

12-2019

QUANTIFYING THE FINANCIAL & CLINICAL IMPACT OF ANNUAL VS. BIENNIAL MAMMOGRAPHY SCREENING

DANMENG HUANG

Follow this and additional works at: https://digitalcommons.library.tmc.edu/uthsph_dissertsopen



Part of the [Community Psychology Commons](#), [Health Psychology Commons](#), and the [Public Health Commons](#)

QUANTIFYING THE FINANCIAL & CLINICAL IMPACT OF ANNUAL VS. BIENNIAL
MAMMOGRAPHY SCREENING

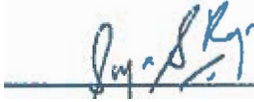
by

DANMENG HUANG, MPA

APPROVED:



STEPHEN H. LINDER, PHD



SUJA S. RAJAN, PHD



DAVID R. LAIRSON, PHD



SHARON H. GIORDANO, MD



WENYAW CHAN, PHD



DEAN, THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH

Copyright
by
Danmeng Huang, MPA, PhD
2019

QUANTIFYING THE FINANCIAL & CLINICAL IMPACT OF ANNUAL VS. BIENNIAL
MAMMOGRAPHY SCREENING

by

DANMENG HUANG
MPA, Cornell University, 2014

Presented to the Faculty of The University of Texas

School of Public Health

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH
Houston, Texas
December, 2019

DEDICATION

To my family, and our fur babies

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my committee for guiding me throughout the journey. They inspired me of my topic, provided me data access, advised me with data analytics, and supported me when I struggled with unsatisfactory results. The timely completion of my dissertation would not have been possible without the support and nurturing of Dr. Scott B. Cantor, who read and edited the very first draft of my work.

I also had great pleasure of working with the statistical analysts in the Department of Health Services Research at The University of Texas MD Anderson Cancer Center, who shared with me valuable experiences and guidance in claims data analysis.

Part of my study was supported by the pre-doctoral fellowship from the University of Texas School of Public Health Cancer Education and Career Development Program (National Cancer Institute/NIH Grant R25 CA57712) and Institute for Health Policy.

I very much appreciate the profound belief in my abilities from my family. Special thanks to my “manager”, Yitian Xu, and my cat, Yesi, for their warm companion.

QUANTIFYING THE FINANCIAL & CLINICAL IMPACT OF ANNUAL VS. BIENNIAL MAMMOGRAPHY SCREENING

Danmeng Huang, PhD, MPA
The University of Texas
School of Public Health, 2019

Dissertation Chair: STEPHEN H. LINDER, PHD

Current guidelines for screening mammography recommend different start ages and intervals for women with average cancer risk. Longer intervals between mammograms allow more time for a tumor to grow, to become symptomatic and clinically detectable, and more likely to be at advanced stage. Existing literature on the associations between mammography screening frequency and risk of more advanced breast cancer have mixed results. Studies have showed that advanced breast cancer costs more to treat, but real-world cost estimates following different mammography screening frequencies are unavailable. To fill the gaps, this dissertation aimed to quantify the clinical and financial impact following annual versus biennial screening.

Commercial claims database provides rich real-world information on diagnoses, medical resources used and the associated costs. To supplement claims database with breast cancer stage information, this study first developed and validated an algorithm with classification and regression tree method using SEER-Medicare data. The performance was measured with sensitivity, specificity, positive predictive value, negative predictive value and area under the receiver operating characteristic curve. The algorithm was then applied to The

MarketScan® Commercial Claims and Encounters Database to estimate breast cancer stage at diagnosis. Incident breast cancer cases identified in the MarketScan database were categorized as annual, biennial and non-screeners based on their pre-diagnosis screenings. Partial proportional odds model was used to estimate the odds ratio of having more advanced breast cancer. Stratified analysis by age was also conducted. For the three screening groups, total healthcare costs, insurer costs, and out-of-pocket costs adjusted by generalized linear model with gamma distribution and log-link function were reported.

The staging algorithm had improved performance than others, especially in the prediction of non-invasive cases, early stage cancers and metastases. Generally, regular screening protects women from more advanced breast cancer. Compared to biennial screening, annual screening was associated with a reduced risk of later stage invasive cancers for both women in their 40s and older. For health insurance payers, there were cost savings in healthcare costs with regular screening, especially for annual screening. Cost reduction was more obvious among women aged 40-49. Compared to insurer's costs, out-of-pocket costs borne by patients were minimal. Although this study showed both clinical and financial benefits in annual versus biennial screening, the optimal screening frequency should be an individual decision weighing both the benefits and harms.

TABLE OF CONTENTS

List of Tables	i
List of Figures	ii
List of Appendices	iii
Background	1
Introduction	1
Literature Review	4
Staging Algorithms	4
Screening Pattern and Stage at Diagnosis	7
Screening Pattern and Healthcare Costs	8
Public Health Significance	9
Hypothesis, Research Question, Specific Aims or Objectives	10
Journal Article	11
A Claim-Based Algorithm Predicting Breast Cancer Stage at Diagnosis	11
Medical Care	11
Journal Article	44
Significant Association of Annual versus Biennial Screening with Later-Stage Breast Cancer at Diagnosis among Women Aged 40-64	44
Journal Article	68
Healthcare Costs of Breast Cancer by Mammography Screening Behavior among Commercially Insured Women Aged 40-64	68
Conclusion	89
References	91

LIST OF TABLES

Table 1: Summary of mammography screening recommendations by organization.	1
Table 1: Patients' Characteristics of Breast Cancer Cohorts.....	19
Table 2: Classification Matrix for Validation Set.....	21
Table 1: Patient's characteristics by screening group.	51
Table 2. Multivariable associations of mammography screening pattern with breast cancer stage at diagnosis.	53
Table 3. Multivariable associations of mammography screening pattern with breast cancer stage at diagnosis by age.	57
Table 1: Patient's characteristics by screening group.	75
Table 2: Model Adjusted Health Care Costs by Group in the First Year After the Index Date*	77
Table 3: Estimated Cost Ratios for Mean Costs in the First Year after the Index Date*	80

LIST OF FIGURES

Figure 1: Study flowchart.	48
Figure 1: Study flowchart.	72

LIST OF APPENDICES

Appendix Table 1: Claims codes to determine selected variables.....	27
Appendix Table 2: NCCN Recommended Treatment Regimens.	31
Appendix Table 3: Final Staging Algorithm.	34

BACKGROUND

Introduction

In United States, breast cancer is overall the most common cancer and the second leading cause of cancer deaths among women (Street, 2019). Mammography, a non-invasive screening tool, causes minimal pain as a result of breast compression. Currently, the U.S. Preventive Services Task Force (USPSTF) recommends biennial screening mammography for average-risk women 50–74 years old and indicates that biennial screening mammography for women younger than 50 should be an individual decision, weighting benefits and harm (Siu, 2016). However, not all organizations agree. For example, the American Cancer Society (ACS) recommends annual mammography for women aged 45–55 (Oeffinger et al., 2015), while others recommend starting at 40 (Table 1).

Table 1: Summary of mammography screening recommendations by organization.

	U.S. Preventive Services Task Force (2016)	American Cancer Society (2015)	American College of Obstetricians and Gynecologists (2011)	American College of Radiology/ SBI (2010)	AMA (2012)	NCCN (2016)	International Agency for Research on Cancer (2015)	American College of Physicians (2015)	American Academy of Family Physicians (2016)
Start Age	50 40-49 Individual decision	45 40-45 individual choice	40	40	40 40-49 Individual decision	40	50	50 40-49 Individual decision	50 40-49 Individual decision
Screening Interval	Biennial	50-54 annual; >55 annual/ biennial	Annual	Annual	Annual	Annual	N/A	Biennial	Biennial

Stop Age	74	When life expectancy is < 10 years	>75 individual choice	When life expectancy is < 5-7 years	When life expectancy is < 10 years	N/A	70	74	74
----------	----	------------------------------------	-----------------------	-------------------------------------	------------------------------------	-----	----	----	----

The conflicting recommendations has created confusion and debate in the public, especially for women aged 40-49. Theoretically, early, intense screening would identify breast cancer sooner and thereby yield better survival rates and lower treatment costs. Randomized, controlled trials and screening program data have shown statistically significant mortality reduction among women older than 40 (Alexander et al., 1999; Andersson et al., 1988; Andersson & Janzon, 1997; Bjurstam et al., 2003; Duffy et al., 2002; Hendrick, Smith, Rutledge III, & Smart, 1997; Lennarth Nyström et al., 2002; L Nyström et al., 1993; Tabar, Faberberg, Day, & Holmberg, 1987). Although breast cancer incidence in women aged 40–49 is only half that of women in their 50s, 34% of total life years lost to breast cancer are in women diagnosed with breast cancer in their 40s (Moss, 2004). While the screening benefit for women aged 40–49 has been established, mammography has higher false-negative and false-positive rates for young women with dense breast tissue (Kerlikowske et al., 2013; Siu, 2016). False-positive rates are also higher in women who have had breast biopsies, family history of breast cancer and who currently take estrogen (Elmore et al., 1998; Hubbard et al., 2011; Marmot et al., 2013). False-positive results can have psychological (stress, anxiety, awareness), behavioral (healthcare hyper-utilization) and economic (unnecessary diagnostic tests and biopsies) impact (Aro, Absetz, van Elderen, van der Ploeg, & van der Kamp, 2000;

Barton et al., 2001; Brewer, Salz, & Lillie, 2007; McGovern et al., 2004; Tosteson et al., 2014). Another known harm is over-diagnosis of noninvasive breast cancers that are unlikely to become clinically evident during the patient's lifetime. Therefore, the weighing of benefits and harms is important when making recommendations.

In supplement to RCTs, modeling studies have investigated the long-term effects of mammography screening in terms of mortality, life years gained and potential harm. For example, six mathematical models developed by the Cancer Intervention and Surveillance Modeling Network of the US National Institutes of Health (NIH) (Mandelblatt et al., 2009) evaluated 20 screening strategies in terms of initiation age and screening interval and found that biennial screening maintained 81% of the benefit of annual screening in mortality while cutting false-positive results by half. Mandelblatt and coworkers also found that beginning screening at age 40 gave only a minimal gain in morality but a larger benefit in life years gained.

Still, simulated findings needs confirmation from observational studies. Although reduced breast cancer-related mortality is the ultimate goal of mammography, it is not always an outcome measure available to researchers. Intermediate measures such as stage at diagnosis are therefore useful to evaluate the effect of screening (Day, 1989; Sant et al., 2003). Treatments, prognosis and survival are largely determined by cancer stage. Specifically, breast cancer stage affects the selection and timing of surgery, radiotherapy, and chemotherapy, as well as overall and disease-free survival (Woodward et al., 2003). In lieu of unethical RCTs, observational studies have been using stage at diagnosis as surrogate endpoint to investigate the clinical impact of mammography screening frequencies.

By changing the cancer stage distribution, mammography screening frequency also has implications on breast cancer treatment costs. The more advanced cancer stage, the higher healthcare costs. Cost analysis provides a different perspective to evaluate the benefit of mammography screening for women and healthcare payers. Studies have estimated the economic burden of breast cancer, including direct and indirect costs in insured and underinsured populations. However, only two modeling studies looked at costs associated with different mammography screening patterns (Carter, Castro, Kessler, & Erickson, 2005; Farley et al., 2015). Therefore, more observational studies are needed to inform the financial impact based on real-world data.

Real-world administrative claims data is well known for the richness in payment related diagnoses and the use of medical resources, but is limited in clinical information such as cancer stage at diagnosis. To overcome this disadvantage, predictive algorithms are developed to infer breast cancer stage at diagnosis based on medical claims.

Literature Review

Staging Algorithms

Few researchers have used medical claims data to develop algorithms to predict breast cancer stage at diagnosis. Cooper et al. (1999) examined the ICD-9 diagnosis codes in Medicare claims and compared the staging with SEER registries (Cooper et al., 1999). They found that Medicare alone overestimated the proportion of localized tumors and underestimated regional stage disease. For regional breast cancer, the sensitivity and positive predictive value (PPV) are 61.6% and 84.6 % respectively; for distant breast cancer, the

sensitivity and PPV are 60.2% and 58.4%. Similarly, Chawla et al. (2014) proved the limited validity of using ICD-9 codes for secondary malignant neoplasms to specific organs and reported the performance for local (sensitivity 98.6%, specificity 43.9%, PPV 78.5%, and negative predictive value (NPV) 93.8%), regional (sensitivity 35.3%, specificity 98.4%, PPV 88.8%, and NPV 80.8%), and distant breast cancer (sensitivity 51%, specificity 98.3%, PPV 65.8%, and NPV 96.9%) (Chawla et al., 2014). Their study re-emphasized the need to validate algorithms using diagnoses, procedures, and drug claims information to identify breast cancer stage more accurately. Partridge et al. (2008) identified codes of early versus advanced breast cancer in the claims and used a scoring system to classify breast cancer cases; their scoring system resulted in many indeterminate cases and was not validated (Partridge et al., 2008). Based on the previous work of Cooper, Smith and colleagues (2010) used SEER-Medicare data and developed two logistic models: identifying metastatic breast cancer (sensitivity 81%, specificity 89%, PPV 24%, and NPV 99%) and distinguishing stages I/II (early) breast cancer from stage III (unconditional sensitivity 83%, specificity 78%, PPV 98%, and NPV 31%) (Smith, Shih, Giordano, Smith, & Buchholz, 2010). Their first model significantly overestimated stage IV cases and therefore affected the accuracy of the second model. Although Smith et al. reported good performance for the stages I/II model, the statistics didn't adjusted for misclassification in the first model. In addition, the researchers only looked at invasive breast cancer, and left DCIS cases out of their study. Also because Smith and coworkers (2010) used SEER-Medicare data from 1992–2002, drug information was not available for algorithm development. Foley and colleagues (2013) recommended in their review that the Smith algorithm be updated with additional codes from Nordstrom and

higher cut points be tested to maximize PPV, but they did not conduct validation (Foley, Shi, Girvan, Ward, & Lipscomb, 2013). Using a different technique-classification and regression tree (CART) models, Nordstrom et al. (2012) tried to identify metastatic breast cancer using outpatient electronic medical record (EMR)-linked claims data. They used diagnoses, procedures, and drug variables from the claims (Nordstrom, Whyte, Stolar, Mercaldi, & Kallich, 2012). The researchers reported sensitivity at 62%, specificity at 97%, PPV at 75%, NPV at 95%, and area under the receiver operating characteristic curve (ROC AUC) at 82%. There were generalizability concerns because of their study design. For example, they applied strict inclusion criteria regarding specialists issuing the diagnosis claim to ensure consistent treatment reporting, which may not be applicable in other claims databases. Second, they only include claims made within 60 days after diagnosis, and therefore missed potential out-of-window predictors. Third, using only outpatient EMR-linked claims data missed the opportunity to improve algorithm performance with claims from other points of service. Moreover, some variables used in those algorithms were database specific (i.e., EMR information may be coded differently than claims). Therefore, the generalizability of Nordstrom's algorithm was limited. Whyte et al. (2015) evaluated generic algorithms (for all cancers) and tumor-specific algorithms (breast cancer only) that identified metastatic disease, and they tested variations of the algorithms in terms of timing, temporal spacing of claims and exclusion of other cancers (Whyte, Engel-Nitz, Teitelbaum, Gomez Rey, & Kallich, 2015). The researchers found that although the generic algorithm had higher sensitivity due to fewer missed patients, its specificity, PPV and NPV were lower compared with tumor-specific algorithms. Breast cancer-specific algorithms with two breast cancer claims that

were more than 30 days apart had the best overall performance. However, the algorithms included only inpatient and outpatient claims but no information of medications. In conclusion, we found that previous studies emphasized on identifying the stage of invasive breast cancer, especially metastatic breast cancer. Only Blumen et al. (2016) tried to identify DCIS cases based on NCCN treatment guidelines, but their algorithm was empirical-based and not validated (Blumen, Fitch, & Polkus, 2016). As Chawla and colleagues pointed out, all diagnoses, procedures, and drug claims should be included when developing algorithms (2014).

Screening Pattern and Stage at Diagnosis

A longer interval between screening mammograms allows more time for a tumor to grow, to become symptomatic and clinically detectable, and therefore, more likely to be at an advanced stage. Freedman et al. (2003) reviewed clinical records of women with breast cancer and reported that those in their 40s were more likely to have DCIS discovered if they underwent annual screening mammography, compared with no or less frequent screening (Freedman et al., 2003). Similarly, White et al. (2004) used data from seven mammogram registries that participate in the National Cancer Institute–funded Breast Cancer Surveillance Consortium (BCSC) and reported the elevated risk of screening biennially versus annually in later stage invasive breast cancer in women aged 40–49 (White et al., 2004). However, Goel et al. (2007) only found higher probability of being diagnosed with advanced breast cancer associated with screening intervals greater than 2.5 years, but not for annual or biennial screening (Goel, Littenberg, & Burack, 2007). Concordant with Goel’s findings, updated BCSC studies found no such association among women older than 40 (Dittus et al., 2013;

Hubbard et al., 2011; Kerlikowske et al., 2013; Miglioretti et al., 2015; O'Meara et al., 2013).

One major limitation of the BCSC studies was that the registries only have women who undergoes regular mammography screening already, which can be a different cohort than the general population. Another limitation was the dichotomization of outcomes (favorable vs. unfavorable), which left out valuable information with the ordinal nature of cancer stages 0-IV.

Screening Pattern and Healthcare Costs

Very limited evidence exists on the financial impact of mammography screening patterns. Farley et al. (2015) used a linear regression model to simulate the effect of screening patterns recommended by the USPSTF and American Community Survey (ACS) on stage, survival and cost of treatment among African-American women in an urban public hospital (Farley et al., 2015). Using breast cancer patients' information obtained from chart review, the researchers estimated new times of cancer diagnosis in USPSTF and ACS scenarios. Based on that estimation, they calculated the stage at diagnosis and likelihood of survival. The average cost of treatment for a given breast cancer stage was used to estimate incremental per-patient costs for all screening scenarios. From the healthcare institution's perspective, Farley (2015) concluded that following the ACS guidelines would have saved \$3745 compared with observed data, and saved \$5528 compared with USPSTF, respectively. The lack of real-world evidence calls for more observational studies in cost analysis.

Public Health Significance

This dissertation updated the breast cancer staging algorithms using diagnoses, procedures, drug claims and recommended treatment regimens. With validated breast cancer stage estimation, administrative healthcare claims data can be more valuable in epidemiologic, health services, and outcomes cancer research.

This dissertation also re-examined the conflicting findings of BCSC studies, using a national insurance claims database, by testing different definitions of screening interval. Compared to clinical trials, claims database provide real-world effectiveness of different mammography screening intervals. In addition, the financial burden was estimated from the perspectives of private insurance payers as well as breast cancer patients who follow different mammography screening patterns, which, to my best knowledge, has not been studied.

Hypothesis, Research Question, Specific Aims or Objectives

Aim 1: Assess the impact of mammography screening patterns on identifying stage of diagnosis among women with breast cancer using estimated staging information in commercial claims data.

Aim 1a: Develop and validate the performance of claim-based algorithms that can predict breast cancer stages from claims-based data.

Aim 1b: Estimate risks of more advanced breast cancer associated with annual and biennial screening mammography.

Hypothesis: Shorter screening interval (annual vs. biennial) detects earlier and more treatable breast cancers.

Aim 2: Assess the financial impact of mammography screening patterns on total healthcare costs among women with breast cancer.

JOURNAL ARTICLE

A Claim-Based Algorithm Predicting Breast Cancer Stage at Diagnosis

Medical Care

Abstract

Background: Administrative claims databases are important sources for epidemiological, healthcare utilization and cost studies. However, the lack of clinical information on cancer stages limits the ability to control for confounders in cancer research.

Objective: To develop and validate a predictive algorithm to identify breast cancer stages using treatment information obtained from claims data.

Research Design: The SEER-Medicare database contains linked “gold standard” cancer stages from SEER cancer registries and claims data from Medicare. We built a classification tree model based on variables identified through diagnosis, procedure and medication codes from inpatient, outpatient, physician and pharmacy claims.

Subjects: Female fee-for-service patients older than 66 and diagnosed with stage0-IV breast cancer between 2008 and 2013 were randomly assigned to training and validation sets. A classification tree model was based on variables identified through diagnosis, procedure and medication codes from inpatient, outpatient, physician and pharmacy claims. The performance of the classification model was measured by sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic curve (AUC).

Results: The algorithm identified stage 0 breast cancer with 85% sensitivity, 97% specificity, 84% PPV, 97% NPV, and 91% AUC; stage I with 83% sensitivity, 75% specificity, 74% PPV, 85% NPV, and 79% AUC; stages I/II with 93% sensitivity, 73% specificity, 90% PPV, 79% NPV, and 83% AUC; stages II/III with 63% sensitivity, 91% specificity, 78% PPV, 83% NPV, and 77% AUC; stage IV with 79% sensitivity, 99% specificity, 67% PPV, 99% NPV, and 89% AUC.

Conclusions: Our algorithm had excellent predictive power for stage 0 and IV breast cancer, and good performance for stage I cases. Stage II and III identification were less successful due to the similarities in treatment recommendations. The overall accuracy significantly improved with combined estimations of stages I and II, as well as stages II and III.

Introduction

Breast cancer stage at diagnosis is a key predictor of prognosis and survival. Generally, non-invasive (stage 0) and early stage invasive breast cancers (stages I and II) have better chance of survival than more advanced stage (stage III and IV). For example, the 5-year breast cancer-specific survival for women with stage I, II and III breast cancers are as high as 99.3% (95% CI 99.2%-99.4%), 94.3% (95% CI 93.6%-94.9%), 88.0% (95% CI 86.8%-89.0%), respectively (Weiss et al., 2018). On the other hand, the 5-year breast cancer-specific survival for stage IV cases is only 35.5% (95% CI 32.7%-38.3%).

Breast cancer stage is also widely used as a covariate or as an inclusion and exclusion criterion in epidemiologic studies of patients with breast cancer. Administrative medical

claims databases, an important source of population-based data, usually lack clinical information such as cancer stage. This major disadvantage has limited the use of these datasets in retrospective outcome-based oncological research.

According to the National Comprehensive Cancer Network (NCCN) breast cancer treatment guidelines (Gradishar et al., 2018), different treatment patterns are recommended based on the Tumor, Node, Metastasis (TNM) staging of breast cancer (Edge & Compton, 2010). Therefore, it is feasible to predict breast cancer stage at diagnosis based on the treatment received. However, predictive models developed based on clinical insight alone can have accuracy less than expected due to the similarities in treatment recommendations between stages, and potential deviations in clinical practice (Chawla et al., 2014).

Only a few claim-based algorithms predict invasive breast cancer stage I-IV at diagnosis. Cooper et al. (1999) and Chawla et al. (2014) both reported limited validity of using ICD-9 diagnosis codes alone in Medicare claims (Chawla et al., 2014; Cooper et al., 1999). Both studies were developed through the linked Surveillance, Epidemiology and End Results (SEER)-Medicare data and used the summary SEER stages as reference standard. However, SEER staging is less helpful in informing survival because their definition of stages are broader. Based on Cooper's single predictor models, Smith et al. (2010) expanded the models by including other predictors including demographic, tumor and treatment characteristics (Smith, Shih, Giordano, Smith, & Buchholz, 2010). Two multinomial logistic regression models were developed to first identify stage IV cases from the entire study cohort, then separate stage I/II from stage III in the remaining cases. Although the second model had decent accuracy, it was conditional on the poorer performance of the first model.

Other algorithms were developed to predict more specific breast cancer stage such as metastatic breast cancer (Nordstrom, Whyte, Stolar, Mercaldi, & Kallich, 2012; Whyte, Engel-Nitz, Teitelbaum, Gomez Rey, & Kallich, 2015), breast cancer bone metastases (Sathiakumar et al., 2017), and early vs. advanced stage with positive Estrogen Receptor (ER) and negative human epidermal growth factor receptor 2 (HER2) breast cancer (Beachler et al., 2019). However, to our knowledge a predictive model for stage 0 breast cancer does not exist.

Previous predictive models developed based on SEER-Medicare data failed to include pharmacy claims, which indicates a missed opportunity to improve the accuracy of predicting breast cancer stage. From a large pool of candidate variables including diagnosis codes, comorbidities, diagnostic and surgical procedures, treatments regimens, and medications, this study aims to develop an improved predictive model for breast cancer stage 0-IV.

Methods

Data Sources

We used the linked SEER-Medicare data to develop and validate an algorithm. SEER-Medicare links cancers cases ascertained through participating cancer registries to claims data from Medicare for people over 65 years old, representing approximately 26% of the US population (Engels et al., 2011). The SEER data contain cancer incidence (month and year of diagnosis, tumor stage at diagnosis), treatment and survival (month and year, and cause of death) information. The Medicare data include claims for hospital inpatient care,

physician and non-institutional providers, outpatient services, prescription medications, home health services, hospice care, and Durable Medical Equipment (DME) services.

Cohort Selection

The initial study cohort includes incidence breast cancer cases diagnosed between January 1, 2008 and December 31, 2013 (N=228,223). Individuals were excluded if they were males (N=2,118), younger than 66 years old (N=82,780), diagnosed at death or autopsy (N=1,081), did not have continuous part A/B enrollment or had HMO enrollment 12 months prior and after diagnosis month (N=61,937), had missing or unknown AJCC cancer stage (N=2,953), and had no claim in the 24 months study period. Our final sample size was 77,273 after applying all inclusion and exclusion criteria.

Study Variables

Dependent Variable: Stage at Diagnosis

To ensure the largest sample size with known cancer stage, the adjusted American Joint Committee on Cancer (AJCC) 6th edition (SEER) was used consistently to derive breast cancer stage at diagnosis as the outcome. The AJCC 6th edition reanalyzes earlier extent of disease (EOD) information collected from 1988 to 2003 to fit into the AJCC 6th definitions and therefore has fewer missing values.

Predictor Variables

The identification of potential predictors were based on a combination of clinical and empirical knowledge. A total of 719 potential indicators of extent of cancer diagnosis,

screening tests, diagnostic tests, pathology tests, surgery, radiation, chemotherapy, hormonal therapy, targeted therapy, hospice services, and comorbidities. The aforementioned indicators were identified using International Classification of Diseases, 9th Revision (ICD-9) diagnosis codes, ICD-9 procedure codes, Common Procedural Terminology (CPT)/The Healthcare Common Procedure Coding System (HCPCS) codes, and National Drug Codes (NDC). We searched through inpatient, outpatient, physician, prescription, home health services, hospice care and DME claims 12 months before and 10 months after the index diagnosis date for the codes used to identify indicators (the study period was selected based on the performance of the final algorithms). Each of the variables was measured both as a binary variable (0 or 1) and as a continuous variable (counts). For each variable constructed based on ICD-9 diagnosis codes, the place where it appeared in the diagnosis list (i.e., as primary, secondary or principal diagnosis) were recorded. The above mentioned codes used to identify the final selected variables are available in appendix table 1. Based on the NCCN Clinical Practice Guidelines (Gradishar et al., 2018) for breast cancer, 66 stage specific treatment patterns were constructed (Appendix Table 2).

Statistical Analysis

The full data set of eligible participants were equally and randomly divided into a training set and a validation set. A Generalized, Unbiased, Interaction Detection and Estimation (GUIDE) machine learning decision-tree algorithm was used to construct the classification tree which grouped patients with breast cancer into different stages (Loh, 2009). GUIDE produces decision trees using a non-parametric technique that combines

variables without complex high-order interaction terms. At each node, GUIDE selects the most significant predictor variable to make a binary split to minimize the Gini index (Lemon, Roy, Clark, Friedmann, & Rakowski, 2003)), which is a measure of the degree of a particular variable being wrongly classified when it is randomly chosen. The splitting process is repeated until specified minimum observations in each node reached. GUIDE has 3-level hierarchical split variable selection rules: first select predictors whose p-value is below Bonferroni threshold in chi-square tests. Unselected predictors are paired up and divided into several regions that are tested against cancer stage using chi-square tests. Again, most significant pair of predictors with p-value below Bonferroni threshold. Predictors not selected in the second step are tested for linear split. For each pair of ordinal predictors, apply the first split select process to its largest linear discriminant coordinates. Otherwise, select most significant predictor from level 1. The classification tree was pruned by penalizing the estimated error using the 0 standard error rule based on the subtree size through 10 cross-validations. Finally, breast cancer stage for each individual in the validation set was determined along with their predictor values. Analysis file preparation was performed with SAS Enterprise Guide 6.1 (SAS Institute Inc, Cary, NC). Classification matrix was constructed showing the predicted and true stages. For each stage, we constructed a 2*2 table to compute sensitivity, specificity, PPV, NPV and AUC.

Results

In 76,898 women there were 17.5% stage 0, 45.8% stage I, 26.1% stage II, 7.6% stage III, and 3.0% stage IV cancers as indicated by the adjusted AJCC stage. Patient

characteristics by selected variables in the final model and cancer stage are presented in Table 1. Selected variables were summarized under diagnosis, work-up tests, breast surgical procedures, radiation, chemotherapy and treatment regimen. Non-invasive cases had significantly more diagnosis code of carcinoma in situ appeared in their medical claims, and much less malignant breast cancer diagnosis codes. Secondary or unspecified lymph node neoplasm and axillary lymph node involvement diagnosis was more commonly seen among stage III patients. A majority of stage IV patients had metastatic breast cancer and secondary neoplasm diagnosis. A history of screening mammography were more common among stage 0 and stage I patients, consistent with the fact that screening mammography identified more non-invasive and early stage breast cancer, while advanced breast cancers were diagnosed more based on symptoms. Advanced stage cases received more PET imaging and CT scan tests. More sentinel lymph node biopsies were performed on stage I and II patients, while axillary lymph node directed surgeries were common for stage I through stage III. Overall, the earlier breast cancer stage, the more breast surgeries were performed. On the other hand, the later cancer stage, the more chemotherapy treatments were administered.

Table 2 shows the counts of predicted stages and their true cancer stages, which allows the assessment of misclassification. Accuracy measures for each stage were listed in Table 3. Our algorithm contains 42 patterns to identify breast cancer stage: 6 for stage 0, 15 for stage I, 14 for stage II, 4 for stage III, and 3 for stage IV (Appendix table 2). The final staging algorithm is provided in Appendix 3.

Table 1: Patients' Characteristics of Breast Cancer Cohorts.

Selected Variables	SEER AJCC Stage				
	Stage 0	Stage I	Stage II	Stage III	Stage IV
	13,447	35,257	20,090	5,809	2,295
	Diagnosis				
Carcinoma in Situ Diagnosis (Principal), %	94.2	23.7	17.4	16.3	11.0
Counts of Carcinoma in Situ Diagnosis (Primary), Mean (SD)	6.1 (6.1)	0.6 (1.6)	0.4 (1.4)	0.3 (1.2)	0.3 (1.3)
Malignant Breast Cancer Diagnosis, %	86.9	98.7	99.0	99.1	98.7
Counts of Malignant Breast Cancer Diagnosis (Primary), Mean (SD)	8.7 (8.8)	19.7 (12.4)	24.4 (16.2)	33.4 (19.5)	27.0 (16.8)
Counts of Malignant Breast Cancer Diagnosis, Mean (SD)	10.5 (10.4)	25.0 (15.7)	32.2 (21.4)	45.9 (26.3)	39.7 (23.0)
Secondary Neoplasm Diagnosis, %	1.4	2.4	5.3	15.9	88.2
Counts of Secondary Neoplasm Diagnosis, Mean (SD)	0.1 (1.3)	0.1 (1.8)	0.3 (2.7)	1.1 (4.7)	17.4 (19.3)
Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary), %	0.8	2.6	35.5	68.8	30.0

Axillary Lymph Node Involvement Diagnosis (Secondary), %	0.4	2.0	33.5	65.9	25.0
Counts of Stage IV Diagnosis, Mean (SD)	0.1 (1.4)	0.2 (1.9)	0.4 (2.9)	1.3 (4.9)	17.5 (18.9)
Work Ups					
Bilateral Screening Mammogram (CPT/HCPCS), %	81.1	75.7	48.7	35.0	23.6
PET Imaging, Skull Base to Mid-Thigh, %	3.2	9.3	22.8	52.1	59.1
Counts of HER-2 Testing, Mean (SD)	2.9 (2.1)	4.9 (2.9)	4.8 (3.1)	4.4 (3.1)	4.3 (3.2)
Breast Procedure					
Needle Biopsy, %	86.2	90.4	88.2	82.7	73.5
Sentinel Lymph Node Biopsy (BLNB), %	24.7	74.1	60.6	29.5	9.2
Surgical Biopsy, %	68.5	86.4	75.3	49.6	22.1
Axillary Lymph Node Directed Surgery (CPT/HCPCS), %	29.4	85.1	85.4	84.0	25.3
Axillary Lymph Node Dissection, %	30.3	85.8	86.6	85.5	26.1
Surgery Specimen Examination, %	84.5	62.6	31.5	17.2	7.0
Lumpectomy, %	84.1	80.6	60.4	34.9	16.0
Breast-Conserving Surgery, %	84.0	80.3	60.0	34.1	15.3
Counts of Breast Surgery, Mean (SD)	11.1 (7.4)	11.2 (7.0)	9.1 (7.3)	6.3 (6.4)	2.2 (4.5)
Chemotherapy					
Any Chemotherapy, %	3.2	12.4	33.5	59.9	58.2
Taxane-Based Chemotherapy, %	0.4	7.6	26.3	50.1	28.9

Regimen for Advanced Breast Cancer ^a , %	Treatment Regimen				
	0.0	0.3	3.9	26.5	3.4

a. Mastectomy and Axillary dissection level I/II followed by radiation and chemotherapy.

Table 2: Classification Matrix for Validation Set

		SEER AJCC Stage				
		Stage 0	Stage I	Stage II	Stage III	Stage IV
Predicted Stage	Stage 0	5,722	923	138	9	1
	Stage I	898	14,573	4,037	256	57
	Stage II	91	2,026	5,012	1,490	129
	Stage III	4	34	672	971	52
	Stage IV	5	88	176	180	905

Tables 3: Performance Measures by Stage

	Stage 0	Stage I	Stage II	Stage III	Stage IV
Sensitivity	85%	83%	50%	33%	79%
Specificity	97%	75%	87%	98%	99%
PPV	84%	74%	57%	56%	67%
NPV	97%	85%	83%	95%	99%
AUC	91%	79%	68%	66%	89%

Discussion

This study developed and validated a claim-based algorithm which predicts breast cancer stages 0-IV using predictive modeling. This algorithm accurately predicts non-invasive and early invasive breast cancer, with poorer performance on stage II and II.

No predictive algorithm has been developed to identify non-invasive stage 0 breast cancer. Stage 0 breast cancer is also called carcinoma in situ, which includes three types: Ductal carcinoma in situ (DCIS), Lobular carcinoma in situ (LCIS) and Paget disease of the nipple. Paget disease is a rare breast cancer in or around the nipple that only accounts less than 5% of United States' breast cancer cases. However, more than 96% of Paget's disease cases also have DCIS or invasive cancer in the same breast (Caliskan et al., 2008; Harris, Lippman, Osborne, & Morrow, 2012; Kanitakis, 2007). Therefore, mastectomy are the standard first treatment for Paget's disease, with breast-conserving surgery and whole breast radiation as alternative for patients with Paget's disease alone (Kanitakis, 2007; Kawase et al., 2005; Marshall et al., 2003). Although LCIS doesn't need treatment, having LCIS does increase the risk of developing invasive breast cancer. Therefore, surgery and other risk reduction hormonal adjuvant therapy may be recommended. DCIS is the most common type of non-invasive cancer. However, the treatment of DCIS is highly controversial due to the lack of scientific evidence on its risk of progression to invasive cancer (Groen et al., 2017; Park & Hwang, 2016). Failing to identify DCIS cases introduces bias into the estimation of cancer recurrence, progression and mortality, leading to erroneous conclusions on cost and benefits which are important to assess overtreatment. However, DCIS identification using medical claims is challenging due to the similarities in treatment recommendations between DCIS and stage I breast cancer (Gradishar et al., 2018). Our algorithm misclassified 13.4% of stage 0 cases as stage I, potentially the DCIS cases were treated more aggressively. Unfortunately, the gold standard DCIS information is not available in SEER-Medicare data, which limits our ability to further identify DCIS from other non-invasive breast cancer (LCIS

and Paget's disease). Future algorithm development studies are needed with a more appropriate dataset.

Previously, breast cancer stages I and II were combined as a single category in Smith's model (Smith et al., 2010). They reported sensitivity of 83%, specificity of 78%, PPV 98%, and NPV of 31% distinguishing stage I/II (early) breast cancer from stage III, with an overall 81.0% AUC. Noted that this model is the second step of a two-stage model, which means its performance is conditional on the performance of a previous model. In their first model identifying metastatic breast cancer, the PPV is only 24%, which means a significant amount of non-metastatic cases were misclassified. However, the information provided in their paper is not enough to assess their extent of misclassification. With our algorithm, we have 93% sensitivity, 73% specificity, 90% PPV, 80% NPV, and 83% AUC. In other words, we successfully identified most of the stage I/II cases and also overestimate them, while the Smith's model underestimated stage I/II cases.

Looking at stage I and II individually, our algorithm has better performance with stage I (sensitivity 83%, specificity 75%, PPV 74%, NPV 84% and AUC 79%). Among the true stage II cases, less than half were misclassified as stage I; while among the predicated stage II cases, 40.2% of them were true stage I and III cancers. This is not unexpected because according to the NCCN guidelines, similar locoregional treatment (surgical and systemic adjuvant) is recommended to clinical stage I (T1, N0, M0), IIA/IIB, and stage IIIA (T3, N1, M0). The ability to distinguish stage II from stage III cases is also limited with our algorithm. More than half of the true cancer were misclassified into stage II category. And among the predicted stage III cases, 38.8% were true stage II. This is because the

preoperative systemic therapy for stage IIA/B is similar to the locally advanced stage IIIA//B/C breast cancers. The similarities in treatment recommendations pose a real challenge to predictive models distinguishing stage II and III based on medical claims. If stage II and III are combined, the sensitivity is 63%, specificity is 99%, PPV is 78%, NPV is 83%, and AUC is 77%. The relatively low sensitivity is mainly because of the true stage II cases that were misclassified into stage I.

Algorithms identifying metastatic breast cancer have been more commonly studied. Chawla et al. 2014 showed that using ICD-9 diagnosis codes related to regional or distant metastases had limited validity (sensitivity 51.0%, specificity 98.3%, and PPV 65.8%) in SEER-Medicare data, which was consistent with Cooper's single predictor. The expanded Smith algorithm had improved performance by adding demographic, tumor, and treatment characteristics (sensitivity of 81%, specificity 89%, positive PPV 24%, and NPV 99%). Their algorithm overestimated metastatic cases, but successfully captured most of the true metastatic and non-metastatic cases. Alternatively, Nordstrom and colleagues (2012) developed an algorithm to identify metastatic breast cancer cases from an outpatient oncology EMR database linked to medical and pharmacy claims data using a classification and regression tree. The reported sensitivity is 62%, specificity 97%, PPV 75%, NPV 95%. One of the major limitations of their study is the sample size. They included a total of 1385 breast cancer patients, out of whom 175 were metastatic. Most of the potential cases were excluded based on very strict criteria: there must be an oncologist issuing a cancer diagnosis within 2 days before or after a patient's index diagnosis date, and continuously reporting claims for 6 months prior and 2 months post the index date. Such a small and highly selective

sample size limits the generalizability of their algorithm. Another limitation is they only used outpatient medical and pharmacy claims to develop to algorithm, and didn't include valuable information from other claim sources such as inpatient services. Finally, 60 days of follow-up may be too short to capture enough information to describe treatment patterns for metastatic cases. Using a similar classification and regression tree approach, our algorithm developed using prescription medication files had a sensitivity of 79%, specificity 99%, PPV 67%, NPV 99%, and AUC 89%, which had a better balance between identifying as much of true cases and misclassification. The most recent breast cancer stage predictive model is the two algorithms developed by Beachler et al., (2019) to identify estrogen receptor positive (ER+) and human epidermal growth factor negative (HER2-) early and advanced stage breast cancer, respectively. In their study, they defined early stages as stage I, II, IIIA, or IIIB, while advance stages include stage IIIC or IV. Based on a validation sample of 3184 ER+ & HER2- early staged and 1436 advanced staged cases, they reported mediocre sensitivity (60% vs. 54%) and high PPV (84% vs. 91%) respectively. Their algorithms are informing being the first validated models to predict both ER/HER2 status, which is not the main focus of our study.

Sensitivity, specificity, PPV and NPV provide different aspects to measure the performance of a predictive algorithm. For example, high sensitivity is needed when the researcher's intent is to capture true cases as many as possible; while high PPV is more important when a pure cohort of true cases are needed for analysis. Compared to the previously published algorithms using logistic regression that requires calculation of

probabilities and selecting cutoff values, our models are easier to use and just need to construct a few variables.

This algorithm development study has important implications for future research. First, the CART approach is appropriate for breast cancer stage prediction, and may be applied to develop and validate predictive models for other cancer stages. Unlike logistic regression models, the final algorithm developed by CART is easier to understand and adopt. Second, our algorithm is most useful when applied to commercial claims databases that are not linked to clinical information. With the predicted breast cancer stages, more epidemiologic and economic studies will become feasible. Third, the prescription medication files are very important in improving the model performance, especially the prediction of stage IV cases. The main reason to include pharmacy claims is that the treatment of metastatic breast cancer is heavily relied on chemotherapy and endocrine therapy with certain metastatic agent.

Some limitations existed in our study. Most importantly, this algorithm is developed on an older population and has not been validated in other databases and populations. Although medical claims coding practice should not change depending on the age of breast cancer patients, researchers should use cautions when applying this algorithm to a different age group. Secondly, since multiple years of data has been compiled to reach a relatively large sample for model development and validation, there may be changes in clinical practice patterns over the years, and approval of new medications. Future studies may explore the feasibility of building year specific model. Third, in this study we used the date of diagnosis from SEER, which is not available in other databases. However, there was nearly 90%

agreement within one month between the SEER diagnosis date and the first Medicare claim with a cancer diagnosis ((Kind, Virnig, & McBean; Lin & Virgo, 2014). Fourth, we required a total of 24 months continuous enrollment (12 months prior and 12 months post the index diagnosis), which limited our ability to make prediction for patients with shorter continuous enrollment. Finally, the accuracy of our algorithm predictive stage IV cases is limited if applied to a breast cancer cohort selected by a validated algorithm (Nattinger, Laud, Bajorunaite, Sparapani, & Freeman, 2004) which is widely used to select breast cancer incident cases from claims data. The reason is that the Nattinger algorithm requires breast surgery for each case identified, which is less common among metastatic breast cancers.

To our best knowledge, the breast cancer stage predictive algorithms we developed and validated are the first attempt to identify stage 0 to stage IV. We expanded the pool of potential predictors by including prescriptive medication files and successfully increased the model performance. Our algorithms are based on classification tree and thus easy for researchers to adapt.

Appendices

Appendix Table 1: Claims codes to determine selected variables.

Selected Variables	ICD-9 Diagnosis Codes	ICD-9 Procedure Codes	CPT/HCPCS Codes
Carcinoma in Situ	233.0		

Malignant Breast Cancer	174		
Secondary Neoplasm	197-199		
Secondary or Unspecified Lymph Node Neoplasm	196		
Axillary Lymph Node Involvement	196.3		
Stage IV Cancer Indicators	196.1 196.2 198.0 198.1 198.82 198.89 196.5-196.6 196.8-196.9 197.0-197.8 198.3-198.8 199.0-199.1		
Bilateral Screening Mammogram			77063 77067 76092 77057 G0202 G0203
PET Imaging, Skull Base to Mid-Thigh			78812 78815
HER-2 Testing			88342 88360 88361 88365 88271 88274 88291 88367 88368 83950
Needle Biopsy			19000 19001 19100 19102 19103 76095 76360 76393 76942 77031 88170 88171 10021 10022 19081-19086
Sentinel Lymph Node Biopsy (BLNB)			38500 38525
Surgical Biopsy			19101 19110 19120 19125 19126 38500 38525 38900 38792

Axillary Lymph Node Directed Surgery			38500 38525 38740 38745 19162 19200 19220 19240 19302 19305 19306 19307
Axillary Lymph Node Dissection		40.3 40.23 40.51 85.43- 85.48	38500 38525 38740 38745 19162 19200 19220 19240 19302 19305 19306 19307
Surgery Specimen Examination			76098
Lumpectomy		85.20 85.21 85.22 85.23 85.25	19120 19125 19126 19160 19162 19301 19302
Breast- Conserving Surgery			19110 19112 19120 19125 19160 19162 19301 19302 19240 19302
Breast Surgery		40.3 40.23 40.51 85.20 85.21 85.22 85.23 85.25 85.41- 85.48	19110 19112 19120 19125 19160 19162 19301 19302 19240 19302 38500 38525 38740 38745 19162 19200 19220 19240 19302 19305 19306 19307 19120 19125 19126 19160 19162 19301 19302 19180 19182 19200 19220 19240 19303 19304 19305 19306 19307 19180-19307

Chemotherapy			G0355 G0359 G0360 G0361 G0362 Q0083 Q0084 Q0085 J7150 J85 J86 J87 J8999 J9 J8510 J8515 J8520 J8521 J8530 J8540 J8560 J8561 J8562 J8565 J8597 J8600 J8610 J8650 J8700 J8705 J8999 C9004 C9012 C9020 C9110 C9127 C9129 C9205 C9207 C9218 C9233 C9235 C9239 C9240 C9253 C9259 C9260 C9262 C9265 C9273 C9276 C9280 C9284 C9287 C9295 C9296 C9298 C9414 C9415 C9417 C9418 C9420 C9421 C9422 C9423 C9424 C9425 C9426 C9427 C9428 C9429 C9431 C9432 C9433 C9436 C9437 C9440 96400-96549 J9000-J9999 Q0083-Q0085 (Exclude J9003 J9035 J8501 J7527 J8561 J3315 J3487 J3488 J9202 J9240 J9295 J9306 J9354 J9355 J9395 J9217 J9218 J9219 J9225 J9226)
Taxane-Based Chemotherapy			J9265 J9264 J9267 C9127 C9431 J9170 J9171
Regimen for Advanced Breast Cancer ^a			
Radiation	V58.0 V66.1 V67.1	92.2 92.3 92.4 92.41 92.20- 92.29 92.30- 92.39	

Chemotherapy	V58.1 V66.2 V67.2 E933.1 E930.7	99.25	G0355 G0359 G0360 G0361 G0362 Q0083 Q0084 Q0085 J7150 J85 J86 J87 J8999 J9 J8510 J8515 J8520 J8521 J8530 J8540 J8560 J8561 J8562 J8565 J8597 J8600 J8610 J8650 J8700 J8705 J8999 C9004 C9012 C9020 C9110 C9127 C9129 C9205 C9207 C9218 C9233 C9235 C9239 C9240 C9253 C9259 C9260 C9262 C9265 C9273 C9276 C9280 C9284 C9287 C9295 C9296 C9298 C9414 C9415 C9417 C9418 C9420 C9421 C9422 C9423 C9424 C9425 C9426 C9427 C9428 C9429 C9431 C9432 C9433 C9436 C9437 C9440 96400-96549 J9000-J9999 Q0083-Q0085 (Exclude J9003 J9035 J8501 J7527 J8561 J3315 J3487 J3488 J9202 J9240 J9295 J9306 J9354 J9355 J9395 J9217 J9218 J9219 J9225 J9226)
Mastectomy		85.41- 85.48	19180 19182 19200 19220 19240 19303 19304 19305 19306 19307 19180-19255
Axillary Lymph Node Dissection (ALND)		40.3 40.23 40.51	38740 38745

a. Mastectomy and Axillary dissection level I/II followed by radiation and chemotherapy.

Appendix Table 2: NCCN Recommended Treatment Regimens.

LCIS:

- Primary treatment
 - Observation
 - bilateral mastectomy
- Risk reduction/Adjuvant therapy
 - +/- tamoxifen or raloxifene

DCIS:

- Primary treatment
 - Total mastectomy w/o lymph node dissection
 - Total mastectomy w/ OR w/o sentinel node biopsy
 - Lumpectomy w/o lymph node dissection
 - Lumpectomy w/o lymph node dissection + (whole breast) RT
 - Excision + RT
 - Excision only
- Risk reduction/Adjuvant therapy
 - +/- tamoxifen or raloxifene

Local stage (stage I/II):

Staging:

- +/- FNA/core biopsy + Axillary dissection level I/II
- FNA/core biopsy + sentinel node mapping and excision +/- Axillary dissection level I/II

Treatment: (sequence is important-endocrine after chemo)

- Lumpectomy + Axillary dissection level I/II + post or concurrent RT +/- Adjuvant chemo
- Mastectomy + Axillary dissection level I/II +/- Adjuvant chemo
- Mastectomy + Axillary dissection level I/II + post or concurrent RT +/- Adjuvant chemo +/- endocrine (hormone) therapy +/- adjuvant chemo
- Sentinel lymph node dissection + surgery
- Axillary lymph node dissection + surgery
- Surgery + Axillary lymph node dissection
- Sentinel lymph node dissection + Axillary lymph node dissection
- Breast conservative surgery + Sentinel lymph node dissection/Axillary lymph node dissection + radiation
- Radiation + surgery
- Neoadjuvant chemo/hormonal therapy + surgery + radiation
- Neoadjuvant chemo/hormonal therapy + surgery (no slnb/alnd) + radiation
- Neoadjuvant chemo/hormonal therapy + surgery + adjuvant chemo/hormonal therapy + radiation

Adjuvant Hormonal Therapy: (AI preferred for postmenopausal)

Pre-menopausal:

- Tamoxifen 2-3 yrs/5 yrs +/- ovarian suppression/ablation + Aromatase Inhibitor (Anastroze/Letrozole/Exemestane) 5 yrs
- Aromatase Inhibitor 5 yrs +/- ovarian suppression/ablation + Aromatase Inhibitor (Anastroze/Letrozole/Exemestane) 5 yrs
- Tamoxifen 5 yrs only if always premenopausal

Post-menopausal:

- Aromatase Inhibitor 5 yrs
- Tamoxifen 2-3 yrs/4.5-6 yrs + Aromatase Inhibitor (Anastroze/Letrozole/Exemestane) 5 yrs
- Tamoxifen 5 yrs
- Tamoxifen 10 yrs

Systemic Therapy:

Breast preserving: (1) Core biopsy +/- FNA (2) Sentinel lymph node procedure

- Breast preserving (1) or (2) + neoadjuvant chemo/hormone therapy (>1 lines) +/- Axillary staging + mastectomy w/Axillary dissection level I/II +/- more individualized chemo + RT
- Breast preserving (1) or (2) + neoadjuvant chemo/hormone therapy (>1 lines) +/- Axillary staging + lumpectomy w/Axillary dissection level I/II +/- more individualized chemo + RT
- Only endocrine systemic therapy
- The use of tamoxifen/anastrozole
- HER2+ patient: use trastuzumab preoperative systemic therapy w/ trastuzumab pertuzumab-containing regimen

Advanced Stage_Stage III/IV:

- Mastectomy + Axillary dissection level I/II + RT + chemo
- Anthracycline-based neoadjuvant chemo +/- tamoxifen + Mastectomy + Axillary dissection level I/II + RT + chemo +/- endocrine therapy (tamoxifen, etc.)
- Anthracycline-based neoadjuvant chemo +/- tamoxifen + lumpectomy + Axillary dissection level I/II + RT + chemo +/- endocrine therapy (tamoxifen, etc.)
- Anthracycline-based neoadjuvant chemo +/- tamoxifen + high dose RT + chemo +/- tamoxifen
- Anthracycline-based neoadjuvant chemo + more systemic chemo
- Anthracycline-based neoadjuvant chemo + preoperative radiation + Mastectomy + Axillary dissection level I/II + RT + chemo +/- endocrine therapy (tamoxifen, etc.)

- Anthracycline-based neoadjuvant chemo + preoperative radiation + lumpectomy + Axillary dissection level I/II + RT + chemo +/- endocrine therapy (tamoxifen, etc.)

Metastatic Stage IV:

- Endocrine therapy (Aromatase Inhibitor OR Antiestrogen) for post-menopausal +/- chemo
- Ovarian ablation/suppression + Aromatase Inhibitor for pre-menopausal +/- chemo
- Ovarian ablation/suppression + Antiestrogen +/- LHRH agonist for pre-menopausal +/- chemo
- Trastuzumab +/- chemo
- Chemo
- [Pertuzumab + Trastuzumab + taxane (preferred)] + [Capecitabine + Lapatinib (preferred)]
- Trastuzumab +/- chemo + Lapatinib
- Trastuzumab +/- chemo

Appendix Table 3: Final Staging Algorithm.

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) = 0 AND
 (Malignant Breast Cancer Diagnosis) = 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) = 0 AND
 (Malignant Breast Cancer Diagnosis) \neq AND
 (Lumpectomy) = 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND

(Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) = 0 AND
 (Malignant Breast Cancer Diagnosis) ~=. AND
 (Lumpectomy) ~=. AND
 (Surgery Specimen Examination) = 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) = 0 AND
 (Malignant Breast Cancer Diagnosis) ~=. AND
 (Lumpectomy) ~=. AND
 (Surgery Specimen Examination) ~= 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) ~= 0 AND
 (Any Chemotherapy) = 0 AND
 (Surgery Specimen Examination) = 0 AND
 (Needle Biopsy) = 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) ~= 0 AND
 (Any Chemotherapy) = 0 AND
 (Surgery Specimen Examination) = 0 AND
 (Needle Biopsy) ~= 0 AND
 (Breast-Conserving Surgery) = 0 AND
 ((Counts of Malignant Breast Cancer Diagnosis (Primary)) <= 13 OR (Counts of
 Malignant Breast Cancer Diagnosis (Primary)) =.) THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
 (Sentinel Lymph Node Biopsy) ~= 0 AND
 (Any Chemotherapy) = 0 AND
 (Surgery Specimen Examination) = 0 AND

(Needle Biopsy) \sim 0 AND
(Breast-Conserving Surgery) = 0 AND
(Counts of Malignant Breast Cancer Diagnosis (Primary)) > 13 THEN Predicted
Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) = 0 AND
(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
(Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
(Sentinel Lymph Node Biopsy) \sim 0 AND
(Any Chemotherapy) = 0 AND
(Surgery Specimen Examination) = 0 AND
(Needle Biopsy) \sim 0 AND
(Breast-Conserving Surgery) \sim 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) = 0 AND
(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
(Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
(Sentinel Lymph Node Biopsy) \sim 0 AND
(Any Chemotherapy) = 0 AND
(Surgery Specimen Examination) \sim 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) = 0 AND
(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
(Bilateral Screening Mammogram (CPT/HCPCS)) = 0 AND
(Sentinel Lymph Node Biopsy) \sim 0 AND
(Any Chemotherapy) \sim 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) = 0 AND
(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
(Bilateral Screening Mammogram (CPT/HCPCS)) \sim 0 AND
(Taxane-Based Chemotherapy) = 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) = 0 AND
(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
(Bilateral Screening Mammogram (CPT/HCPCS)) \sim 0 AND
(Taxane-Based Chemotherapy) \sim 0 AND
(Sentinel Lymph Node Biopsy) = 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) ~= 0 AND
 (Taxane-Based Chemotherapy) ~= 0 AND
 (Sentinel Lymph Node Biopsy) ~= 0 AND
 (Surgery Specimen Examination) = 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) = 0 AND
 (Bilateral Screening Mammogram (CPT/HCPCS)) ~= 0 AND
 (Taxane-Based Chemotherapy) ~= 0 AND
 (Sentinel Lymph Node Biopsy) ~= 0 AND
 (Surgery Specimen Examination) ~= 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) ~= 0
AND
 (Regimen for Advanced Breast Cancer) = 0 AND
 (Surgical Biopsy) = 0 AND
 (PET Imaging, Skull Base to Mid-Thigh) = 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) ~= 0
AND
 (Regimen for Advanced Breast Cancer) = 0 AND
 (Surgical Biopsy) = 0 AND
 (PET Imaging, Skull Base to Mid-Thigh) ~= 0 THEN Predicted Stage = 3;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND
 (Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) ~= 0
AND
 (Regimen for Advanced Breast Cancer) = 0 AND
 (Surgical Biopsy) ~= 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
 (Carcinoma in Situ Diagnosis (Principal)) = 0 AND

(Secondary or Unspecified Lymph Node Neoplasm Diagnosis (Secondary)) \sim 0
AND
(Regimen for Advanced Breast Cancer) \sim 0 THEN Predicted Stage = 3;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) \sim 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 2 AND
(Axillary Lymph Node Dissection) = 0 AND
(Surgery Specimen Examination) = 0 AND
(Counts of Malignant Breast Cancer Diagnosis (Primary)) \leq 6 THEN Predicted
Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) \sim 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 2 AND
(Axillary Lymph Node Dissection) = 0 AND
(Surgery Specimen Examination) = 0 AND
(Counts of Malignant Breast Cancer Diagnosis (Primary)) $>$ 6 THEN Predicted Stage
= 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) \sim 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 2 AND
(Axillary Lymph Node Dissection) = 0 AND
(Surgery Specimen Examination) \sim 0 THEN Predicted Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) \sim 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 2 AND
(Axillary Lymph Node Dissection) \sim 0 AND
(Counts of Malignant Breast Cancer Diagnosis) \leq 10 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 1 OR (Counts of Carcinoma
in Situ Diagnosis (Primary)) =.) THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) \sim 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) \leq 2 AND
(Axillary Lymph Node Dissection) \sim 0 AND

(Counts of Malignant Breast Cancer Diagnosis) <= 10 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) > 1 THEN Predicted Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Counts of Carcinoma in Situ Diagnosis (Primary)) <= 2 AND
(Axillary Lymph Node Dissection) ~= 0 AND
((Counts of Malignant Breast Cancer Diagnosis) > 10 OR (Counts of Malignant Breast Cancer Diagnosis) =.) THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) = 0 THEN Predicted Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) ~= 0 AND
((Counts of Malignant Breast Cancer Diagnosis) <= 15 OR (Counts of Malignant Breast Cancer Diagnosis) =.) THEN Predicted Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) ~= 0 AND
(Counts of Malignant Breast Cancer Diagnosis) > 15 AND
(Surgery Specimen Examination) = 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) ~= 0 AND

(Counts of Malignant Breast Cancer Diagnosis) > 15 AND
(Surgery Specimen Examination) ~= 0 AND
((Counts of Malignant Breast Cancer Diagnosis) <= 25 OR (Counts of Malignant Breast Cancer Diagnosis) =.) AND
((Counts of HER-2 Testing) <= 4 OR (Counts of HER-2 Testing) =.) THEN
Predicted Stage = 0;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) ~= 0 AND
(Counts of Malignant Breast Cancer Diagnosis) > 15 AND
(Surgery Specimen Examination) ~= 0 AND
((Counts of Malignant Breast Cancer Diagnosis) <= 25 OR (Counts of Malignant Breast Cancer Diagnosis) =.) AND
(Counts of HER-2 Testing) > 4 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
((Counts of Carcinoma in Situ Diagnosis (Primary)) > 2 OR (Counts of Carcinoma in Situ Diagnosis (Primary)) =.) AND
(Axillary Lymph Node Directed Surgery (CPT/HCPCS)) ~= 0 AND
(Counts of Malignant Breast Cancer Diagnosis) > 15 AND
(Surgery Specimen Examination) ~= 0 AND
(Counts of Malignant Breast Cancer Diagnosis) > 25 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) ~= 0 AND
(Surgery Specimen Examination) = 0 AND
(Sentinel Lymph Node Biopsy) = 0 THEN Predicted Stage = 3;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) ~= 0 AND
(Surgery Specimen Examination) = 0 AND
(Sentinel Lymph Node Biopsy) ~= 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) = 0 AND
(Carcinoma in Situ Diagnosis (Principal)) ~= 0 AND

(Axillary Lymph Node Involvement Diagnosis (Secondary)) \sim 0 AND
(Surgery Specimen Examination) \sim 0 THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
(Counts of Breast Surgery) \leq 0 THEN Predicted Stage = 4;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
((Counts of Breast Surgery) $>$ 0 OR (Counts of Breast Surgery) =.) AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Surgery Specimen Examination) = 0 AND
((Counts of Secondary Neoplasm Diagnosis) \leq 5 OR (Counts of Secondary
Neoplasm Diagnosis) =.) THEN Predicted Stage = 2;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
((Counts of Breast Surgery) $>$ 0 OR (Counts of Breast Surgery) =.) AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Surgery Specimen Examination) = 0 AND
(Counts of Secondary Neoplasm Diagnosis) $>$ 5 THEN Predicted Stage = 4;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
((Counts of Breast Surgery) $>$ 0 OR (Counts of Breast Surgery) =.) AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) = 0 AND
(Surgery Specimen Examination) \sim 0 THEN Predicted Stage = 1;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
((Counts of Breast Surgery) $>$ 0 OR (Counts of Breast Surgery) =.) AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) \sim 0 AND
((Counts of Stage IV Diagnosis) \leq 7 OR (Counts of Stage IV Diagnosis) =.) THEN
Predicted Stage = 3;

IF (Secondary Neoplasm Diagnosis) \sim 0 AND
((Counts of Breast Surgery) $>$ 0 OR (Counts of Breast Surgery) =.) AND
(Axillary Lymph Node Involvement Diagnosis (Secondary)) \sim 0 AND
(Counts of Stage IV Diagnosis) $>$ 7 THEN Predicted Stage = 4;

References

- Beachler, D. C., de Luise, C., Yin, R., Gangemi, K., Cochetti, P. T., & Lanes, S. (2019). Predictive model algorithms identifying early and advanced stage ER+/HER2- breast cancer in claims data. *Pharmacoepidemiology and drug safety*, 28(2), 171-178.
- Caliskan, M., Gatti, G., Sosnovskikh, I., Rotmensz, N., Botteri, E., Musmeci, S., . . . Luini, A. (2008). Paget's disease of the breast: the experience of the European Institute of Oncology and review of the literature. *Breast cancer research and treatment*, 112(3), 513.
- Chawla, N., Yabroff, K. R., Mariotto, A., McNeel, T. S., Schrag, D., & Warren, J. L. (2014). Limited validity of diagnosis codes in Medicare claims for identifying cancer metastases and inferring stage. *Annals of epidemiology*, 24(9), 666-672. e662.
- Cooper, G. S., Yuan, Z., Stange, K. C., Amini, S. B., Dennis, L. K., & Rimm, A. A. (1999). The utility of Medicare claims data for measuring cancer stage. *Medical care*, 706-711.
- Edge, S. B., & Compton, C. C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Annals of surgical oncology*, 17(6), 1471-1474.
- Engels, E. A., Pfeiffer, R. M., Ricker, W., Wheeler, W., Parsons, R., & Warren, J. L. (2011). Use of surveillance, epidemiology, and end results-medicare data to conduct case-control studies of cancer among the US elderly. *American journal of epidemiology*, 174(7), 860-870.
- Gradishar, W. J., Anderson, B. O., Balassanian, R., Blair, S. L., Burstein, H. J., Cyr, A., . . . Giordano, S. H. (2018). Breast cancer, version 4.2017, NCCN clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network*, 16(3), 310-320.
- Groen, E. J., Elshof, L. E., Visser, L. L., Emiel, J. T., Winter-Warnars, H. A., Lips, E. H., & Wesseling, J. (2017). Finding the balance between over-and under-treatment of ductal carcinoma in situ (DCIS). *The Breast*, 31, 274-283.
- Harris, J. R., Lippman, M. E., Osborne, C. K., & Morrow, M. (2012). *Diseases of the Breast*: Lippincott Williams & Wilkins.
- Kanitakis, J. (2007). Mammary and extramammary Paget's disease. *Journal of the European Academy of Dermatology and Venereology*, 21(5), 581-590.
- Kawase, K., DiMaio, D. J., Tucker, S. L., Buchholz, T. A., Ross, M. I., Feig, B. W., . . . Ames, F. C. (2005). Paget's disease of the breast: there is a role for breast-conserving therapy. *Annals of surgical oncology*, 12(5), 391-397.
- Kind, S., Virnig, B., & McBean, M. Retrieved from <https://healthcaredelivery.cancer.gov/seermedicare/considerations/date.html>
- Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine*, 26(3), 172-181.
- Lin, C. C., & Virgo, K. S. (2014). Diagnosis date agreement between SEER and Medicare claims data: Impact on treatment. *Medical care*, 52(1), 32-37.

- Loh, W.-Y. (2009). Improving the precision of classification trees. *The Annals of Applied Statistics*, 3(4), 1710-1737.
- Marshall, J. K., Griffith, K. A., Haffty, B. G., Solin, L. J., Vicini, F. A., McCormick, B., . . . Pierce, L. J. (2003). Conservative management of Paget disease of the breast with radiotherapy: 10 - and 15 - year results. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 97(9), 2142-2149.
- Nattinger, A. B., Laud, P. W., Bajorunaite, R., Sparapani, R. A., & Freeman, J. L. (2004). An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health services research*, 39(6p1), 1733-1750.
- Nordstrom, B. L., Whyte, J. L., Stolar, M., Mercaldi, C., & Kallich, J. D. (2012). Identification of metastatic cancer in claims data. *Pharmacoepidemiology and drug safety*, 21, 21-28.
- Park, T. S., & Hwang, E. S. (2016). Current trends in the management of ductal carcinoma in situ. *Oncology*, 30(9).
- Sathiakumar, N., Delzell, E., Yun, H., Jooste, R., Godby, K., Falkson, C., . . . Kilgore, M. L. (2017). Accuracy of Medicare Claim-based Algorithm to Detect Breast, Prostate, or Lung Cancer Bone Metastases. *Medical care*, 55(12), e144-e149.
- Smith, G. L., Shih, Y.-C. T., Giordano, S. H., Smith, B. D., & Buchholz, T. A. (2010). A method to predict breast cancer stage using Medicare claims. *Epidemiologic Perspectives & Innovations*, 7(1), 1.
- Weiss, A., Chavez-MacGregor, M., Lichtensztajn, D. Y., Yi, M., Tadros, A., Hortobagyi, G. N., . . . Mittendorf, E. A. (2018). Validation study of the American Joint Committee on Cancer eighth edition prognostic stage compared with the anatomic stage in breast cancer. *JAMA oncology*, 4(2), 203-209.
- Whyte, J. L., Engel-Nitz, N. M., Teitelbaum, A., Gomez Rey, G., & Kallich, J. D. (2015). An evaluation of algorithms for identifying metastatic breast, lung, or colorectal cancer in administrative claims data. *Medical care*, 53(7), e49-e57.

JOURNAL ARTICLE

Significant Association of Annual versus Biennial Screening with Later-Stage Breast Cancer at Diagnosis among Women Aged 40-64

Journal of National Cancer Institute

Abstract

Importance— Mixed results exist on the controversial topic of mammography screening frequency, especially for women aged 40-49.

Objective— To evaluate the risk of later-staged disease in women with breast cancer had no screening, annual and biennial screening.

Design—Retrospective observational study from 1999-2013.

Setting—MarketScan commercial claims database.

Participants— 65,025 female incident breast cancer cases with continuous enrollment 33 months prior to and 12 months post the index diagnosis date were identified.

Exposure— Eligible patients were categorized based on their screening patterns during the 33 months pre-diagnosis period as non-screeners, annual and biennial screeners.

Main outcomes and measures— Breast cancer stage 0-IV were predicted by a validated algorithm using diagnostic and treatment information from claims.

Results— Annual screening was associated with less risk of stage IV cancers (Odds ratio [OR]=0.63, 95% confidence interval [CI]=0.51-0.77), stage III-IV cancers (OR=0.75, 95% CI=0.66-0.82) and stage II-IV cancers (OR=0.75, 95% CI=0.72-0.78) compared to biennial screening. Similar association was observed both among women in their 40s and older.

Women older than 50 did not present with more invasive (stage I-IV) than non-invasive cancer (stage 0) if received annual rather than biennial screening.

Conclusions and relevance— Both annual and biennial screening protect women against later-staged breast cancer at diagnosis. For women aged 40-64, annual screeners are less likely to have later-staged disease than biennial screeners.

Introduction

Mammography screening has been considered gold standard technique for early breast cancer detection, supported by randomized controlled trials ((Gøtzsche & Jørgensen, 2013; Nelson et al., 2009). Women at average-risk of breast cancer are considered who don't have personal history of breast cancer, previous diagnosis of a high-risk breast lesion, *BRCA1/2* or other gene mutation related to breast cancer, or no exposure to chest radiation in childhood (Oeffinger et al., 2015; Siu, 2016). And the ideal mammography screening interval for average risk women has been controversial, according to various screening recommendations by governmental organization (Beyers et al., 2018; Lauby-Secretan et al., 2015; Siu, 2016) and medical societies (Mango, Bryce, Morris, Gianotti, & Pinker, 2018; Monticciolo et al., 2017; Oeffinger et al., 2015; Physicians & Physicians, 2012; Wilt, Harris, & Qaseem, 2015). In summary, the recommendation discrepancies center on the age of starting mammography screening, and the screening interval for women at average risk.

A longer interval between screening mammograms allows more time for a tumor to grow, to become symptomatic and clinically detectable, and therefore, more likely to be at an advanced stage. A randomized controlled trial (RCT) found smaller tumor size for breast

cancers detected through annual screening compared to triennial screening (The Breast Screening Frequency Trial Group, 2002). Annual versus biennial screening has never been studied in any RCT up to date. Moreover, the protocol-driven care of RCTs undermines their generalizability and doesn't reflect the real-world pattern mammography screening.

Empirical studies using real-world data looked at the association between mammography screening interval and breast cancer stage at diagnosis. Freedman et al. (2003) reviewed clinical records of women with breast cancer and reported that women in their 40s were more likely to have DCIS discovered if they had underwent annual screening mammography, compared with no or less frequent screening (Freedman et al., 2003). Similarly, White et al. (2004) used data from seven participating mammogram registries in the National Cancer Institute-funded Breast Cancer Surveillance Consortium (BCSC) and reported the elevated risk of screening biennially versus annually in later stage invasive breast cancer among women aged 40–49 (White et al., 2004). In a follow-up study with the BCSC, Goel et al. (2007) only found higher probability of being diagnosed with advanced breast cancer associated with screening intervals greater than 2.5 years, but not for annual or biennial screening (Goel, Littenberg, & Burack, 2007). Concordant with Goel's findings, updated BCSC studies found no such association among women older than 40 (Dittus et al., 2013; Hubbard et al., 2011; Kerlikowske et al., 2013; Miglioretti et al., 2015; O'Meara et al., 2013). Although BCSC is a large population-based database, it only consists of women with regular screening behavior, and among them only a small proportion developed breast cancer. Therefore, the comparison between screeners and non-screener is not available for BCSC studies, and small sample size has been one of the limitations (Braithwaite et al.,

2013; Dittus et al., 2013). Finally, the BCSC studies reported their findings on dichotomized cancer outcomes (favorable or not), which may have missed valuable information in specific breast cancer stage 0-IV.

Based on a large population-based database with numerous incident breast cancer cases, our study aimed to re-examine the association between the extent of mammography screening (i.e., no screening, annual screening and biennial screening) and breast cancer stage at diagnosis (stage 0 through stage IV) among women aged 40-64.

Methods

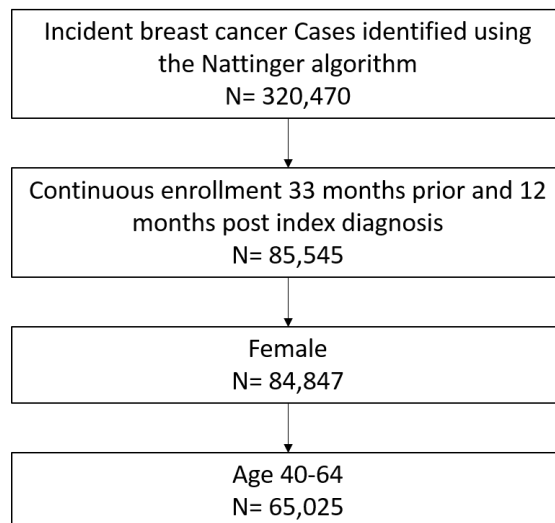
Data Source

The MarketScan® Commercial Claims and Encounters Database includes 200 million employees from approximately 45 large employers and covered by more than 100 private health insurance companies in the United States. Allowing access to large numbers of population over expansive geographic regions, MarketScan has been widely used for epidemiological, effectiveness, healthcare utilization and cost research (Sarrazin & Rosenthal, 2012). The database contains individual level information on diagnosis codes, procedure codes, medication codes, and payments for healthcare utilization.

Incident breast cancer cases were selected using the Nattinger algorithm ((Nattinger, Laud, Bajorunaite, Sparapani, & Freeman, 2004) based on the following criteria: patients diagnosed between Jan 1, 1999 and December 31, 2013, continuous enrollment 33 month prior and 12 months after the index diagnosis date, females, age at diagnosis between 40-64 (Figure 1). The index diagnosis date was the first date of definitive breast cancer surgery

according to the Nattinger algorithm. We identified claims of screening and diagnostic mammography for each patient using Common Procedural Terminology (CPT)/The Healthcare Common Procedure Coding System (HCPCS) codes. We then applied the Fenton algorithm (Fenton et al., 2014) to exclude mammograms for diagnostic purpose. Three groups of screening pattern were defined using the interval(s) between screening mammograms with 33 months period before the index diagnosis date: non-screeners had no screening record; annual screeners received 2 screening mammograms which were 9-18 months apart; similarly, biennial screeners had 1 screening mammograms or 2 screenings but the interval was 19-30 months (Braithwaite et al., 2013; Dittus et al., 2013; Hubbard et al., 2011; Kerlikowske et al., 2013; O'Meara et al., 2013; White et al., 2004). In sensitivity analysis, we applied a more restrictive definition for annual and biennial screeners with screening intervals of 11-14 months and 23-26 months, respectively (Miglioretti et al., 2015).

Figure 1: Study flowchart.



Measures of Outcome and Covariates

Due to the lack of clinical information in administrative claims data, breast cancer stage was predicted based on the diagnoses, medical procedures, and prescriptive medication received by each eligible patient in the 22 months study period (12 months prior and 10 months post the index diagnosis date) using the validated algorithm developed in the first journal article of this dissertation. Based on the validated Fenton algorithm that determined the detecting method of an incidence breast cancer, we categorized incidence cases diagnosed within one year of the most recent screening mammogram as detected by screening (Fenton et al., 2014). We measured primary care utilization as routine health management by counting the number of claims with specialty codes for a primary care physician and the place of service being outpatient clinic or office (Krause, Ganduglia-Cazaban, Piller, & Venkataraman) from the 3rd month and back to the 33rd month prior to index diagnosis date (30 months in total). The counts of claims were further grouped into quantiles. The age at diagnosis of patients was also categorized into 5-year intervals. Patients were determined to be post-menopausal if aged 55 or older at diagnosis, ever used fulvestrant or aromatase inhibitors (i.e., anastrozole, letrozole, or exemestane) as treatment, ever received bilateral oophorectomy, or had an ICD-9 diagnosis code related to postmenopausal status (Li et al., 2016; Miglioretti et al., 2015). The patient's comorbidity was measured using the Charlson Comorbidity Index (CCI) (Charlson, Pompei, Ales, & MacKenzie, 1987)

using ICD-9 diagnosis and surgical codes from claims 12 months prior to the index diagnosis date. In addition, we further categorized CCI scores into 3 classes (0, 1, and ≥ 2).

Statistical Analysis

Patient characteristics at diagnosis of comparison groups were reported descriptively and compared using the chi-square test. We fitted a partial proportional odds model because the proportional odds assumption that assuming the same relationship between each pair of outcome groups was violated for predictive variables. The analyses were further stratified into women in ages 40-49 and 50-64. The data preparation and analyses were performed with SAS Enterprise Guide 6.1 (SAS Institute Inc, Cary, NC).

Results

Patient characteristics by screening group are presented in Table 1. Among 65,025 patients with breast cancer, 21.5% had no previous screening in the previous 33 months before their diagnosis, 38.1% were classified as annual screeners, and 40.4% were considered biennial screeners. Using the more restrictive definition of screening intervals, 11,335 patients were excluded because their screening pattern did not match our definitions. As a result, the proportions of biennial and annual screeners decreased (34.6% and 39.4% respectively). Compared to the no screening group, biennial screeners were slightly older while annual screeners were much older. Among the 3 groups, annual screeners used the most primary care, followed by biennial screeners, and finally the non-screeners, which was consistent with their screening behavior. Compared to biennial screeners, annual screeners

had slightly lower CCI scores. In fact, CCI scores were comparable between women with no screening and biennial screening. For screeners, more than half of the cancers were detected by screening. Particularly, there were more screening detected cancers among the annual than biennial screeners. More women in the annual screening group were post-menopausal compared to the other groups. In general, the distribution of variables were similar between our primary and the alternative and more restricted definition of screening groups, except the proportions of screening detected cancer decreased more than 10% for both screening groups.

The staging algorithm estimated the majority of breast cancer cases to be stages I and II. The overestimation of stages I and II was the result of applying the Nattinger algorithm to identify incident breast cancer and the staging algorithm to estimate breast cancer stage at diagnosis. The Nattinger algorithm can only identify breast cancer cases who received breast surgery. The staging algorithm misclassifies less than half of the true stage II cases as stage I, and stage III as stage II. However, we assumed that the misclassification pattern to be the same across different screening groups we didn't expect that patients with different screening behavior prior to the diagnosis were treated differently.

Table 1: Patient's characteristics by screening group.

N	Primary Definition of Screening Pattern				Alternative Definition of Screening Pattern			
	No Screening	Biennial	Annual	p-value	No Screening	Biennial	Annual	p-value
	13,998	26,272	24,755		13,998	21,133	18,559	
Age, %								

40-44	13.2	11.0	6.3	<.000 1	13.2	12.1	6.0	<.000 1
45-49	18.2	17.6	14.7		18.2	17.8	14.2	
50-54	21.8	22.8	22.2		21.8	22.6	21.8	
55-59	23.9	24.8	27.4		23.9	24.5	27.7	
60-64	22.8	23.9	29.4		22.8	23.1	30.2	
Primary Care Visits, %								
0-2	38.5	29.9	20.4	<.000 1	38.5	30.7	19.7	<.000 1
3-6	25.0	27.4	26.2		25.0	27.1	26.2	
7-10	16.8	19.7	23.4		16.8	19.5	23.5	
>10	19.6	23.0	30.1		19.6	22.7	30.6	
Charlson Comorbidity Index								
0	85.7	84.9	86.1	<.000 1	85.7	85.1	86.2	<.000 1
1	10.9	11.8	11.3		10.9	11.6	11.2	
>=2	3.4	3.3	2.7		3.4	3.4	2.6	
Screening Detected Cancer, %								
Yes	0.0	57.7	65.2	<.000 1	0.0	47.5	52.5	<.000 1
Post-Menopausal Status, %								
Yes	74.7	75.8	81.7	<.000 1	74.7	74.8	82.2	<.000 1
Estimated Cancer Stage, %								
Stage 0	2.3	3.4	3.0	<.000 1	2.3	3.4	2.9	<.000 1
Stage I	34.5	64.9	72.5		34.5	62.3	72.4	
Stage II	51.9	26.4	20.7		51.9	28.6	20.8	
Stage III	8.3	4.3	3.2		8.3	4.6	3.3	
Stage IV	2.9	1.0	0.6		2.9	1.1	0.6	

We calculated the odds ratios of stage IV versus stage 0-III, stage III-IV versus stage 0-II, stage II-IV versus stage 0-I, and stage I-IV versus stage 0 for women diagnosed with breast cancer with no screening, annual screening or biennial screening in Table 2. The

partial proportional odds models were adjusted for age at diagnosis, primary care visits, CCI scores, method of cancer detection, and menopausal status.

Generally, we found that annual and biennial screening substantially reduced the odds of having more advanced stage at diagnosis compared to no screening. In addition, the protective effect of screening against more advanced breast cancer was more considerable for later stages. Annual screening did not increase the odds of developing invasive breast cancer than no screening. Moreover, neither annual nor biennial screening was associated with having invasive versus non-invasive disease when we applied the narrow screening interval definitions. Among the regular screeners, more frequent screening decreased the odds of having more advanced (stage II and above) cancer. However, no association was observed between screening interval and invasive stage I through IV breast cancer. In the cohort identified with alternative screening pattern definition, the above associations were consistently observed.

Table 2. Multivariable associations of mammography screening pattern with breast cancer stage at diagnosis.

		Primary Definition of Screening Pattern			Primary Definition of Screening Pattern		
		OR	95% Confidence Interval	P-value	OR	95% Confidence Interval	P-value
Stage IV vs. Stage 0-III							
Screening Pattern	Annual vs. No Screening	0.333	(0.271 , 0.409)	<.0001	0.32	(0.255 , 0.403)	<.0001
	Biennial vs. No Screening	0.529	(0.445 , 0.629)	<.0001	0.542	(0.454 , 0.647)	<.0001

	Annual vs. Biennial	0.629	(0.514 , 0.77)		0.591	(0.472 , 0.741)	
Age	40-44	Reference			Reference		
	45-49	1.234	(0.932 , 1.634)	0.1426	1.248	(0.936 , 1.664)	0.131
	50-54	1.297	(0.99 , 1.7)	0.0588	1.289	(0.976 , 1.702)	0.0733
	55-59	1.104	(0.828 , 1.473)	0.4998	1.161	(0.863 , 1.561)	0.323
	60-64	0.963	(0.718 , 1.294)	0.8043	0.975	(0.719 , 1.322)	0.8681
Primary Care Visits	0-2	Reference			Reference		
	3-6	1.009	(0.966 , 1.054)	0.6857	0.997	(0.951 , 1.046)	0.9155
	7-10	0.982	(0.936 , 1.03)	0.4474	0.957	(0.909 , 1.007)	0.093
	>10	0.982	(0.938 , 1.028)	0.4363	0.955	(0.909 , 1.004)	0.07
Screening Detected vs. Symptomatic Cancer		0.384	(0.315 , 0.47)	<.0001	0.428	(0.344 , 0.534)	<.0001
Post- vs. Pre-Menopausal		1.23	(1.011 , 1.496)	0.0387	1.185	(0.967 , 1.452)	0.1025
Charlson Comorbidity Index	0	Reference			Reference		
	1	1.389	(1.138 , 1.696)	0.0012	1.403	(1.14 , 1.728)	0.0014
	>=2	1.943	(1.434 , 2.634)	<.0001	1.936	(1.408 , 2.664)	<.0001
Stage III-IV vs. Stage 0-II							
Screening Pattern	Annual vs. No Screening	0.467	(0.425 , 0.514)	<.0001	0.47	(0.424 , 0.522)	<.0001
	Biennial vs. No Screening	0.619	(0.57 , 0.673)	<.0001	0.642	(0.589 , 0.699)	<.0001
	Annual vs. Biennial	0.754	(0.662 , 0.821)		0.733	(0.666 , 0.807)	
Age	40-44	Reference			Reference		
	45-49	0.982	(0.869 , 1.11)	0.7701	0.971	(0.854 , 1.105)	0.6583
	50-54	0.898	(0.796 , 1.012)	0.0779	0.911	(0.803 , 1.033)	0.1446
	55-59	0.715	(0.629 , 0.812)	<.0001	0.75	(0.656 , 0.858)	<.0001

	60-64	0.622	(0.546 , 0.709)	<.0001	0.653	(0.569 , 0.75)	<.0001
Primary Care Visits	0-2	Reference			Reference		
	3-6	1.009	(0.966 , 1.054)	0.6857	0.997	(0.951 , 1.046)	0.9155
	7-10	0.982	(0.936 , 1.03)	0.4474	0.957	(0.909 , 1.007)	0.093
	>10	0.982	(0.638 , 1.028)	0.4363	0.955	(0.909 , 1.004)	0.07
Screening Detected vs. Symptomatic Cancer		0.525	(0.483 , 0.571)	<.0001	0.535	(0.487 , 0.587)	<.0001
Post- vs. Pre-Menopausal		1.412	(1.288 , 1.549)	<.0001	1.359	(1.232 , 1.498)	<.0001
Charlson Comorbidity Index	0	Reference			Reference		
	1	1.288	(1.166 , 1.421)	<.0001	1.296	(1.166 , 1.439)	<.0001
	>=2	1.491	(1.261 , 1.763)	<.0001	1.525	(1.279 , 1.818)	<.0001
Stage II-IV vs. Stage 0-I							
Screening Pattern	Annual vs. No Screening	0.334	(0.317 , 0.351)	<.0001	0.341	(0.323 , 0.361)	<.0001
	Biennial vs. No Screening	0.444	(0.423 , 0.465)	<.0001	0.476	(0.543 , 0.5)	<.0001
	Annual vs. Biennial	0.753	(0.723 , 0.784)		0.717	(0.685 , 0.751)	
Age	40-44	Reference			Reference		
	45-49	0.896	(0.838 , 0.959)	0.0015	0.884	(0.822 , 0.95)	0.0008
	50-54	0.737	(0.69 , 0.788)	<.0001	0.731	(0.681 , 0.785)	<.0001
	55-59	0.608	(0.567 , 0.653)	<.0001	0.609	(0.564 , 0.75)	<.0001
	60-64	0.558	(0.519 , 0.599)	<.0001	0.558	(0.517 , 0.603)	<.0001
Primary Care Visits	0-2	Reference			Reference		
	3-6	1.009	(0.966 , 1.054)	0.6857	0.997	(0.951 , 1.046)	0.9155
	7-10	0.982	(0.936 , 1.03)	0.4474	0.957	(0.909 , 1.007)	0.093
	>10	0.982	(0.938 , 1.028)	0.4363	0.955	(0.909 , 1.004)	0.07
Screening Detected vs. Symptomatic Cancer		0.4	(0.385 , 0.417)	<.0001	0.391	(0.374 , 0.409)	<.0001

Post- vs. Pre-Menopausal		1.607	(1.528 , 1.69)	<.0001	1.603	(1.518 , 1.693)	<.0001
Charlson Comorbidity Index	0						
	1	1.153	(1.092 , 1.217)	<.0001	1.173	(1.105 , 1.245)	<.0001
	>=2	1.152	(1.044 , 1.271)	0.005	1.175	(1.056 , 1.307)	0.0032
Stage I-IV vs. Stage 0							
Screening Pattern	Annual vs. No Screening	0.962	(0.825 , 1.123)	0.6262	0.988	(0.838 , 1.165)	0.8874
	Biennial vs. No Screening	0.885	(0.763 , 1.026)	0.1059	0.853	(0.733 , 0.993)	0.0406
	Annual vs. Biennial	1.087	(0.984 , 1.202)		1.158	(1.032 , 1.3)	
Age	40-44	Reference			Reference		
	45-49	1.26	(1.077 , 1.475)	0.004	1.207	(1.016 , 1.435)	0.0322
	50-54	1.187	(1.016 , 1.386)	0.0303	1.154	(0.972 , 1.368)	0.1011
	55-59	1.046	(0.87 , 1.258)	0.6319	0.969	(0.79 , 1.188)	0.7635
	60-64	1.25	(1.034 , 1.51)	0.0209	1.203	(0.974 , 1.486)	0.0859
Primary Care Visits	0-2	Reference			Reference		
	3-6	1.009	(0.966 , 1.054)	0.6857	0.997	(0.951 , 1.046)	0.9155
	7-10	0.982	(0.936 , 1.03)	0.4474	0.957	(0.909 , 1.007)	0.093
	>10	0.982	(0.938 , 1.028)	0.4363	0.955	(0.909 , 1.004)	0.07
Screening Detected vs. Symptomatic Cancer		0.655	(0.588 , 0.729)	<.0001	0.674	(0.598 , 0.759)	<.0001
Post- vs. Pre-Menopausal		2.238	(1.967 , 2.545)	<.0001	2.327	(2.015 , 2.688)	<.0001
Charlson Comorbidity Index	0	Reference			Reference		
	1	0.973	(0.841 , 1.125)	0.7122	0.9	(0.769 , 1.055)	0.1944
	>=2	1.717	(1.211 , 2.436)	0.0024	1.666	(1.133 , 2.45)	0.0095

Stratification by age revealed additional patterns (Table 3). Among women in their 40s, annual screeners were more likely to have invasive breast cancer than biennial screeners. Additionally, with narrow screening intervals, the advantage of annual over biennial screening in avoiding stage II and IV cancers was no longer significant. However, the odds of being diagnosed with stage II-IV cancer was still smaller with annual screening compared to biennial screening. The findings from women aged 50-64 years were similar to the main model, since they consisted of two thirds of our study cohort.

Table 3. Multivariable associations of mammography screening pattern with breast cancer stage at diagnosis by age.

		Primary Definition of Screening Pattern			Primary Definition of Screening Pattern		
		OR	95% Confidence Interval	P-value	OR	95% Confidence Interval	P-value
Age 40-49							
Stage IV vs. Stage 0-III							
Screening Pattern	Annual vs. No Screening	0.306	(0.196 , 0.479)	<.0001	0.365	(0.227 , 0.588)	<.0001
	Biennial vs. No Screening	0.614	(0.445 , 0.847)	0.003	0.641	(0.461 , 0.89)	0.0025
	Annual vs. Biennial	0.498	(0.325 , 0.765)		0.57	(0.359 , 0.905)	
Stage III-IV vs. Stage 0-II							
Screening Pattern	Annual vs. No Screening	0.505	(0.422 , 0.604)	<.0001	0.534	(0.437 , 0.652)	<.0001

	Biennial vs. No Screening	0.693	(0.599 , 0.803)	0.0006	0.726	(0.626 , 0.843)	0.0079
	Annual vs. Biennial	0.728	(0.62 , 0.855)		0.735	(0.611 , 0.885)	
Stage II-IV vs. Stage 0-I							
Screening Pattern	Annual vs. No Screening	0.311	(0.282 , 0.343)	<.0001	0.317	(0.285 , 0.353)	<.0001
	Biennial vs. No Screening	0.424	(0.388 , 0.463)	<.0001	0.454	(0.415 , 0.497)	<.0001
	Annual vs. Biennial	0.733	(0.677 , 0.794)		0.699	(0.638 , 0.766)	
Stage I-IV vs. Stage 0							
Screening Pattern	Annual vs. No Screening	0.945	(0.72 , 1.243)	0.6909	0.908	(0.678 , 1.216)	0.5178
	Biennial vs. No Screening	0.774	(0.602 , 0.995)	0.0455	0.707	(0.549 , 0.912)	0.0075
	Annual vs. Biennial	1.222	(1.022 , 1.463)		1.284	(1.042 , 1.582)	
Age 50-64							
Stage IV vs. Stage 0-III							
Screening Pattern	Annual vs. No Screening	0.337	(0.267 , 0.425)	<.0001	0.31	(0.239 , 0.403)	<.0001
	Biennial vs. No Screening	0.493	(0.401 , 0.606)	<.0001	0.504	(0.407 , 0.624)	<.0001
	Annual vs. Biennial	0.684	(0.543 , 0.862)		0.615	(0.474 , 0.798)	
Stage III-IV vs. Stage 0-II							

Screening Pattern	Annual vs. No Screening	0.446	(0.399 , 0.5)	<.0001	0.443	(0.392 , 0.502)	<.0001
	Biennial vs. No Screening	0.581	(0.524 , 0.644)	<.0001	0.6	(0.54 , 0.667)	<.0001
	Annual vs. Biennial	0.768	(0.695 , 0.85)		0.739	(0.66 , 0.828)	
Stage II-IV vs. Stage 0-I							
Screening Pattern	Annual vs. No Screening	0.344	(0.324 , 0.366)	<.0001	0.352	(0.33 , 0.375)	<.0001
	Biennial vs. No Screening	0.453	(0.428 , 0.48)	<.0001	0.487	(0.459 , 0.517)	<.0001
	Annual vs. Biennial	0.76	(0.725 , 0.796)		0.722	(0.685 , 0.761)	
Stage I-IV vs. Stage 0							
Screening Pattern	Annual vs. No Screening	0.966	(0.801 , 1.165)	0.7162	1.018	(0.833 , 1.244)	0.8591
	Biennial vs. No Screening	0.954	(0.794 , 1.148)	0.62	0.948	(0.783 , 1.147)	0.5811
	Annual vs. Biennial	1.012	(0.896 , 1.143)		1.075	(0.933 , 1.238)	

Discussion

This observational study found substantial protective effects against being diagnosed with more advanced breast cancer with regular and more frequent mammography screening. However, those associations were not significant when comparing the odds of having

invasive cancer and non-invasive cancer except for the subgroup analysis among women aged 40-49.

In supplement to previous studies among women older than 65 (Badgwell et al., 2008; Galit, Green, & Lital, 2007; McCarthy et al., 2000; Vyas, Madhavan, & Sambamoorthi, 2014), we found that regular and more frequent mammography screening were associated with earlier cancer stage for women aged 40-64. However, the BCSC studies conducted including women in the younger age showed mixed results. The first BCSC finding (White et al., 2004) of the elevated risk of later stage invasive breast cancer associated with biennially versus annually was no longer statistically significant in the later BCSC studies (Dittus et al., 2013; Hubbard et al., 2011; Kerlikowske et al., 2013; Miglioretti et al., 2015; O'Meara et al., 2013). However, the similar trend observed in the later studies cannot rule out that the small number of incident cancers led to the broad confidence interval of their estimates (Hubbard et al., 2011). The BCSC studies first identified women with use of screening mammography, then followed them up for breast cancer development, which only yield a few thousands of incident breast cancer cases. Unlike the BCSC study design, our study started with women with incident breast cancer, and worked backwards to identify their mammography screening pattern. Therefore, our study cohort included over 60,000 incident breast cancer cases from 14 years of data, which explained our clinically significant findings between annual versus no screening, biennial versus no screening, and also annual versus biennial screening. Another feature of the BCSC cohort was that it included primarily more health conscious women receiving screening mammography in community practice.

Therefore, the BCSC database was not suitable to study the risk of later breast cancer stage comparing regular screeners to non-screeners.

Changing from wider screening intervals to classify women as annual or biennial screener to narrower ones (9-18 months vs. 11-14 for annual screening; 19-30 months vs. 23-26 biennial screening) didn't affect the conclusions. The intention of using alternative definitions was to evaluate a subgroup of women who were more closely adhere to the screening schedule. Consistent with the most recent BCSC study, our findings did not change with the definitions of screening intervals.

Our study found that women aged 40-49 would still benefit from regular and more intense mammography screening, which was in contrast to other BCSC studies (Hubbard et al., 2011; Miglioretti et al., 2015). However, Kerlikowske et al. (2013) did find increased risk of advanced-stage breast cancer among women with extremely dense breasts receive biennial instead of annual screening.

One interesting finding in our stratified analyses by age was the reversed effect of that annual screeners were more likely to be diagnosed with invasive versus non-invasive cancers than biennial screeners among women aged 40-49. This might be the result of missing confounders in our model. For example, women in their 40s might have self-selected to have more frequent screening because they were at higher risk of breast cancer. Unfortunately, with medical claims data, one is unable to accurately identify the risk breast cancer of patients based on genetic testing results and childhood chest radiation. Another possibility is

that we only find the association by chance, which explains the wide confidence interval close to 1.

Our findings are subject to several limitations. Due to the lack of clinical information in the commercial medical claims database, we have to predict breast cancer stage at diagnosis based on a validated algorithm using information extracted from claims including diagnosis codes, diagnostic and surgical procedures, treatment regimens and prescriptive medication. The algorithm has high accuracy predicting stage 0 and I cancers. However, it has less accuracy distinguishing stage II from stage I cases, as well as stage III from stage II because of the similarities in the treatment recommendations from the NCCN guidelines (Bevers et al., 2018). Therefore, the predicted disease stages are expected to be skewed towards stage I and II, as what we have seen in our stage distributions. However, we assume that the predicting error is systematic that similar misclassification is expected in every screening group. Because it is reasonable to believe that the medical practice of breast cancer diagnosis and treatment, as well as claims documentation don't change with a patient's previous mammography screening behavior.

The generalizability of our study is limited because of the eligibility criteria we use to select our study cohort. First, incident breast cancer cases are identified using Nattinger algorithm, which requires breast directed surgery to be eligible. Our study cohort therefore are missing a portion of non-invasive and metastatic cancers because they are not candidate for surgery. Second, continuous enrollment of 45 months is one of our eligibility criteria, which means patients need to be stably employed for almost 4 years. The fact that

commercial claims database consists of individuals covered by private insurance makes our findings less generalizable to population with low socioeconomic status.

Another limitation is, factors such as body mass index (BMI), race/ethnicity, and family history of cancer can potentially confound our results. Previous studies have identified elevated risk of more advanced stage associated with biennial versus annual screening among Hispanic women aged 50-74 years, Asian women aged 40-49 (O'Meara et al., 2013), premenopausal women (Miglioretti et al., 2015), premenopausal obese women (Dittus et al., 2013), and women with extremely dense breast (Kerlikowske et al., 2013). As a secondary analysis of a retrospective cohort using administrative claims data, lacking of clinical information has always been a limitation. However, considering the magnitude of association found in our study, it is unlikely that the difference in risk can be fully explained by BMI, race/ethnicity, and family history.

Although our analyses have shown substantial benefits of regular and more frequent against more advanced breast cancer, the decision about the appropriate screening interval should be an individual after weighting associated benefits and harms. Other factors for consideration include false-positive mammograms, diagnostic tests and surgical procedures, costs and psychological anxiety. Ideally, our findings should be further confirmed with another large population-based commercial claims database linked with verified clinical information.

Conclusion

Women with regular and more intense mammography screening (annual vs. biennial) are less likely to be diagnosed at later disease stage for women younger or older than 50. In the absence of head to head RCT comparing annual to biennial screening, one important use of our results is in decision models that predict the effectiveness or cost-effectiveness of various screening strategies (Mandelblatt et al., 2009; O'Donoghue, Eklund, Ozanne, & Esserman, 2014; Schousboe, Kerlikowske, Loh, & Cummings, 2011; Stout et al., 2014; Van Ravesteyn et al., 2012). The results of current study add to the evidence of the potential benefits and harms of mammography screening. Our findings can also help healthcare professionals in guideline development, decision makers and advocate groups in screening campaign planning, and women in decision making on screening intervals.

References

- Badgwell, B. D., Giordano, S. H., Duan, Z. Z., Fang, S., Bedrosian, I., Kuerer, H. M., . . . Babiera, G. (2008). Mammography before diagnosis among women age 80 years and older with breast cancer. *Journal of clinical oncology*, 26(15), 2482-2488.
- Bevers, T. B., Helvie, M., Bonaccio, E., Calhoun, K. E., Daly, M. B., Farrar, W. B., . . . Greenup, R. (2018). Breast cancer screening and diagnosis, version 3.2018, NCCN clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network*, 16(11), 1362-1389.
- Braithwaite, D., Zhu, W., Hubbard, R. A., O'Meara, E. S., Miglioretti, D. L., Geller, B., . . . Mandelblatt, J. (2013). Screening outcomes in older US women undergoing multiple mammograms in community practice: does interval, age, or comorbidity score affect tumor characteristics or false positive rates? *Journal of the National Cancer Institute*, 105(5), 334-341.
- BreastScreeningFrequencyTrialGroup. (2002). The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. *European Journal of Cancer*, 38(11), 1458-1464.
- Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, 40(5), 373-383.

- Dittus, K., Geller, B., Weaver, D. L., Kerlikowske, K., Zhu, W., Hubbard, R., . . . Consortium, B. C. S. (2013). Impact of mammography screening interval on breast cancer diagnosis by menopausal status and BMI. *Journal of general internal medicine*, 28(11), 1454-1462.
- Fenton, J. J., Zhu, W., Balch, S., Smith-Bindman, R., Fishman, P., & Hubbard, R. A. (2014). Distinguishing screening from diagnostic mammograms using Medicare claims data. *Medical care*, 52(7), e44.
- Freedman, G. M., Anderson, P. R., Goldstein, L. J., Hanlon, A. L., Cianfrocca, M. E., Millenson, M. M., . . . Nicolaou, N. (2003). Routine mammography is associated with earlier stage disease and greater eligibility for breast conservation in breast carcinoma patients age 40 years and older. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 98(5), 918-925.
- Galit, W., Green, M. S., & Lital, K.-B. (2007). Routine screening mammography in women older than 74 years: a review of the available data. *Maturitas*, 57(2), 109-119.
- Goel, A., Littenberg, B., & Burack, R. C. (2007). The association between the pre-diagnosis mammography screening interval and advanced breast cancer. *Breast cancer research and treatment*, 102(3), 339-345.
- Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane database of systematic reviews*(6).
- Hubbard, R. A., Kerlikowske, K., Flowers, C. I., Yankaskas, B. C., Zhu, W., & Miglioretti, D. L. (2011). Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Annals of internal medicine*, 155(8), 481-492.
- Kerlikowske, K., Zhu, W., Hubbard, R. A., Geller, B., Dittus, K., Braithwaite, D., . . . Consortium, B. C. S. (2013). Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. *JAMA internal medicine*, 173(9), 807-816.
- Krause, T. M., Ganduglia-Cazaban, C., Piller, L. B., & Venkataraman, V. Comparison of utilization of urgent care and primary care 2011-2015.
- Lauby-Secretan, B., Scoccianti, C., Loomis, D., Benbrahim-Tallaa, L., Bouvard, V., Bianchini, F., & Straif, K. (2015). Breast-cancer screening—viewpoint of the IARC Working Group. *New England Journal of Medicine*, 372(24), 2353-2358.
- Li, N., Hao, Y., Kageleiry, A., Peebles, M., Fang, A., Koo, V., . . . Guérin, A. (2016). Time on treatment of everolimus and chemotherapy among postmenopausal women with hormone-receptor-positive/human-epidermal-growth-factor-receptor-2-negative metastatic breast cancer: a retrospective claims study in the US. *Current medical research and opinion*, 32(2), 385-394.
- Mandelblatt, J. S., Cronin, K. A., Bailey, S., Berry, D. A., De Koning, H. J., Draisma, G., . . . Plevritis, S. K. (2009). Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Annals of internal medicine*, 151(10), 738-747.

- Mango, V., Bryce, Y., Morris, E. A., Gianotti, E., & Pinker, K. (2018). Commentary ACOG Practice Bulletin July 2017: Breast Cancer Risk Assessment and Screening in Average-Risk Women. *The British journal of radiology*, 91(1090), 20170907.
- McCarthy, E. P., Burns, R. B., Freund, K. M., Ash, A. S., Shwartz, M., Marwill, S. L., & Moskowitz, M. A. (2000). Mammography use, breast cancer stage at diagnosis, and survival among older women. *Journal of the American Geriatrics society*, 48(10), 1226-1233.
- Miglioretti, D. L., Zhu, W., Kerlikowske, K., Sprague, B. L., Onega, T., Buist, D. S., . . . Smith, R. A. (2015). Breast tumor prognostic characteristics and biennial vs annual mammography, age, and menopausal status. *JAMA oncology*, 1(8), 1069-1077.
- Monticciolo, D. L., Newell, M. S., Hendrick, R. E., Helvie, M. A., Moy, L., Monsees, B., . . . Sickles, E. A. (2017). Breast cancer screening for average-risk women: recommendations from the ACR Commission on Breast Imaging. *Journal of the American College of Radiology*, 14(9), 1137-1143.
- Nattinger, A. B., Laud, P. W., Bajorunaite, R., Sparapani, R. A., & Freeman, J. L. (2004). An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health services research*, 39(6p1), 1733-1750.
- Nelson, H. D., Tyne, K., Naik, A., Bougatsos, C., Chan, B. K., & Humphrey, L. (2009). Screening for breast cancer: an update for the US Preventive Services Task Force. *Annals of internal medicine*, 151(10), 727-737.
- O'Donoghue, C., Eklund, M., Ozanne, E. M., & Esserman, L. J. (2014). Aggregate cost of mammography screening in the United States: comparison of current practice and advocated guidelines. *Annals of internal medicine*, 160(3), 145-153.
- O'Meara, E. S., Zhu, W., Hubbard, R. A., Braithwaite, D., Kerlikowske, K., Dittus, K. L., . . . Miglioretti, D. L. (2013). Mammographic screening interval in relation to tumor characteristics and false - positive risk by race/ethnicity and age. *Cancer*, 119(22), 3959-3967.
- Oeffinger, K. C., Fontham, E. T., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y.-C. T., . . . LaMonte, S. J. (2015). Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *Jama*, 314(15), 1599-1614.
- Physicians, A. A. o. F., & Physicians, A. A. o. F. (2012). Summary of recommendations for clinical preventive services. *Leawood, KS: American Academy of Family Physicians*.
- Sarrazin, M. S. V., & Rosenthal, G. E. (2012). Finding pure and simple truths with administrative data. *Jama*, 307(13), 1433-1435.
- Schousboe, J. T., Kerlikowske, K., Loh, A., & Cummings, S. R. (2011). Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Annals of internal medicine*, 155(1), 10-20.
- Siu, A. L. (2016). Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*, 164(4), 279-296.
- Stout, N. K., Lee, S. J., Schechter, C. B., Kerlikowske, K., Alagoz, O., Berry, D., . . . De Koning, H. J. (2014). Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. *JNCI: Journal of the National Cancer Institute*, 106(6).

- Van Ravesteyn, N. T., Miglioretti, D. L., Stout, N. K., Lee, S. J., Schechter, C. B., Buist, D. S., . . . Alagoz, O. (2012). Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. *Annals of internal medicine*, 156(9), 609-617.
- Vyas, A., Madhavan, S., & Sambamoorthi, U. (2014). Association between persistence with mammography screening and stage at diagnosis among elderly women diagnosed with breast cancer. *Breast cancer research and treatment*, 148(3), 645-654.
- White, E., Miglioretti, D. L., Yankaskas, B. C., Geller, B. M., Rosenberg, R. D., Kerlikowske, K., . . . Buist, D. S. (2004). Biennial versus annual mammography and the risk of late-stage breast cancer. *Journal of the National Cancer Institute*, 96(24), 1832-1839.
- Wilt, T. J., Harris, R. P., & Qaseem, A. (2015). Screening for cancer: advice for high-value care from the American College of Physicians. *Annals of internal medicine*, 162(10), 718-725.

JOURNAL ARTICLE

Healthcare Costs of Breast Cancer by Mammography Screening Behavior among Commercially Insured Women Aged 40-64 Health Affairs

Abstract

Mammography screening saves lives, but the appropriate screening interval for younger women remains controversial. Little is known about the financial burden borne by health plan payers and women with breast cancer. From the perspective of insurers and patients, our study provides dollar estimates of healthcare costs from patients with different screening behavior (annual, biennial and none) identified 33 months before cancer diagnosis based on the MarketScan commercial claims database in the years 1999-2014. A generalized linear model was used to adjust for potential confounders. The first-year healthcare costs from payer's perspective for no screening, annual screening and biennial screening were \$117,317, \$99,615, and \$103,364, respectively. And the out-of-pocket costs paid by patients were \$7,237, \$6,660, and \$6,569. The reported estimates appear to be much higher than previously documented. For health plan payers, there are cost savings associated with regular and more frequent screening. For patients, regular screening reduces out-of-pocket costs but the difference under different screening intervals is minimal. While regular mammography screening substantially reduces healthcare costs, the decision of annual versus biennial screening should be based on individual weighing of the benefits and harms.

Introduction

Breast cancer is the most common cancer and second leading cause of cancer death among women in the United States (Street 2019). Over the years of 1989 to 2016, the female breast cancer mortality has been decreasing steadily due to early detection (Howlader et al., 2019). Screening mammography is considered the gold standard in breast cancer detection. Controversially, large organizations like the US Preventive Services Task Force (USPSTF) and the American Cancer Society (ACS) endorse different mammography screening intervals for average-risk women, based on different interpretations on the existing evidences: the USPSTF suggests biennial screening (Siu, 2016), while the ACS recommended annual screening for women ages 40-54 and biennial screening for women older than 55 (Oeffinger (Oeffinger et al., 2015). In line with the controversies, the Affordable Care Act requires that all new health insurance plans to fully cover the costs mammography screening for women older than 40 every 1-2 years with no deductibles or copayments or otherwise share costs since 2010 (Cassidy, 2010). So far, there is no real-world evidence revealing the financial implications for both insurance payers and breast cancer patients following annual or biennial screening recommendations.

Multiple simulation models have been built to predict the cost, effectiveness, and cost-effectiveness of different breast cancer screening strategies (Lindfors & Rosenquist, 1995; Rosenquist & Lindfors, 1998; Stout et al., 2014). With the modeling approach, the short- or long-term per patient healthcare costs were predicted based on the number of breast cancer cases detected, and the corresponding treatment costs by stage. These estimations

were mainly based on various modeling assumptions and may not represent of the real-world financial burdens. Moreover, most of the modeling studies were conducted from a societal or healthcare payer perspective, but rarely from a patient's perspective. Farley et al. (2015) reported the observed healthcare costs by stage using Medicaid claims data linked to the Georgia Comprehensive Cancer Registry (Farley et al., 2015). They also simulated the cost of treatment under the USPSTF and ACS screening guidelines respectively and compared the observed cost to the simulated ones. Their findings suggested that per patient healthcare costs under the ACS guidelines was lower than the observations (\$4,171), while the USPSTF scenario resulted in a higher cost than observed (\$2,188). However, their cost estimation was based on a limited sample of 274 African-America female breast cancer patients from the Georgia Cancer Registry, which restricted the generalizability of their findings. In addition, the use of the Georgia Medicaid claims data only informed the costs of breast cancer treatment incurred to the state-level Medicaid program covering low-income and disadvantaged population.

The objective of this study was to estimate the per patient healthcare costs in the initial and continuing treatment phases of incident breast cancer associated with different mammography screening intervals, from a large U.S. commercial insurance claims database. The real-world cost estimates-as opposed to model simulated numbers-provide more realistic evidence to help health insurance, healthcare providers, employees, and healthcare organizations to advocate and promote more appropriate mammography screening.

Methods

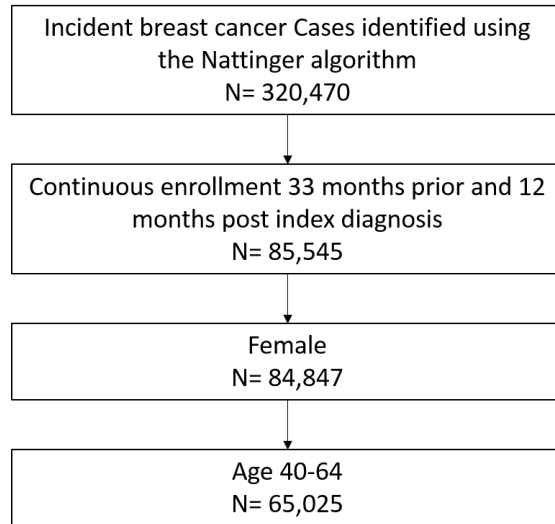
Data Source and Study Population

In this retrospective case-control study, the MarketScan® Commercial Claims and Encounters and Medicare Supplemental and Coordination of Benefits Database was used to identify the study population and perform cost analyses. This study was based on all healthcare costs because we took the perspective of healthcare payers. The MarketScan claims database includes de-identified employees, spouses and dependents in employer-sponsored private insurance in the U.S., tracked across healthcare systems. With over 300 employers and 40 health plans, the MarketScan claims database contains fully paid and adjudicated service level claims for 150 million unique persons. The claims database captures information on enrollment, healthcare utilization and expenditures for prescription drugs and outpatient and inpatient encounters. This claims and encounters database contains detailed tables of enrollment information, inpatient admissions, facility information, inpatient and outpatient services, outpatient pharmaceutical claims, aggregated populations information, and annual enrollment summaries. Therefore, the MarketScan database enables the analysis of real-world treatment patterns and costs. The analysis of this de-identified database was exempt from institutional human subjects review.

From the MarketScan database, we selected female breast cancer incident cases of women older than 40 at time of diagnosis between the years of 1/1/1999-10/1/2014 using a modified algorithm (Nattinger, Laud, Bajorunaite, Sparapani, & Freeman, 2004). According to the algorithm, the index diagnosis date for each case was defined as the date of first breast cancer surgery. Eligible women should be continuously enrolled in health insurance plan 33

months before and until 12 months after diagnosis index date to allow observation of biennial screening (30 months) plus diagnosis period (3 months), and to capture all healthcare costs.

Figure 1: Study flowchart.



Eligible women after inclusion/exclusion criteria were categorized based on their mammography screening behavior before breast cancer diagnosis. Screening mammography was identified and distinguished from diagnostic mammography using a 3-step claims-based algorithm (Fenton et al., 2014). The algorithm is expected to identify screening mammograms with 97.1% sensitivity, 69.4% specificity, and a positive predictive value of 94.9%. The categories of screening pattern were defined as follows: non-screeners had 0 mammogram in the 33 months before the index diagnosis date; annual screeners had 2 screening mammograms in 33 months from the index diagnosis date, which are 9-18 months apart; biennial screeners had 1 screening mammogram 33 months from the index diagnosis

date, or have 2 screening mammograms in 33 months from the index diagnosis date, which are 19-30 months apart.

Costs Estimation and Statistical Analysis

Baseline characteristics were compared among the 3 groups of breast cancer incident cases based on screening pattern. The primary outcomes were the mean model-adjusted healthcare costs among the 3 groups in the 3 months prior to the index date and 1 year after (27 months in total) discussed in the following paragraph. We included the costs 3 months before the index date to cover any possible neoadjuvant therapy before definitive breast surgery. The healthcare costs in our study were assessed from 3 different perspectives: insurance payer perspective (insurer costs), patient perspective (out-of-pocket costs), and payer plus patient perspective (allowed amounts). The total amounts paid by health insurance plans were the main outcome of interest because these numbers were the most accurate in commercial administrative claims data. In addition, we reported the total allowed amounts payable to providers includes payments from the health insurance plans, patients, and any other sources. Out-of-pocket costs were payments made by patients that included copayment, coinsurance and deductibles.

We measured and included the following variables as covariates: age at diagnosis, type of insurance plan, employment status, geographic area, the Charlson Comorbidity Index (CCI) score (Charlson, Pompei, Ales, & MacKenzie, 1987), psychiatric diagnostic groups (PDG) score (Ashcraft et al., 1989), counts of primary care visits in the past year, whether cancer was screening detected, and menopausal status. Incident cancer cases following a

screening mammography less than a year were considered screening detected by a validated algorithm (Fenton et al., 2016). Women older than 55 years, who used fulvestrant or aromatase inhibitors treatment, had bilateral oophorectomy or an International Classification of Diseases (ICD)-9 diagnosis code related to postmenopausal status were considered postmenopausal (Li et al., 2016; Miglioretti et al., 2015).

The 3 types of healthcare costs were adjusted by different generalized linear models with log-link function and gamma distribution (Glick, Doshi, Sonnad, & Polsky, 2014). Akaike Information Criterion (AIC) value (Akaike, 1981) was used to find the best fit of model. All healthcare costs were inflated to 2018 US dollars using the Medical Care component of the U.S. Bureau of Labor Statistics Consumer Price Index (CPI) (www.bls.gov; accessed September 30, 2019). The model adjusted costs were reported by 3 screening groups, and further stratified by age groups 40-49 and 50-64.

Results

We identified 13,998 non-screeners, 24,755 annual screeners, and 26,272 biennial screeners (Table 1). On average, annual screeners were older, post-menopausal at diagnosis, with more cancer detected by screening mammography, had more primary visits but the lowest pre-diagnosis healthcare costs. Biennial screeners had the highest proportion of full time employees, followed by annual screener and then non-screeners. Overall, more than half of the patients were covered by a preferred provider organization. Highest percentage of patients resided in the East North Central region (Illinois, Indiana, Michigan, Ohio, and

Wisconsin) and South Atlantic region (Delaware, Maryland, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, and the District of Columbia).

Table 1: Patient's characteristics by screening group.

N	Primary Definition of Screening Pattern			
	No Screening	Annual	Biennial	p-value
	13,998	24,755	26,272	
Enrollment duration, months				
Mean	50.8	46.3	47.6	<.0001
SD	35.5	31.1	32.6	
Age at Diagnosis, years				
Mean	53.2	55.0	53.7	<.0001
SD	6.8	6.1	6.6	
Charlson Comorbidity Index				
Mean	0.2	0.2	0.2	<.0001
SD	0.6	0.5	0.6	
Psychiatric diagnostic groups scores				
Mean	0.1	0.1	0.1	<.0001
SD	0.4	0.4	0.5	
Primary Care Visits, counts				
Mean	6.3	8.5	7.1	<.0001
SD	8.1	8.0	7.5	
Screening Detected Cancer, %				
Yes	0.0	65.2	57.7	<.0001
Post-Menopausal Status, %				
Yes	74.7	81.7	75.8	<.0001
Health insurance plan type, %				
Comprehensive	9.7	7.2	7.4	<.0001
EPO/unknown	1.1	1.0	1.2	

HMO/POS with capitation/ POS	26.6	27.5	30.9	
PPO	57.7	56.8	54.2	
CDHP/HDHP	4.9	7.6	6.3	
Employment status, %				
Active full time	47.2	49.1	51.9	<.0001
Active part time or seasonal	0.8	1.0	1.0	
Retired	21.2	23.6	20.7	
COBRA beneficiary (continue)/ long-term disability/ surviving spouse/dependent	1.4	0.9	1.0	
Other/unknown	29.4	25.5	25.4	
Geographic area (division), %				
New England	3.3	6.3	5.1	<.0001
Middle Atlantic	9.7	8.7	8.4	
East North Central	19.4	21.0	20.0	
West North Central	3.7	5.1	4.2	
South Atlantic	27.7	21.9	22.5	
East South Central	11.0	9.9	9.2	
West South Central	8.9	8.5	8.2	

Mountain	3.5	4.2	4.5	
Pacific/missing	12.8	14.4	17.8	

Table 2 presents the model adjusted costs if the first year after the diagnosis by screening group and age group. Consistently, the allowed amounts, insurer costs and out-of-pocket costs per patient are highest in the no screening group, followed by biennial screening group, and finally the annual screening group. Despite the differences in groups, only the difference (around \$10,000) between non- and annual screeners in total allowed amounts was statistically significant (p-values in table 3). Insurer costs for non-screeners were significantly higher than biennial and annual screeners. But the small difference between the screening groups was not significant. For out-of-pocket costs, annual screeners spent less than biennial and non-screeners.

Table 2: Model Adjusted Health Care Costs by Group in the First Year After the Index Date*

No Screening			Biennial Screening		Annual Screening		P-value
Cost Estimate	Mean	(SD)	Mean	(SD)	Mean	(SD)	
All Age Groups							
Total Allowed Amounts	168,786	(4,739)	165,814	(3,435)	157,950	(3,612)	0.0027
Insurer Costs	128,921	(3,060)	118,496	(2,167)	113,254	(2,226)	<.0001
Out-of-pocket Costs	7,101	(203)	6,687	(147)	6,178	(146)	<.0001
Age 40-49							

Total Allowed Amounts	193,103	(10,940)	177,505	(8,742)	151,975	(8,675)	0.0079
Insurer Costs	145,235	(7,203)	131,066	(5,709)	114,517	(5,790)	0.0011
Out-of-pocket Costs	7,403	(447)	7,140	(368)	6,064	(370)	0.0629
Age 50-64							
Total Allowed Amounts	162,910	(5,871)	155,023	(3,882)	154,648	(3,993)	0.0027
Insurer Costs	126,181	(3,785)	114,255	(2,419)	110,012	(2,418)	<.0001
Out-of-pocket Costs	7,081	(255)	6,622	(169)	6,006	(157)	<.0001

*Values in table are mean costs per patient in U.S. dollars adjusted using the 2018 Consumer Price Index.

Stratified analyses showed substantial differences between age groups. Generally, younger women (40-49 years old) had higher costs than older women (50-64 years old). Additionally, we observed that the differences between age groups decreased with more intense screening. In other words, more frequent screening closed up the gaps between age groups. Noted that the total allowed amounts per patient among younger annual screeners were less than the older ones. Although we did not observed the same pattern for insurer costs and out-of-pocket costs, the differences among annual screeners between age groups were much smaller than biennial and non-screeners.

For the total allowed amounts and insurer costs, the savings from screenings were larger in younger women. Annual screening saved much more than biennial screening in

younger women when compared to non-screeners. On the other hand, the savings from annual and biennial screening were similar among older women.

Cost ratios were also reported in table 3 because relative costs were likely to be more stable than absolute costs across different populations. Cost ratios can be easily interpreted as a ratio of adjusted costs between two groups. In addition, a ratio informs the incremental aspect of the cost analysis. The results showed that for all age groups, both annual and biennial screening saved costs for insurance payers. Especially, annual screening significantly reduced total allowed amounts, insurer costs, and out-of-pocket costs when compared to no screening. However, when comparing annual to biennial screening, the cost reduction was only marginally significant for insurers. After we stratified the analysis by age group, annual screening significantly reduced healthcare costs for both insurance payers and women aged 40-49 compared to biennial screening. On the other hand, costs for younger biennial screeners were not significantly different from non-screeners. Consistently, annual screeners had less out-of-pocket costs than biennial and non-screeners across age groups.

Table 3: Estimated Cost Ratios for Mean Costs in the First Year after the Index Date*.

	Biennial Screening vs. No Screening			Annual Screening vs. No Screening			Annual Screening vs. Biennial Screening		
Cost Estimate	Mean	95% CI	P-value	Mean	95% CI	P-value	Mean	95% CI	P-value
All Age Groups									
Total Allowed Amounts	0.95	(0.89 , 1.01)	0.1256	0.94	(0.88 , 1.00)	0.0454	0.98	(0.93 , 1.04)	0.5346
Insurer Costs	0.92	(0.87 , 0.97)	0.0019	0.88	(0.83 , 0.93)	<.0001	0.96	(0.91 , 1.00)	0.0531
Out-of-pocket Costs	0.94	(0.88 , 1.00)	0.0693	0.87	(0.81 , 0.93)	<.0001	0.92	(0.87 , 0.98)	0.0057
Age 40-49									
Total Allowed Amounts	0.92	(0.81 , 1.05)	0.2104	0.79	(0.68 , 0.91)	0.0011	0.86	(0.75 , 0.98)	0.0207
Insurer Costs	0.90	(0.80 , 1.01)	0.0808	0.79	(0.69 , 0.89)	0.0002	0.87	(0.78 , 0.98)	0.0226
Out-of-pocket Costs	0.96	(0.84 , 1.11)	0.6117	0.82	(0.70 , 0.96)	0.0109	0.85	(0.74 , 0.98)	0.0218
Age 50-64									
Total Allowed Amounts	0.95	(0.88 , 1.03)	0.2232	0.95	(0.88 , 1.03)	0.2062	1.00	(0.94 , 1.06)	0.9392
Insurer Costs	0.91	(0.85 , 0.97)	0.0034	0.87	(0.82 , 0.93)	<.0001	0.96	(0.91 , 1.01)	0.1566
Out-of-pocket Costs	0.94	(0.86 , 1.01)	0.1028	0.85	(0.78 , 0.92)	<.0001	0.91	(0.85 , 0.97)	0.0025

*Values in table are mean costs per patient in U.S. dollars adjusted using the 2018 Consumer Price Index.

Discussion

To the best of our knowledge, total healthcare costs has not been reported by mammography screening frequency from real-world data. The savings in total healthcare costs of patients with breast cancer following regular mammography screening is substantial compared to no screening to the healthcare payers. On the other hand, the amounts of savings from more frequent screening (annual vs. biennial) is much smaller. The differences in financial burden of patients with and without screening are minimal (out-of-pocket costs less than \$1,000 in the first 2 years). And no statistically significant cost saving in out-of-pocket costs is found between annual and biennial screening.

We find higher financial burden to patients among advanced than early stage cases. However, patients with no screening mammography only paid a relatively small amount out-of-pocket (less than \$1,000 in the first 2 years). Very limited evidence is available on the financial burden of patients diagnosed with breast cancer. In a small sample of older patients with non-metastatic breast cancer, 40% of them self-reported the out-of-pocket costs related to breast cancer below \$500; 25% of the patients spent \$500-\$2,000; and 28% reported spent between \$2,001 and \$10,000. Patients spent more than \$10,000 only accounted for 7% (Jagsi et al., 2014). In our study, the median out-of-pocket costs is \$5,572 in the no screening group, \$5,318 in the annual screening group, and \$5,095 in the biennial screening group in the first year after diagnosis. Overall, the reported out-of-pocket costs in our study are much higher than previously reported even after adjustment for inflation. Potential explanation is that the patients recruited in the prior study were diagnosed between year 2005 and 2007. Surveying patients a long time from their diagnosis may lead to recall bias when reporting

out-of-pocket costs. Using medical claims data, our study captures payments from patients to the providers with higher accuracy.

Although we have seen treatment costs reduction in annual versus biennial screening groups, other costs need to be considered. For example, the cost of diagnostic workups following false-positive mammograms can be as high as \$766-\$852 per beneficiary according to a published study (Ong & Mandl, 2015). Since false-positive mammography are more common among women in their 40s (Armstrong, Moye, Williams, Berlin, & Reynolds, 2007; Carney et al., 2003; Elmore et al., 1998), the accumulated costs in population level can be large for annual versus biennial screening, and therefore outweigh the savings in all healthcare costs. However, studies have shown that there are multiple approaches to reduce false-positive rates such as the use of more advanced digital breast tomosynthesis (Ciatto et al., 2013; Haas et al., 2013) and screening women with dense breasts using ultrasound (Drukteinis, Mooney, Flowers, & Gatenby, 2013). A study also reported that women who screened previously reduced by at least half the incidence of false-positive result (Kleit & Ruiz, 2003).

In addition, the cost attributed to overdiagnosis can be substantial. Overdiagnosis refers to screening detected cancers that are never destined to cause symptoms or result in death. For breast cancers that grow and progress slowly, treatments may be unnecessary and harmful (Woloshin & Schwartz, 2010). Ong & Mandl (2015) estimated the annual national costs of overdiagnosis ranging from 0.5 billion to more than 1.5 billion using published national overdiagnosis rates. Future study is needed to evaluate whether the cost savings associated with annual screening still exists when considering the costs of overdiagnosis.

Besides financial consequences, the harms of mammography screening also include psychological effects following false-positive mammograms. According to a systematic review (Bond et al., 2013), women with false-positive results had elevated stress because of diagnostic procedures and were less likely to return for next screening, and therefore more likely to have interval cancer. Therefore, decisions regarding screening should be based on individual's weighing of benefits and harms.

Limitations

There are some limitations to be considered when interpreting our results. First, we used the Nattinger algorithm to identify incident breast cancer cases which requires definitive breast surgery. Therefore, our results only show the estimates and trend of healthcare costs among patients who were previously treated with surgery. In addition, the MarketScan data only comprised claims for individuals younger than 65 years old. Although our results can only be generalized to younger women with breast cancer, this population is the main interest in our study. As mentioned earlier, the controversy around the frequency of mammography centers around women aged 40-49. Therefore, this database is appropriate for our research objectives. Third, this study only reported total allowed amounts, insurer costs and out-of-pocket costs in the first year after the diagnosis. We chose to look the initial phase of breast cancer because the highest healthcare costs occurred during the first year after the diagnosis and the last year of life. Existing evidence showed that the annualized mean net costs of care in the initial year was 10 times more the continuing phase (Mariotto, Robin Yabroff, Shao, Feuer, & Brown, 2011). Moreover, unlike Medicare beneficiaries, the enrollees of

commercial healthcare plans have much shorter continuous enrollment. Therefore, estimating healthcare costs in the first year after diagnosis is appropriate to quantify the difference among women who had annual, biennial or none mammography screening before their cancer diagnosis. Finally, this cost analysis was based on total healthcare costs instead of breast cancer attributable costs. Since this study was conducted from an all healthcare payers' perspective, total healthcare costs were considered more relevant to the payers. It is intuitive that later breast cancer stage drives not only costs associated with breast cancer treatment, but also other healthcare costs which are part of payers' responsibility as well. Therefore, the estimates of total healthcare costs can better inform payers the financial burden following different mammography screening patterns.

Conclusion

Our study shows that commercial healthcare payers can save substantially in breast cancer treatment costs if women are regularly screened between age 40 and 65. In addition, there will be extra savings if women opt for annual versus biennial screening. From the perspective of patients, the difference between annual and biennial screening is negligible. From insurance payer's perspective, there is financial incentives to promote regular mammography screening because of the large cost savings. While the savings of annual over biennial is much smaller, the payers need to decide whether the cost of false-positive recalls and overdiagnosis can be offset. From the patient's perspective, there is limited financial incentive according to the cost sharing mechanism under commercial health plans. Although

regular mammography screening is strongly recommended, the decision on screening frequency should be based on individual perceptions of the benefits and harms of screening.

References

- Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of econometrics*, 16(1), 3-14.
- Armstrong, K., Moye, E., Williams, S., Berlin, J. A., & Reynolds, E. E. (2007). Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Annals of internal medicine*, 146(7), 516-526.
- Ashcraft, M. L., Fries, B. E., Nerenz, D. R., Falcon, S. P., Srivastava, S. V., Lee, C. Z., . . . Errera, P. (1989). A psychiatric patient classification system: an alternative to diagnosis-related groups. *Medical care*, 543-557.
- Bond, M., Pavey, T., Welch, K., Cooper, C., Garside, R., Dean, S., & Hyde, C. (2013). Systematic review of the psychological consequences of false-positive screening mammograms.
- Carney, P. A., Miglioretti, D. L., Yankaskas, B. C., Kerlikowske, K., Rosenberg, R., Rutter, C. M., . . . Dignan, M. (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.
- Cassidy, A. (2010). Preventive services without cost sharing. *Health Affairs website*. <http://www.healthaffairs.org/healthpolicybriefs/brief.php>.
- Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, 40(5), 373-383.
- Ciatto, S., Houssami, N., Bernardi, D., Caumo, F., Pellegrini, M., Brunelli, S., . . . Valentini, M. (2013). Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *The lancet oncology*, 14(7), 583-589.
- Drukteinis, J. S., Mooney, B. P., Flowers, C. I., & Gatenby, R. A. (2013). Beyond mammography: new frontiers in breast cancer screening. *The American journal of medicine*, 126(6), 472-479.
- Elmore, J. G., Barton, M. B., Mocer, V. M., Polk, S., Arena, P. J., & Fletcher, S. W. (1998). Ten-year risk of false positive screening mammograms and clinical breast examinations. *New England Journal of Medicine*, 338(16), 1089-1096.
- Farley, C., Friedman, D., Habtes, I., Raskind-Hood, C., Adams, E. K., Becker, E. R., . . . Gabram-Mendola, S. (2015). Screening mammography in a public hospital serving predominantly African-American women: a stage-survival-cost model. *Women's Health Issues*, 25(4), 322-330.
- Fenton, J. J., Onega, T., Zhu, W., Balch, S., Smith-Bindman, R., Henderson, L., . . . Hubbard, R. A. (2016). Validation of a Medicare claims-based algorithm for identifying breast cancers detected at screening mammography. *Medical care*, 54(3), e15.
- Fenton, J. J., Zhu, W., Balch, S., Smith-Bindman, R., Fishman, P., & Hubbard, R. A. (2014). Distinguishing screening from diagnostic mammograms using Medicare claims data. *Medical care*, 52(7), e44.

- Glick, H. A., Doshi, J. A., Sonnad, S. S., & Polsky, D. (2014). *Economic evaluation in clinical trials*: OUP Oxford.
- Haas, B. M., Kalra, V., Geisel, J., Raghu, M., Durand, M., & Philpotts, L. E. (2013). Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology*, 269(3), 694-700.
- Howlander, N., Noone, A., Krapcho, M., Miller, D., Brest, A., Yu, M., . . . Lewis, D. (2019). SEER Cancer Statistics Review, 1975-2016, National Cancer Institute. Bethesda, MD: National Cancer Institute. Bethesda, MD.
- Jagsi, R., Pottow, J. A., Griffith, K. A., Bradley, C., Hamilton, A. S., Graff, J., . . . Hawley, S. T. (2014). Long-term financial burden of breast cancer: experiences of a diverse cohort of survivors identified through population-based registries. *Journal of clinical oncology*, 32(12), 1269.
- Kleit, A. N., & Ruiz, J. F. (2003). False positive mammograms and detection controlled estimation. *Health services research*, 38(4), 1207-1228.
- Li, N., Hao, Y., Kageleiry, A., Peeples, M., Fang, A., Koo, V., . . . Guérin, A. (2016). Time on treatment of everolimus and chemotherapy among postmenopausal women with hormone-receptor-positive/human-epidermal-growth-factor-receptor-2-negative metastatic breast cancer: a retrospective claims study in the US. *Current medical research and opinion*, 32(2), 385-394.
- Lindfors, K. K., & Rosenquist, C. J. (1995). The cost-effectiveness of mammographic screening strategies. *Jama*, 274(11), 881-884.
- Mariotto, A. B., Robin Yabroff, K., Shao, Y., Feuer, E. J., & Brown, M. L. (2011). Projections of the cost of cancer care in the United States: 2010–2020. *Journal of the National Cancer Institute*, 103(2), 117-128.
- Miglioretti, D. L., Zhu, W., Kerlikowske, K., Sprague, B. L., Onega, T., Buist, D. S., . . . Smith, R. A. (2015). Breast tumor prognostic characteristics and biennial vs annual mammography, age, and menopausal status. *JAMA oncology*, 1(8), 1069-1077.
- Nattinger, A. B., Laud, P. W., Bajorunaite, R., Sparapani, R. A., & Freeman, J. L. (2004). An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health services research*, 39(6p1), 1733-1750.
- Oeffinger, K. C., Fontham, E. T., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y.-C. T., . . . LaMonte, S. J. (2015). Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *Jama*, 314(15), 1599-1614.
- Ong, M.-S., & Mandl, K. D. (2015). National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at \$4 billion a year. *Health affairs*, 34(4), 576-583.
- Rosenquist, C. J., & Lindfors, K. K. (1998). Screening mammography beginning at age 40 years: a reappraisal of cost - effectiveness. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 82(11), 2235-2240.
- Siu, A. L. (2016). Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*, 164(4), 279-296.
- Stout, N. K., Lee, S. J., Schechter, C. B., Kerlikowske, K., Alagoz, O., Berry, D., . . . De Koning, H. J. (2014). Benefits, harms, and costs for breast cancer screening after US

- implementation of digital mammography. *JNCI: Journal of the National Cancer Institute*, 106(6).
- Street, W. (2019). Cancer Facts & Figures 2019. *American Cancer Society: Atlanta, GA, USA*.
- Woloshin, S., & Schwartz, L. M. (2010). The benefits and harms of mammography screening: understanding the trade-offs. *Jama*, 303(2), 164-165.

CONCLUSION

An improved breast cancer staging algorithm based on medical claims data was developed and validated. This algorithm supplements claims data with clinical stage information, which expands the use of claims data in epidemiological studies. To my best knowledge, this is the first attempt to identify stage 0 cases from invasive cancers. As a one-step solution, this algorithm is able to predict stages 0-IV at once. Unlike other logistic models, this algorithm is based on CART method and easy for researchers to understand and adapt. Overall, the algorithm has very good predictive power for early stages as well as metastases because of the large pool of potential predictors from diagnoses, procedures and medication use. Noted that this algorithm was developed and validated in SEER-Medicare population, therefore the generalizability was limited considering the differences in treatment practice among older compared to younger patients. The use of SEER diagnosis month and year as index also raises concerns about the inconsistencies between the SEER and the Medicare diagnoses, although 90% of the SEER diagnosis time falls within a month of the first Medicare claim with a cancer diagnosis (CMS website; Lin & Virgo 2014). This algorithm was developed using multiple years of breast cancer cases and associated medical claims. The changes in clinical practice and the approval of new therapies over the years should be also noted. This algorithm requires a total of 22 months of continuous enrollment to capture needed information and therefore is not applicable for patients with shorter enrollment period. When this staging algorithm is applied to a pre-selected population (e.g. incident breast cancer cases selected by the Nattinger algorithm), the accuracy will be

affected because of the other selection criteria used. As the next step, a similar algorithm development study should be conducted using a claims database linked with clinical information that consists of women younger than 65.

With the estimated staging information, protective effect against more advanced breast cancer is found among women who screen regularly and on annual basis. The inverse association exists for women in their 40s and also for older women. The findings contribute to the mixed literature on whether women aged 40-64 benefit from annual screening. Although only estimated staging information is available, the use of different cut points to dichotomize cancer stages take into consideration of the uncertainty around the stage estimation. This association is only applicable for commercially insured women with at least 45 months of continuous enrollment. Another study with “gold standard” staging information is needed to confirm the findings among women younger than 65.

Finally, this dissertation reports real-world healthcare costs, insurer costs, and out-of-pocket costs by mammography screening pattern. There are substantial cost savings for insurers with regular and annual screenings in particular. The savings are larger for women aged 40-49 than older women. Overall younger women have higher costs, but more frequent screening reduces the difference between age groups. On the other hand, patient only save marginally from more intense screening. This costing study only reports total healthcare costs instead of breast cancer associated costs because all healthcare costs are insurer’s responsibilities from the payer’s perspective. Given the limited scope of the costs borne by women, future studies are needed to better capture the other indirect and productivity costs under different mammography screening pattern.

REFERENCES

- Alexander, F., Anderson, T., Brown, H., Forrest, A., Hepburn, W., Kirkpatrick, A., . . . Smith, A. (1999). 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *The Lancet*, 353(9168), 1903-1908.
- Andersson, I., Aspegren, K., Janzon, L., Landberg, T., Lindholm, K., Linell, F., . . . Sigfusson, B. (1988). Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *Bmj*, 297(6654), 943-948.
- Andersson, I., & Janzon, L. (1997). Reduced breast cancer mortality in women under age 50: updated results from the Malmö Mammographic Screening Program. *JNCI Monographs*, 1997(22), 63-67.
- Aro, A. R., Absetz, S. P., van Elderen, T. M., van der Ploeg, E., & van der Kamp, L. T. (2000). False-positive findings in mammography screening induces short-term distress—breast cancer-specific concern prevails longer. *European Journal of Cancer*, 36(9), 1089-1097.
- Barton, M. B., Moore, S., Polk, S., Shtatland, E., Elmore, J. G., & Fletcher, S. W. (2001). Increased patient concern after false-positive mammograms. *Journal of general internal medicine*, 16(3), 150-156.
- Bjurstam, N., Björneld, L., Warwick, J., Sala, E., Duffy, S. W., Nyström, L., . . . Hafström, L. O. (2003). The Gothenburg breast screening trial. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 97(10), 2387-2396.
- Blumen, H., Fitch, K., & Polkus, V. (2016). Comparison of treatment costs for breast cancer, by tumor stage and type of service. *American health & drug benefits*, 9(1), 23.
- Brewer, N., Salz, T., & Lillie, S. (2007). Systematic Review: The Long-Term Effects of False-Positive Mammograms. *Annals of internal medicine*, 146(7), 502-510.
- Carter, K. J., Castro, F., Kessler, E., & Erickson, B. A. (2005). Simulation of begin and end ages for mammography screening. *Journal for Healthcare Quality*, 27(1), 40-47.
- Chawla, N., Yabroff, K. R., Mariotto, A., McNeel, T. S., Schrag, D., & Warren, J. L. (2014). Limited validity of diagnosis codes in Medicare claims for identifying cancer metastases and inferring stage. *Annals of epidemiology*, 24(9), 666-672. e662.
- Cooper, G. S., Yuan, Z., Stange, K. C., Amini, S. B., Dennis, L. K., & Rimm, A. A. (1999). The utility of Medicare claims data for measuring cancer stage. *Medical care*, 706-711.
- Day, N. (1989). Quantitative approaches to the evaluation of screening programs. *World journal of surgery*, 13(1), 3-8.
- Dittus, K., Geller, B., Weaver, D. L., Kerlikowske, K., Zhu, W., Hubbard, R., . . . Consortium, B. C. S. (2013). Impact of mammography screening interval on breast cancer diagnosis by menopausal status and BMI. *Journal of general internal medicine*, 28(11), 1454-1462.
- Duffy, S. W., Tabár, L., Chen, H. H., Holmqvist, M., Yen, M. F., Abdsalah, S., . . . Hedborg - Melander, C. (2002). The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties: a collaborative

- evaluation. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 95(3), 458-469.
- Elmore, J. G., Barton, M. B., Mocer, V. M., Polk, S., Arena, P. J., & Fletcher, S. W. (1998). Ten-year risk of false positive screening mammograms and clinical breast examinations. *New England Journal of Medicine*, 338(16), 1089-1096.
- Farley, C., Friedman, D., Habtes, I., Raskind-Hood, C., Adams, E. K., Becker, E. R., . . . Gabram-Mendola, S. (2015). Screening mammography in a public hospital serving predominantly African-American women: a stage-survival-cost model. *Women's Health Issues*, 25(4), 322-330.
- Foley, K., Shi, N., Girvan, A., Ward, K., & Lipscomb, J. (2013). CLAIMS DATA ALGORITHMS FOR IDENTIFYING INCIDENT COLORECTAL CANCER (CRC) CASES AND CANCER DISEASE STAGE: A CRITICAL REVIEW OF THE LITERATURE. *VALUE IN HEALTH*, 1(6), A1-A298.
- Freedman, G. M., Anderson, P. R., Goldstein, L. J., Hanlon, A. L., Cianfrocca, M. E., Millenson, M. M., . . . Nicolaou, N. (2003). Routine mammography is associated with earlier stage disease and greater eligibility for breast conservation in breast carcinoma patients age 40 years and older. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 98(5), 918-925.
- Goel, A., Littenberg, B., & Burack, R. C. (2007). The association between the pre-diagnosis mammography screening interval and advanced breast cancer. *Breast cancer research and treatment*, 102(3), 339-345.
- Hendrick, R. E., Smith, R. A., Rutledge III, J. H., & Smart, C. R. (1997). Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. *JNCI Monographs*, 1997(22), 87-92.
- Hubbard, R. A., Kerlikowske, K., Flowers, C. I., Yankaskas, B. C., Zhu, W., & Miglioretti, D. L. (2011). Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Annals of internal medicine*, 155(8), 481-492.
- Kerlikowske, K., Zhu, W., Hubbard, R. A., Geller, B., Dittus, K., Braithwaite, D., . . . Consortium, B. C. S. (2013). Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. *JAMA internal medicine*, 173(9), 807-816.
- Mandelblatt, J. S., Cronin, K. A., Bailey, S., Berry, D. A., De Koning, H. J., Draisma, G., . . . Plevritis, S. K. (2009). Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Annals of internal medicine*, 151(10), 738-747.
- Marmot, M. G., Altman, D., Cameron, D., Dewar, J., Thompson, S., & Wilcox, M. (2013). The benefits and harms of breast cancer screening: an independent review. *British journal of cancer*, 108(11), 2205.
- McGovern, P. M., Gross, C. R., Krueger, R. A., Engelhard, D. A., Cordes, J. E., & Church, T. R. (2004). False-positive cancer screens and health-related quality of life. *Cancer Nursing*, 27(5), 347-352.

- Miglioretti, D. L., Zhu, W., Kerlikowske, K., Sprague, B. L., Onega, T., Buist, D. S., . . . Smith, R. A. (2015). Breast tumor prognostic characteristics and biennial vs annual mammography, age, and menopausal status. *JAMA oncology*, 1(8), 1069-1077.
- Moss, S. (2004). Should women under 50 be screened for breast cancer? *British journal of cancer*, 91(3), 413.
- Nordstrom, B. L., Whyte, J. L., Stolar, M., Mercaldi, C., & Kallich, J. D. (2012). Identification of metastatic cancer in claims data. *Pharmacoepidemiology and drug safety*, 21, 21-28.
- Nyström, L., Andersson, I., Bjurstam, N., Frisell, J., Nordenskjöld, B., & Rutqvist, L. E. (2002). Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *The Lancet*, 359(9310), 909-919.
- Nyström, L., Wall, S., Rutqvist, L., Lindgren, A., Lindqvist, M., Rydén, S., . . . Frisell, J. (1993). Breast cancer screening with mammography: overview of Swedish randomised trials. *The Lancet*, 341(8851), 973-978.
- O'Meara, E. S., Zhu, W., Hubbard, R. A., Braithwaite, D., Kerlikowske, K., Dittus, K. L., . . . Miglioretti, D. L. (2013). Mammographic screening interval in relation to tumor characteristics and false - positive risk by race/ethnicity and age. *Cancer*, 119(22), 3959-3967.
- Oeffinger, K. C., Fontham, E. T., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y.-C. T., . . . LaMonte, S. J. (2015). Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *Jama*, 314(15), 1599-1614.
- Partridge, A. H., LaFountain, A., Mayer, E., Taylor, B. S., Winer, E., & Asnis-Alibozek, A. (2008). Adherence to initial adjuvant anastrozole therapy among women with early-stage breast cancer. *Journal of clinical oncology*, 26(4), 556-562.
- Sant, M., Allemani, C., Capocaccia, R., Hakulinen, T., Aareleid, T., Coebergh, J. W., . . . Bell, J. (2003). Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *International journal of cancer*, 106(3), 416-422.
- Siu, A. L. (2016). Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*, 164(4), 279-296.
- Smith, G. L., Shih, Y.-C. T., Giordano, S. H., Smith, B. D., & Buchholz, T. A. (2010). A method to predict breast cancer stage using Medicare claims. *Epidemiologic Perspectives & Innovations*, 7(1), 1.
- Street, W. (2019). Cancer Facts & Figures 2019. *American Cancer Society: Atlanta, GA, USA*.
- Tabar, L., Faberberg, G., Day, N., & Holmberg, L. (1987). What is the optimum interval between mammographic screening examinations—An analysis based on the latest results of the Swedish two-county breast cancer screening trial. *British journal of cancer*, 55(5), 547.
- Tosteson, A. N., Fryback, D. G., Hammond, C. S., Hanna, L. G., Grove, M. R., Brown, M., . . . Pisano, E. D. (2014). Consequences of false-positive screening mammograms. *JAMA internal medicine*, 174(6), 954-961.
- White, E., Miglioretti, D. L., Yankaskas, B. C., Geller, B. M., Rosenberg, R. D., Kerlikowske, K., . . . Buist, D. S. (2004). Biennial versus annual mammography and

- the risk of late-stage breast cancer. *Journal of the National Cancer Institute*, 96(24), 1832-1839.
- Whyte, J. L., Engel-Nitz, N. M., Teitelbaum, A., Gomez Rey, G., & Kallich, J. D. (2015). An evaluation of algorithms for identifying metastatic breast, lung, or colorectal cancer in administrative claims data. *Medical care*, 53(7), e49-e57.
- Woodward, W. A., Strom, E. A., Tucker, S. L., McNeese, M. D., Perkins, G. H., Schechter, N. R., . . . Hunt, K. K. (2003). Changes in the 2003 American Joint Committee on Cancer staging for breast cancer dramatically affect stage-specific survival. *Journal of clinical oncology*, 21(17), 3244-3248.