Title

Diagnostic Mammography Performance across Racial and Ethnic Groups in a National Network of Community-Based Breast Imaging Facilities.

Journal

Cancer Epidemiology Biomarkers & Prevention, 31(7)

ISSN

1055-9965

Authors

Nyante, Sarah J
Abraham, Linn
Aiello Bowles, Erin J
et al.

Publication Date

2022-07-01

DOI

10.1158/1055-9965.epi-21-1379

Peer reviewed

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2022 July 01; 31(7): 1324–1333.

doi:10.1158/1055-9965.EPI-21-1379.

Diagnostic mammography performance across racial and ethnic groups in a national network of community-based breast imaging facilities

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, 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 differences in diagnostic mammography performance based on women's race/ethnicity.

Methods: This cohort study included 267,868 diagnostic mammograms performed to evaluate screening mammogram findings at 98 facilities in the Breast Cancer Surveillance Consortium between 2005–2017. Mammogram assessments were recorded prospectively and breast cancers occurring within one year were ascertained. Performance statistics were calculated with 95%

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: Research grant to institution, GE Healthcare; 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

confidence intervals (CI) for each racial/ethnic group. Multivariable regression was used to control for personal characteristics and imaging facility.

Results: Among non-Hispanic White (70%), non-Hispanic Black (13%), Asian/Pacific Islander (10%) and Hispanic (7%) women, the invasive cancer detection rate (iCDR, per 1000 mammograms) and positive predictive value (PPV2) were highest among non-Hispanic White women (iCDR=35.8 [95% CI=35.0,36.7]; PPV2=27.8 [95% CI=27.3,28.3]) and lowest among Hispanic women (iCDR=22.3 [95% CI=20.2,24.6]; PPV2=19.4 [95% CI=18.0,20.9]). Short interval follow-up recommendations were most common among non-Hispanic Black women (31.0% [95% CI=30.6%,31.5%] versus other groups, range 16.6% to 23.6%). False-positive biopsy recommendations were most common among Asian/Pacific Islander women (per 1000 mammograms: 169.2 [95% CI=164.8,173.7] versus other groups, range 126.5 to 136.1). Some differences were explained by adjusting for receipt of diagnostic ultrasound or magnetic resonance imaging for iCDR and imaging facility for short-interval follow-up. Other differences changed little after adjustment.

Conclusions: Diagnostic mammography performance varied across racial/ethnic groups. Addressing characteristics related to imaging facility and access, rather than personal characteristics, may help reduce some of these disparities.

Impact: Diagnostic mammography performance studies should include racially and ethnically diverse populations to provide an accurate view of the population-level effects.

Keywords

mammography; breast cancer; race; ethnicity; BCSC

Introduction

Racial and ethnic differences in breast cancer incidence and mortality are well documented. In the US, breast cancer incidence is highest among non-Hispanic White and non-Hispanic Black women and breast cancer mortality is highest for non-Hispanic Black women, whereas incidence and mortality rates are lowest for Hispanic and Asian/Pacific Islander women (1). Differences are also apparent in the types of breast cancers diagnosed. Non-Hispanic Black women are more likely to be diagnosed with tumors that have poorer prognostic characteristics (*e.g.*, large size, high grade, hormone receptor negative) compared with other groups (1). Hispanic women are more commonly diagnosed with larger breast tumors and regional or distant stage disease when compared with non-Hispanic White women (2–4). Efforts to determine the root causes of these differences have identified racial/ethnic differences in risk factors, socioeconomic status, health insurance coverage, provider actions/inaction, access to timely treatment, and access to high-quality care as contributors (5,6). However, differences are not eliminated completely after controlling for these factors (4,7–9), suggesting that additional causes may be involved.

Mammography plays a major role in breast cancer detection as a screening and diagnostic tool, but its potential contributions to racial differences in breast cancer characteristics and mortality have not been fully established. While several studies have addressed racial/ethnic differences in the timeliness of breast cancer screening and receipt of care

after an abnormal screening mammogram (3,10–15), little is known about differences in other parts of the diagnostic work-up continuum, particularly outcomes following a diagnostic mammography examination. Studies addressing this have analyzed film screen mammography (a technology that is no longer used) and suggest that diagnostic mammography performance varies across racial/ethnic groups, which could affect the characteristics of cancers diagnosed. Diagnostic film mammography sensitivity is higher for Black women compared with White women and specificity is lower for Black women compared with White women (16). Other data show that rates of abnormal interpretation and cancer detection differ for digital vs. film screen mammography (17,18). However, these data were not reported for racial/ethnic groups, leaving the question of whether the racial and ethnic differences reported for diagnostic performance with film screen mammography persist for digital mammography.

The goal of this study was to evaluate racial/ethnic differences in the performance of diagnostic digital mammography conducted following a recent screening mammogram. Using prospectively collected data from six population-based breast imaging registries, we evaluated diagnostic mammography performance statistics among non-Hispanic White, non-Hispanic Black, Asian/Pacific Islander, and Hispanic women. Examining differences in diagnostic digital mammography performance and tumor characteristic outcomes by race and ethnicity may help to understand why disparities in cancer detection and quality of care may persist for some demographic groups.

Materials and Methods

Study Population

This study was conducted using data from the Breast Cancer Surveillance Consortium (BCSC) (19), a collaborative network of breast imaging registries that collects data on breast imaging examinations, breast procedures, and cancer diagnoses occurring among women seen at participating facilities. This analysis included data from digital diagnostic mammograms (including full-field digital mammograms [FFDM] and digital breast tomosynthesis [DBT]) performed between 2005–2017 for women aged ≥ 18 years at imaging facilities in six registries: Carolina Mammography Registry, Kaiser Permanente Washington Registry, Metropolitan Chicago Breast Cancer Registry, New Hampshire Mammography Network, San Francisco Mammography Registry, and Vermont Breast Cancer Surveillance system. Each BCSC registry and the Statistical Coordinating Center 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. Overall, diagnostic exams from 98 imaging facilities were included in our study.

Mammograms

This study included diagnostic digital mammograms (FFDM or DBT) that occurred up to 90 days after a screening mammogram and had an indication of “additional evaluation of a recent mammogram” (Figure 1). In addition, we included diagnostic mammograms with

a missing indication if they occurred up to 90 days after a screening mammogram that had a Breast Imaging and Reporting Data Systems [BI-RADS®] assessment of 0 (needs additional imaging evaluation), 3 (probably benign), 4 (suspicious for malignancy), or 5 (highly suggestive of malignancy).

The final diagnostic BI-RADS assessment was used to determine the diagnostic mammogram outcome. Diagnostic mammograms assigned a BI-RADS assessment of 4 or 5 were considered positive and mammograms with an assessment of 1 (negative), 2 (benign finding), or 3 were considered negative. If an assessment was BI-RADS 0 or missing, then the diagnostic mammogram was followed for 90 days to determine the final assessment using methods described previously (17,18). For the small number of mammograms (N=1487, <1%) where the final BI-RADS assessment during the 90-day period was 0, we imputed a positive/negative result based on age, facility, reader and cancer outcome (20).

Covariates

Data on demographics and health history were self-reported at each imaging visit 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. Whether women received a diagnostic ultrasound and/or magnetic resonance imaging (MRI) within 90 days after the additional evaluation diagnostic mammogram was determined based on registry imaging records. BI-RADS® breast density was classified by the interpreting radiologist as almost entirely fatty (a), scattered fibroglandular densities (b), heterogeneously dense (c), or extremely dense (d). Responses to health history questions were used to calculate the predicted risk of developing breast cancer during the next 5 years using the validated BCSC Risk Calculator (21–23). All information was reported at the time of the diagnostic mammogram, except time since a woman's last mammogram and breast density, which were reported at the screening mammogram immediately prior to the diagnostic mammogram.

Breast Cancer Ascertainment

Invasive breast carcinoma and ductal carcinoma *in situ* (DCIS) diagnoses occurring within one year of the diagnostic mammogram were ascertained through linkage with state or regional cancer registries; Surveillance, Epidemiology, and End Results registries; and pathology databases. Anatomic stage at diagnosis was classified using the AJCC 8th edition definitions (24). Tumor size, tumor grade, axillary lymph node status, and hormone receptor status were tabulated among invasive breast cancers only, due to the low frequency with which these characteristics were reported for DCIS. A combined estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 (HER2) variable was constructed to identify triple-negative breast cancers. Categorizations were chosen based on prior studies that have used immunohistochemistry markers to approximate breast cancer molecular subtypes (25,26).

Analytic Population

A detailed schematic of how the final study population was identified is in Figure 1. Briefly, we included diagnostic mammograms performed between January 2005 and June 2017 if: (1) the participant was ≥ 18 years old; (2) the diagnostic mammogram met our criteria for additional evaluation of a recent screening mammogram (described above); and (3) there was at least one year of follow-up time after the diagnostic mammogram. Diagnostic mammograms were excluded if: (1) the individual had a prior breast cancer diagnosis, mastectomy, or breast implants; (2) the individual had a diagnostic mammogram in the preceding 90 days; (3) the screening mammogram prior to the diagnostic mammogram was not a FFDM or DBT exam; or (4) the diagnostic mammogram had a final BI-RADS assessment of 6 (biopsy-proven cancer) or missing. Mammograms with race classified as American Indian/Alaska Native, other race, or mixed race were excluded due to the small sample size in each group. Examinations were also excluded if race/ethnicity was missing.

Statistical Analysis

Race and Hispanic ethnicity were combined into a single variable and mammogram and breast cancer characteristics were tabulated by racial/ethnic group. Means and standard deviations (SD) were calculated for continuous variables. Differences in categorical variable distributions were assessed using the chi-square test. Diagnostic mammography performance metrics were assessed as recommended by the American College of Radiology and are defined in Table 1 (27). Sensitivity, cancer detection rate, and false-negative rate were also calculated considering invasive breast cancers only, to explore the influence of DCIS diagnoses on the results. Wald 95% confidence intervals (CIs) were calculated for each statistic using previously described methods (17,27).

To explore the degree to which between-group differences in performance statistics may be influenced by personal or examination-related characteristics, we used logistic regression to model the probability of a given outcome with race/ethnicity as the independent variable and adjusting for additional factors. Regression models were constructed sequentially, first adjusting for mammography registry and age (continuous and age-squared), then additionally adjusting for family history of breast cancer, breast density, prior biopsy, type of diagnostic mammogram (2-D vs. 3-D), type of screening mammogram (2-D vs. 3-D), time since last mammogram, receipt of a diagnostic ultrasound and/or MRI. Imaging facility was entered into the model as a fixed effect. Models included all observations with non-missing data at each step, such that the number of observations in each sequential model varied.

Sensitivity Analysis

2.6% of the exams in this analysis were conducted among women who reported a breast problem (*e.g.*, pain, lump, nipple discharge, other problem not specified) at the time of their screening mammogram. We estimated performance statistics excluding these exams to determine whether their inclusion influenced the results. Additionally, we constructed logistic regression models using only the subset of observations with non-missing data for all covariates to determine whether variation in the sample used for each model influenced the observed racial/ethnic differences in performance.

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 (19) upon reasonable request. Restrictions apply to the availability of these data due to patient privacy requirements.

Results

Population Characteristics

A total of 267,868 digital diagnostic mammograms performed among 234,818 unique women were included in this study (Figure 1). The population included exams from individuals who identified as non-Hispanic White (70%), non-Hispanic Black (13%), Asian/Pacific Islander (10%) or Hispanic (7%), with a mean age at mammography of 55 years (SD: 11 years) (Table 2). For about half of the exams, ultrasound was used in addition to mammography to determine the BI-RADS assessment, with receipt highest among Hispanic women (55%) and lowest among non-Hispanic Black women (43%) (Table 2).

Diagnostic Mammography Assessments

The proportion of exams with a negative assessment was similar across racial/ethnic groups after adjustment for mammography registry (BI-RADS 1 or 2 assessment - 55–61%, Table 3). An assessment of BI-RADS 3, typically used to denote an abnormality that should be followed with additional imaging within 6 months, was most common among non-Hispanic Black women (31%) and least common among Asian/Pacific Islander women (17%). 48,017 (18%) mammograms were assigned a BI-RADS assessment that indicated a suspicion of cancer, the vast majority of which were BI-RADS 4 assessments.

Diagnostic Mammography Performance

Exam performance statistics varied by racial/ethnic group (Table 4). A diagnostic mammogram resulting in a recommendation for biopsy was most common among Asian/Pacific Islander women (22.4%, versus other groups range 16.9% – 17.5%). Biopsy receipt was slightly lower among Hispanic women compared with other women, but the median time between a positive mammogram and biopsy was similar across groups (Supplementary Table S1). Cancer yield among positive mammograms, as measured by PPV2 and PPV3, was highest among non-Hispanic White women (PPV2: non-Hispanic White 27.8% versus other groups range 19.4% – 24.3%; PPV3: non-Hispanic White 30.5% versus other groups range 20.9% – 25.9%). All three statistics were lowest among Hispanic women.

The cancer detection rate was highest among Asian/Pacific Islander women (54.3 per 1000) and lowest among Hispanic women (32.8 per 1000). When considering invasive cancers only, the cancer detection rate was highest among non-Hispanic White women (35.8 per 1000 mammograms) and lowest among Hispanic women (22.3 per 1000). Mammograms from Asian/Pacific Islander women also had the highest sensitivity (overall and invasive

cancers only)and highest false-positive rate. The false-negative rate (overall and invasive cancers only) was highest among non-Hispanic Black women.

A short interval follow-up recommendation was most common among non-Hispanic Black women (31.0%, versus other groups range 16.6% – 23.6%). However, among those with mammograms recommended for short interval follow-up, the short interval follow-up cancer yield was similar for non-Hispanic Black, non-Hispanic White, and Asian/Pacific Islander women and lowest among Hispanic women.

When we examined the impact of patient-level and facility-level characteristics on differences in performance statistics across racial/ethnic groups, we found that the greatest influences were receipt of ultrasound and/or MRI during the diagnostic work-up and imaging facility where women received their diagnostic mammogram (Figure 2 and Supplementary Table S2). Adjusting for ultrasound and/or MRI reduced the Black-White difference in cancer detection rates (model 7 odds ratio [OR] 0.84 (95% confidence interval [CI] 0.79 – 0.90] versus model 8 OR 0.98 (95% CI 0.91 – 1.05)). Adjusting for imaging facility attenuated Black-White differences in the short interval follow-up recommendation proportion (model 8 OR 1.39 [95% CI 1.35 – 1.43] versus model 9 OR 1.03 [95% CI 0.99 – 1.07]) and Hispanic-White differences in biopsy recommendation proportion (model 8 OR 0.90 [95% CI 0.86 – 0.94] versus model 9 OR 1.01 [95% CI 0.96 – 1.06]). Adjustment for other factors did little to explain racial/ethnic differences in mammography performance. The data were similar when regression models included only the subset of observations with non-missing data for all covariates (Supplementary Table S3) and when women who reported a breast problem at the time of screening were excluded (Supplementary Table S4).

Breast Cancer Characteristics

A total of 9027 invasive breast cancers and 3633 DCIS were detected by diagnostic mammography (Supplementary Table S5). Asian/Pacific Islander women had the highest proportion of DCIS diagnoses (43.0%, versus other groups range 27.6% – 33.5%). When considering only invasive breast cancers, non-Hispanic Black women had the highest proportions of advanced stage at diagnosis, high tumor grade, hormone receptor negative, triple-negative, and lymph node positive tumors when compared with other racial/ethnic groups. Non-Hispanic White and Asian/Pacific Islander women had more favorable tumor characteristic profiles, with a smaller mean invasive tumor size and similarly high proportions of lymph node negative and low-grade cancers. However, HER2+ tumors were more common among Asian/Pacific Islander women (14.5%, versus other groups range 10.4% – 12.1%). Hispanic women were more likely to have tumors >20 mm in size compared with other women (25.9%, versus other groups range 20.1% – 24.3%).

Discussion

The role that diagnostic mammography may play in contributing to racial and ethnic differences in breast cancer incidence and breast cancer survival is understudied. We evaluated diagnostic mammography performance among four racial/ethnic groups to clarify how differences in diagnostic work-up after a recent screening may influence those disparities. We found variation in diagnostic mammography performance by racial/

ethnic group, and in most cases the variation was not explained by women's individual characteristics. Imaging facility and concurrent use of ultrasound or MRI during the diagnostic process were the only two factors that explained differences in performance. While access to care has been identified as a barrier to breast cancer care in other settings, it is important to note that the women in this study had undergone breast cancer screening and were engaged with a breast imaging facility. These data suggest interventions that target the imaging facility and use of additional imaging modalities may have success in reducing some diagnostic disparities.

Our data suggest that non-Hispanic Black women may experience more potential harms related to diagnostic mammography than other groups. Exams conducted among non-Hispanic Black women had the lowest sensitivity and highest false-negative rate, indicating that this group may experience a delayed diagnosis more often. Others have shown that delayed diagnosis due to longer intervals between screening examinations results in higher proportions of advanced stage, high grade, and lymph node positive cancers among non-Hispanic Black women compared to other racial/ethnic groups (3). Consistent with that finding, the non-Hispanic Black women in our study had the highest proportion of tumors that were diagnosed at an advanced stage or were high grade or lymph node positive. Non-Hispanic Black women also had the highest proportion of recommendations for short interval follow-up. This assessment requires women to return for additional imaging in approximately 6 months. However, the short interval follow-up cancer yield for non-Hispanic Black women was not higher than any other group, suggesting that the additional financial costs, time, and stress associated with a short interval follow-up assessment were not balanced by benefits in additional cancers detected.

Our model-based analysis suggests that imaging facility characteristics may play a role in the higher frequency of short interval follow-up that was observed among non-Hispanic Black women. We did not have detailed information on the characteristics of each imaging facility in this study. However, prior studies have shown that Black women are less likely to obtain mammograms at facilities accredited by the National Consortium of Breast Centers or American College of Radiology's Breast Imaging Center of Excellence programs and therefore may receive a lower quality of care (28). Facility characteristics, including whether the facility was at an academic institution, a center of breast excellence, and had dedicated radiologists were stronger predictors of a recently screened woman presenting with breast cancer symptoms than socioeconomic factors such as income, education level, and insurance coverage (29). Structural factors have also been shown to have a negative association with breast cancer mortality, though this association is likely mediated by more than just the diagnostic process. Additional efforts are needed to identify the specific structural inequities that lead to disparities in diagnostic mammography performance. Enacting system-level interventions to address those inequities has the potential to reduce some of the disparities we observed in this study. However, because Black women experience delays at every step of the detection, diagnosis, treatment, and care pathway (30), larger coordinated efforts may be needed to see a meaningful impact on known mortality disparities.

Many performance statistics were lowest among Hispanic women. This group had the lowest biopsy recommendation proportion, PPVs, and cancer detection rate, indicating

that Hispanic women were least likely to receive a biopsy recommendation, and that if they did receive a biopsy recommendation, they were least likely to have breast cancer. Adjustment for imaging facility partially explained Hispanic-White differences in cancer detection rates, but detection rates remained significantly lower for the Hispanic group in the final multivariable model. Sensitivity for Hispanic women was comparable to that of non-Hispanic White women and higher than that of non-Hispanic Black women, suggesting that the low cancer detection rate and PPVs are not due to a failure to detect breast cancer when it is present. It is possible that there are simply fewer breast cancers to be detected among Hispanic women compared to the other women in this study. This hypothesis is supported by the fact that the Hispanic women were younger and had the lowest predicted breast cancer risk of the four groups. However, Hispanic women had the largest mean tumor size and the second highest proportion of advanced stage tumors. This result suggests a need for earlier breast cancer detection among Hispanic women, although any efforts to increase early detection must be balanced with the potential risks of over-detection and overdiagnosis.

In the BCSC, Asian/Pacific Islander women were the most likely to receive a recommendation for a biopsy following a diagnostic mammogram. While cancer was often diagnosed (Asian/Pacific Islander women had the highest cancer detection rate and sensitivity), other statistics suggest that the high frequency of biopsies may result in an unnecessary work-up. Asian/Pacific Islander women had the highest false-positive rate and the highest proportion of DCIS among cancers diagnosed. When considering invasive cancers only, Asian/Pacific Islander women no longer had the highest detection rate. The reasons for these patterns among Asian/Pacific Islander women are unclear. High rates of breast cancer screening do not explain the high proportion of DCIS diagnoses — data from a nationally representative survey show that Asian/Pacific Islander women do not participate in screening more often than non-Hispanic White or non-Hispanic Black women (31). Furthermore, all participants in our analysis received a screening mammogram prior to the diagnostic mammogram and thus had made the individual decision to seek preventive care.

Our analysis had limitations that may affect the interpretation of the results. Due to the time period covered by our study, DBT examinations were a small proportion of the mammograms studied. Thus, generalizability of these results to diagnostic performance at the growing number of imaging facilities that use DBT today may be limited. Data from studies of screening mammography suggest that DBT results in higher positive predictive values and higher cancer detection rates when compared with 2-D digital mammography alone (32–36). However, whether these differences are similar for diagnostic imaging or whether they vary by race or ethnic group remains to be seen. Diagnostic mammography indication was not reported in detail for some registries. For these records, we inferred the indication of additional evaluation of a recent mammogram based on the sequence and findings of prior exams. This approach may have failed to classify some women as having an additional evaluation if details of the screening exam findings were incomplete. However, only 3.7% of exams without a specific indication did not meet the criteria of being an additional evaluation, thus the number of potentially misclassified exams is low. We required that exams have a record of a prior screening mammogram, which may have excluded women who received their screening mammogram at an imaging facility not affiliated

with a BCSC registry. We lacked information about why some participants did not receive recommended diagnostic procedures. Additional research that directly elicits patient and provider feedback on why care was not provided is needed to fully understand the causes of the differences described in this study. We were not able to compare characteristics of breast cancers diagnosed after a false-negative mammogram due to the small numbers of cases, particularly among Asian/Pacific Islander and Hispanic women. We did not adjust for interpreting radiologist in our multivariable models. Future studies should examine the effect of the interpreting radiologist characteristics, which may be distinct from facility effects.

The strengths of this study include the analysis of data from more than 260,000 diagnostic mammograms, which enabled precise estimation of performance statistics, even for less populous racial/ethnic groups such as Asian/Pacific Islander and Hispanic women. Ninety-eight breast imaging facilities from across the US contributed data, increasing the generalizability of our results. Participant, radiologist, and examination data were collected using a standardized format, ensuring that the interpretation of each data element was comparable across registries. Data were collected prospectively, eliminating the risk of recall bias, and included patient-reported and radiologist-reported items, allowing for a comprehensive evaluation of the diagnostic work-up process. We included information on diagnostic breast ultrasound and MRI imaging that may have contributed to the final assessment, providing additional insight into the differential results for statistics related to short-term follow-up and cancer detection. Examination outcomes were based on information from multiple sources, including pathology databases and cancer registries, to increase the chances that all newly diagnosed breast cancers were ascertained.

In conclusion, diagnostic mammography performance varied among the four racial/ethnic groups in this study. Our findings indicate that factors associated with the imaging facility, rather than individual characteristics, may explain some of these differences, particularly since the study population consisted of women who had already received a screening mammogram. Performance did not fall below the minimal acceptable standards proposed by Carney et al. (37) for any group, but the differences suggest that approaches for maximizing early detection of cancer while limiting the burden of false-positive exams and overdiagnosis differ based on women’s race and ethnicity. This underscores the need to ensure that the studies that provide the evidence base for mammography recommendations include populations with sufficient racial and ethnic diversity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the participating individuals, mammography facilities, and radiologists for the data they have 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, D. L. Miglioretti, K. Kerlikowske), and R50CA211115 (E. J. A. Bowles). More information about the Breast Cancer Surveillance Consortium may be found at: http://www.bcsc-research.org/. Data collection was

additionally supported by funding from the Patient-Centered Outcomes Research Institute under award number PCS-1504-30370 (D. L. Miglioretti, K. Kerlikowske, L. M. Henderson, B. L. Sprague) and the Agency for Health Research and Quality under award number R01HS018366 (G. Rauscher). Cancer data collection was supported by several state public health departments and cancer registries. 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, the National Institutes of Health, or the Agency for Health Research and Quality.

Financial support: National Cancer Institute (P01CA154292, R01CA266377, U54CA163303, and R50CA211115), Patient-Centered Outcomes Research Institute (PCS-1504-30370), and the Agency for Health Research and Quality (R01 HS018366-01A1)

References

1. 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]
2. American Cancer Society. Cancer Facts & Figures for Hispanics/Latinos 2018–2020. Atlanta: American Cancer Society, Inc.; 2018.
3. Smith-Bindman R, Miglioretti DL, Lurie N, Abraham L, Barbash RB, Strzelczyk J, et al. Does utilization of screening mammography explain racial and ethnic differences in breast cancer? Ann Intern Med 2006;144:541–53. [PubMed: 16618951]
4. 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]
5. Blackman DJ, Masi CM. Racial and ethnic disparities in breast cancer mortality: are we doing enough to address the root causes? J Clin Oncol 2006;24:2170–8. [PubMed: 16682736]
6. Gerend MA, Pai M. Social determinants of Black-White disparities in breast cancer mortality: a review. Cancer Epidemiol Biomarkers Prev 2008;17:2913–23. [PubMed: 18990731]
7. Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA. Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. J Clin Oncol 2006;24:1342–9. [PubMed: 16549828]
8. Albain KS, Unger JM, Crowley JJ, Coltman CA, Jr., Hershman DL. Racial disparities in cancer survival among randomized clinical trials patients of the Southwest Oncology Group. J Natl Cancer Inst 2009;101:984–92. [PubMed: 19584328]
9. Parise CA, Caggiano V. Disparities in race/ethnicity and socioeconomic status: risk of mortality of breast cancer patients in the California Cancer Registry, 2000–2010. BMC Cancer 2013;13:449. [PubMed: 24083624]
10. Caplan LS, Helzlsouer KJ, Shapiro S, Wesley MN, Edwards BK. Reasons for delay in breast cancer diagnosis. Prev Med 1996;25:218–24. [PubMed: 8860288]
11. Chang SW, Kerlikowske K, Nápoles-Springer A, Posner SF, Sickles EA, Pérez-Stable EJ. Racial differences in timeliness of follow-up after abnormal screening mammography. Cancer 1996;78:1395–402. [PubMed: 8839544]
12. Jones BA, Dailey A, Calvocoressi L, Reams K, Kasl SV, Lee C, et al. Inadequate follow-up of abnormal screening mammograms: findings from the race differences in screening mammography process study (United States). Cancer Causes Control 2005;16:809–21. [PubMed: 16132791]
13. McCarthy BD, Yood MU, Boohaker EA, Ward RE, Rebner M, Johnson CC. Inadequate follow-up of abnormal mammograms. Am J Prev Med 1996;12:282–8. [PubMed: 8874693]
14. Rojas M, Mandelblatt J, Cagney K, Kerner J, Freeman H. Barriers to follow-up of abnormal screening mammograms among low-income minority women. Cancer Control Center of Harlem. Ethn Health 1996;1:221–8. [PubMed: 9395566]
15. Jones BA, Kasl SV, Curnen MG, Owens PH, Dubrow R. Can mammography screening explain the race difference in stage at diagnosis of breast cancer? Cancer 1995;75:2103–13. [PubMed: 7697601]
16. Yankaskas BC, Gill KS. Diagnostic mammography performance and race: outcomes in Black and White women. Cancer 2005;104:2671–81. [PubMed: 16288489]

17. 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]
18. 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]
19. Breast Cancer Surveillance Consortium (BCSC). Available from <https://www.bcsc-research.org/>.
20. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30:377–99. [PubMed: 21225900]
21. 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]
22. 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]
23. 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]
24. 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.
25. 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]
26. 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]
27. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology; 2013.
28. 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]
29. Mortel M, Rauscher GH, Murphy AM, Hoskins K, Warnecke RB. Racial and Ethnic Disparity in Symptomatic Breast Cancer Awareness despite a Recent Screen: The Role of Tumor Biology and Mammography Facility Characteristics. Cancer Epidemiol Biomarkers Prev 2015;24:1599–606. [PubMed: 26199340]
30. Elmore JG, Nakano CY, Linden HM, Reisch LM, Ayanian JZ, Larson EB. Racial inequities in the timing of breast cancer detection, diagnosis, and initiation of treatment. Medical care 2005;43:141–8. [PubMed: 15655427]
31. Sauer AG, Siegel RL, Jemal A, Fedewa SA. Updated Review of Prevalence of Major Risk Factors and Use of Screening Tests for Cancer in the United States. Cancer Epidemiol Biomarkers Prev 2017;26:1192–208. [PubMed: 28515109]
32. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol 2016;17:1105–13. [PubMed: 27345635]
33. Cohen EO, Weaver OO, Tso HH, Gerlach KE, Leung JWT. Breast Cancer Screening via Digital Mammography, Synthetic Mammography, and Tomosynthesis. Am J Prev Med 2020;58:470–72. [PubMed: 31732323]
34. Hofvind S, Holen ÅS, Aase HS, Houssami N, Sebuødegård S, Moger TA, et al. Two-view digital breast tomosynthesis versus digital mammography in a population-based breast cancer screening programme (To-Be): a randomised, controlled trial. Lancet Oncol 2019;20:795–805. [PubMed: 31078459]

35. Houssami N, Bernardi D, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer detection using single-reading of breast tomosynthesis (3D-mammography) compared to double-reading of 2D-mammography: Evidence from a population-based trial. Cancer Epidemiol 2017;47:94–99. [PubMed: 28192742]

36. Libesman S, Zackrisson S, Hofvind S, Seidler AL, Bernardi D, Lång K, et al. An individual participant data meta-analysis of breast cancer detection and recall rates for digital breast tomosynthesis versus digital mammography population screening. Clin Breast Cancer 2022.

37. Carney PA, Parikh J, Sickles EA, Feig SA, Monsees B, Bassett LW, et al. Diagnostic mammography: identifying minimally acceptable interpretive performance criteria. Radiology 2013;267:359–67. [PubMed: 23297329]

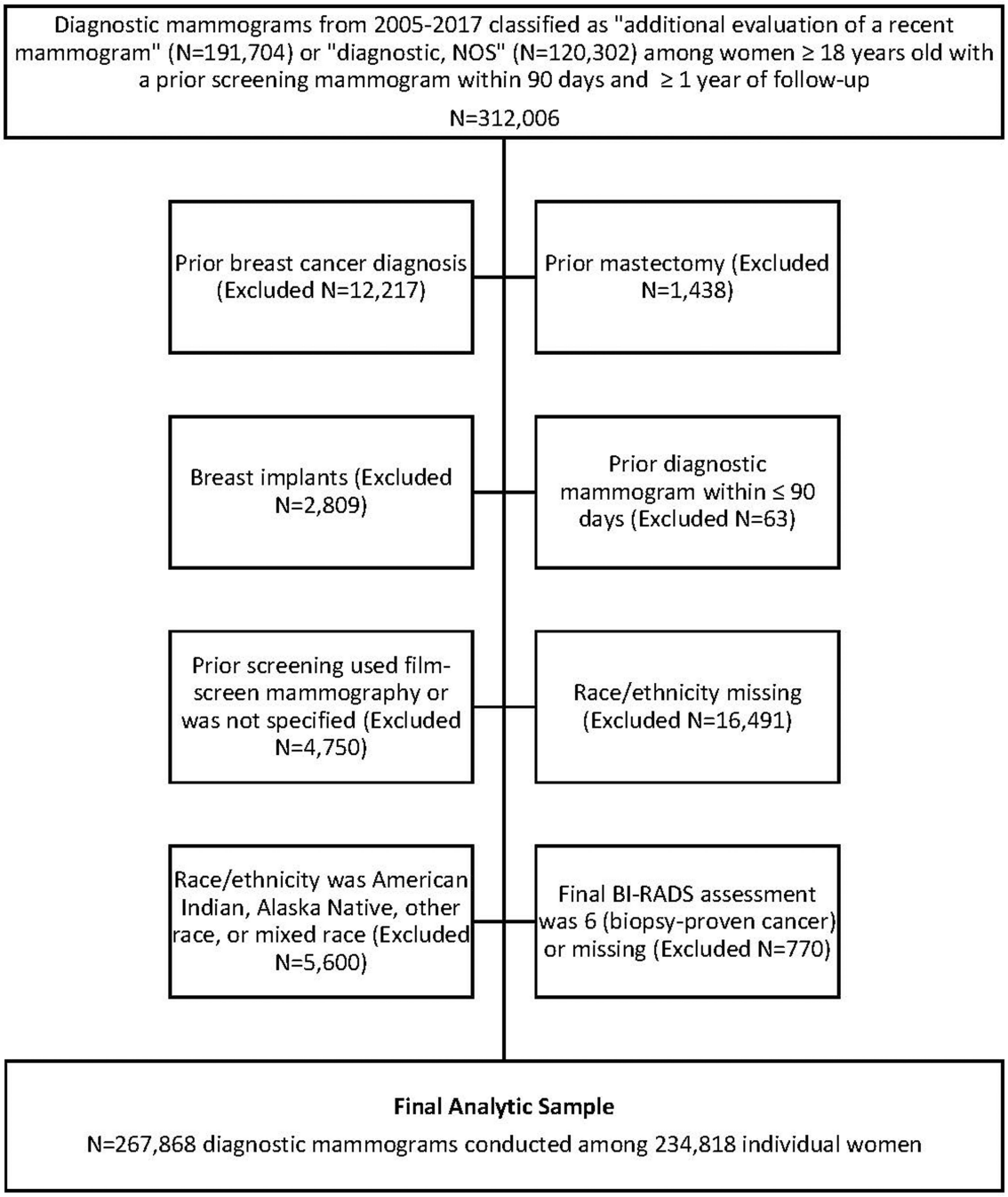

Figure 1. Diagnostic mammogram inclusion and exclusion criteria.

This study included digital diagnostic mammograms performed at an imaging facility affiliated with the Breast Cancer Surveillance Consortium (BCSC) between 2005 and 2017. All diagnostic mammograms were follow-ups to a digital screening mammogram that had occurred within the prior 90 days. NOS, not otherwise specified.

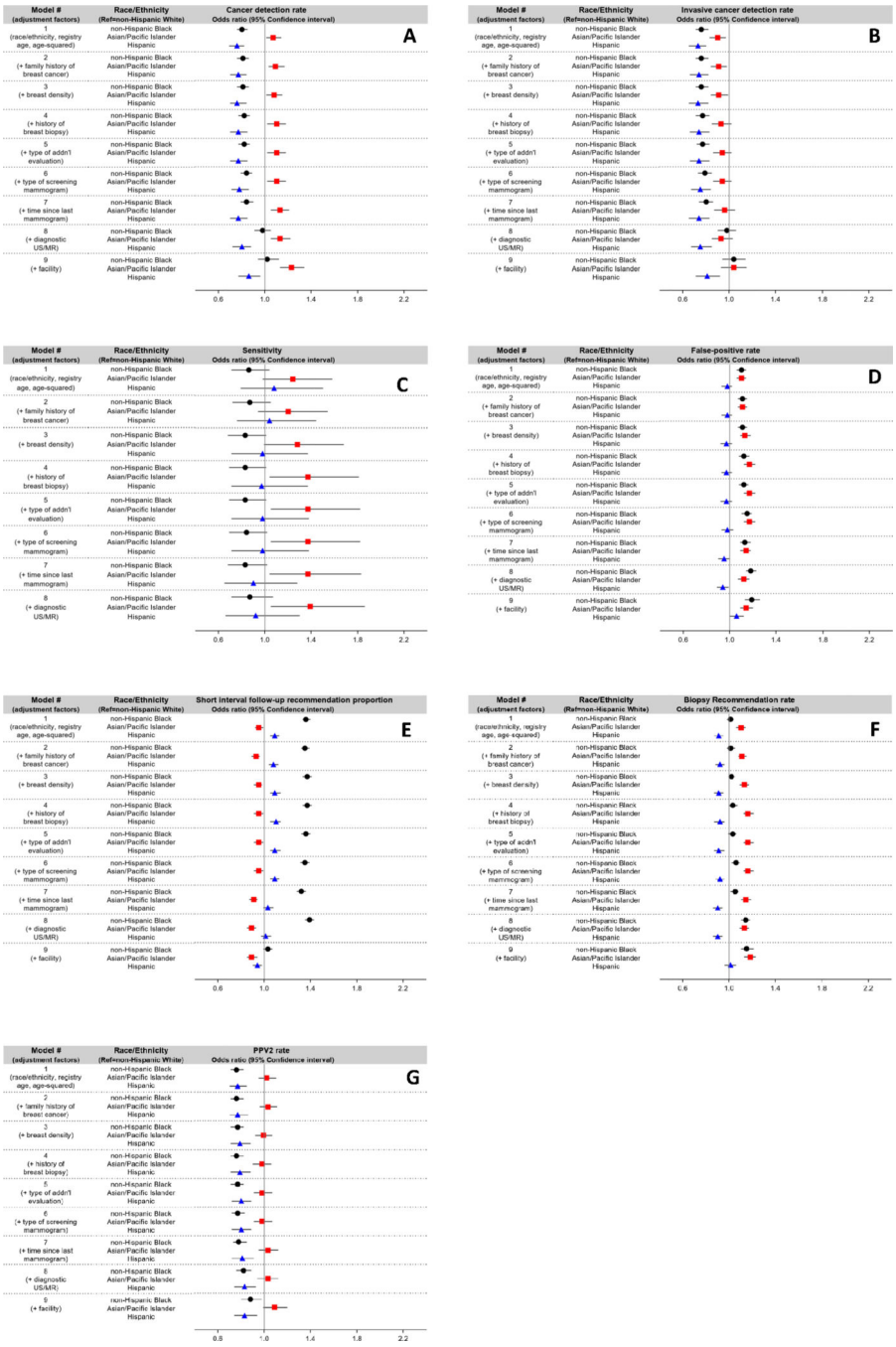

Figure 2. Adjusted diagnostic mammography performance statistics by race/ethnicity for 267,868 diagnostic mammograms performed among 234,818 women in the BCSC (2005–2017).

We compared cancer detection rate (panel A), invasive cancer detection rate (panel B), sensitivity (panel C), false-positive rate (panel D), short interval follow-up recommendation proportion (panel E), biopsy recommendation proportion (panel F), and PPV2 (panel G) among racial/ethnic groups while adjusting for patient characteristics to evaluate whether the characteristics might explain some of the differences present in the study population. Full definitions of each outcome are in Table 1. In each model, the performance statistic for non-Hispanic Black (circle), Asian/Pacific Islander (square), and Hispanic (triangle)

groups was compared with the performance statistic for White individuals (reference group). Adjustment was performed sequentially, where each model also adjusted for the factors listed in the prior model.

Table 1.

Definitions of diagnostic mammography performance statistics.

Statistic	Definition
Biopsy recommendation proportion	Proportion of mammograms that had a positive assessment (BI-RADS[a] 4 or 5)
Positive predictive value 2 (PPV2)	Proportion of mammograms recommended for biopsy or surgical consultation that were followed by a breast cancer diagnosis within 1 year
Positive predictive value 3 (PPV3)	Proportion of biopsies performed following a positive mammogram that were followed by a diagnosis of breast cancer within 1 year
Cancer detection rate	Number of breast cancers diagnosed following a positive mammogram assessment, per 1000 mammograms performed
Sensitivity	Number of breast cancers detected by mammography divided by the total number of breast cancers diagnosed
False-positive rate	Number of positive mammograms that were not followed by a breast cancer diagnosis within 1 year, per 1000 mammograms
False-negative rate	Number of negative mammograms that were followed by a breast cancer diagnosis within 1 year, per 1000 mammograms
Short interval follow-up recommendation proportion	Proportion of mammograms with a BI-RADS[a] assessment of 3
Short interval follow-up cancer yield	Proportion of mammograms recommended for short interval follow-up that were followed by a breast cancer diagnosis within 1 year

[a]BI-RADS – Breast Imaging Reporting and Data Systems

Table 2.

Personal characteristics associated with 267,868 diagnostic mammograms performed as an additional evaluation of a prior screening mammogram among 234,818 women in the Breast Cancer Surveillance Consortium, 2005–2017.

Characteristic	Overall (N=267,868)	Race and Ethnicity			
		Non-Hispanic White (N=188,268)	Non-Hispanic Black (N=34,573)	Asian/Pacific Islander (N=27,391)	Hispanic (N=17,636)
	N (%)	N (%)	N (%)	N (%)	N (%)
Age (years)					
< 40	7771 (2.9)	5305 (2.8)	1070 (3.1)	723 (2.6)	673 (3.8)
40–49	88,649 (33.1)	60,820 (32.3)	10,189 (29.5)	9688 (35.4)	7952 (45.1)
50–59	80,741 (30.1)	56,112 (29.8)	10,460 (30.3)	9195 (33.6)	4974 (28.2)
60–69	57,777 (21.6)	41,606 (22.1)	7821 (22.6)	5581 (20.4)	2769 (15.7)
70–79	25,984 (9.7)	19,044 (10.1)	4064 (11.8)	1830 (6.7)	1046 (5.9)
≥ 80	6946 (2.6)	5381 (2.9)	969 (2.8)	374 (1.4)	222 (1.3)
Mean (SD)	55.3 (11.3)	55.6 (11.4)	56.2 (11.5)	53.9 (10.2)	52.0 (10.4)
First-degree family history of breast cancer					
No	212,929 (82.8)	145,724 (81.0)	28,656 (83.7)	23,592 (90.1)	14,957 (88.2)
Yes	44,383 (17.2)	34,188 (19.0)	5600 (16.3)	2590 (9.9)	2005 (11.8)
Unknown	*10556*	*8356*	*317*	*1209*	*674*
History of breast biopsy or aspiration					
No	20,1632 (78.4)	137,152 (76.4)	27,709 (80.7)	22,241 (84.2)	14,530 (84.6)
Yes	55,679 (21.6)	42,266 (23.6)	6615 (19.3)	4163 (15.8)	2635 (15.4)
Unknown	*10557*	*8850*	*249*	*987*	*471*
Type of additional evaluation					
Digital (2-D)	257,930 (96.3)	180,026 (95.6)	33,914 (98.1)	26,785 (97.8)	17,205 (97.6)
DBT[a] (3-D)	9938 (3.7)	8242 (4.4)	659 (1.9)	606 (2.2)	431 (2.4)
Type of previous screening mammogram					
Digital (2-D)	255,336 (95.3)	177,527 (94.3)	34,081 (98.6)	26,650 (97.3)	17,078 (96.8)
DBT[a] (3-D)	12,532 (4.7)	10,741 (5.7)	492 (1.4)	741 (2.7)	558 (3.2)
Time since last mammogram (months)					
No prior mammogram	30,760 (12.4)	17,917 (10.2)	4600 (15.2)	4678 (18.4)	3565 (22.1)
< 12	6408 (2.6)	4531 (2.6)	810 (2.7)	691 (2.7)	376 (2.3)
12–23	148,535 (59.8)	110,318 (62.5)	17,327 (57.1)	13,121 (51.6)	7769 (48.2)
24–35	31270 (12.6)	21,952 (12.4)	3463 (11.4)	3729 (14.7)	2126 (13.2)
> 35	31,413 (12.6)	21,746 (12.3)	4152 (13.7)	3228 (12.7)	2287 (14.2)
Unknown	*19482*	*11804*	*4221*	*1944*	*1513*
Diagnostic US/MRI[b] within 90 days					

Characteristic	Overall (N=267,868)	Race and Ethnicity			
		Non-Hispanic White (N=188,268)	Non-Hispanic Black (N=34,573)	Asian/Pacific Islander (N=27,391)	Hispanic (N=17,636)
	N (%)	N (%)	N (%)	N (%)	N (%)
None	139,024 (51.9)	97,314 (51.7)	19,439 (56.2)	14,403 (52.6)	7868 (44.6)
US only	126,620 (47.3)	89,255 (47.4)	14,848 (42.9)	12,846 (46.9)	9671 (54.8)
MR (with or without US[C])	2224 (0.8)	1699 (0.9)	286 (0.8)	142 (0.5)	97 (0.6)
Breast density					
Almost entirely fat	14,986 (6.0)	11,064 (6.3)	2179 (6.5)	637 (2.6)	1106 (6.7)
Scattered fibroglandular densities	102,455 (40.9)	73,641 (41.7)	15,784 (47.3)	6645 (27.3)	6385 (38.8)
Heterogeneously dense	113,398 (45.2)	78,209 (44.3)	14,057 (42.1)	13,285 (54.6)	7847 (47.7)
Extremely dense	19,938 (8.0)	13,705 (7.8)	1354 (4.1)	3751 (15.4)	1128 (6.9)
Unknown	*17,091*	*11,649*	*1199*	*3073*	*1170*
BCSC 5-year predicted breast cancer risk (%)					
<1	84,929 (36.3)	50,655 (30.9)	11,324 (36.8)	12,561 (53.6)	10,389 (65.5)
1.00 – 1.66	85,279 (36.4)	59,987 (36.6)	11,894 (38.6)	9037 (38.6)	4361 (27.5)
1.67 – 2.49	43,090 (18.4)	34,877 (21.3)	5747 (18.7)	1537 (6.6)	929 (5.9)
2.50 – 3.99	18,157 (7.8)	16,044 (9.8)	1652 (5.4)	281 (1.2)	180 (1.1)
≥ 4.00	2657 (1.1)	2444 (1.5)	176 (0.6)	26 (0.1)	11 (0.1)
Unknown	*33,756*	*24261*	*3780*	*3949*	*1766*
Mean (SD)	1.4 (0.8)	1.5 (0.9)	1.3 (0.7)	1.0 (0.5)	0.9 (0.5)

[a]DBT – digital breast tomosynthesis

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

[c]379 exams with MRI and US

Table 3.

Comparison of final diagnostic mammography assessments for 267,868 additional evaluations of a recent screening mammogram by race/ethnicity, Breast Cancer Surveillance Consortium 2005–2017

BI-RADS assessment category[a,b]	Race and Ethnicity											
	Non-Hispanic White			Non-Hispanic Black			Asian/Pacific Islander			Hispanic		
	N	Raw %[c]	Adjusted %[d]	N	Raw %[c]	Adjusted %[d]	N	Raw %[c]	Adjusted %[d]	N	Raw %[c]	Adjusted %[d]
Negative (1)	48563	26.0	25.1	5093	14.8	22.0	10034	36.8	29.5	3930	22.3	23.9
Benign (2)	63897	34.2	34.8	12607	36.6	33.4	6612	24.2	29.5	6524	37.1	37.1
Probably benign (3)	41681	22.3	22.1	10733	31.1	25.5	4535	16.6	20.8	4155	23.6	23.0
Suspicious (4)	30792	16.5	16.9	5763	16.7	18.3	5795	21.2	19.3	2837	16.1	15.2
Highly suggestive of malignancy (5)	2112	1.1	1.1	265	0.8	0.8	315	1.2	0.9	138	0.8	0.8

[a] BI-RADS, Breast Imaging Reporting and Data Systems

[b] 1,487 mammograms with a final assessment of BI-RADS 0 are not shown

[c] Column percentage, unadjusted

[d] Column percentage, adjusted for mammography registry using direct standardization

Table 4.

Comparison of mammography performance for 267,868 diagnostic mammograms performed for additional evaluation of a recent screening mammogram among 234,818 women in the Breast Cancer Surveillance Consortium by race and ethnicity, 2005–2017.

Performance Statistic	Non-Hispanic White	Non-Hispanic Black	Asian/Pacific Islander	Hispanic
Biopsy				
Biopsy recommendation proportion, % (95% CI[a])	17.5 (17.4, 17.7)	17.5 (17.1, 17.9)	22.4 (21.9, 22.8)	16.9 (16.3, 17.5)
# positive[b] mammograms	33,003	6034	6122	2979
Total # mammograms	188,268	34,573	27,391	17,636
PPV2[c], % (95% CI[a])	27.8 (27.3, 28.3)	23.4 (22.3, 24.5)	24.3 (23.2, 25.4)	19.4 (18.0, 20.9)
# true-positive mammograms	9182	1412	1488	579
# positive[b] mammograms	33,003	6034	6122	2979
PPV3[c], % (95% CI[a])	30.5 (30.0, 31.1)	25.9 (24.8, 27.2)	25.5 (24.3, 26.6)	20.9 (19.4, 22.5)
# true-positive mammograms with biopsies performed	8811	1374	1376	567
# positive[b] mammograms with biopsies performed	28,873	5295	5405	2710
Cancer Detection				
Cancer detection rate, per 1000 mammograms (95% CI[a])	48.8 (47.8, 49.8)	40.8 (38.8, 43.0)	54.3 (51.7, 57.1)	32.8 (30.2, 35.6)
# true-positive mammograms	9182	1412	1488	579
Total # mammograms	188,268	34,573	27,391	17,636
Invasive cancer detection rate, per 1000 mammograms (95% CI[a])	35.8 (35.0, 36.7)	29.5 (27.8, 31.4)	31.6 (29.6, 33.8)	22.3 (20.2, 24.6)
# true-positive (invasive) mammograms	6746	1021	866	394
Total # mammograms	188,268	34,573	27,391	17,636
Sensitivity, % (95% CI[a])	92.3 (91.8, 92.9)	89.8 (88.2, 91.3)	94.2 (93.0, 95.3)	92.8 (90.5, 94.7)
# true-positive mammograms	9182	1412	1488	579
# mammograms followed by breast cancer	9943	1572	1579	624

Performance Statistic	Non-Hispanic White	Non-Hispanic Black	Asian/Pacific Islander	Hispanic
Sensitivity for invasive cancer, % (95% CI[a])	92.4 (91.7, 93.0)	89.7 (87.8, 91.4)	94.4 (92.8, 95.8)	92.9 (90.1, 95.2)
# true-positive (invasive) mammograms	6746	1021	866	394
# mammograms followed by invasive breast cancer	7304	1138	917	424
False-positive rate, per 1000 mammograms (95% CI[a])	126.5 (125.0, 128.0)	133.7 (130.1, 137.3)	169.2 (164.8, 173.7)	136.1 (131.1, 141.2)
# false-positive mammograms	23,821	4622	4634	2400
Total # mammograms	188,268	34,573	27,391	17,636
False-negative rate, per 1000 mammograms (95% CI[a])	4.0 (3.8, 4.3)	4.6 (3.9, 5.4)	3.3 (2.7, 4.1)	2.6 (1.9, 3.4)
# false-negative mammograms	761	160	91	45
Total # mammograms	188,268	34,573	27,391	17,636
False-negative rate for invasive cancer, per 1000 mammograms (95% CI[a])	3.0 (2.7, 3.2)	3.4 (2.8, 4.1)	1.9 (1.4, 2.4)	1.7 (1.1, 2.4)
# false-negative mammograms (invasive)	558	117	51	30
Total # mammograms	188,268	34,573	27,391	17,636
Short Interval Follow-Up				
Short interval follow-up recommendation proportion, % (95% CI[a])	22.1 (22.0, 22.3)	31.0 (30.6, 31.5)	16.6 (16.1, 17.0)	23.6 (22.9, 24.2)
# mammograms recommended for short interval follow-up	41,681	10,733	4535	4155
Total # mammograms	188,268	34,573	27,391	17,636
Short interval follow-up cancer yield, % (95% CI[a])	1.3 (1.2, 1.4)	1.1 (0.9, 1.3)	1.4 (1.1, 1.8)	0.8 (0.6, 1.1)
# short-interval follow-up mammograms with breast cancer during follow-up	537	113	63	34
# mammograms recommended for short interval follow-up	41,681	10,733	4535	4155

[a] CI – confidence interval

[b] Mammograms with a final BI-RADS assessment of 4 and 5 were classified as positive. Mammograms with a final assessment of 1, 2, or 3 were classified as negative. Imputed positive/negative results for BI-RADS 0 mammograms were used in calculations.

[c] PPV – positive predictive value