estimates of the overdiagnosis and mortality risks are consistent with the values reported in previous studies in the literature and the derived mortality risk estimation based on the model proposed in Chapter 3.

In addition, we derive the (near) optimal policies with minimum overdiagnosis risk that are guaranteed to have a mortality risk lower than the average breast cancer mortality risk reported by the American Cancer Society (1 in 37 lifetime risk of dying from breast cancer). We consider two possible distributions for cancer sojourn time: exponential and lognormal distributions. The derived (near) optimal policies mainly recommend more screenings at early ages (40-65) when the cancer sojourn time is shorter and the probability of dying from competing causes are lower.

There are more dispersed examinations after age 65, and no recommended screening after age 74 and 81 for exponential and lognormal sojourn time, respectively. Results show that more efficient policies with fewer screening examinations and lower overdiagnosis and mortality risks than in-practice policies can be obtained.

References

1. Heywang-Köbrunner, S. H., Hacker, A., & Sedlacek, S. (2011). Advantages and disadvantages of mammography screening. Breast Care, 6(3), 199-207.

2. Welch, H. G., & Black, W. C. (2010). Overdiagnosis in cancer. Journal of the National Cancer Institute, 102(9), 605-613.

3. Black W. C. 2000. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. J Natl Cancer Inst; 92:1280-2.

4. Jørgensen, K. J., & Gøtzsche, P. C. (2009). Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends.Bmj, 339.

5. Welch GH. Should I be tested for cancer?. 2004. California: University of California Press.

6. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. 2007. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol. 8:1129–1138

7. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH (2008) Complexities in the estimation of overdiagnosis in breast cancer screening. Br J Cancer 99: 557–562

8. Duffy, S. W., & Parmar, D. (2013). Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time. *Breast Cancer Res*, *15*(3), R41.

9. Carter, J. L., Coletti, R. J., & Harris, R. P. 2015. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. *BMJ*, *350*, g7773.

10. Davidov, O., & Zelen, M. (2004). Overdiagnosis in early detection programs. Biostatistics, 5(4), 603-613.

11. Gunsoy, N. B., Garcia-Closas, M., & Moss, S. M. (2014). Estimating breast cancer mortality reduction and overdiagnosis due to screening for different strategies in the United Kingdom. British journal of cancer.

12. Seigneurin, A., François, O., Labarère, J., Oudeville, P., Monlong, J., & Colonna, M. 2011. Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. *BMJ*, *343*.

13. Shen, Y., & Zelen, M. (2001). Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. Journal of Clinical Oncology, 19(15), 3490-3499

14. Caumo, F., Vecchiato, F., Strabbioli, M., Zorzi, M., Baracco, S., & Ciatto, S. (2010). Interval cancers in breast cancer screening: comparison of stage and biological characteristics with screen-detected cancers or incident cancers in the absence of screening. Tumori, 96(2), 198-201.

15. Parmigiani G., Skates S. (2001). Estimating distribution of age of the onset of detectable asymptomatic cancer. Mathematical and Computer Modelling, 33: 1347-1360.

16. Shen, Y., & Zelen, M. (2005). Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. Biostatistics, 6(4), 604-614.

17. Tabar L, Duffy SW, Vitak B, Chen H. H., Prevost T.C., (1999). The natural history of breast carcinoma. Cancer 86:449-462, 1999

18. Chen, H. H., Duffy, S. W., & Tabar, L. (1997). A mover-stayer mixture of Markov chain models for the assessment of dedifferentiation and tumour progression in breast cancer. Journal of Applied Statistics, 24(3), 265-278.

19. Michaelson, J., Satija, S., Moore, R., Weber, G., Halpern, E., Garland, A., & Kopans, D. B. (2003). Estimates of breast cancer growth rate and sojourn time from screening database information. Journal of Womens Imaging, 5, 11-19.

20. Spratt, J. S., Greenberg, R. A., & Heuser, L. S. (1986). Geometry, growth rates, and duration of cancer and carcinoma in situ of the breast before detection by screening. Cancer research, 46(2), 970-974.

21. Peer, P. G., Van Dijck, J. A., Verbeek, A. L., Hendriks, J. H., & Holland, R. (1993). Age-dependent growth rate of primary breast cancer. *Cancer*, *71*(11), 3547-3551.

22. Center for Disease Control and Prevention, ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/61_03/Table03.xls, accessed September 2014.

23. Surveillance, epidemiology and end results, http://seer.cancer.gov/archive/csr/1975_2009_pops09/results_merged/topic_lifetime_risk.pdf , accessed September 2014.

24. Schairer, C., Mink, P. J., Carroll, L., & Devesa, S. S. (2004). Probabilities of death from breast cancer and other causes among female breast cancer patients. Journal of the National Cancer Institute, 96(17), 1311-1321.

25. Zhang S., (2011). Modeling the Complexity of Breast Cancer under Conditions of Uncertainty. Ph.D. dissertation, Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina

26. Bleyer, A., & Welch, H. G. (2012). Effect of three decades of screening mammography on breast-cancer incidence. New England Journal of Medicine, 367(21), 1998-2005.

27. Plevritis, S. K., Salzman, P., Sigal, B. M., & Glynn, P. W. (2007). A natural history model of stage progression applied to breast cancer. *Statistics in medicine*, *26*(3), 581-595.

28. Kerlikowske, K., Carney, P. A., Geller, B., Mandelson, M. T., Taplin, S. H., Malvin, K., Ernster V., ... & Ballard-Barbash, R. (2000). Performance of screening mammography among women with and without a first-degree relative with breast cancer. Annals of Internal Medicine, 133(11), 855-863.

29. Roberts, C.P., Casells, G. (2013). Monte Carlo Statistical Methods. Springer Science & Business Media

30. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of overdiagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. BMJ 2006;332:689-91.

31. De Gelder R, Fracheboud J, Heijnsdijk EAM, den Heeten G, Verbeek ALM, Broeders MJM, et al. Digital mammography screening: weighing reduced mortality against increased overdiagnosis. Prev Med 2011;53:134-40.

32. De Gelder R, Heijnsdijk EAM, Van Ravesteyn NT, Fracheboud J, Draisma G, De Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 2011;33:111-21.

33. Kennedy, J. (2010). Particle swarm optimization. In Encyclopedia of Machine Learning (pp. 760-766). Springer US.

34. American Cancer Society, http://www.cancer.org/cancer/cancerbasics/lifetime-probability-of-developing-or-dying-from-cancer, accessed May 2015

35. Rezaee Jordehi, A., & Jasni, J. (2013). Parameter selection in particle swarm optimisation: a survey. Journal of Experimental & Theoretical Artificial Intelligence, 25(4), 527-542.

5 Estimation of Utilization and Usage Time of Servers Stored in a Shared Stack in an Unreliable Queueing System

5.1 Introduction

It is very important to keep service systems operating normally in this service oriented era. Failing to operate a system continuously will inevitably result in waste and inefficiency, including downtime cost, loss of customers, and delay in providing services to customers or production of goods. That is the reason why preventive maintenance is of great importance in modem complex systems. Preventive maintenance policies enable us to reduce operating costs as well as the risk of breakdowns.

In general, reliability and maintenance literature can be categorized into two broad classes: single component models and multi-component models. Since almost every system in the real world is multi-component, there has been a growing interest in the optimization of maintenance of multi-components systems. Multi-component systems are specifically important because of the interactions between components in the system. These interactions between components can be classified into three different types: economic dependence, structural dependence, and stochastic dependence. Economic dependence implies that costs can be saved when several components are jointly maintained instead of separately (economies of scales can be obtained). Stochastic dependence occurs if the condition of components influences the lifetime distribution of other components. Structural dependence applies if components structurally form a part, so that maintenance of a failed component implies maintenance of working components [1]. As an example of economically dependent components, assume a system in which the maintenance of each component requires set-up work that can be shared when several components are

maintained simultaneously. Set-up costs can be saved when maintenance activities on different components are executed simultaneously, since execution of a group of activities requires only one set-up. This can yield considerable cost savings, and therefore, the development of optimization models for multiple components is an important research issue.

In recent years, a large amount of research has been devoted to finding optimal maintenance policies for multi-component systems under various assumptions. Comprehensive reviews of the literature on maintenance of multi-component systems are given in Dekker and Wildeman [1], Nicolai and Dekker [2], and Wang [3]. Recently, the relatively new problem of optimal maintenance policies for "*queues with unreliable servers*" caught the interest of researchers in the maintenance optimization area. There are a considerable number of studies addressing the problem of unreliable queuing systems. However, most of these studies deal with single unreliable server queueing systems [4-8]. A detailed literature review on unreliable queues with single server was given by Choudhury and Tadj [4]. However, there are a limited number of studies on unreliable multi-server queueing systems. The very first study on unreliable multi-server queues is by Mitrany and Avi-Itzhak [9]. They studied a steady-state $M/M/N$ queuing system where each server is subject to random breakdowns of exponentially distributed duration. They derived the explicit form of moment generating function of the queue size when the number of servers in the system is less than 3, and proposed a numerical method for systems with larger number of servers. Neuts and Lucantoni [10] studied a queue with N identical servers in parallel that may break down and require repairs. They assumed that there are c repairmen in the system and the service time, time to failure and repair time are exponential. They derived the stationary distributions of various waiting times in the system. Wang and Chang [11] considered a queue with balking, reneging, and server breakdowns in which arrival,

service times, breakdown times and repair times of the servers are assumed to follow a negative exponential distribution. They developed a cost model to determine the optimum number of servers in the system. Wu and Ke [12] derived the necessary and sufficient condition of system equilibrium in an *M*/*M*/*c* queueing system with balking, reneging, and server breakdowns. Gharbi and Dutheillet [13] developed a Generalized Stochastic Petri Nets (GSPNs) model to analyze multi-server finite-source retrial systems with unreliable servers subject to breakdowns and repairs. They formulated the main stationary performance and reliability indices as a function of the number of servers, the size of the customer source and the stationary probabilities of the systems.

There are very few papers addressing the maintenance optimization aspect in unreliable queueing systems. Chapin [14], in his Master's thesis, addressed the problem of periodic replacement for a multi-server queueing system in which each server is subject to degradation as a function of the mean service rate and a stochastic and dynamic environment. Liu [15] developed a specific class of *m*-failure group replacement policy for an $M/M/c$ queueing system where repair is started as soon as the number of failed machines reaches a predetermined level m. He assumed that the servers are unreliable with identically exponentially distributed failure times and the repair cost consists of a fixed cost and a variable cost proportional to the number of machines repaired. In addition, they assumed that there is a holding cost for each customer in the system per unit of time. In another study, Liu [16] developed three *m*-failure group maintenance models for $M/M/c$ queues. The first model is the basic model with positive repair time and allows server failures during maintenance. The second one is a modified model with positive repair time, but does not allow server failures during maintenance. The last model is considering instantaneous repair. They proved that there exists an optimal group maintenance parameter m^* which can be

used to find the minimal average cost for all three models. Kocuk et al. [17] studied an infinite server queue that is subject to randomly occurring shocks. They assumed that shocks increase the service time of all servers, and can cause further service deterioration. They identified the optimal repair rate that minimizes the expected long-run average cost incurred due to delay and repairs. Wu et al. [18] introduced a controllable repair policy for an infinite-capacity multi-server queueing system in which the servers are assumed unreliable and may fail at any time. The failed servers will be sent to the repair facility only when the number of failed machines in the system achieves a preset threshold value. They developed a cost model to determine the optimal values of the number of servers, service rate and repair rate simultaneously at a minimal total expected cost per unit time.

The common assumption of all of these studies is that the servers (components) are operating in parallel and are selected randomly for service. This implies that the servers (components) are used in a homogeneous manner leading to independent usage and transient deterioration behavior. However, there exist systems in which servers (components) are not being used homogeneously. In such systems, some servers (components) are more likely to be used because of various possible reasons. For example, the layout of the servers (components) in the system makes some components more accessible than others. In this case, there is stochastic dependence among components since the system layout and conditions of components affect the utilization and lifetime distributions of components.

In 2011, Gandhi et al. [19] introduced a Most-Recently-Busy (MRB) routing policy in which an arriving customer is routed to the idle server that was most recently busy. Note that in a system with the MRB routing policy, the usage of a server is stochastically dependent on its and the

other servers status in the system. In this chapter, we consider a queueing system with a similar routing policy as MRB. In particular, we consider a queueing system with n identical servers that are used intermittently and, when not in use, servers are stored in a shared stack. In such cases, customers may find it more convenient to select the server that is on the top of the stack. Similarly, customers finished with a server may find it more convenient to place the server back to the top of the stack. As a result, servers that begin at or near the top of the stack are used more frequently than the servers that are placed lower in the stack. In such a system, a server's utilization is stochastically dependent on its initial location in the stack and the usage of the other servers. Note that although we consider a queueing system with stacked servers, the outcomes of this study are applicable to any queuing systems with the MRB routing policy.

An everyday example of a queueing system with stacked servers or MRB routing policy is a set of shopping carts in a retail store. Often, shopping carts (the servers in this example) are arranged in stacks, so that the customers typically choose the cart at the front (the top in this example) of the stack. After usage, carts are often returned at or near the front of the stack.

The goal of this study is to compute the cumulative, transient utilization and usage time of a server as well as the expected complete reshuffling time in such a queuing system based on the number of servers in the stack, the initial position of the server in the stack, and the arrival and service rates. The expected complete reshuffling time is defined as the minimum time it takes for all the servers in the system to be used. We assume that the servers are identical and numbered based on their initial position in the stack. We also assume that the demand for servers occurs according to a Poisson process and the duration of a single service instance follows an exponential distribution. Customers arriving to a full system balk. This study provides a tool to

investigate the utilization and aging patterns of components in such a system. The outcomes of this study can assist decision makers to have an accurate estimate on the deterioration behavior of components in such a system, and thus maintain a more effective maintenance policy for components and have a more robust estimate of reliability measure of such systems.

The remainder of this chapter is organized as follows. In Section 5.2, the proposed continuous-time Markov chain model is presented. Section 5.3 presents the computational results. Finally, Section 5.4 summarizes the findings and discusses future directions.

5.2 Proposed Model

Consider a queueing system with n identical servers stored in a stack. Let server i $(i = 1, 2, \dots, n)$ be the server whose initial position in the stack is i. We develop a continuous-time Markov chain (CTMC) model to compute the cumulative, transient utilization, and usage time of server i and the expected reshuffling time of the system. The states of the system are represented by $X(t) = (p, q)$, where $p \in \{0, 1, \dots, n\}$ represents the server's position in the stack and $q \in \{0, 1, 2, \dots, n\}$ represents the total number of busy servers at time t. For the first element of the states $p = 0$ represents the situation when the server is busy and is not stored in the stack, and $p = k$ $(k = 1, 2, \dots, n)$ represents the situation when the server is stored in the k^{th} position in the stack. Note that in the remainder of this chapter, the i^{th}server refers to the server with initial position i in the system. We assume that at time zero the system is empty, i.e., $X(t = 0) = (i, 0)$. The total number of states in the proposed CTMC is $N = \frac{n(n+1)}{2} + n$. Figure 5-1 shows the transition diagram of the system along with the transition rates.

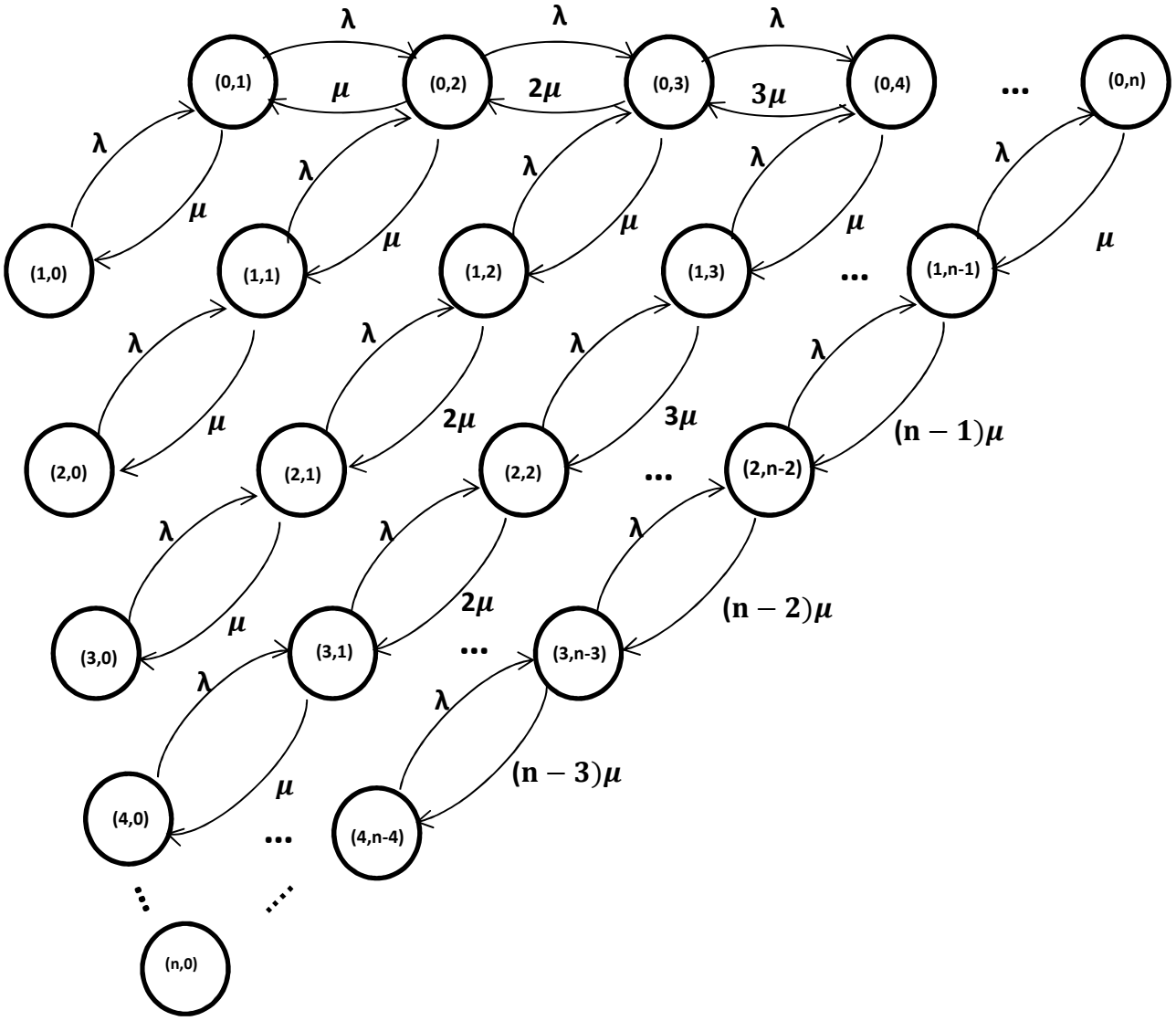


Figure 5-1 Transition diagram of the proposed Markov chain model

5.2.1 Servers' Utilization and Usage Time

Let V_n be the "in use" set, or the set of states in which the server is busy when there are n servers in the system (the states in the first row of the transition diagram in Figure 5-1).

$$V_n = \{(0, i), i = 1, 2, \dots, n\}$$

Let $U_i^n(t)$ be the utilization of the server up to time t. $U_i^n(t)$ is the proportion of time that the server is in the "in use" set, or set V_n, when starting from state $(i, 0)$, as defined in Equation (5-1)

$$U_i^n(t) = \frac{\sum_{s \in V_n} M_{(i,0),s}(t)}{t}, \quad 5\text{-}1$$

where $M_{(i,0),s}(t)$ is the occupancy time of state s up to time t when starting from state $(i, 0)$, and is calculated as

$$M_{(i,0),s}(t) = \int_0^t P_{(i,0)s}(v)dv, \quad 5\text{-}2$$

where $P_{(i,0)s}(v)$ is the transition probability from state $(i, 0)$ to state s at time v. $P_{(i,0)s}(v)$ can be calculated using different methods: e.g., Chapman-Kolmogrov equations, Laplace transform and eigenvalue decomposition of the infinitesimal generator matrix Q of the CTMC model [20]. In eigenvalue decomposition method, the transition probability matrix of a CTMC with infinitesimal generator matrix Q is given by

$$P(t) = Xe^{Dt}X^{-1}, \quad 5\text{-}3$$

where D is a diagonal matrix for which the elements on the diagonal are the eigenvalues of matrix Q, and X is an invertible matrix whose columns are the eigenvectors of matrix Q.

Note that since we consider balking in our queuing system, the effective arrival rate of the system is $\lambda_b = (1 - P_b) \cdot \lambda$, where P_b is the probability that an arriving customer balks and is presented in Equation 5-4.

$$P_b = \lim_{t \to \infty} P_{s(0,n)}(t), \forall s \quad 5\text{-}4$$

Therefore, as shown in Equation 5-5, the server's long run utilization is $\frac{\lambda_b}{n\mu}$, which is smaller than the stationary utilization level in a non-balking system. That is,

$$\lim_{t\to\infty} U_i^n(t) = \frac{\lambda_b}{n\mu} < \frac{\lambda}{n\mu}. \qquad 5\text{-}5$$

A main characteristic of the system that is critical in identifying an optimal maintenance policy is a server's age at time t. To model the deterioration pattern and failure probabilities of the servers, it is critical to know for how long a server has been used in the system. The usage time of a server *i* is the area under the utilization curve up to time t, as presented in Equation 5-6.

$$\mathrm{A}_i^n(t) = \int_0^t U_i^n(v)\mathrm{d}v \qquad 5\text{-}6$$

5.2.2 Complete Reshuffling Time

Reshuffling time of the first k servers in the stack is defined as the time it takes for the system to use the server with initial position k in the stack for the first time. Equivalently, reshuffling time of the first k servers is the first passage time to state $(0, k)$ in the proposed Markov model. We specifically are interested in the "complete reshuffling time" which is the time that it takes for the system to have a complete reshuffling of all servers in the stack, or the time to hit state $(0, n)$ in the model for the first time. This measure is important since once we have a complete reshuffling, all servers become identical from that point forward. In other words, the transient effect of their starting positions in the stack is effectively over.

Let $\mathrm{T}_{ss'}$ be the time to visit state s' for the first time when starting from state s. Reshuffling time of the first k servers in the system or time to visit state $(0, k)$ for the first time is defined in Equation 5-7.

$$T_{(i,0)(0,k)} \equiv \min\{t \geq 0 \,|X(t) = (0,k),\, X(0) = (i,0)\}, i = 1, 2, \dots, k \qquad 5\text{-}7$$

Note that the reshuffling time is independent of i (the initial position of the server in the system) since we assume that the system is empty at time $t = 0$.

Let m_s be the expected first passage time to state $(0, n)$ when starting from state s, i.e.,

$$m_s = E\left[T_{s(0,n)}\right]. \quad 5\text{-}8$$

Then $m = [m_1, m_2, \dots]'$ can be calculated by solving this system of equations

$$Mm + \mathbf{1} = 0, \quad 5\text{-}9$$

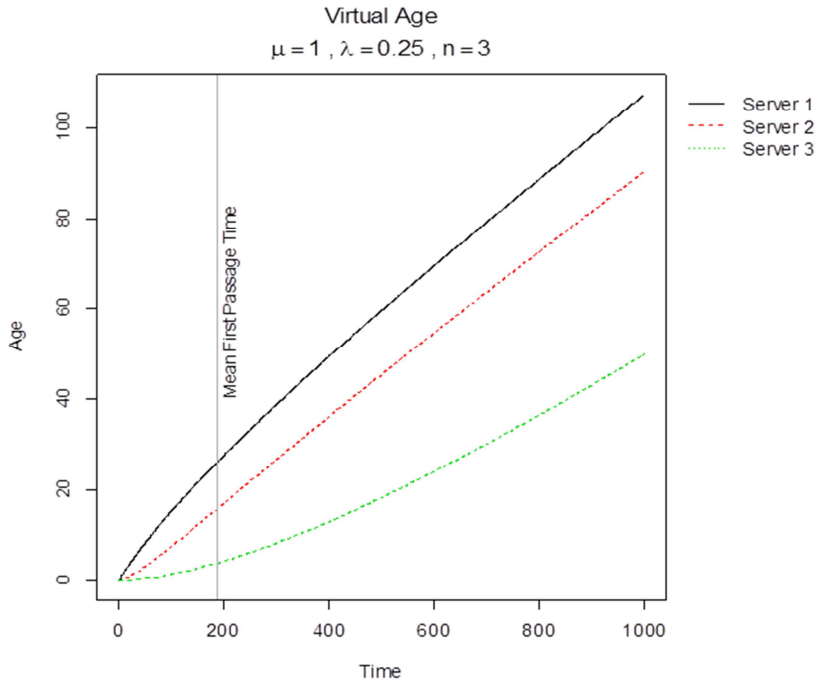
where M is obtained by deleting the row and column corresponding to state $(0, n)$ in matrix Q [20]. Note that when $s = (i, 0), i = 1, 2, \dots, n$, the corresponding value m_s presents the expected complete reshuffling time.

5.3 Numerical Examples

This section illustrates the cumulative transient utilization, virtual age and complete reshuffling time for different queuing systems. Note that in all numerical examples without loss of generality we assume that $\mu = 1$.

Figures 5-2 and 5-3 present the utilization and usage time plots for all servers in the stack along with the expected complete reshuffling time for a queueing system with $n = 3$ servers and arrival rates $\lambda = 0.25, \text{and } \lambda = 1$, respectively.

a) Usage time ($\lambda = 0.25$)



b) Utilization ($\lambda = 0.25$)

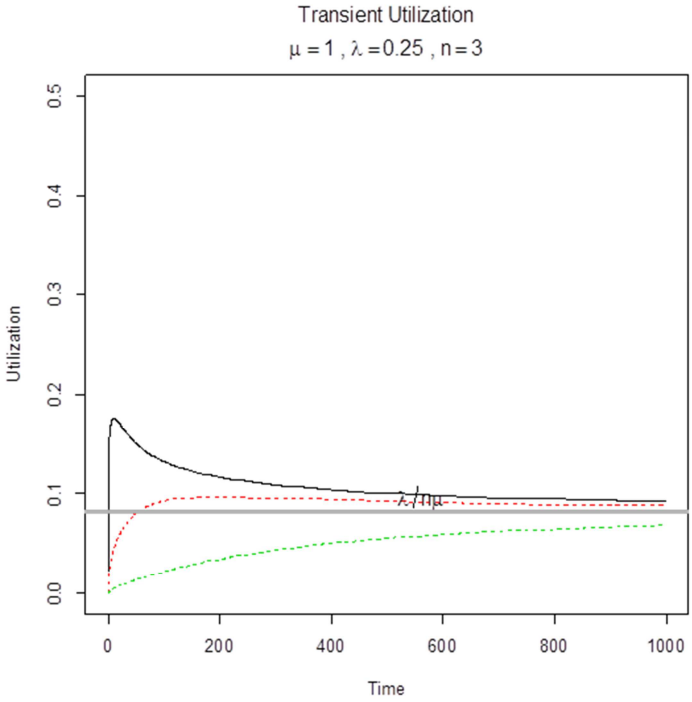
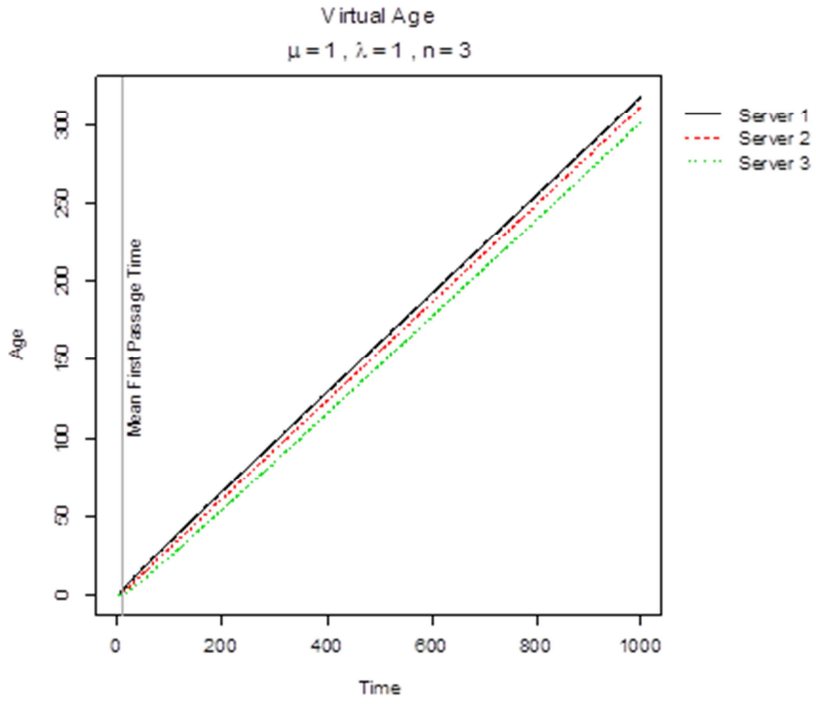


Figure 5-2 a) Usage time and b) utilization of servers in a queuing system with $n = 3$ and arrival rate 0.25

a) Usage time ($\lambda = 1$)



b) Utilization ($\lambda = 1$)

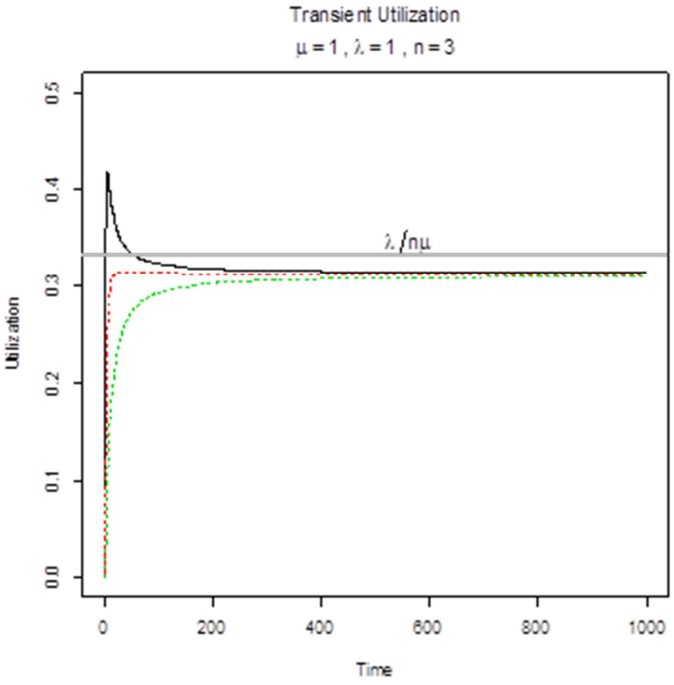
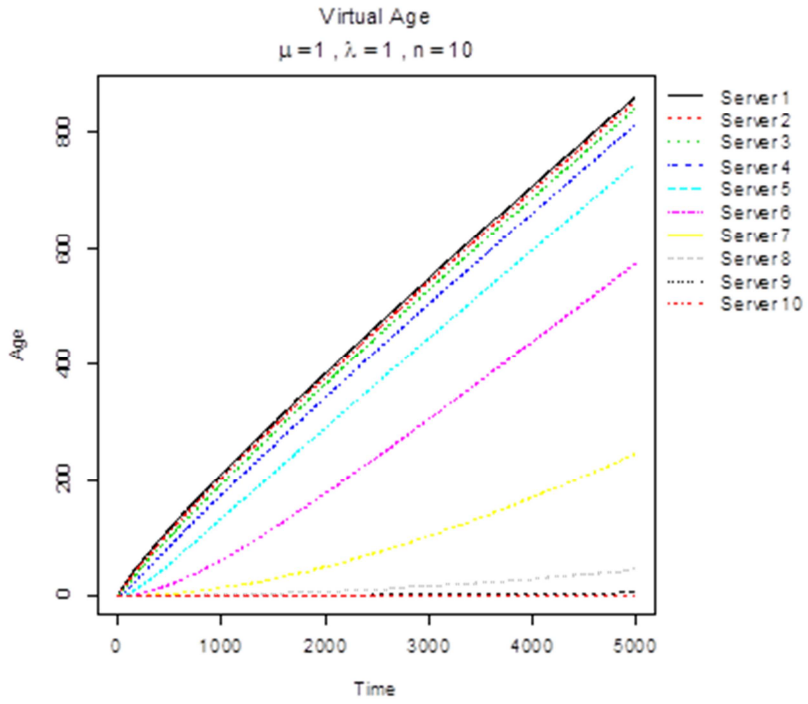


Figure 5-3 a) Usage time and b) utilization of servers in a queuing system with $n = 3$ and arrival rate 1

The corresponding balking probabilities for the two different arrival rates are 0.002, and 0.062, respectively. The horizontal dark gray lines in the utilization plots show the servers' stationary utilization in a non-balking system, which, as expected, is larger than the servers' stationary utilization in the system under study. Note that in the first case where $\lambda = 0.25$, the system has not yet reached the steady state after 1000 time units and it has a significantly longer mean complete reshuffling time than the case with arrival rate $\lambda = 1$ (188 time units versus 8 for the second case).

Figures 5-4 to 5-6 present the results for a queueing system with $n = 10$ servers and arrival rates $\lambda = 1,\ \lambda = 2, \text{and } \lambda = 3$) for the first 5000 time units, respectively. For none of the cases, the system has reached the steady state after 5000 unites of time. All the three cases have negligible balking probabilities. The complete reshuffling times for these cases are 1,112,083, 3,410.56, and 192.94, respectively. Note that for the first case with $\lambda = 1$ at time $t = 5{,}000$, the ages of the first four servers in the stack are over 800 each. However, the last two servers in the stack have not been used yet. In the case with $\lambda = 3$, however, the servers have a more consistent aging pattern.

a) Usage time ($\lambda = 1$)



b) Utilization ($\lambda = 1$)

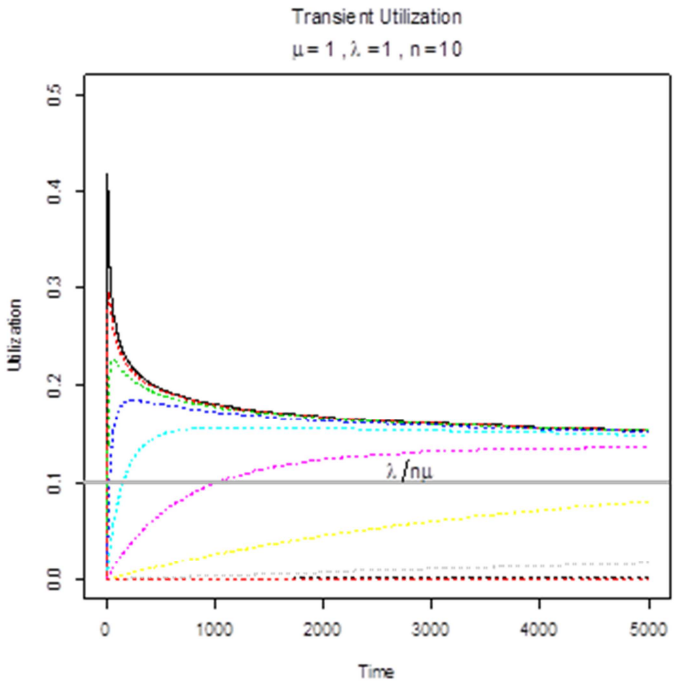
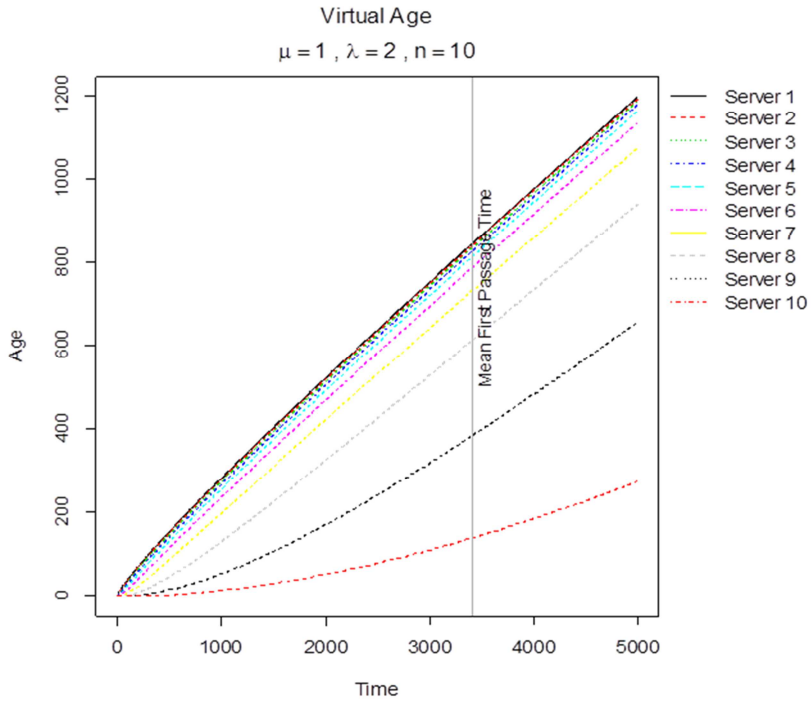


Figure 5-4 a) Usage time and b) utilization of servers in a queuing system with $n = 10$ and arrival rate 1

a) Usage time ($\lambda = 2$)



b) Utilization ($\lambda = 2$)

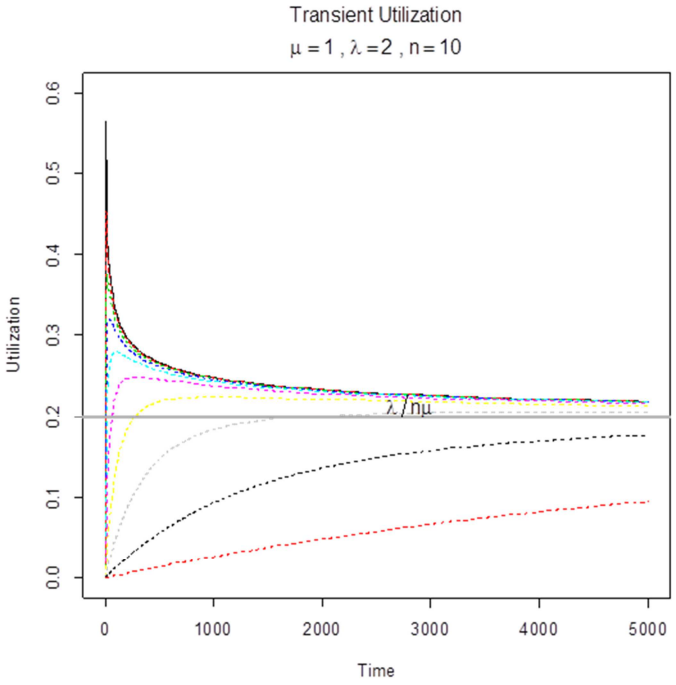
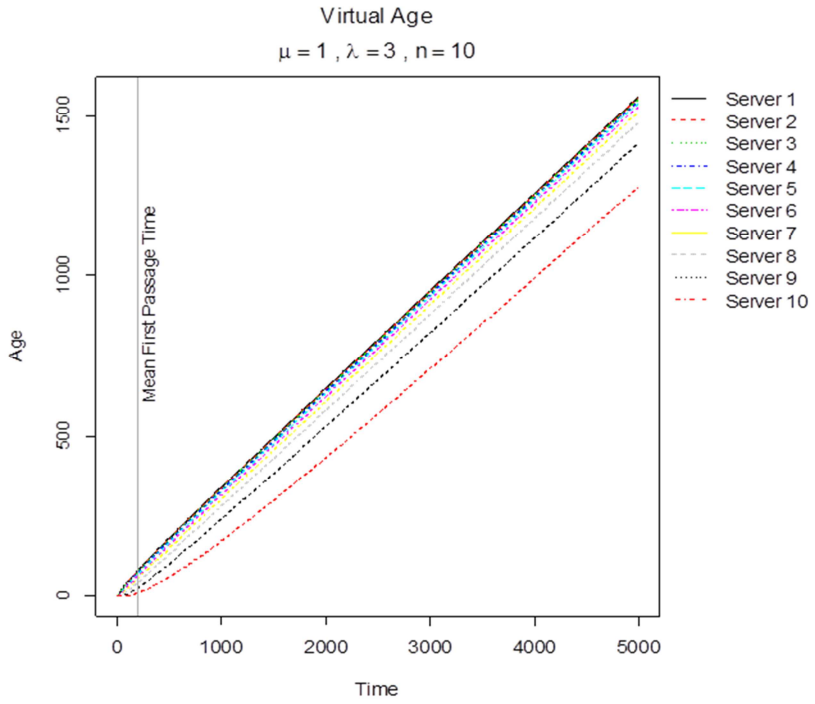


Figure 5-5 a) Usage time and b) utilization of servers in a queuing system with $n = 10$ and arrival rate 2

a) Usage time ($\lambda = 3$)



b) Utilization ($\lambda = 3$)

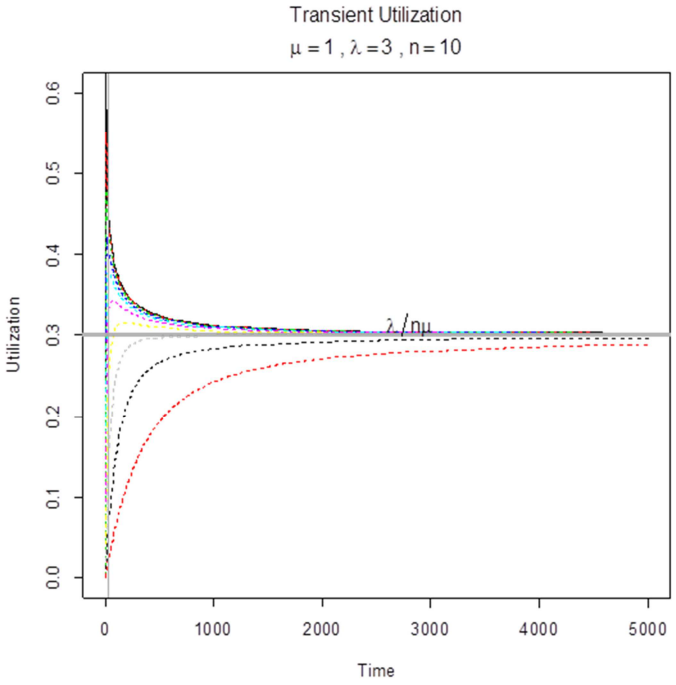


Figure 5-6 a) Usage time and b) utilization of servers in a queuing system with $n = 10$ and arrival rate 3

5.4 Conclusion

All the previous studies on unreliable multi-server queuing systems assume that servers' utilizations are independent from each other and an arriving customer selects one of the available servers randomly. However, there exist systems in which the customers routing pattern is non-random and customers select a server according to a defined pattern. An example of a non-random routing policy is the Most Recently Busy (MRB) policy, in which an arriving customer selects the idle server that was the most recently busy.

In this chapter, we considered an unreliable queuing system with the MRB routing policy. In particular, we considered a queueing system for which the servers are stored in a stack when they are not in use. In such systems, an arriving customer picks the idle servers on the top of the stack for service. Similarly, when the service is over, the server is returned to the top of the stack. In such a system servers have heterogeneous utilization and deterioration behaviors because of the stochastic dependence of a server usage on its position in the stack and the utilization of the other servers in the system.

To maintain such systems more effectively, the decision maker needs to have an understanding of the utilization behavior and deterioration pattern of each server in the system. In this chapter we proposed a model to characterize and quantify the critical feature in estimating the deterioration of each server. We developed a continues-time Markov chain (CTMC) model to measure the transient utilization and usage time of each server based on the initial position of the server, the arrival rate and the service rate in the system. We also proposed a formula to compute the system's expected complete reshuffling time, i.e., the expected time to use the last server in the stack for the first time. This measure is of interest since when the last server in the stack is

used for the first time, all servers become identical in terms of their possible position in the stack from that point on.

This study provides insight to the decision maker to have a more efficient maintenance policy for a queueing system with the MRB routing policy, or more specifically a queueing system with stacked servers. The model presented in this study helps the decision maker to effectively estimate the viral age of each server in the system, and accordingly approximate the deterioration of a specific server and derive the optimal maintenance policy based on the estimated deterioration.

A future direction for this research is to study different individual and group maintenance policies such as age replacement, block replacement, m–failure, or policy $(m - T)$ policy (the combination of block replacement and m–failure policy) , and perform a cost analysis to derive the optimal maintenance policy for such systems.

References

1. Dekker, R., Wildeman, R. E., & van der Duyn Schouten, F. A. (1997). A review of multi-component maintenance models with economic dependence. Mathematical Methods of Operations Research, 45(3), 411-435.

2. Nicolai, R. P., & Dekker, R. (2008). Optimal maintenance of multi-component systems: a review (pp. 263-286). Springer London

3. Wang, H. (2002). A survey of maintenance policies of deteriorating systems. European journal of operational research, 139(3), 469-489.

4. Choudhury, G., & Tadj, L. (2011). The optimal control of an Mx/G/1 unreliable server queue with two phases of service and Bernoulli vacation schedule. Mathematical and Computer Modelling, 54(1), 673-688.

5. Federgruen, A., & So, K. C. (1990). Optimal maintenance policies for single-server queueing systems subject to breakdowns. Operations Research, 38(2), 330-343.

6. Almási, B., Roszik, J., & Sztrik, J. (2005). Homogeneous finite-source retrial queues with server subject to breakdowns and repairs. Mathematical and Computer Modelling, 42(5), 673-682.

7. Sztrik, J., Almási, B., & Roszik, J. (2006). Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs. Journal of Mathematical Sciences, 132(5), 677-685.

8. Falin, G. I. (2008). The M/M/1 retrial queue with retrials due to server failures. Queueing Systems, 58(3), 155-160.

9. Mitrany, I. L., & Avi-Itzhak, B. (1968). A many-server queue with service interruptions. Operations Research, 16(3), 628-638.

10. Neuts, M. F., & Lucantoni, D. M. (1979). A Markovian queue with N servers subject to breakdowns and repairs. Management Science, 25(9), 849-861.

11. Wang, K. H., & Chang, Y. C. (2002). Cost analysis of a finite M/M/R queueing system with balking, reneging, and server breakdowns. Mathematical Methods of Operations Research, 56(2), 169-180.

12. Wu, C. H., & Ke, J. C. (2010). Computational algorithm and parameter optimization for a multi-server system with unreliable servers and impatient customers. Journal of computational and applied mathematics, 235(3), 547-562.

13. Gharbi, N., & Dutheillet, C. (2011). An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. Computers & Mathematics with Applications, 62(6), 2535-2546.

14. Chapin, P. (2004). Age replacement and service rate control of stochastically degrading queues, (Master's thesis), Air Force Institute of Technology.

15. Liu, G. S. (2004). M-failure group replacement model for queueing systems with unreliable and repairable servers. Journal of the Chinese Institute of Industrial Engineers, 21(5), 423-431.

16. Liu, G. S. (2012). Three m-failure group maintenance models for M/M/N unreliable queuing service systems. Computers & Industrial Engineering, 62(4), 1011-1024.

17. Kocuk, B., Sivrikaya, Ö. E., & Baykal-Gürsoy, M. (2013). Maintenance of infinite-server service systems subjected to random shocks. Computers & Industrial Engineering, 66(1), 86-94.

18. Wu, C. H., Lee, W. C., Ke, J. C., & Liu, T. H. (2014). Optimization analysis of an unreliable multi-server queue with a controllable repair policy. Computers & Operations Research, 49, 83-96.

19. Gandhi, A., Harchol-Balter, M., Raghunathan, R., & Kozuch, M. A. (2011). Distributed, robust auto-scaling policies for power management in compute intensive server farms. In Open Cirrus Summit (OCS), 2011 Sixth (pp. 1-5). IEEE.

20. Kulkarni, V.G. (2010). Modeling and analysis of stochastic systems. 2nd ed. Chappman & Hall