

# False-Positive Mammography and Its Association With Health Service Use

Christine M. Gunn, PhD; Barbara Bokhour, PhD; Tracy A. Battaglia, MD, MPH; Rebecca A. Silliman, MD, PhD; and Amresh Hanchate, PhD

**D**espite debate about the benefits and harms of screening mammography,<sup>1-3</sup> it remains the best available tool to detect breast cancer early and prevent morbidity and mortality. As such, the majority of women participate in mammography screening. In 2013, the CDC estimated that 69.1% of women older than 50 years underwent mammography<sup>4</sup> as recommended by the United States Preventive Services Task Force.<sup>5</sup> Identified harms of mammography include a high prevalence of false positives, need for additional imaging and biopsies, potential overdiagnosis of breast cancer, radiation exposure, and masking bias in dense breast tissue.<sup>1</sup> These issues have received extensive attention in the literature and are important concepts for women to understand when considering screening.

False-positive mammograms and the resulting psychosocial harms have garnered attention in recent years.<sup>6,7</sup> Over a 10-year screening period, the cumulative probability of a false-positive mammogram is 42% to 61%, depending on screening interval.<sup>8,9</sup> Anxiety and distress are common among those experiencing a false-positive mammogram, with effects persisting in the months following the test.<sup>6,10</sup> The prevalence of false positives and documented negative psychosocial outcomes are, in part, the basis for not routinely screening women younger than 50 years, who are more frequently affected by these outcomes.<sup>5,11</sup> Guidelines instead promote shared decision making for these women, although this ideal is not often realized in practice.<sup>12-14</sup> Research has shown that women are not often aware of the harms of mammography, including the potential for false-positive findings and future testing.<sup>15,16</sup>

Having a false-positive mammogram is further associated with increases in individuals' risk perceptions, despite no actual change in risk or health threat,<sup>10,17</sup> that can persist long after the abnormal test and affect health service use.<sup>18,19</sup> Although there is interest in determining the multidimensional benefits and harms associated with screening and risk communication, the effects of designating individuals as "at risk" on healthcare utilization measures beyond the downstream diagnostic workup are understudied. This is a critical gap in knowledge at a time when improving cancer screening

## ABSTRACT

**OBJECTIVES:** A false-positive mammogram can result in anxiety, distress, and increased perceptions of breast cancer risk, potentially changing how women utilize healthcare. This study examined whether having an abnormal mammogram, considered a proxy for elevated risk perception, was associated with greater future health service use (outpatient visits and referrals).

**STUDY DESIGN:** A retrospective cohort study using electronic health record data, spanning 2008 to 2012, from Boston Medical Center, a safety-net hospital.

**METHODS:** We grouped 3920 women aged 40 to 75 years receiving primary care and who had a mammogram between 2010 and 2011 into 3 categories: false-positive mammogram at index date; previous false positive, but normal index mammogram; and no history of false-positive mammograms. We contrasted the longitudinal changes in outpatient visits and provider referrals, before versus after the index mammogram, between women with false-positive mammogram and those without using Poisson regression models with a difference-in-differences specification. Clinical, visit, and demographic data were obtained from the institutional clinical data warehouse.

**RESULTS:** Adjusting for baseline differences in sociodemographic characteristics across risk groups and for secular changes between pre- and postindex periods, a current false-positive mammogram was associated with an 18% increase in overall outpatient visits (incidence rate ratio [IRR], 1.18; 95% CI, 1.07-1.51), but no corresponding increase in provider referrals (IRR, 1.15; 95% CI, 0.99-1.34), relative to never having a false positive. A previous false-positive mammogram had no associated change in outpatient utilization (IRR, 0.99; 95% CI, 0.91-1.07).

**CONCLUSIONS:** Providers should discuss the implications of mammography findings at the time of screening to help mitigate potential detrimental effects and promote appropriate engagement in health services.

*Am J Manag Care.* 2018;24(3):131-138

## TAKEAWAY POINTS

We used a difference-in-differences approach to measure changes in healthcare utilization associated with false-positive mammography. A current false positive was associated with an 18% increase in outpatient visits, but no increase in provider referrals, relative to women with no false positive.

- ▶ Trends encouraging mammography adherence increase opportunities for false positives, which we have shown to impact short-term health service use.
- ▶ Counseling about potential false positives and their meaning prior to screening may mitigate the anxiety experienced at the time of an abnormal mammogram.
- ▶ Interventions to ensure appropriate utilization are needed to address women's concerns while promoting evidence-based care.

rates is emphasized through payment schemes intended to improve quality of care.<sup>20</sup> Yet the process of undergoing a mammogram and receiving results may change the way a woman conceptualizes her health and engages with the medical system. Understanding the impact of these tests on how women utilize healthcare services over time is needed to improve the quality of care.

This study examined whether having an abnormal mammogram, considered a proxy for elevated risk perception,<sup>10,18</sup> was associated with greater future health service use. Previous studies have examined breast-specific health service use in relation to other clinical markers of risk status, such as genetic mutations or family history.<sup>21-23</sup> This study expanded the measure of healthcare utilization beyond breast-specific services among a population of women undergoing mammography. We hypothesized that women who experienced a recent or previous false positive would have higher postmammogram visit and referral rates relative to women without false-positive mammograms. Although we were unable to fully characterize the purpose and content of these outpatient visits and referrals, this study sought to deliver an initial assessment of healthcare utilization changes following a false-positive test result, establishing the basis for a detailed exploration of such visits and assessment of whether additional service use represents value-based care.

## METHODS

This retrospective cohort study examined whether experiencing a false-positive mammogram recently or in the past was associated with healthcare utilization by comparing rates of outpatient visits and referrals in the 2 years prior to an index mammogram with 1 year following. This study was determined to be exempt by the Boston University Medical Campus Institutional Review Board.

## Participants

Utilization data came from Boston Medical Center (BMC), an urban safety-net hospital. Included women: 1) were aged 40 to 75 years, 2) received primary care at BMC, 3) had a screening mammogram at BMC in 2010 or 2011 that did not result in a cancer diagnosis, and 4) had their first mammogram at 40 years or older. Women

receiving primary care at BMC were identified by a documented visit with a designated provider in the preceding 2 years. Prior work has demonstrated that these patients use the hospital system as their main source of care for all conditions, including emergency and specialist care.<sup>24</sup> The first mammogram during 2010 or 2011 was defined as the index mammogram. Women with a prior cancer or cancer diagnosed during the study period or those who died during the study were excluded.

## Data Collection

Administrative and clinical data were obtained through the clinical data warehouse, a comprehensive database that aggregates data from multiple electronic hospital sources. The use of retrospective clinical data provided a measure of actual utilization rates, eliminating bias from self-reported data. Data collected spanned calendar years 2008 to 2012.

## Study Design

We used a difference-in-differences (DID) approach, which measures the difference in utilization before and after the index mammogram for women who experienced a false-positive mammogram and contrasts this difference with the corresponding difference for women who experienced no false-positive mammogram.<sup>25</sup> This approach adjusts for pre-index differences in utilization between the 2 groups and secular changes in utilization between pre-index and postindex periods. We defined the 2-year period before and 1-year period after the index mammogram visit as the pre-index and postindex periods, respectively. This provided a stable measure of utilization prior to the index mammogram and allowed for assessment of bias that may have threatened internal validity.<sup>26</sup>

## Comparison Groups

An abnormal mammogram was defined using criteria of the *Breast Imaging Reporting and Data System (BI-RADS) Atlas*, 4th edition.<sup>27</sup> A normal finding included BI-RADS 1 (normal) or 2 (benign) classifications. BI-RADS scores of 0 (incomplete), 3 (probably benign), 4 (suspicious), or 5 (highly suggestive of malignancy) were included in the abnormal mammogram group, as they required a follow-up or further testing. False positives were defined as those abnormal findings that did not result in a cancer diagnosis as assessed by the Tumor Registry.

BI-RADS results were used to create 3 groups of women that reflected levels of risk designation<sup>7,10,17,18</sup>: 1) current false positive (high risk): women experiencing a false-positive mammogram at the index date and indicative of high risk perception; 2) previous false positive (intermediate risk): women who had a normal mammogram result on the index date with a false-positive result

in the past; and 3) no false positives (low risk): women who had a normal index mammogram result and no previous false-positive results. We distinguished current from previous false-positive mammograms to test if changes in utilization following a current false positive would persist over time.

### Measures

We examined 2 utilization indicators as outcomes: referrals and outpatient visits. The number of referrals measured provider-initiated orders for health services and was totaled for each time period. Referrals were captured as “orders,” and all types (ie, specialty visits, laboratory) were compiled to measure the total count documented in the medical record over each time frame. Counts of all outpatient visits attended for each 12-month period were also measured and included all primary care, specialty, emergency department, laboratory, and procedure visits. Diagnostic tests included in the follow-up for an abnormal mammogram were excluded.

Demographic covariates included age, race/ethnicity, insurance, education, and primary language. Baseline comorbidities were measured by identifying all unique diagnosis codes (*International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM]*) in the pre-index period and grouping them into 30 categories using the Elixhauser classification.<sup>28,29</sup> A dichotomous indicator variable for time (pre-index vs post index) was created and later interacted with the 3 groups to assess differences over time by group.<sup>25</sup>

### Comparison of Pre-Index Longitudinal Trends Between Comparison Cohorts

A critical assumption underlying the appropriateness of the DID design was that the longitudinal trends in utilization for the 2 comparison cohorts, women with a current or previous false-positive mammogram versus those without, diverged from a parallel trajectory only after the false-positive mammogram event; that is, prior to this event, the longitudinal trends between the 2 groups should have followed similar trends (ie, were parallel).<sup>30</sup> We tested this “parallel trends” test by using only pre-index event data, dividing the 2 pre-index years of utilization into 2 single-year utilization periods; estimated a similarly specified DID model limited to only the pre-index data; and tested if the changes in utilization between the 2 years were similar between the 2 comparison groups (current vs never; previous vs never). Confirmation of parallel trends enhanced the ability of the main DID model (with complete data) to attribute postindex utilization changes to false-positive mammograms.

### Propensity Score Weighting

Another potential source of confounding was systematic differences in the prevalence of the observed demographic and socioeconomic covariates and pre-index utilization between the groups of women compared. One approach to adjust for potential imbalance was the

use of propensity score weighting known as standardized mortality ratio weighting (SMRW).<sup>31</sup> In this 2-step method, we first estimated a logistic regression of the indicator of false-positive mammogram (0/1) on age, race, language, education, insurance, Elixhauser categories, and pre-index visits (or referrals) to obtain the predicted probability of false-positive mammogram (propensity score); in the second step, we estimated the main regression model using SMRW, wherein women with false-positive mammograms were assigned a weight of 1 and women without false-positive mammograms were assigned a weight equal to the ratio of the propensity score to 1 minus the propensity score.<sup>32</sup> The propensity score weighting was used to produce comparable groups with parallel pre-visit trends.

### Statistical Analysis

We performed descriptive analyses and examined patient characteristics by group. Our core analysis estimated the change in each utilization outcome associated with a current (at index mammogram) or previous false-positive mammogram (prior to index, with normal index mammogram) using a weighted Poisson regression model with a DID specification.<sup>25,33</sup> Multivariable models adjusted for covariates, including comorbidities present at the index mammogram. Two models for each outcome were constructed: one to compare the current versus never false-positive groups and the other to compare the previous versus never false positive-groups. Thus, 4 final DID models were run to account for selection of appropriate comparison groups via propensity score weighting.

The primary independent variable of interest in the DID specification was the interaction between time and having a false-positive mammogram result, which compared the average difference in utilization in the pre- versus postindex mammogram periods for the current and previous false-positive mammogram groups compared with those with no false-positive mammograms. We hypothesized that a current false-positive would result in higher rates of referrals and outpatient visits in the follow-up period, as women might feel that they are at higher risk.

To perform the parallel trends test described above, we estimated an identically specified weighted Poisson DID model, with the data limited to the pre-index period and with utilization in the 2 pre-index years as the 2 periods compared. All analyses were performed in SAS version 9.3 (SAS Institute Inc; Cary, North Carolina). Estimates are reported as incidence rate ratios (IRRs) with 95% CIs. Statistical significance was set at  $\alpha = .05$  for all models.

### Sensitivity Analysis

False-positive rates may be confounded by age due to differences in the properties of mammography and breast tissue structure.<sup>34</sup> We conducted a sensitivity analysis to assess the influence of age, stratifying the sample into 2 groups (40-49 and 50-75 years). As Poisson models may not capture overdispersion in outcomes, we performed estimation using negative binomial distribution models.<sup>35</sup>

**TABLE 1.** Demographic Characteristics of Sample by False-Positive Group (N = 3920)

Variable	n (%)	Current False Positive (n = 500)	Previous False Positive (n = 594)	No False Positive (n = 2826)	P
Age group, years					.0002 <sup>a</sup>
<50	1678 (43)	256 (51)	227 (38)	1195 (42)	
50-64	1668 (42)	188 (38)	280 (47)	1200 (42)	
65-75	574 (15)	56 (11)	87 (15)	431 (15)	
Race/ethnicity					.059
Non-Hispanic white	595 (15)	90 (18)	87 (15)	418 (15)	
Hispanic	533 (14)	59 (12)	99 (17)	375 (13)	
Black	2035 (52)	257 (51)	311 (52)	1467 (52)	
Other/declined	757 (19)	94 (19)	97 (16)	566 (20)	
Language					.18
English	2410 (61)	301 (60)	374 (63)	1735 (61)	
Spanish	381 (10)	45 (9)	72 (12)	264 (9)	
Haitian Creole	556 (14)	79 (16)	73 (12)	404 (14)	
Other	573 (15)	75 (15)	75 (13)	423 (15)	
Primary insurance					.37
Commerical/private	1008 (26)	124 (25)	156 (26)	728 (26)	
Public	2137 (54)	270 (54)	338 (57)	1529 (54)	
Other/uninsured	775 (20)	106 (21)	100 (17)	569 (20)	
Education					.02 <sup>a</sup>
Less than high school	1610 (41)	213 (43)	272 (46)	1125 (40)	
High school/GED diploma	1110 (28)	132 (26)	151 (25)	827 (29)	
Some college/associate degree	381 (10)	37 (7)	65 (11)	279 (10)	
College/postgrad	539 (14)	79 (16)	76 (13)	384 (14)	
Other/missing	280 (7)	39 (8)	30 (5)	211 (7)	
Comorbidities					.03 <sup>a</sup>
≥1 Elixhauser comorbidity	2971 (75.8)	357 (71.4)	477 (80.3)	2137 (75.6)	

GED indicates General Educational Development.  
<sup>a</sup>Indicates P < .05.

## RESULTS

The sample included 3920 women who met inclusion criteria (Table 1). Almost half (42%) were women for whom routine screening was universally recommended (aged 50-64 years), 52% were black, and 61% spoke English as their primary language. Most had public insurance (55%), and 41% had less than a high school education. Seventy-six percent of women had at least 1 Elixhauser-defined comorbid condition. The previous false-positive group had the highest prevalence of 1 or more comorbidities at 80.3% versus 75.6% in the no false-positive group and 71.4% in the current false-positive group (details in Appendix Table [eAppendices available at [ajmc.com](http://ajmc.com)]).

## Associations Between Demographics and Risk Groups

Sample characteristics identified the demographic correlates of false-positive mammograms. The current false-positive group was younger than the other 2 groups (P = .0002). Education (P = .02) was the only other characteristic significantly associated with risk group. Unadjusted visit rates over time by risk group are displayed in Table 2. Testing differences in these trends across the pre-index period using the SMRW propensity score method resulted in nonsignificant associations, suggesting that the parallel trends assumption was satisfied.

## Changes in Utilization Associated With a False-Positive Mammogram

Adjusting for baseline differences in socio-demographic characteristics across groups and for secular changes between the pre- and postindex periods, a current false-positive mammogram was associated with an 18% increase in overall outpatient visits (IRR, 1.18; 95% CI, 1.07-1.51) (Table 3). In contrast, a previous false-positive mammogram had no associated change in utilization (IRR, 0.99; 95% CI, 0.91-1.07). Educational achievement, insurance coverage type, and baseline comorbidities were associated with systematic differences in outpatient visits in the pre- and postindex periods.

Contrary to the hypothesis, women with current false positives saw no increase in referral rates (IRR, 1.15; 95% CI, 0.99-1.34). Those with previous false positives also did not have a significant change in referral rates (IRR, 0.99; 95% CI, 0.83-1.20). Although point estimates for referrals were similar to the visit outcome,

differing baseline outcome rates resulted in varying magnitudes, affecting the significance of the association. Nonwhite race, having public insurance, and baseline comorbidities were associated with higher referral rates.

## Sensitivity Analyses

A stratified analysis by age was conducted to account for the lower sensitivity and specificity of mammography for younger women<sup>34</sup> and found slight variations by age (Table 4). The increase in visits associated with a false positive appeared to be driven by women younger than 50 years, with no significant change in utilization for older women (IRR, 1.13; 95% CI, 0.97-1.31); however, the general trend was similar. Stratifying by age, older women (aged 50-74 years) with

current false positives received significantly fewer referrals compared with those without false positives (IRR, 0.85; 95% CI, 0.79-0.92), whereas there was no significant difference in referral rates across false-positive groups among younger women (aged 40-49 years). Estimates of change for the current and previous false-positive groups were robust when re-estimating the models using a negative binomial model.

## DISCUSSION

Having a false-positive mammogram was associated with significantly increased rates of outpatient visits in the year following the test, suggesting that utilization may be driven by patients seeking care, but is likely not generated by the health system as there was no corresponding increase in referrals for care. Increased utilization dissipated over time: We found comparable visit and referral rates between those with a previous false positive and a recent normal mammogram and those with no false-positive mammograms.

One striking finding from this study was the marked increase in outpatient visits following an abnormal mammogram. Although increased breast-specific utilization is expected in order to rule out a cancer diagnosis,<sup>36</sup> the reasons for increases in other types of care are less obvious. Our findings could indicate that women may conceptualize themselves as less healthy than prior to their false-positive mammogram<sup>37</sup> or that they may experience increased anxiety, leading to additional visits that increase reassurance through monitoring. Prior study findings have confirmed that increases in utilization may be related to the negative psychosocial outcomes of having a false-positive mammogram, showing increases in mental health services following the receipt of an abnormal mammogram.<sup>19</sup>

Undergoing imaging and receiving a normal result has been characterized as a “gift of knowing,” providing reassurance to those who engage in screening.<sup>38</sup> This reassurance may be driving decreases in utilization in the previous false-positive group. This group was possibly reassured of good health with the normal mammogram result, thus removing

**TABLE 2.** Unadjusted Median Outpatient Visit and Referral Rates Over Time by False-Positive Group

	Outpatient Visits Median (IQR)			Referrals Median (IQR)		
	2 Years Pre-Index	1 Year Pre-Index	1 Year Post Index	2 Years Pre-Index	1 Year Pre-Index	1 Year Post Index
Current false positive	4 (0-11)	9 (5-17)	11 (6-21)	0 (0-2)	2 (0-4)	2 (0-4)
Previous false positive	12 (7-20)	12 (7-21)	10 (5-19)	2 (0-4)	2 (0-4)	2 (0-4)
No false positive	7 (0-16)	10 (6-20)	10 (4-21)	0 (0-4)	2 (0-5)	2 (0-4)

IQR indicates interquartile range.

**TABLE 3.** Poisson Regression Model Estimates for DID Analysis for Outpatient Visits and Referrals

	Outpatient Visits		Referrals	
	IRR Current vs Never	IRR Previous vs Never	IRR Current vs Never	IRR Previous vs Never
False-positive category <sup>a</sup>				
Ref: no abnormal (none)	0.99	0.89 <sup>b</sup>	0.99	0.85 <sup>b</sup>
Time (ref: 1 year pre-index)				
Post index	1.01	0.92 <sup>b</sup>	0.82 <sup>b</sup>	0.83 <sup>b</sup>
2 years pre-index	0.61 <sup>b</sup>	0.90 <sup>b</sup>	0.53 <sup>b</sup>	0.93
False-positive category × time				
False positive × post index	1.18 <sup>b</sup>	0.99	1.15	1.07
False positive × 2 years pre-index	0.98	1.08	0.99	1.13
Age group, years (ref: 50-64)				
40-50	1.06	1.05	1.00	0.93
65-75	1.07	0.99	0.98	0.92
Race/ethnicity (ref: non-Hispanic white)				
Black	1.18 <sup>b</sup>	0.98	1.17 <sup>b</sup>	1.04
Hispanic	1.14	1.16	1.43 <sup>b</sup>	1.20
Other	1.19	0.99	1.21 <sup>b</sup>	0.97
Language (ref: English)				
Haitian Creole	0.90	0.86 <sup>b</sup>	0.81 <sup>b</sup>	0.84 <sup>b</sup>
Spanish	1.07	0.88	0.80	0.92
Other	1.00	1.02	1.00	1.08
Education (ref: college/postgrad)				
Less than high school	1.12	1.13	1.04	1.06
High school/GED diploma	1.22 <sup>b</sup>	1.24 <sup>b</sup>	1.10	1.30 <sup>b</sup>
Some college/associate degree	1.15	1.08	0.99	0.99
Other/missing	0.88	1.02	0.82 <sup>b</sup>	1.06
Primary insurance (ref: private)				
Public	1.26 <sup>b</sup>	1.32 <sup>b</sup>	1.22 <sup>b</sup>	1.24 <sup>b</sup>
Other	1.12	1.00	1.12	0.95

(continued)

**TABLE 3.** (Continued) Poisson Regression Model Estimates for DID Analysis for Outpatient Visits and Referrals

	Outpatient Visits		Referrals	
	IRR Current vs Never	IRR Previous vs Never	IRR Current vs Never	IRR Previous vs Never
<b>Comorbidities</b>				
Congestive heart failure	1.36 <sup>b</sup>	1.43 <sup>b</sup>	0.98	1.09
Cardiac arrhythmias	1.32 <sup>b</sup>	1.56 <sup>b</sup>	1.17	0.87
Valvular disease	1.22	0.97	0.98	0.95
Pulmonary circulation disorders	2.36 <sup>b</sup>	1.60 <sup>b</sup>	1.58	1.51 <sup>b</sup>
Peripheral vascular disorders	1.40 <sup>b</sup>	1.29 <sup>b</sup>	1.48 <sup>b</sup>	1.10
Hypertension, uncomplicated	1.11 <sup>b</sup>	1.16 <sup>b</sup>	1.03	1.11 <sup>b</sup>
Hypertension, complicated	1.31	2.04 <sup>b</sup>	0.68	1.17
Paralysis	0.83	0.70 <sup>b</sup>	1.20	0.72
Other neurological disorders	1.43 <sup>b</sup>	1.32 <sup>b</sup>	1.26 <sup>b</sup>	1.40 <sup>b</sup>
Chronic pulmonary disease	1.32 <sup>b</sup>	1.20 <sup>b</sup>	1.19 <sup>b</sup>	1.10
Diabetes, uncomplicated	1.30 <sup>b</sup>	1.64 <sup>b</sup>	1.42 <sup>b</sup>	1.35 <sup>b</sup>
Diabetes, complicated	1.96 <sup>b</sup>	1.45 <sup>b</sup>	1.77 <sup>b</sup>	1.50 <sup>b</sup>
Hypothyroidism	1.30 <sup>b</sup>	1.19 <sup>b</sup>	1.24 <sup>b</sup>	1.03
Renal failure	1.65 <sup>b</sup>	1.62 <sup>b</sup>	1.14	1.07
Liver disease	1.78 <sup>b</sup>	1.28	1.66 <sup>b</sup>	1.35 <sup>b</sup>
Peptic ulcer disease, excluding bleeding	0.94	1.04	0.81	0.80
Rheumatoid arthritis/collagen vascular disorders	1.22	1.23 <sup>b</sup>	1.27 <sup>b</sup>	1.18
Coagulopathy	1.09	1.36	1.19	1.48 <sup>b</sup>
Obesity	1.17 <sup>b</sup>	1.25 <sup>b</sup>	1.21 <sup>b</sup>	1.34 <sup>b</sup>
Weight loss	1.29 <sup>b</sup>	0.99	1.22 <sup>b</sup>	1.09
Fluid/electrolyte disorders	1.26 <sup>b</sup>	1.39 <sup>b</sup>	1.01	1.20
Deficiency anemias	1.34 <sup>b</sup>	1.23 <sup>b</sup>	1.34 <sup>b</sup>	1.40 <sup>b</sup>
Alcohol abuse	1.17	1.03	1.08	0.88
Drug abuse	1.30 <sup>b</sup>	1.33 <sup>b</sup>	1.09	1.14
Psychoses	1.35 <sup>b</sup>	1.32	1.05	0.89
Depression	1.30 <sup>b</sup>	1.31 <sup>b</sup>	1.37 <sup>b</sup>	1.34 <sup>b</sup>
<b>Index date</b>				
2011 index	0.99	0.94	0.97	0.99

DID indicates difference-in-differences; GED, General Educational Development; IRR, incidence rate ratio; ref, reference.

\*Categories: "Current" indicates current false positive at index date; "Previous," previous false positive with normal index mammogram; "Never," no false positives at any time.

<sup>b</sup>IRR estimates are significant at  $\alpha = .05$ .

the impetus to continue higher rates of utilization. The increase in immediate post-false-positive visits suggests that communication at the time of mammogram or in the immediate postmammogram period may be required to mitigate short-term anxiety and increased utilization. This may be particularly true for underserved populations similar to those studied here: We found that women with a high school education and black women had significantly more

these changes prior to the index mammogram. Study results have also shown that the changes in guidelines have produced minimal provider practice changes over time.<sup>42</sup> Further, 1 year after the guideline changes, a large survey of women found that less than half were even aware of these changes,<sup>43</sup> leading us to conclude that these guidelines likely had a minimal impact on the trend for younger women to increase utilization at a higher rate following a false positive.

utilization associated with a false positive. This finding is similar to others that have noted that minority and low-literacy women experience higher distress and lower understanding when presented with abnormal mammogram results relative to white and high-literacy women.<sup>39,40</sup> Interventions targeted toward these populations, such as patient navigation, have been shown to decrease anxiety following a false positive,<sup>41</sup> but navigator effects on utilization beyond the diagnostic follow-up are limited. Prospectively understanding the role of anxiety generated by a false-positive test in subsequent healthcare utilization is required to tailor interventions to promote appropriate, high-quality care to women undergoing mammography screening.

**Limitations**

Using a proxy for risk perception presumed that women adopted a risk identity concordant with our classification and assumed that women with normal mammograms did not have perceptions of increased risk. We only included women who had their first mammogram after the age of 40 years, omitting those with very high risk perceptions due to family history, as these women usually begin screening earlier and might have higher risk perceptions. However, the study design did not fully characterize risk or women's interpretation of the abnormal finding. Future studies are required to describe behavior patterns based on a more sensitive measure of risk perception.

By using both current and previous false-positive groups, this study showed the timing of the mammogram may be essential in establishing influences on patterns of care. Age-related differences discovered in the sensitivity analysis, although minimal, may have been influenced by increased anxiety due to changes in guidelines recommending less mammography for women younger than 50 years that were implemented in 2009. However, all groups were exposed to

Study exclusion criteria also eliminated women without a primary care visit in the 2 years prior to the index mammogram and did not include utilization outside of the safety-net health system in which patients had their mammogram, limiting generalizability to those women who were connected to primary care in an urban safety-net setting. Further work is needed to understand how women without a regular source of care and those who seek care in other settings respond to a false-positive screening test.

Finally, we were unable to fully characterize or control for the purpose of outpatient visits. We chose to exclude outpatient services related to the diagnostic workup (ie, biopsies), but there was the potential for residual confounding if a significant number of outpatient visits were indirectly related to the abnormal mammogram. This would require data abstraction and/or prospective characterization of the visits, which was beyond the scope of this study. It is also unknown whether these additional visits were necessary or represented overuse of care. The utility of these additional visits is an important area of future investigation to understand how mammography may impact the provision of value-based care.

## CONCLUSIONS

This study provides evidence that abnormal screening tests can heighten risk perceptions that may result in higher rates of health service utilization. It expands previous work to include services beyond breast-specific follow-up testing. However, prospective mixed-methods approaches that combine utilization data with assessment of patient psychosocial outcomes are warranted to accurately represent both patient perceptions and the timing of events. Current trends encouraging adherence to mammography increase opportunities to designate women as “at risk,” which we have shown to impact short-term health service use. Counseling women about potential false positives and their meaning prior to screening may mitigate the anxiety experienced at the time of the abnormal mammogram. Further research is needed to better characterize the drivers of increased utilization, the types of visits that increase the most, and the value that additional care provides. This will present opportunities for intervention to increase appropriate utilization and develop tools for providers in addressing women’s concerns while providing evidence-based care. ■

### Acknowledgments

The authors would like to thank Na Wang for her contribution to the analysis of the data that contributed to this manuscript. This research was conducted without financial support.

**Author Affiliations:** Women’s Health Unit, Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center and Women’s Health Interdisciplinary Research Center (CMG, TAB), Boston, MA; Department of Health Law, Policy, and Management (CMG), and Department of Epidemiology (RAS), Boston University School of Public Health, Boston, MA; Department of Medicine, Section of Geriatrics (RAS), and Section of General Internal Medicine, Boston University School of Medicine (AH), Boston, MA.

**TABLE 4.** DID Estimates for Sensitivity Analyses by Age and Distribution for Outpatient Visits and Referrals

	DID IRR (95% CI) for Current False Positive	DID IRR (95% CI) for Previous False Positive
<b>Age Analysis</b>		
Outpatient visits		
Total	1.18 (1.07-1.31) <sup>a</sup>	0.99 (0.91-1.07)
40-49 years	1.24 (1.09-1.41) <sup>a</sup>	0.96 (0.81-1.14)
50-75 years	1.13 (0.97-1.31)	1.00 (0.90-1.12)
Referrals		
Total	1.15 (0.99-1.34)	1.07 (0.94-1.23)
40-49 years	1.23 (0.98-1.54)	1.41 (0.96-2.08)
50-75 years	0.85 (0.79-0.92) <sup>a</sup>	1.01 (0.86-1.18)
<b>Distribution Analysis</b>		
Outpatient visits		
Poisson model	1.18 (1.07-1.31) <sup>a</sup>	0.99 (0.91-1.07)
Negative binomial model	1.17 (1.04-1.31) <sup>a</sup>	1.02 (0.92-1.12)
Referrals		
Poisson model	1.15 (0.99-1.34)	1.07 (0.94-1.23)
Negative binomial model	1.12 (0.96-1.30)	1.05 (0.90-1.23)

DID indicates difference-in-differences; IRR, incidence rate ratio.

<sup>a</sup>Significant at  $P < .05$ .

**Source of Funding:** None.

**Author Disclosures:** The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

**Authorship Information:** Concept and design (CMG, BB, TAB, RAS, AH); acquisition of data (CMG, TAB); analysis and interpretation of data (CMG, BB, TAB, RAS, AH); drafting of the manuscript (CMG, TAB); critical revision of the manuscript for important intellectual content (TAB, RAS, AH); statistical analysis (CMG, AH); provision of patients or study materials (TAB); and supervision (BB, TAB, AH).

**Address Correspondence to:** Christine M. Gunn, PhD, Women’s Health Unit, Boston Medical Center, 801 Massachusetts Ave, 1st Fl, Boston, MA 02118. Email: cgunn@bu.edu.

## REFERENCES

- Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L; US Preventive Services Task Force. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2009;151(10):727-737. WZ37-WZ42. doi: 10.7326/0003-4819-151-10-200911170-00009.
- Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer.* 2013;108:2205-2240. doi: 10.1038/bjc.2013.177.
- Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. *BMJ.* 2014;348:g366. doi: 10.1136/bmj.g366.
- National Center for Health Statistics. Health, United States, 2015: with special feature on racial and ethnic health disparities. CDC website. cdc.gov/nchs/data/atus/hus15.pdf#070. Published May 2016. Updated June 22, 2017. Accessed January 17, 2017.
- Siu AL; US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force Recommendation Statement [erratum in *Ann Intern Med.* 2016;164(6):448. doi: 10.7326/L16-0404]. *Ann Intern Med.* 2016;164(4):279-296. doi: 10.7326/M15-2886.
- Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med.* 2007;146(7):502-510. doi: 10.7326/0003-4819-146-7-200704030-00006.
- Watson EK, Henderson BJ, Brett J, Bankhead C, Austoker J. The psychological impact of mammographic screening on women with a family history of breast cancer—a systematic review. *Psychooncology.* 2005;14(11):939-948. doi: 10.1002/pon.903.

8. Hubbard RA, Kertlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography [erratum in *Ann Intern Med*. 2014;160(9):658]. *Ann Intern Med*. 2011;155(8):481-W-147.
9. Christiansen CL, Wang F, Barton MB, et al. Predicting the cumulative risk of false-positive mammograms. *J Natl Cancer Inst*. 2000;92(20):1657-1666.
10. Aro AR, Pitvikki Absetz S, van Elderen TM, van der Ploeg E, van der Kamp LJT. False-positive findings in mammography screening induces short-term distress—breast cancer-specific concern prevails longer. *Eur J Cancer*. 2000;36(9):1089-1097. doi: 10.1016/S0959-8049(00)00065-4.
11. Oeffinger KC, Fontham ET, Etzioni R, et al; American Cancer Society. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society [erratum in *JAMA*. 2016;315(13):1406]. doi: 10.1001/jama.2016.3404. *JAMA*. 2015;314(15):1599-1614. doi: 10.1001/jama.2015.12783.
12. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. *Patient Educ Couns*. 2006;60(3):301-312. doi: 10.1016/j.pec.2005.06.010.
13. Hoffman RM, Lewis CL, Pignone MP, et al. Decision-making processes for breast, colorectal, and prostate cancer screening: the DECISIONS survey. *Med Decis Making*. 2010;30(suppl 5):S53-S64. doi: 10.1177/0272989X10378701.
14. Gunn CM, Soley-Bori M, Battaglia TA, Cabral H, Kazis L. Shared decision making and the use of screening mammography in women under 50. *J Health Commun*. 2015;20(9):1060-1066. doi: 10.1080/10810730.2015.1018628.
15. Wegwarth O, Gigerenzer G. "There is nothing to worry about": gynecologists' counseling on mammography. *Patient Educ Couns*. 2011;84(2):251-256. doi: 10.1016/j.pec.2010.07.025.
16. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. *BMJ*. 2000;320(7250):1635-1640. doi: 10.1136/bmj.320.7250.1635.
17. Rosenbaum L. Invisible risks, emotional choice—mammography and medical decision making. *N Engl J Med*. 2014;371(16):1549-1552. doi: 10.1056/NEJMs1409003.
18. Lindberg LG, Svendsen M, Dømggaard M, Brodersen J. Better safe than sorry: a long-term perspective on experiences with a false-positive screening mammography in Denmark. *Health Risk Soc*. 2013;15(8):699-716. doi: 10.1080/13698575.2013.848845.
19. Barton MB, Moore S, Polk S, Shtatland E, Elmore JG, Fletcher SW. Increased patient concern after false-positive mammograms: clinician documentation and subsequent ambulatory visits. *J Gen Intern Med*. 2001;16(3):150-156. doi: 10.1111/j.1525-1497.2001.00329.x.
20. Centers for Medicare and Medicaid Innovation. Accountable care organization 2015 program analysis quality performance standards: narrative measure specifications. CMS website. cms.gov/medicare/medicare-fee-for-service-payment/sharesavingsprogram/downloads/ry2015-narrative-specifications.pdf. Published January 9, 2015. Accessed September 14, 2016.
21. Haber G, Ahmed NU, Pekovic V. Family history of cancer and its association with breast cancer risk perception and repeat mammography. *Am J Public Health*. 2012;102(12):2322-2329. doi: 10.2105/AJPH.2012.300786.
22. Wainberg S, Husted J. Utilization of screening and preventive surgery among unaffected carriers of a *BRCA1* or *BRCA2* gene mutation. *Cancer Epidemiol Biomarkers Prev*. 2004;13(12):1989-1995.
23. Larouche G, Bouchard K, Chiquette J, Desbiens C, Simard J, Dorval M. Self-reported mammography use following *BRCA1/2* genetic testing may be overestimated. *Fam Cancer*. 2012;11(1):27-32. doi: 10.1007/s10689-011-9490-6.
24. Kapoor A, Battaglia TA, Isabelle AP, et al. The impact of insurance coverage during insurance reform on diagnostic resolution of cancer screening abnormalities. *J Health Care Poor Underserved*. 2014;25(suppl 1):109-121. doi: 10.1353/hpu.2014.0063.
25. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA*. 2014;312(22):2401-2402. doi: 10.1001/jama.2014.16153.
26. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company; 2002.
27. American College of Radiology. *Breast Imaging and Reporting Data System Atlas*. 4th ed. Reston, VA: American College of Radiology; 2003.
28. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27.
29. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-1139.
30. Ryan AM, Burgess JF Jr, Dimick JB. Why we should not be indifferent to specification choices for difference-in-differences. *Health Serv Res*. 2015;50(4):1211-1235. doi: 10.1111/1475-6773.12270.
31. Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research. *Circ Cardiovasc Qual Outcomes*. 2013;6(5):604-611. doi: 10.1161/CIRCOUTCOMES.113.000359.
32. Keating NL, Kouri EM, He Y, West DW, Winer EP. Effect of Massachusetts health insurance reform on mammography use and breast cancer stage at diagnosis. *Cancer*. 2013;119(2):250-258. doi: 10.1002/cncr.27757.
33. Bertrand M, Dufo E, Mullainathan S. How much should we trust differences-in-differences estimates? *Q J Econ*. 2004;119(1):249-275. doi: 10.1162/003355304772839588.
34. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography [erratum in *Ann Intern Med*. 2003;138(9):771]. *Ann Intern Med*. 2003;138(3):168-175. doi: 10.7326/0003-4819-138-3-200302040-00008.
35. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.
36. Welch HG, Fisher ES. Diagnostic testing following screening mammography in the elderly. *J Natl Cancer Inst*. 1998;90(18):1389-1392.
37. Gillespie C. The experience of risk as "measured vulnerability": health screening and lay uses of numerical risk. *Social Health Illn*. 2012;34(2):194-207. doi: 10.1111/j.1467-9566.2011.01381.x.
38. Kenen RH. The at-risk health status and technology: A diagnostic invitation and the "gift" of knowing. *Soc Sci Med*. 1996;42(11):1545-1553. doi: 10.1016/0277-9536(95)00248-0.
39. Molina Y, Hohl SD, Ko LK, Rodriguez EA, Thompson B, Beresford SA. Understanding the patient-provider communication needs and experiences of Latina and non-Latina white women following an abnormal mammogram. *J Cancer Educ*. 2014;29:781-789. doi: 10.1007/s13187-014-0654-6.
40. Karliner LS, Kaplan CP, Juarbe T, Pasick R, Pérez-Stable EJ. Poor patient comprehension of abnormal mammography results. *J Gen Intern Med*. 2005;20(5):432-437. doi: 10.1111/j.1525-1497.2005.40281.x.
41. Ferrante JM, Chen PH, Kim S. The Effect of patient navigation on time to diagnosis, anxiety, and satisfaction in urban minority women with abnormal mammograms: a randomized controlled trial. *J Urban Health*. 2008;85(1):114-124. doi: 10.1007/s11524-007-9228-9.
42. Haas JS, Sprague BL, Klabunde CN, et al; PROSPR (Population-based Research Optimizing Screening through Personalized Regimens) Consortium. Provider attitudes and screening practices following changes in breast and cervical cancer screening guidelines. *J Gen Intern Med*. 2016;31(1):52-59. doi: 10.1007/s11606-015-3449-5.
43. Kiviniemi MT, Hay JL. Awareness of the 2009 US Preventive Services Task Force recommended changes in mammography screening guidelines, accuracy of awareness, sources of knowledge about recommendations, and attitudes about updated screening guidelines in women ages 40-49 and 50+. *BMC Public Health*. 2012;12:899. doi: 10.1186/1471-2458-12-899.

**eAppendix Table.** Distribution of Comorbidities by False-Positive Group

	<b>Total (n = 3920)</b>	<b>Current False Positive (n = 500)</b>	<b>Previous False Positive (n = 594)</b>	<b>No False Positive (n = 2826)</b>	<b><i>P</i></b>
<b>Comorbidity Group</b>	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>	
Congestive heart failure	69 (1.8)	4 (0.80)	8 (1.35)	57 (2.0)	.11
Cardiac arrhythmias	96 (2.4)	11 (2.2)	14 (2.4)	71 (2.5)	.91
Valvular disease	50 (1.3)	3 (0.60)	4 (0.67)	43 (1.5)	.09
Pulmonary circulation disorders	40 (1.0)	3 (0.60)	8 (1.35)	29 (1.0)	.48
Peripheral vascular disorders	71 (1.8)	7 (1.40)	17 (2.9)	47 (1.7)	.10
Hypertension, uncomplicated	1801 (45.9)	182 (36.4)	314 (52.9)	1305 (46.2)	<.0001
Hypertension, complicated	17 (0.4)	2 (0.40)	1 (0.17)	14 (0.5)	.67
Paralysis	14 (0.4)	3 (0.60)	1 (0.17)	10 (0.4)	.51
Other neurological disorders	85 (2.2)	10 (2.0)	16 (2.7)	59 (2.1)	.63
Chronic pulmonary disease	562 (14.3)	73 (14.6)	109 (18.4)	380 (13.4)	.008
Diabetes, uncomplicated	152 (3.9)	19 (3.8)	31 (5.2)	102 (3.6)	.18
Diabetes, complicated	169 (4.3)	21 (4.2)	20 (3.4)	128 (4.5)	.44
Hypothyroidism	336 (8.6)	48 (9.6)	69 (11.6)	219 (7.7)	.006
Renal failure	68 (1.7)	6 (1.20)	12 (2.0)	50 (1.8)	.56
Liver disease	54 (1.4)	5 (1.00)	10 (1.7)	39 (1.4)	.63
Peptic ulcer disease, excluding bleeding	30 (0.8)	2 (0.40)	9 (1.52)	19 (0.7)	.08
Rheumatoid arthritis/collagen vascular disorders	88 (2.2)	9 (1.80)	17 (2.9)	62 (2.2)	.47
Coagulopathy	42 (1.1)	5 (1.00)	6 (1.01)	31 (1.1)	.97
Obesity	228 (5.8)	31 (6.2)	20 (3.4)	177 (6.3)	.022
Weight loss	131 (3.3)	15 (3.0)	21 (3.5)	95 (3.4)	.88
Fluid/electrolyte disorders	65 (1.7)	10 (2.0)	13 (2.2)	42 (1.5)	.39
Deficiency anemias	270 (6.9)	31 (6.2)	36 (6.1)	203 (7.2)	.50
Alcohol abuse	91 (2.3)	9 (1.80)	11 (1.9)	71 (2.5)	.44
Drug abuse	194 (4.9)	30 (6.0)	31 (5.2)	133 (4.7)	.44
Psychoses	58 (1.5)	6 (1.20)	10 (1.7)	42 (1.5)	.80
Depression	1140 (29.1)	133 (26.6)	195 (32.8)	812 (28.7)	.06
≥1 Elixhauser comorbidity	2971 (75.8)	357 (71.4)	477 (80.3)	2137 (75.6)	.003