Title

Racial and Ethnic Variation in Diagnostic Mammography Performance among Women Reporting a Breast Lump.

Journal

Cancer Epidemiology Biomarkers & Prevention, 32(11)

ISSN

1055-9965

Authors

Nyante, Sarah J

Abraham, Linn

Aiello Bowles, Erin J

et al.

Publication Date

2023-11-01

DOI

10.1158/1055-9965.epi-23-0289

Peer reviewed

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2023 November 01; 32(11): 1542–1551.

doi:10.1158/1055-9965.EPI-23-0289.

Racial and ethnic variation in diagnostic mammography performance among women reporting a breast lump

Sarah J. Nyante[1,2], **Linn Abraham**[3], **Erin J. Aiello Bowles**[3], **Christoph I. Lee**[4], **Karla Kerlikowske**[5], **Diana L. Miglioretti**[6,3], **Brian L. Sprague**[7], **Louise M. Henderson**[1,2], **Breast Cancer Surveillance Consortium**

[1]Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC

[2]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

[3]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle, WA

[4]Department of Radiology, University of Washington School of Medicine; Department of Health Services, University of Washington School of Public Health; Fred Hutchinson Cancer Center, Seattle, WA

[5]Departments of Medicine and Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA

[6]Department of Public Health Sciences, University of California, Davis, Davis, CA

[7]Department of Surgery and University of Vermont Cancer Center, University of Vermont, Burlington, VT

Abstract

Background: We evaluated diagnostic mammography among women with a breast lump to determine whether performance varied across racial and ethnic groups.

Methods: This study included 51,014 diagnostic mammograms performed between 2005–2018 in the Breast Cancer Surveillance Consortium among Asian/Pacific Islander (12%), Black (7%), Hispanic/Latina (6%), and White (75%) women reporting a lump. Breast cancers occurring within one year were ascertained from cancer registry linkages. Multivariable regression was used to

Corresponding author: Sarah J. Nyante, CB#7515, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7515; sarah_nyante@med.unc.edu; phone: (919) 445-3009.

Conflict of Interest Disclosures:
Nyante: No conflicts of interest to disclose.
Abraham: No conflicts of interest to disclose.
Bowles: No conflicts of interest to disclose.
Lee: Personal fees from GRAIL, Inc. for service on data safety monitoring board; textbook royalties from McGraw Hill, Oxford University press, Wolters-Kluwer; personal fees from the American College of Radiology for *Journal of the American College of Radiology* Deputy Editor duties.
Kerlikowske: Non-paid consultant with Grail on the STRIVE study.
Miglioretti: Textbook royalties from Elsevier.
Sprague: No conflicts of interest to disclose.
Henderson: No conflicts of interest to disclose.

adjust performance statistic comparisons for breast cancer risk factors, mammogram modality, demographics, additional imaging, and imaging facility.

Results: Cancer detection rates were highest among Asian/Pacific Islander (per 1000 exams, 84.2 [95%CI 72.0–98.2]) and Black women (81.4 95%CI 69.4–95.2]) and lowest among Hispanic/Latina women (42.9 [95%CI 34.2–53.6]). Positive predictive values (PPV) were higher among Black (37.0% [95%CI 31.2%–43.3%]) and White (37.0% [95%CI 30.0%–44.6%]) women and lowest among Hispanic/Latina women (22.0% [95%CI 17.2%–27.7%]). False-positive results were most common among Asian/Pacific Islander women (per 1000 exams, 183.9 [95%CI 126.7–259.2]) and lowest among White women (112.4 [95%CI 86.1–145.5]). After adjustment, false-positive and cancer detection rates remained higher for Asian/Pacific Islander and Black women (vs. Hispanic/Latina and White). Adjusted PPV was highest among Asian/Pacific Islander women.

Conclusions: Among women with a lump, Asian/Pacific Islander and Black women were more likely to have cancer detected and more likely to receive a false-positive result compared with White and Hispanic/Latina women.

Impact: Strategies for optimizing diagnostic mammography among women with a lump may vary by racial/ethnic group, but additional factors that influence performance differences need to be identified.

Keywords

mammography; breast neoplasms; racial groups; ethnicity; breast symptom; symptom assessment; Breast Cancer Surveillance Consortium

Introduction

Breast palpable abnormalities, or breast lumps, are one of the most common breast symptoms reported by women (1,2). Multiple studies show that mammograms of women with a palpable lump have higher cancer detection rates and positive predictive values (PPVs) and lower specificity than mammograms of women without a lump (1–8). The presence of a breast lump also significantly affects outcomes following a breast cancer diagnosis. Women with a lump have significantly higher rates of interval breast cancers, death due to breast cancer, and death due to any cause than women without a breast lump (5,9).

Given these worse outcomes, the ability to identify lumps with a high likelihood of malignancy is a key task for breast diagnostic imaging. Data suggest that mammography performance can differ across racial and ethnic groups (10). We recently reported that, among women who received diagnostic imaging at facilities affiliated with the Breast Cancer Surveillance Consortium (BCSC), cancer detection rates and false-positive rates were highest among Asian/Pacific Islander women and PPVs were highest among White women (11). That study described mammography performance patterns among women receiving diagnostic mammography following an abnormal screening mammogram. How those differences translate to those presenting symptomatically with a breast lump, which increases the likelihood of cancer, is unknown (1).

In the US, there are marked differences in breast cancer prognosis and survival according to race and ethnicity (12–14). Diagnostic mammography is used in the clinical workup of almost all instances of suspected breast cancer. However, whether digital diagnostic mammography performance varies across racial and ethnic groups among women with a breast lump and the impact of such performance differences on tumor prognostic characteristics is unclear. Therefore, we evaluated the performance of digital diagnostic mammography among women in four racial and ethnic groups who reported having a breast lump at the time of the mammogram. In addition, we used regression modeling to explore the extent to which racial and ethnic differences in mammography performance may be influenced by between-group differences in demographics, breast cancer risk factors, mammogram characteristics, and use of additional breast imaging.

Based on our prior work (11), we hypothesized that there is significant variation in performance by race and ethnicity, with the greatest diagnostic mammography accuracy among White women. Due to the overall worse prognosis of breast cancers diagnosed among women with a lump and higher breast cancer mortality among Black women (14,15), understanding how digital diagnostic mammography performance, including measures of sensitivity, false-positive results, and PPVs, varies across racial and ethnic groups among women with a lump is critical for identifying potential strategies for reducing racial disparities in breast cancer survival.

Materials and Methods

This study included data from women who received a diagnostic mammogram at a breast imaging facility affiliated with the BCSC. Each registry received institutional review board approval for all study procedures. All procedures were Health Insurance Portability and Accountability Act compliant. All registries and the Statistical Coordinating Center received a Federal Certificate of Confidentiality and other protections for the identities of individuals, physicians, and facilities that contributed to this research.

Data were from five BCSC regional registries that collect specific indications for diagnostic mammograms: the Carolina Mammography Registry, Kaiser Permanente Washington Registry, New Hampshire Mammography Network, San Francisco Mammography Registry, and Vermont Breast Cancer Surveillance system. This study included diagnostic digital mammograms (full-field digital mammograms [FFDM] or digital breast tomosynthesis [DBT]) performed at 72 facilities affiliated with these registries between January 2005 and August 2018 that had an indication of “evaluation of a breast problem” with the presence of a lump reported at the time of the mammogram.

A flow diagram of study exclusion criteria is in Figure 1. Mammograms were excluded for: (1) prior breast cancer diagnosis, mastectomy, or breast implants (N=7301); (2) diagnostic mammogram in the preceding 90 days (N=1582); or (3) a missing Breast Imaging Reporting and Data System (BI-RADS®) assessment (N=172). Due to the study’s focus on comparing performance among racial and ethnic groups, mammograms were excluded if race and ethnicity were missing (N=5265), or if race was classified as American Indian/Alaska Native

(N=202), other race (N=534), or mixed race (N=1098) due to the small sample sizes of these groups.

Data on demographics and health history were self-reported at the time of the diagnostic mammogram or collected from electronic health records. These data included self-reported race and ethnicity, age, first-degree family history of breast cancer, history of breast procedures, and time since last mammogram. BI-RADS breast density was classified by the interpreting radiologist. Responses to health history questions were used to calculate the predicted risk of developing breast cancer during the next 5 years using the BCSC risk calculator (16–18). Participants' residential addresses were classified according to rural-urban commuting area codes (19,20). Education and income levels within women's zip codes were obtained from the 2007–2011 American Community Survey (21). Whether women received a diagnostic ultrasound or magnetic resonance imaging (MRI) within the 90 days before or after the diagnostic mammogram was determined based on the presence of BCSC registry imaging records.

All women were followed for diagnoses of invasive breast cancer and ductal carcinoma *in situ* (DCIS) in the year following the diagnostic mammogram using cancer registry and pathology database linkages, as described previously (11).

Statistical Analysis

Race and Hispanic/Latina ethnicity were combined into a single variable for analysis, with classifications of non-Hispanic Asian/Pacific Islander (Asian/Pacific Islander), non-Hispanic Black (Black), Hispanic/Latina (any race) or non-Hispanic White (White). Patient and exam characteristics were tabulated according to race and ethnicity. Means and standard deviations were calculated for continuous variables. Diagnostic mammography performance was evaluated based on guidelines from the American College of Radiology (22). Definitions of each statistic evaluated are detailed in Nyante et al. (11) and the raw numbers used in each calculation are in Supplementary Table S1. Positive mammograms were defined as those with a final BI-RADS assessment of 4 or 5 and negative mammograms were defined as those with a final assessment of 1, 2, or 3. If a diagnostic mammogram assessment was 0 or missing, then the mammogram was followed for 90 days to determine the final assessment using methods described previously (1,23). For the 0.77% of exams with a final BI-RADS assessment of 0 during the 90-day period, we imputed a positive or negative mammogram result based on age, mammography registry, facility, reader and cancer outcome for use in calculating performance statistics (24). We computed 95% confidence intervals (CIs) using generalized estimated equations with a working independence correlation structure to account for clustering within facilities.

Logistic regression models with a binary outcome for each performance statistic as the dependent variable and race and ethnicity as the independent variable were constructed to determine how racial and ethnic differences in performance were influenced by other personal, demographic, or examination-related characteristics. The models were constructed by first adjusting for mammography registry and age (continuous age and age-squared), then adjusting sequentially for rural/urban location of residence, geocoded education level, geocoded income level, first-degree family history of breast cancer, BI-RADS breast density,

prior breast biopsy, type of diagnostic mammogram (FFDM vs. DBT), time since last mammogram, receipt of a diagnostic ultrasound and/or MRI within 90 days of the diagnostic mammogram, and imaging facility. Imaging facility was entered into the model as a fixed effect, and only observations from facilities with at least 10 breast cancer events were included. Models included all observations with non-missing data at each step, so the number of observations in each sequential model varied. As a sensitivity analysis to ensure that odds ratio differences between models were not solely due to differences in the study sample, we also evaluated the regression models for the cancer detection rate, false-positive rate, and positive predictive value using data from the subset of women who had non-missing data for all adjustment factors.

To characterize the breast cancers that were detected by diagnostic mammography, stage at diagnosis, tumor size, tumor grade, positive lymph node status, and hormone receptor status were tabulated among invasive cancers. Pathologic prognostic stage at diagnosis was classified using the American Joint Committee on Cancer (AJCC) 8th edition definitions (25). Advanced stage at diagnosis was defined as pathologic prognostic stage II or higher (25). If pathologic prognostic stage was missing, advanced stage was defined based on AJCC anatomic stage IIB or higher (26). A combined estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) variable was constructed for the analysis of triple-negative breast cancers (27,28). We tested for differences in the distribution of tumor characteristics by race and ethnicity using the chi-square test.

Statistical analyses were conducted using SAS v9.4 (SAS Institute, Cary, NC). All statistical tests were two-sided and P-values less than 0.05 were considered statistically significant.

Data Availability

The data analyzed in this study may be accessed through the BCSC upon reasonable request. Restrictions apply to the availability of these data due to patient privacy requirements.

Results

Population characteristics

This study included 51,014 diagnostic mammograms performed among 45,571 women (12% Asian/Pacific Islander, 7% Black, 6% Hispanic/Latina, 75% White) (Table 1). The mean age at exam was 48.4 years (SD 12.9 years) and 52.8% of women reported having had a mammogram in the preceding 23 months. Most diagnostic mammograms conducted were FFDM, with DBT use ranging from 1.6% among Asian/Pacific Islander women to 11.9% among White women. Additional use of diagnostic breast ultrasound and/or diagnostic MRI was common — use exceeded 90% for Asian/Pacific Islander and Hispanic/Latina women but was lower for Black women (72.0%) and White women (85.7%). Differences in ultrasound and MRI use were similar when the population was stratified by age (Supplementary Table S2).

Diagnostic mammography performance among women reporting a lump

In this study, the biopsy recommendation rate was highest among Asian/Pacific Islander women (26.8%, 95% CI: 20.6%, 34.0%) and lowest among White women (17.8%, 95% CI: 15.2%, 20.9%) (Table 2). However, positive predictive values for cancer diagnosis in one year following a positive mammogram (PPV2) and cancer diagnosis in one year following a breast biopsy (PPV3) were highest among Black and White women and lowest among Hispanic/Latina women. The cancer detection rate was highest among Asian/Pacific Islander (per 1000 mammograms, 84.2, 95% CI: 72.0, 98.2) and Black (81.4, 95% CI: 69.4, 95.2) women and lowest among Hispanic women (42.9, 95% CI: 34.2, 53.6). Trends were similar when considering the detection of invasive breast cancers only (Table 2).

Exams conducted among Asian/Pacific Islander women had the highest false-positive rate (per 1000 mammograms, 183.9, 95% CI: 126.7, 259.2) whereas the false-positive rate among White women was the lowest (112.4, 95% CI: 86.1, 145.5). Exam sensitivity was slightly higher among Asian/Pacific Islander women but similar for Black, Hispanic/Latina, and White women. The false-negative rate ranged between 3.1 and 6.3 per 1000 mammograms across all racial and ethnic groups. The majority (82.5%) of the 229 diagnoses following a false-negative exam were invasive breast cancer (Supplementary Table S3). The proportion of exams that were recommended for short interval follow-up (BI-RADS 3) was highest among Asian/Pacific Islander (9.7%, 95% CI: 5.9%, 15.7%) and Black women (9.3%, 95% CI: 7.6%, 11.4%) and lowest among Hispanic/Latina women (6.4%, 95% CI: 4.8%, 8.4%).

Contribution of personal and clinical characteristics to differences in mammography performance

Some performance differences were attenuated after adjusting for mammographic registry, including differences in the odds ratios for biopsy recommendation and PPV2 for Asian/Pacific Islander women (vs. White women) and the odds ratios for false-positives for Asian/Pacific Islander and Hispanic women (vs. White women) (Table 3). Other differences became more pronounced, such as the odds ratios for short-interval follow-up recommendation for Asian Pacific Islander (vs. White) women and false-positives for Black (vs. White) women.

Additional adjustment for age resulted in higher odds ratios for the cancer detection rate and PPV2 for Asian/Pacific Islander (vs. White women) and odds ratios closer to 1 for the cancer detection rate and PPV2 for Hispanic women (vs. White women). Adjustment for additional covariates beyond age and registry resulted in small changes to odds ratios with no consistent patterns were identified, except for adjustment for facility. Controlling for facility appeared to account for large differences in short-interval follow-up recommendation between Asian/Pacific Islander and White women, but also resulted in higher odds ratios for differences in cancer detection and biopsy recommendation for Asian/Pacific Islander women (Table 3 and Supplementary Table S4). Overall, following adjustment there were no statistically significant differences between Hispanic and White women for any performance statistics examined. Furthermore, no statistically significant difference was detected in sensitivity by race or ethnicity in any models. However, almost all adjusted odds ratios

were greater than 1 for Asian/Pacific Islander and Black women compared to White women. Regression model results were similar when the data were analyzed using the subset of participants with non-missing data for all adjustment covariates (Supplementary Table S5).

Characteristics of breast cancers diagnosed within 1 year of the diagnostic mammogram

A total of 3233 invasive breast carcinomas and 220 DCIS cases were diagnosed after one year of follow-up of a positive diagnostic mammogram assessment. The median time from the diagnostic mammogram to the cancer diagnosis ranged from 3 days for White women to 7 days for Black women (Table 4). Asian/Pacific Islander women had a higher proportion of DCIS and HER2-positive tumors and lower proportion of advanced cancers (prognostic stage II-IV) compared to other racial or ethnic groups. Black women had higher proportions of tumors that were high grade, ER-negative, PR-negative, or triple negative compared with women in other racial and ethnic groups (all $p<0.02$; Table 4). In contrast, no statistically significant differences were observed in tumor size distribution ($p=0.61$) or positive lymph node status ($p=0.06$). For all racial and ethnic groups, most tumors were >20 mm and lymph node positivity ranged from 38% to 48% (Table 4).

Discussion

We evaluated diagnostic mammography performance to determine whether there are differences by racial and ethnic group when women present with a breast lump. A comparison of unadjusted statistics showed significant differences by race and ethnicity for multiple measures, including cancer detection rate, false-positive rate, and PPVs. Mammography registry (which correlates with geography) and age were the two factors with the greatest influence on differences in diagnostic mammography performance. After accounting for registry and age differences, further adjustment for ecologic factors (such as rural/urban residence, area-level education, and income), breast cancer risk factors, mammogram modality, and time since last mammogram had little effect on differences in diagnostic mammogram performance.

After controlling for differences among groups in age distribution and mammography registry, diagnostic mammography performance was similar for Hispanic and White women. In contrast, age and registry-adjusted odds ratios were above 1 for Asian/Pacific Islander and Black women compared to White women, indicating that diagnostic mammography resulted in more follow-up exams, more biopsies, more cancers detected, and more false-positive results. One way to interpret this pattern is that under current diagnostic mammography practice for women with a breast lump, Asian/Pacific Islander and Black women are most likely to experience both cancer detection and additional testing. This result complicates the optimization of diagnostic processes to detect more cancers with fewer unnecessary procedures. Race is a social construct and serves as a proxy for an array of social, environmental, and structural processes. The elevated odds ratios after adjustment suggest that social and structural factors that were not evaluated in this study may account for the observed differences in mammography performance. For example, reports indicate that receiving care at an imaging facility that lacks accreditation is associated with breast cancer diagnostic delays and attending an unaccredited facility is more common among

Black women (29). Whether accreditation is associated with the performance measures we evaluated is unclear, however. Additionally, radiologist reading volume (30) and workflow patterns (31) are associated with mammography interpretation accuracy, but we did not have access to this information in this study and were unable to investigate the extent to which they differ among facilities that serve women of different racial and ethnic groups. Identifying additional relevant factors, particularly modifiable ones, is essential for developing interventions to eliminate racial and ethnic differences in mammography performance.

Several of the racial and ethnic performance differences for diagnostic mammography among women with a lump were similar to differences among women whose diagnostic mammogram was to evaluate a problem identified through screening (11). For example, in both populations, Asian/Pacific Islander women had the highest cancer detection rate and biopsy recommendation proportion, Hispanic/Latina women had the lowest cancer detection rate and PPVs, and White women had the lowest false-positive rate. Imaging facility was associated with differences in recommendation for short-interval follow-up in both studies, but among women with a lump, facility seemed to contribute to differences between Asian/Pacific Islander and White women, but not Black and White women as was seen for diagnostic workup following an abnormal screening mammogram (11).

Unlike diagnostic exams to evaluate a suspicious imaging finding on a screening mammogram (11), the high cancer detection rate among Asian/Pacific Islander women with a lump remained after excluding DCIS diagnoses. Additionally, the proportions of exams recommended for short-interval follow-up were similar among Asian/Pacific Islander, Black, and White women with a lump. In contrast, the short-interval follow-up recommendations were starkly higher for Black women compared to other groups undergoing diagnostic mammography after an abnormal screening mammogram (11). Finally, among women with a lump, adjustment for facility affected odds ratios for multiple performance statistics, whereas among women evaluated for a screening abnormality, the facility influence was limited to the recommendation for short-interval follow-up.

In this population, Black and White women were less likely to receive an ultrasound alongside their diagnostic mammogram when compared with Asian and Hispanic women. This was unexpected because ultrasound is considered a standard addition to the clinical workup of younger women and women with a breast lump (32–36). Our study does not have access to information that might explain why ultrasound was not performed on some women. However, the results of our regression models suggest that differences in the prevalence of ultrasound during diagnostic workup had little influence on racial and ethnic differences in mammography performance observed in this study.

Although cancer detection was similarly high among Asian/Pacific Islander and Black women, the prognostic profiles of the detected tumors were noticeably different between the groups. Black women were more likely to be diagnosed with tumors that were advanced stage, high grade, hormone receptor negative or triple-negative, or node-positive compared with Asian/Pacific Islander women. This result mirrored patterns of stage at diagnosis and hormone receptor reported for the general population (13,14) and women evaluated

for a screening abnormality (11). These data highlight the disparities in tumor prognosis regardless of the reason for a diagnostic mammogram. Improving disparities in tumor prognostic factors will likely require interventions before diagnostic mammography and may require fundamental changes in exposure to risk factors that determine tumor biology.

This study must be interpreted considering several limitations. We excluded 21% of diagnostic mammograms due to missing information on self-reported lump, usually because the facility did not provide the information to the BCSC. This creates the potential for selection bias, particularly if the demographics of facilities that did and did not report lump information differ. The number of DBT diagnostic exams during the study period was low, so the degree to which the findings apply to diagnostic DBT are unclear. Although DBT has become prevalent across the U.S., a prior BCSC study showed that women who are Asian/Pacific Islander, living in small rural areas, or have less than a high school education have less access to DBT (37). Until this access gap narrows, studies that include FFDM are still relevant for understanding how imaging may be related to breast cancer disparities. In a screening setting, DBT has been associated with lower rates of abnormal interpretation and higher rates of cancer detection when compared with full-field digital mammography (38–40). It is unclear if these same qualities hold true for diagnostic DBT and FFDM to evaluate a lump, but if they do, it is possible that the racial and ethnic gap in false-positive rates may be diminished in a population managed with diagnostic DBT. Information on ultrasound and MRI use was based on having a registry record of the exam and some facilities may not capture procedures conducted at facilities not affiliated with the registry. However, we expect that most imaging facilities with diagnostic mammography can also perform ultrasound (41).Therefore, this potential misclassification would affect only receipt of MRI, which is not recommended for routine diagnostic imaging of a lump (41). Body mass index (BMI) was not available for all mammography registries included in this analysis. Therefore, we did not account for racial and ethnic group differences in BMI, which is lowest in Asian women and highest in Black women (42). Elevated BMI increases cancer and advanced breast cancer rates (43), which could account for some differences in diagnostic performance across racial and ethnic groups.

These limitations were balanced by several strengths. The BCSC is one of the largest U.S. sources of detailed mammography data linked to cancer outcomes, which allowed us to analyze Asian/Pacific Islander and Hispanic women as distinct groups alongside Black and White women. The data included patient-reported breast problems, which allowed us to identify exams associated with a symptomatic lump. The data were collected prospectively, minimizing the possibility of recall bias in symptom reporting among women who were eventually diagnosed with cancer. We used American Community Survey data to adjust for community-level income and education level to evaluate how socioeconomic status may be associated with differences in performance, but we saw little influence of income, education, and other patient characteristics on between-group differences in performance.

In summary, age and BCSC registry explained most of the differences in diagnostic mammography performance between Hispanic and White women. However, the higher rates of biopsy recommendation, cancer detection, and false-positive results seen among Asian/Pacific Islander and Black women were not explained by any of the factors measured

in this study, including the additional use of ultrasound in the presence of a lump. Further research that considers multiple performance statistics is needed to determine the degree to which ultrasound is an effective adjunct to diagnostic mammography in community settings. Future research is also needed to address the role of factors that we were unable to evaluate, including imaging facility characteristics, BMI, and other social determinants of health, that may explain the performance differences for Black and Asian women. Additionally, exploration of racial and ethnic differences in diagnostic mammography performance by age groups, breast density category, and other factors known to affect mammography performance may further clarify some of the patterns we observed in this study. Despite similarly high rates of cancer detection for Asian/Pacific Islander and Black women, the tumor profiles of Black women displayed a poorer prognosis, similar to previously published patterns. Thus, efforts to reduce racial and ethnic disparities in breast cancer survival will require targeted efforts to reduce the risk of poor prognosis tumors in addition to improving the accuracy of cancer detection among women with a breast lump.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the participating individuals, mammography facilities, and radiologists for the data they provided for this study. We also thank Chris Tachibana, PhD, of Kaiser Permanente Washington for assistance with manuscript editing. This work was supported by the Breast Cancer Surveillance Consortium with funding from the National Cancer Institute of the National Institutes of Health under award numbers P01CA154292 (K. Kerlikowske, D. L. Miglioretti, L. A. Abraham, E. J. A. Bowles, C. I. Lee, B. L. Sprague, L. M. Henderson), R01CA266377 (C. I. Lee, D. L. Miglioretti, K. Kerlikowske, L. M. Henderson, B. L. Sprague), U54CA163303 (B. L. Sprague), and R50CA211115 (E. J. A. Bowles). More information about the Breast Cancer Surveillance Consortium may be found at: http://www.bcsc-research.org/. Research reported in this article was partially funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (PCS-1504-30370 - D. L. Miglioretti, K. Kerlikowske, L. M. Henderson, B. L. Sprague). Cancer data collection was supported by several state public health departments and cancer registries throughout the US. For a full description of these sources, please see: https://www.bcsc-research.org/about/work-acknowledgement. All statements in this manuscript, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors or Methodology Committee, nor those of the National Cancer Institute or the National Institutes of Health.

Financial support:

National Cancer Institute of the National Institutes of Health (P01CA154292 - K. Kerlikowske, D. L. Miglioretti, L. A. Abraham, E. J. A. Bowles, C. I. Lee, B. L. Sprague, L. M. Henderson; R01CA266377 - C. I. Lee, D. L. Miglioretti, K. Kerlikowske, L. M. Henderson, B. L. Sprague; U54CA163303 - B. L. Sprague; and R50CA211115 - E. J. A. Bowles) and Patient-Centered Outcomes Research Institute (PCS-1504-30370 - D. L. Miglioretti, K. Kerlikowske, L. M. Henderson, B. L. Sprague)

References

1. Sprague BL, Arao RF, Miglioretti DL, Henderson LM, Buist DS, Onega T, et al. National performance benchmarks for modern diagnostic digital mammography: Update from the Breast Cancer Surveillance Consortium. Radiology 2017;283:59–69. [PubMed: 28244803]
2. Barton MB, Elmore JG, Fletcher SW. Breast symptoms among women enrolled in a health maintenance organization: Frequency, evaluation, and outcome. Ann Intern Med 1999;130:651–7. [PubMed: 10215561]

3. Barlow WE, Lehman CD, Zheng Y, Ballard-Barbash R, Yankaskas BC, Cutter GR, et al. Performance of diagnostic mammography for women with signs or symptoms of breast cancer. J Natl Cancer Inst 2002;94:1151–9. [PubMed: 12165640]
4. Ryerson AB, Miller J, Eheman CR. Reported breast symptoms in the National Breast and Cervical Cancer Early Detection Program. Cancer Causes Control 2015;26:733–40. [PubMed: 25754109]
5. Singh D, Miettinen J, Duffy S, Malila N, Pitkäniemi J, Anttila A. Association of symptoms and interval breast cancers in the mammography-screening programme: Population-based matched cohort study. Br J Cancer 2018;119:1428–35. [PubMed: 30401890]
6. Aiello EJ, Buist DS, White E, Seger D, Taplin SH. Rate of breast cancer diagnoses among postmenopausal women with self-reported breast symptoms. J Am Board Fam Pract 2004;17:408–15. [PubMed: 15575032]
7. Singh D, Pitkäniemi J, Malila N, Anttila A. Cumulative risk of false positive test in relation to breast symptoms in mammography screening: A historical prospective cohort study. Breast Cancer Res Treat 2016;159:305–13. [PubMed: 27496148]
8. Singh D, Malila N, Pokhrel A, Anttila A. Association of symptoms and breast cancer in population-based mammography screening in finland. Int J Cancer 2015;136:E630–7. [PubMed: 25160029]
9. Brouckaert O, Schoneveld A, Truyers C, Kellen E, Van Ongeval C, Vergote I, et al. Breast cancer phenotype, nodal status and palpability may be useful in the detection of overdiagnosed screening-detected breast cancers. Ann Oncol 2013;24:1847–52. [PubMed: 23680691]
10. Yankaskas BC, Gill KS. Diagnostic mammography performance and race: Outcomes in black and white women. Cancer 2005;104:2671–81. [PubMed: 16288489]
11. Nyante SJ, Abraham L, Aiello Bowles EJ, Lee CI, Kerlikowske K, Miglioretti DL, et al. Diagnostic mammography performance across racial and ethnic groups in a national network of community-based breast imaging facilities. Cancer Epidemiol Biomarkers Prev 2022;31:1324–33. [PubMed: 35712862]
12. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, et al. Ethnicity and breast cancer: Factors influencing differences in incidence and outcome. J Natl Cancer Inst 2005;97:439–48. [PubMed: 15770008]
13. Ellis L, Canchola AJ, Spiegel D, Ladabaum U, Haile R, Gomez SL. Racial and ethnic disparities in cancer survival: The contribution of tumor, sociodemographic, institutional, and neighborhood characteristics. J Clin Oncol 2018;36:25–33. [PubMed: 29035642]
14. Iqbal J, Ginsburg O, Rochon PA, Sun P, Narod SA. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. Jama 2015;313:165–73. [PubMed: 25585328]
15. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69:438–51. [PubMed: 31577379]
16. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast density and benign breast disease: Risk assessment to identify women at high risk of breast cancer. J Clin Oncol 2015;33:3137–43. [PubMed: 26282663]
17. Frank RD, Winham SJ, Vierkant RA, Frost MH, Radisky DC, Ghosh K, et al. Evaluation of 2 breast cancer risk models in a benign breast disease cohort. Cancer 2018;124:3319–28. [PubMed: 29932456]
18. Tice JA, Bissell MCS, Miglioretti DL, Gard CC, Rauscher GH, Dabbous FM, et al. Validation of the Breast Cancer Surveillance Consortium model of breast cancer risk. Breast Cancer Res Treat 2019;175:519–23. [PubMed: 30796654]
19. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. American journal of public health 2005;95:1149–55. [PubMed: 15983270]
20. Morrill R, Cromartie J, Hart G. Metropolitan, urban, and rural commuting areas: Toward a better depiction of the united states settlement system. Urban geography 1999;20:727–48.
21. 2007–2011 American Community Survey 5-year estimate. US Census Bureau. Available from <https://www.census.gov/programs-surveys/acs/data.html>.
22. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas. Reston, VA: American College of Radiology; 2013.

23. Sickles EA, Miglioretti DL, Ballard-Barbash R, Geller BM, Leung JW, Rosenberg RD, et al. Performance benchmarks for diagnostic mammography. Radiology 2005;235:775–90. [PubMed: 15914475]
24. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30:377–99. [PubMed: 21225900]
25. Amin MB, Edge S, Greene F, Byrd DR, Brookland RK, Washington MK, et al., editors. AJCC cancer staging manual. 8th ed. New York: Springer International Publishing; 2017.
26. Kerlikowske K, Bissell MCS, Sprague BL, Buist DSM, Henderson LM, Lee JM, et al. Advanced breast cancer definitions by staging system examined in the Breast Cancer Surveillance Consortium. J Natl Cancer Inst 2020.
27. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. Jama 2006;295:2492–502. [PubMed: 16757721]
28. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010;16:5222–32. [PubMed: 20837693]
29. Molina Y, Silva A, Rauscher GH. Racial/ethnic disparities in time to a breast cancer diagnosis: The mediating effects of health care facility factors. Med Care 2015;53:872–8. [PubMed: 26366519]
30. Théberge I, Chang SL, Vandal N, Daigle JM, Guertin MH, Pelletier E, et al. Radiologist interpretive volume and breast cancer screening accuracy in a Canadian organized screening program. J Natl Cancer Inst 2014;106:djt461. [PubMed: 24598715]
31. Buist DS, Anderson ML, Smith RA, Carney PA, Miglioretti DL, Monsees BS, et al. Effect of radiologists' diagnostic work-up volume on interpretive performance. Radiology 2014;273:351–64. [PubMed: 24960110]
32. Chen Y, Chou SS, Blaschke EM, Specht MC, Lehman CD. Value of mammography for women 30–39 years old presenting with breast symptoms. AJR Am J Roentgenol 2018;211:1416–24. [PubMed: 30300000]
33. Houssami N, Irwig L, Simpson JM, McKessar M, Blome S, Noakes J. Sydney breast imaging accuracy study: Comparative sensitivity and specificity of mammography and sonography in young women with symptoms. AJR Am J Roentgenol 2003;180:935–40. [PubMed: 12646432]
34. Lehman CD, Lee CI, Loving VA, Portillo MS, Peacock S, DeMartini WB. Accuracy and value of breast ultrasound for primary imaging evaluation of symptomatic women 30–39 years of age. AJR Am J Roentgenol 2012;199:1169–77. [PubMed: 23096195]
35. Britton P, Warwick J, Wallis MG, O'Keeffe S, Taylor K, Sinnatamby R, et al. Measuring the accuracy of diagnostic imaging in symptomatic breast patients: Team and individual performance. Br J Radiol 2012;85:415–22. [PubMed: 21224304]
36. McCavert M, O'Donnell ME, Aroori S, Badger SA, Sharif MA, Crothers JG, et al. Ultrasound is a useful adjunct to mammography in the assessment of breast tumours in all patients. Int J Clin Pract 2009;63:1589–94. [PubMed: 19686337]
37. Lee CI, Zhu W, Onega T, Henderson LM, Kerlikowske K, Sprague BL, et al. Comparative access to and use of digital breast tomosynthesis screening by women's race/ethnicity and socioeconomic status. JAMA Netw Open 2021;4:e2037546. [PubMed: 33606032]
38. Honig EL, Mullen LA, Amir T, Alvin MD, Jones MK, Ambinder EB, et al. Factors impacting false positive recall in screening mammography. Acad Radiol 2019;26:1505–12. [PubMed: 30772138]
39. Lowry KP, Coley RY, Miglioretti DL, Kerlikowske K, Henderson LM, Onega T, et al. Screening performance of digital breast tomosynthesis vs digital mammography in community practice by patient age, screening round, and breast density. JAMA Netw Open 2020;3:e2011792. [PubMed: 32721031]
40. Skaane P, Sebuødegård S, Bandos AI, Gur D, Østerås BH, Gullien R, et al. Performance of breast cancer screening using digital breast tomosynthesis: Results from the prospective population-based oslo tomosynthesis screening trial. Breast Cancer Res Treat 2018;169:489–96. [PubMed: 29429017]

41. Lee CI, Bogart A, Hubbard RA, Obadina ET, Hill DA, Haas JS, et al. Advanced breast imaging availability by screening facility characteristics. Acad Radiol 2015;22:846–52. [PubMed: 25851643]

42. Bissell MCS, Kerlikowske K, Sprague BL, Tice JA, Gard CC, Tossas KY, et al. Breast cancer population attributable risk proportions associated with body mass index and breast density by race/ethnicity and menopausal status. Cancer Epidemiol Biomarkers Prev 2020;29:2048–56. [PubMed: 32727722]

43. Kerlikowske K, Walker R, Miglioretti DL, Desai A, Ballard-Barbash R, Buist DS. Obesity, mammography use and accuracy, and advanced breast cancer risk. J Natl Cancer Inst 2008;100:1724–33. [PubMed: 19033562]

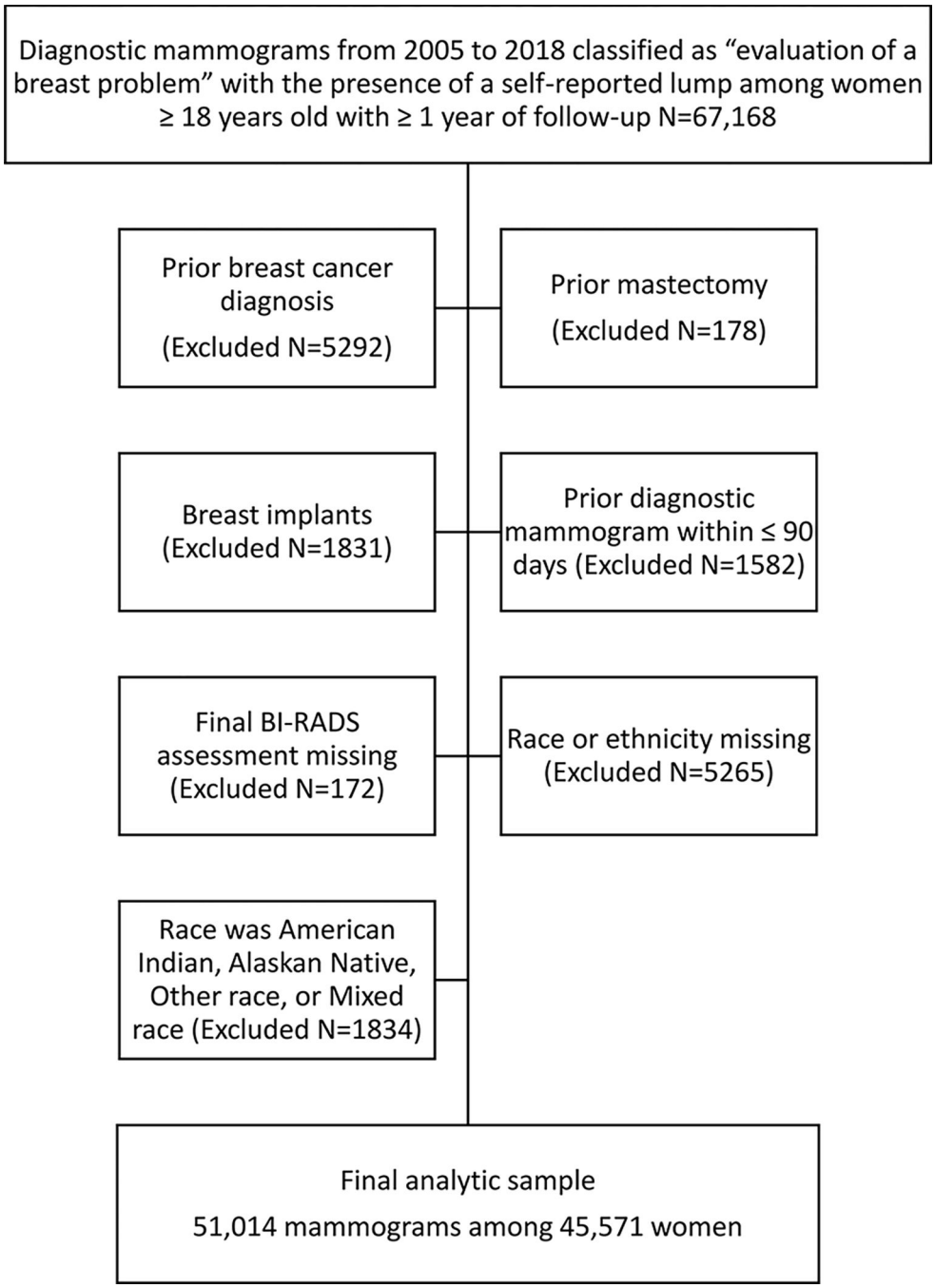

Figure 1. Diagnostic mammogram inclusion criteria.

This study included digital diagnostic mammograms performed at an imaging facility affiliated with the Breast Cancer Surveillance Consortium (BCSC) among women who reported a breast lump. Eligible mammograms were performed between 2005 and 2018 and did not have another diagnostic mammogram in the preceding 90 days.

Table 1.

Characteristics for 51,014 diagnostic mammograms performed to evaluate a breast lump among 45,571 women in the Breast Cancer Surveillance Consortium, 2005–2018.

		Race and Ethnicity			
Characteristic	**Overall**	**Asian/Pacific Islander**	**Black**	**Hispanic/Latina**	**White**
Total mammograms	**51014**	**6271**	**3355**	**3242**	**38146**
	N (%)	N (%)	N (%)	N (%)	N (%)
Age (years)					
< 40	13826 (27.1)	1844 (29.4)	920 (27.4)	1126 (34.7)	9936 (26.0)
40–49	17035 (33.4)	2352 (37.5)	1059 (31.6)	1234 (38.1)	12390 (32.5)
50–59	10475 (20.5)	1204 (19.2)	690 (20.6)	568 (17.5)	8013 (21.0)
60–69	5825 (11.4)	520 (8.3)	431 (12.8)	218 (6.7)	4656 (12.2)
70–79	2548 (5.0)	237 (3.8)	177 (5.3)	69 (2.1)	2065 (5.4)
≥ 80	1305 (2.6)	114 (1.8)	78 (2.3)	27 (0.8)	1086 (2.8)
Mean (SD)	48.4 (12.9)	46.7 (11.9)	48.6 (13.2)	44.7 (10.8)	49.0 (13.2)
Probability of college or higher education in zip code[a,b]					
0–55%	9587 (19.9)	1439 (24.0)	1433 (44.4)	532 (17.7)	6183 (17.2)
56–70%	13270 (27.6)	1568 (26.1)	1100 (34.1)	747 (24.8)	9855 (27.5)
71–80%	13458 (28.0)	1858 (31.0)	377 (11.7)	1012 (33.6)	10211 (28.5)
81–100%	11810 (24.5)	1135 (18.9)	318 (9.9)	721 (23.9)	9636 (26.9)
Missing	*2889 (5.7)*	*271 (4.3)*	*127 (3.8)*	*230 (7.1)*	*2261 (5.9)*
Median family income based in zip code[a]					
≤ $60,000	11691 (24.3)	847 (14.1)	1950 (60.4)	550 (18.3)	8344 (23.3)
$60,001-$80,000	13524 (28.1)	1507 (25.1)	722 (22.4)	758 (25.2)	10537 (29.4)
$80,001-$100,000	10830 (22.5)	2163 (36.1)	314 (9.7)	856 (28.4)	7497 (20.9)
> $100,000	12038 (25.0)	1482 (24.7)	242 (7.5)	848 (28.2)	9466 (26.4)
Missing	*2931 (5.7)*	*272 (4.3)*	*127 (3.8)*	*230 (7.1)*	*2302 (6.0)*
Rural/urban residence					
Urban	39416 (80.7)	5972 (98.8)	2490 (75.3)	2814 (92.0)	28140 (77.2)
Large rural	3756 (7.7)	32 (0.5)	622 (18.8)	120 (3.9)	2982 (8.2)
Small rural	2399 (4.9)	19 (0.3)	125 (3.8)	39 (1.3)	2216 (6.1)
Isolated rural	3298 (6.7)	19 (0.3)	71 (2.1)	86 (2.8)	3122 (8.6)
Missing	*2145 (4.2)*	*229 (3.7)*	*47 (1.4)*	*183 (5.6)*	*1686 (4.4)*
First-degree family history of breast cancer					
No	40136 (83.8)	5360 (90.9)	2678 (85.4)	2598 (87.3)	29500 (82.2)
Yes	7775 (16.2)	537 (9.1)	456 (14.6)	377 (12.7)	6405 (17.8)
Unknown	*3103 (6.1)*	*374 (6.0)*	*221 (6.6)*	*267 (8.2)*	*2241 (5.9)*
History of breast biopsy or aspiration					
No	36456 (72.0)	4689 (75.0)	2507 (74.8)	2397 (74.3)	26863 (71.1)

		Race and Ethnicity			
Characteristic	**Overall**	**Asian/Pacific Islander**	**Black**	**Hispanic/Latina**	**White**
Yes	14170 (28.0)	1567 (25.0)	844 (25.2)	829 (25.7)	10930 (28.9)
Unknown	*388 (0.8)*	*15 (0.2)*	*4 (0.1)*	*16 (0.5)*	*353 (0.9)*
Diagnostic mammogram modality					
Digital	45856 (89.9)	6169 (98.4)	3012 (89.8)	3052 (94.1)	33623 (88.1)
DBT[c]	5158 (10.1)	102 (1.6)	343 (10.2)	190 (5.9)	4523 (11.9)
Time since last mammogram					
No prior mammogram	11045 (23.6)	1720 (29.4)	558 (20.3)	929 (31.4)	7838 (22.3)
< 12 months	12086 (25.8)	1306 (22.3)	668 (24.3)	581 (19.7)	9531 (27.1)
12–23 months	12620 (27.0)	1683 (28.7)	696 (25.3)	754 (25.5)	9487 (26.9)
24–36 months	3960 (8.5)	438 (7.5)	279 (10.2)	268 (9.1)	2975 (8.4)
> 36 months	7061 (15.1)	713 (12.2)	547 (19.9)	423 (14.3)	5378 (15.3)
Unknown	*4242 (8.3)*	*411 (6.6)*	*607 (18.1)*	*287 (8.9)*	*2937 (7.7)*
Breast density					
Almost entirely fat	3972 (8.7)	150 (2.9)	316 (11.0)	204 (7.0)	3302 (9.5)
Scattered fibroglandular densities	13831 (30.2)	768 (14.8)	1215 (42.3)	861 (29.4)	10987 (31.6)
Heterogeneously dense	19335 (42.3)	2262 (43.5)	1071 (37.3)	1358 (46.4)	14644 (42.2)
Extremely dense	8591 (18.8)	2023 (38.9)	269 (9.4)	504 (17.2)	5795 (16.7)
Unknown	*5285 (10.4)*	*1068 (17.0)*	*484 (14.4)*	*315 (9.7)*	*3418 (9.0)*
BCSC 5-year predicted breast cancer risk (%)					
< 1	18692 (48.4)	2770 (61.7)	1189 (50.0)	1740 (71.0)	12993 (44.4)
1.00 – 1.66	11290 (29.3)	1390 (31.0)	763 (32.1)	503 (20.5)	8634 (29.5)
1.67 – 2.49	5615 (14.5)	272 (6.1)	297 (12.5)	170 (6.9)	4876 (16.7)
2.50 – 3.99	2568 (6.7)	53 (1.2)	117 (4.9)	36 (1.5)	2362 (8.1)
≥ 4.00	429 (1.1)	2 (0.0)	12 (0.5)	2 (0.1)	413 (1.4)
mean (sd)	1.2 (0.9)	0.9 (0.5)	1.1 (0.7)	0.8 (0.6)	1.3 (0.9)
Unknown	*12420 (24.3)*	*1784 (28.4)*	*977 (29.1)*	*791 (24.4)*	*8868 (23.2)*
Diagnostic US/MRI[d] within +/− 90 days of exam					
No	7045 (13.8)	387 (6.2)	938 (28.0)	245 (7.6)	5475 (14.4)
US only	43600 (85.5)	5850 (93.3)	2400 (71.5)	2983 (92.0)	32367 (84.9)
MRI (with and without US)[e]	369 (0.7)	34 (0.5)	17 (0.5)	14 (0.4)	304 (0.8)

[a] Based on 2007–2011 American Community Survey data

[b] Average probability of college or higher education

[c] DBT – digital breast tomosynthesis

[d] US – ultrasound; MRI – magnetic resonance imaging

[e] 348 exams with MRI and US

Table 2.

Mammography performance estimates for 51,014 diagnostic mammograms performed to evaluate a breast lump in the Breast Cancer Surveillance Consortium by race and ethnicity, 2005–2018.

	Race and Ethnicity			
Measure	**Asian/Pacific Islander**	**Black**	**Hispanic/Latina**	**White**
Total number of mammograms	6271	3355	3242	38146
Biopsy				
Biopsy recommendation rate (%)	26.8 (20.6, 34.0)	22.0 (19.3, 24.9)	19.5 (16.4, 23.1)	17.8 (15.2, 20.9)
PPV[a] 2 - biopsy recommended (%)	31.4 (23.6, 40.4)	37.0 (31.2, 43.3)	22.0 (17.2, 27.7)	37.0 (30.0, 44.6)
PPV3 - biopsy performed (%)	34.7 (28.6, 41.4)	43.9 (37.4, 50.7)	25.9 (20.8, 31.6)	44.5 (37.0, 52.3)
Cancer Detection				
Cancer detection rate (per 1000 exams)	84.2 (72.0, 98.2)	81.4 (69.4, 95.2)	42.9 (34.2, 53.6)	66.0 (58.9, 74.0)
Invasive cancer detection rate (per 1000 exams)	74.5 (63.1, 87.6)	76.0 (64.0, 90.0)	39.5 (31.2, 49.9)	62.6 (55.6, 70.4)
False-positive rate (per 1000 exams)	183.9 (126.7, 259.2)	138.3 (113.8, 167.1)	152.4 (122.7, 187.6)	112.4 (86.1, 145.5)
Sensitivity (%)	96.2 (94.4, 97.4)	92.9 (88.5, 95.6)	93.3 (88.3, 96.2)	93.4 (91.5, 95.0)
Sensitivity - invasive cancers (%)	96.7 (94.6, 98.0)	93.4 (88.9, 96.2)	94.8 (88.9, 97.7)	94.2 (92.3, 95.6)
False-negative rate (per 1000 exams)	3.3 (2.3, 4.9)	6.3 (3.7, 10.6)	3.1 (1.9, 5.1)	4.6 (3.5, 6.2)
False-negative rate - invasive cancers (per 1000 exams)	2.6 (1.6, 4.2)	5.4 (3.0, 9.5)	2.2 (1.1, 4.3)	3.9 (2.9, 5.2)
Short-interval follow-up				
Short-interval follow-up recommendation proportion (%)	9.7 (5.9, 15.7)	9.3 (7.6, 11.4)	6.4 (4.8, 8.4)	8.5 (6.5, 11.0)
False-positive short-interval follow-up rate (per 1000 exams)	96.3 (57.8, 156.2)	91.2 (74.3, 111.5)	63.5 (48.0, 83.7)	83.6 (64.0, 108.5)
PPV – short-interval follow-up (%)	1.0 (0.6, 1.7)	1.9 (1.1, 3.4)	0.5 (0.1, 2.6)	1.1 (0.7, 1.7)

[a]PPV, positive predictive value

Table 3.

Comparison of diagnostic mammography performance statistics among women being evaluated for a breast lump by race and ethnicity in (2005–2018), adjusted for individual and examination characteristics.

Model adjustment variables[a,b,c]	**Sensitivity (modeling probability of a positive mammogram given a cancer)**	**False-positive rate (modeling probability of false-positive mammogram)**	**Cancer detection rate (modeling probability of a true positive mammogram)**	**Invasive cancer detection rate (modeling probability of a true positive [invasive] mammogram)**	**Short interval follow-up recommendation rate (modeling probability of a BI-RADS 3 assessment)**	**Biopsy recommendation rate (modeling probability of a positive mammogram)**	**PPV2 (modeling probability of a true positive given a positive mammogram)**
	(3688 exams with 3459 events)	**(51014 exams with 6399 events)**	**(51014 exams with 3459 events)**	**(51014 exams with 3237 events)**	**(51014 exams with 4353 events)**	**(51014 exams with 9858 events)**	**(9858 exams with 3459 events)**
	OR (95% CI)	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**
Unadjusted Model							
White	ref	ref	ref	ref	ref	ref	ref
Asian/Pacific Islander	1.77 (1.14, 2.88)	1.78 (1.66, 1.91)	1.30 (1.18, 1.43)	1.21 (1.09, 1.33)	1.17 (1.06, 1.28)	1.69 (1.58, 1.79)	0.78 (0.70, 0.87)
Black	0.91 (0.58, 1.50)	1.27 (1.14, 1.40)	1.25 (1.10, 1.42)	1.23 (1.08, 1.41)	1.11 (0.98, 1.25)	1.30 (1.19, 1.41)	1.00 (0.85, 1.17)
Hispanic/Latina	0.98 (0.53, 2.01)	1.42 (1.28, 1.57)	0.63 (0.53, 0.75)	0.62 (0.51, 0.74)	0.74 (0.64, 0.85)	1.12 (1.02, 1.22)	0.48 (0.39, 0.58)
P-value[d]	*(p=0.065)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*
Model 1 [e]							
White	ref	ref	ref	ref	ref	ref	ref
Asian/Pacific Islander	1.32 (0.81, 2.25)	1.29 (1.19, 1.39)	1.36 (1.22, 1.51)	1.29 (1.16, 1.45)	1.81 (1.63, 2.01)	1.35 (1.27, 1.45)	1.09 (0.96, 1.24)
Black	1.36 (0.84, 2.31)	1.50 (1.34, 1.67)	1.22 (1.06, 1.40)	1.18 (1.02, 1.36)	1.17 (1.02, 1.33)	1.43 (1.31, 1.57)	0.80 (0.67, 0.96)
Hispanic/Latina	0.76 (0.40, 1.59)	1.15 (1.04, 1.27)	0.64 (0.53, 0.76)	0.63 (0.52, 0.75)	0.98 (0.84, 1.13)	0.96 (0.88, 1.06)	0.58 (0.47, 0.70)
P-value	*(p=0.3148)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*
Model 2 [f]							
White	ref	ref	ref	ref	ref	ref	ref
Asian/Pacific Islander	1.36 (0.83, 2.31)	1.24 (1.15, 1.34)	1.63 (1.46, 1.81)	1.55 (1.38, 1.73)	1.76 (1.58, 1.95)	1.39 (1.30, 1.49)	1.33 (1.16, 1.52)
Black	1.43 (0.88, 2.44)	1.50 (1.34, 1.67)	1.32 (1.15, 1.52)	1.28 (1.10, 1.48)	1.16 (1.02, 1.32)	1.47 (1.34, 1.61)	0.86 (0.71, 1.04)
Hispanic/Latina	0.83 (0.44, 1.75)	1.06 (0.95, 1.17)	0.86 (0.72, 1.03)	0.85 (0.71, 1.03)	0.92 (0.79, 1.07)	1.01 (0.92, 1.11)	0.84 (0.67, 1.03)
P-value	*(p=0.27)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*

Model adjustment variables[a,b,c]	Sensitivity (modeling probability of a positive mammogram given a cancer)	False-positive rate (modeling probability of false-positive mammogram)	Cancer detection rate (modeling probability of a true positive mammogram)	Invasive cancer detection rate (modeling probability of a true positive [invasive] mammogram)	Short interval follow-up recommendation rate (modeling probability of a BI-RADS 3 assessment)	Biopsy recommendation rate (modeling probability of a positive mammogram)	PPV2 (modeling probability of a true positive given a positive mammogram)
	(3688 exams with 3459 events)	(51014 exams with 6399 events)	(51014 exams with 3459 events)	(51014 exams with 3237 events)	(51014 exams with 4353 events)	(51014 exams with 9858 events)	(9858 exams with 3459 events)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Model 11 [g]							
White	ref	ref	ref	ref	ref	ref	ref
Asian/Pacific Islander	1.05 (0.58, 1.99)	1.11 (1.00, 1.23)	1.37 (1.18, 1.59)	1.29 (1.11, 1.50)	1.66 (1.45, 1.89)	1.22 (1.12, 1.33)	1.27 (1.06, 1.53)
Black	1.18 (0.64, 2.27)	1.56 (1.35, 1.79)	1.34 (1.11, 1.61)	1.27 (1.05, 1.53)	1.17 (0.99, 1.38)	1.55 (1.37, 1.74)	0.83 (0.64, 1.06)
Hispanic/Latina	0.61 (0.29, 1.44)	1.01 (0.89, 1.15)	0.90 (0.72, 1.11)	0.89 (0.71, 1.10)	0.91 (0.75, 1.08)	0.99 (0.88, 1.10)	0.84 (0.64, 1.08)
P-value	*(p=0.62)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p=0.0048)*
Model 12 [h]							
White	ref	ref	ref	ref	ref	ref	ref
Asian/Pacific Islander	---[i]	1.26 (1.13, 1.40)	1.66 (1.41, 1.95)	1.59 (1.34, 1.88)	0.87 (0.73, 1.04)	1.44 (1.30, 1.58)	1.31 (1.07, 1.60)
Black	---[i]	1.44 (1.24, 1.68)	1.26 (1.03, 1.53)	1.18 (0.96, 1.45)	1.17 (0.98, 1.40)	1.42 (1.25, 1.61)	0.82 (0.63, 1.06)
Hispanic/Latina	---[i]	1.05 (0.92, 1.20)	0.90 (0.72, 1.12)	0.89 (0.70, 1.11)	0.95 (0.78, 1.14)	1.02 (0.90, 1.14)	0.81 (0.62, 1.06)
P-value		*(p < 0.001)*	*(p < 0.001)*	*(p < 0.001)*	*(p=0.12)*	*(p < 0.001)*	*(p=0.0036)*

[a] Models included observations with non-missing data for each adjustment factor (i.e., number of observations varied by model).

[b] Models 3–10 are shown in Supplementary Table S4

[c] The numbers of exams and events by racial and ethnic group are shown in Supplementary Table S1

[d] Chi-square-based P-value comparing differences across all four groups

[e] Adjusted for mammography registry

[f] Adjusted for mammography registry, age, and age-squared

[g] Adjusted for mammography registry, age, age-squared, rural/urban residence, education, income, family history of breast cancer, breast density, history of breast biopsy, diagnostic mammogram modality, time since last mammogram, and receipt of diagnostic ultrasound or MRI

[h] Adjusted for mammography registry, age, age-squared, rural/urban residence, education, income, family history of breast cancer, breast density, history of breast biopsy, diagnostic mammogram modality, time since last mammogram, receipt of diagnostic ultrasound or MRI, and imaging facility

[i] Could not be computed due to model non-convergence

Table 4.

Characteristics of 3453 breast cancers detected by diagnostic mammography among women being evaluated for a breast lump in the Breast Cancer Surveillance Consortium, 2005–2018.

		Race and Ethnicity				
Characteristic	**Overall**	**Asian/Pacific Islander**	**Black**	**Hispanic**	**White**	**P-value**
# mammographically detected cancers	3453	527	273	139	2514	
Median days between diagnostic mammogram and cancer diagnosis (interquartile range)	4 (0, 9)	6 (0, 14)	7 (1, 14)	5 (1, 9)	3 (0, 8)	
Cancer stage at diagnosis [a]						< 0.001
DCIS (0)	220 (6.5)	61 (11.8)	18 (6.8)	11 (8.0)	130 (5.3)	
Early (I)	2085 (61.8)	317 (61.2)	132 (50.0)	81 (59.1)	1555 (63.3)	
Advanced (II-IV)	1071 (31.7)	140 (27.0)	114 (43.2)	45 (32.8)	772 (31.4)	
Unknown	*77 (2.2)*	*9 (1.7)*	*9 (3.3)*	*2 (1.4)*	*57 (2.3)*	
Invasive tumor size (mm) [b]						0.606
1–10	319 (10.1)	53 (11.7)	22 (8.9)	9 (7.3)	235 (10.1)	
>10–20	967 (30.7)	141 (31.1)	69 (28.0)	36 (29.0)	721 (31.0)	
>20	1865 (59.2)	260 (57.3)	155 (63.0)	79 (63.7)	1371 (58.9)	
Unknown	*82 (2.5)*	*12 (2.6)*	*9 (3.5)*	*4 (3.1)*	*57 (2.4)*	
Mean (SD)	29.3 (19.3)	27.5 (16.6)	32.4 (21.6)	30.6 (19.2)	29.2 (19.5)	
Tumor grade[b]						0.019
Low	504 (16.4)	66 (14.9)	25 (10.5)	23 (18.9)	390 (17.2)	
Intermediate	1287 (41.9)	186 (42.0)	91 (38.2)	46 (37.7)	964 (42.6)	
High	1277 (41.6)	191 (43.1)	122 (51.3)	53 (43.4)	911 (40.2)	
Unknown	*165 (5.1)*	*23 (4.9)*	*17 (6.7)*	*6 (4.7)*	*119 (5.0)*	
Estrogen receptor[b]						< 0.001
Positive	2434 (78.3)	349 (76.9)	145 (63.0)	98 (80.3)	1842 (80.0)	
Negative	675 (21.7)	105 (23.1)	85 (37.0)	24 (19.7)	461 (20.0)	
Unknown	*124 (3.8)*	*12 (2.6)*	*25 (9.8)*	*6 (4.7)*	*81 (3.4)*	
Progesterone receptor[b]						< 0.001
Positive	2170 (69.9)	317 (69.8)	119 (52.0)	92 (75.4)	1642 (71.4)	
Negative	936 (30.1)	137 (30.2)	110 (48.0)	30 (24.6)	659 (28.6)	
Unknown	*127 (3.9)*	*12 (2.6)*	*26 (10.2)*	*6 (4.7)*	*83 (3.5)*	
Positive lymph node status[b]						0.059
No positive lymph nodes	1811 (57.1)	284 (62.0)	128 (52.0)	69 (54.8)	1330 (56.8)	
≥ 1 positive lymph nodes	1360 (42.9)	174 (38.0)	118 (48.0)	57 (45.2)	1011 (43.2)	
Unknown	*62 (1.9)*	*8 (1.7)*	*9 (3.5)*	*2 (1.6)*	*43 (1.8)*	
HER2[b]						0.002
Positive	483 (17.6)	106 (23.7)	29 (16.1)	15 (12.5)	333 (16.7)	

		Race and Ethnicity				
Characteristic	**Overall**	**Asian/Pacific Islander**	**Black**	**Hispanic**	**White**	**P-value**
Negative	2262 (82.4)	341 (76.3)	151 (83.9)	105 (87.5)	1665 (83.3)	
Unknown	*488 (15.1)*	*19 (4.1)*	*75 (29.4)*	*8 (6.3)*	*386 (16.2)*	
ER/PR/HER2[b]						< 0.001
ER+ and/or PR+, HER2−	*1862 (68.0)*	281 (63.0)	101 (56.1)	86 (71.7)	1394 (69.9)	
ER+ and/or PR+, HER2+	*333 (12.2)*	69 (15.5)	19 (10.6)	12 (10.0)	233 (11.7)	
ER−, PR−, HER2−	*395 (14.4)*	59 (13.2)	50 (27.8)	19 (15.8)	267 (13.4)	
ER−, PR−, HER2+	*150 (5.5)*	37 (8.3)	10 (5.6)	3 (2.5)	100 (5.0)	
Unknown	493 (15.2)	*20 (4.3)*	*75 (29.4)*	*8 (6.3)*	*390 (16.4)*	

[a]Pathologic prognostic stage, AJCC 8th edition. If pathologic prognostic stage was missing, advanced cancer status was defined based on anatomic stage IIB or higher.

[b]Tabulated among invasive breast cancer diagnoses only

Abbreviations: DCIS – ductal carcinoma in situ; SD – standard deviation; HER2 – human epidermal growth factor receptor 2; ER – estrogen receptor; PR – progesterone receptor.