Radiology

ORIGINAL RESEARCH • BREAST IMAGING

Mammographic Screening in Routine Practice: Multisite Study of Digital Breast Tomosynthesis and Digital Mammography Screenings

Emily F. Conant, MD • Melinda M. Talley, MD • Chirag R. Parghi, MD, MBA • Bryant C. Sheh, MD • Su-Ying Liang, PhD • Scott Pohlman, MS • Amey Rane, MS, BPharm • Yoojin Jung, PhD • Lauren A. S. Stevens, MPH • Jessica K. Paulus, ScD • Nila Alsheik, MD

From the Department of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104 (E.F.C.); Sanford Health, Sioux Falls, SD (M.M.T.); Solis Mammography, Houston, Tex (C.R.P.); Sutter Health, Fremont, Calif (B.C.S.); Sutter Health, Palo Alto, Calif (S.Y.L.); Hologic, Marlborough, Mass (S.P., A.R.); OM1, Boston, Mass (Y.J., L.A.S.S., J.K.P.); and Department of Radiology, Advocate Caldwell Breast Center, Park Ridge, Ill (N.A.). Received July 1, 2022; revision requested September 8; revision received January 6, 2023; accepted January 31. **Address correspondence to** E.F.C. (email: *Emily.Conant@pennmedicine.upenn.edu).*

Conflicts of interest are listed at the end of this article.

See also the editorial by Bae and Seo in this issue.

Radiology 2023; 307(3):e221571 • https://doi.org/10.1148/radiol.221571 • Content code: BR

Background: The use of digital breast tomosynthesis (DBT) is increasing over digital mammography (DM) following studies demonstrating lower recall rates (RRs) and higher cancer detection rates (CDRs). However, inconsistent interpretation of evidence on the risks and benefits of mammography has resulted in varying screening mammography recommendations.

Purpose: To evaluate screening outcomes among women in the United States who underwent routine DM or DBT mammographic screening.

Materials and Methods: This retrospective cohort study included women aged 40–79 years who underwent DM or DBT screening mammograms between January 2014 and December 2020. Outcomes of RR, CDR, positive predictive value of recall (PPV1), biopsy rate, and positive predictive value of biopsy (PPV3) were compared between DM and DBT with use of adjusted multivariable logistic regression models.

Results: A total of 2 528 063 screening mammograms from 1 100 447 women (mean age, 57 years ± 10 [SD]) were included. In crude analyses, DBT (1 693 727 screening mammograms vs 834 336 DM screening mammograms) demonstrated lower RR (10.3% [95% CI: 10.3, 10.4] for DM vs 8.9% [95% CI: 8.9, 9.0] for DBT; $P < .001$) and higher CDR (4.5 of 1000 screening mammograms [95% CI: 4.3, 4.6] vs 5.3 of 1000 [95% CI: 5.2, 5.5]; $P < .001$), PPV1 (4.3% [95% CI: 4.2, 4.5] vs 5.9% [95% CI: 5.7, 6.0]; $P < .001$), and biopsy rates (14.5 of 1000 screening mammograms [95% CI: 14.2, 14.7] vs 17.6 of 1000 [95% CI: 17.4, 17.8]; $P < .001$). PPV3 was similar between cohorts (30.0% [95% CI: 29.2, 30.9] for DM vs 29.3% [95% CI: 28.7, 29.9] for DBT; $P = .16$). After adjustment for age, breast density, site, and index year, associations remained stable with respect to statistical significance.

Conclusion: Women undergoing digital breast tomosynthesis had improved screening mammography outcomes compared with women who underwent digital mammography.

Supplemental material is available for this article.

Breast cancer is the most common cancer among women in the United States, accounting for 30% of new cancer cases in 2021 (1). While incidence rates continue to increase by approximately 0.5% per year, breast cancer mortality rates have declined by 41% since 1989 due to improvements in early detection and treatment. Despite this reduction in mortality, breast cancer remains the leading cause of cancer death among women aged 20–59 years, and almost 44 000 women were estimated to die of breast cancer in 2021 (1).

Mammography is the standard of care for the early detection of breast cancer and reduces mortality from breast cancer (2). The reconstructed quasi–three-dimensional data acquired with digital breast tomosynthesis (DBT) improves lesion conspicuity, allowing improved detection, characterization, and localization of lesions (3,4). Screening mammography using DBT is replacing digital mammography (DM) as the preferred imaging modality following studies demonstrating lower recall rates (RRs) and higher cancer detection rates (CDRs) (5–15).

Mammography, like all screening methods, has limitations, including failure to detect some cancers (false-negative results), detection of cancers that may never cause harm (overdiagnosis), and detection of abnormalities that turn out to be benign (false-positive results). Inconsistent interpretation of evidence for the risks and benefits of routine mammographic screening in different patient subgroups, including younger and older women, has led to variation in recommendations for mammographic screening from the United States Preventive Services Task Force, the American Cancer Society, and the American College of Radiology (16–18). Importantly, variation in mammography screening outcomes within subgroups defined by key patient characteristics (eg, breast density, age) may also depend on the imaging modality used. For example,

Abbreviations

BI-RADS = Breast Imaging Reporting and Data System, CDR = cancer detection rate, DBT = digital breast tomosynthesis, DM = digital mammography, OR = odds ratio, PPV1 = positive predictive value of recall, PPV3 = positive predictive value of biopsy, RR = recall rate

Summary

Compared with digital mammography, digital breast tomosynthesis showed lower recall and higher positive predictive value of recall, cancer detection rate, and biopsy rate, but similar positive predictive value of biopsy.

Key Results

- A retrospective study of 2 528 063 screening mammograms demonstrated improved outcomes with use of digital breast tomosynthesis (DBT) over digital mammography (DM).
- Compared with DM, DBT had a lower recall rate (10.3% vs 8.9%; $P < .001$) and higher positive predictive value of recall (4.3% vs 5.9%; $P < .001$) and cancer detection rate (4.5 of 1000 screening mammograms vs 5.3 of 1000; $P < .001$).
- The biopsy rate was higher with DBT versus DM (17.6 of 1000 screening mammograms vs 14.5 of 1000; $P < .001$); however, the positive predictive value of biopsy did not differ (29.3% vs 30.0%; $P = .16$).

mammography screening with DBT has been found to have superior cancer detection compared with DM, particularly for younger women (19) and those with dense breasts (8,20).

To date, many studies evaluating the potential benefits DBT over DM have been conducted with a focus on specific risk factors, such as breast density, or on the detection of interval or advanced cancers (21–24). The purpose of our study was to evaluate screening outcomes among a large cohort of women in the United States who underwent routine DM or DBT mammographic screening.

Materials and Methods

Study Design

This retrospective study was conducted in compliance with the Health Insurance Portability and Accountability Act and approved by a central institutional review board with a waiver of consent to use a database containing standardized and integrated electronic medical record, radiology information system, and tumor registry data from five large health care systems (University of Pennsylvania, Sanford Health, Advocate Health Care, Sutter Health, and Solis Mammography). The inclusion criteria included all screening mammograms performed among women aged 40–79 years who underwent at least one screening mammogram from January 2014 through December 2020. The current study expands on previously published research that evaluated key screening outcomes among 385 503 women with known race data from three health care systems (11) to identify potential disparities by screening modality (DM vs DBT).

In the current study, each screening mammogram date defined an index date. Women with multiple screening mammograms contributed multiple index dates to the DM and DBT groups. Women with breast cancer on or before each index date were excluded (Fig 1). Screening intervals were calculated as the time between the index and the prior screening mammogram. When no prior screening mammogram was recorded, the screening interval was defined as "unknown," indicating that women may have had earlier screening mammograms not captured because the examination was performed at another institution and/or occurred outside the study period. Therefore, the presence of only a single mammogram in the data set did not equate to a baseline screening mammogram.

Mammogram Screening Outcomes

The primary outcomes were RR, CDR, positive predictive value of recall (PPV1), biopsy rate, and positive predictive value of biopsy (PPV3). A recall was defined as an index screening examination with an initial Breast Imaging Reporting and Data System (BI-RADS) score of 0 (incomplete test, need for additional imaging), 4 (suspicious findings or abnormalities), or 5 (highly suspicious findings), requiring additional follow-up. The RR (per 1000 screening mammograms) was calculated as the proportion of screening mammograms resulting in recalls. The CDR (per 1000 screening mammograms) was calculated for women with at least 6 months of follow-up as the number of screening-detected cancers identified within 6 months of index (Fig S1) divided

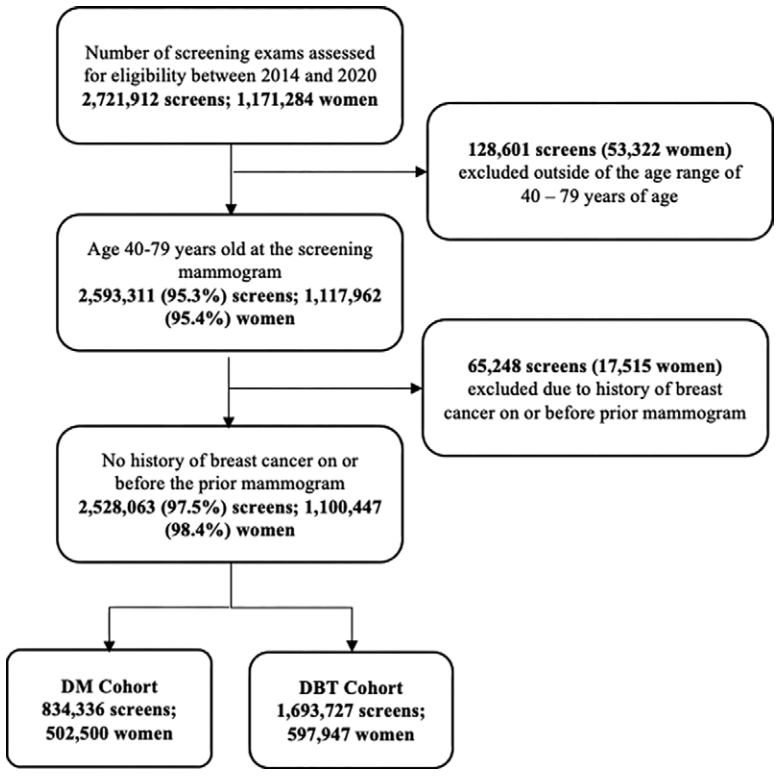

Figure 1: Study cohort creation. Flowchart outlines each inclusion and exclusion criterion required to be met for inclusion in the analysis. Women aged between 40–79 years at the time screening and without a history of breast cancer on or before the prior mammogram were eligible for inclusion.

by the total number of screening mammograms. Because of variations in tumor registry reporting by health system, registry completeness was assessed by calculating the ratio of reported cancers to conducted screening mammograms per month across the observation period and deemed incomplete after a 2-month period where the ratio dropped by more than 30% compared with the prior month. Only screening mammograms conducted at least 6 months prior to the 2-month 30% drop were included. PPV1 was defined as the proportion of women diagnosed with screening-detected breast cancer within 6 months of recall.

The biopsy rate was defined as the number of biopsies performed after index mammogram divided by the number of valid screening mammograms, regardless of screening BI-RADS score, presented per 1000 screening mammograms. It was calculated for women with at least 3 months of postindex follow-up as the number of biopsies performed within 3 months of an index screening examination with an initial BI-RADS category of 0, 4, or 5 divided by the number of screening mammograms. Only biopsies conducted at least 3 months before the 2-month 30% drop in the cancer-to-screening ratio were included. PPV3 was calculated as the proportion of women diagnosed with breast cancer within 6 months of an index screening examination resulting in biopsy within 90 days (with an initial BI-RADS score of 0, 4, or 5).

Statistical Analysis

Analyses were performed at the screening mammogram level. Patient characteristics were described at each index screening examination, overall and by screening modality, including risk factors of age, race, breast density, short-term risk status, index year (not reported), number of screening mammograms before and including the index screen (one vs two or more), and supplemental screening. Supplemental screening mammograms were defined as bilateral complete US or MRI examinations occurring between 2 days and 9 months after a screening mammogram (with a final BI-RADS score of 1 or 2 and no additional imaging between index and supplemental screening examinations). Data on race and breast density were collected in accordance with local procedure and used to characterize the cohort more fully. Breast density was defined as almost entirely fatty, scattered fibroglandular densities, heterogeneously dense, extremely dense, or unknown. Short-term risk status was determined with use of the Gail model 5-year risk score, where elevated risk was defined as a score of 1.66 or higher. As the outcomes of biopsy and CDR required 3 or 6 months of follow-up, respectively, characteristics of women meeting these requirements were described.

Outcomes were summarized descriptively overall, by modality, and by select patient characteristics. Logistic regression models were used to estimate unadjusted and adjusted odds ratios (ORs) and 95% CIs for each outcome. Adjusted multivariable models included select risk factors based on a priori subject matter expertise: age category, breast density (with indicator for "unknown"), system, and index year. Statistical analysis was performed by using SAS software, version 9.4 (SAS Institute) with two-tailed tests and an alpha of .05.

Results

Patient Characteristics

A total of 2 528 063 screening mammograms (DM, 834 336; DBT, 1 693 727) among 1 100 447 women (DM, 502 500; DBT, 597 947) were included after the exclusion of 128 601 screening mammograms in 53 322 women outside of the required age range of 40–79 years and 65 248 screening mammograms in 17 515 women with a history of breast cancer. The mean age at screening was 57 ± 10 years (SD) for women undergoing DM and 57 ± 10 years for those undergoing DBT, with 73% of the overall cohort being aged 50 years or older (Table 1). Most screening mammograms were in women who had undergone at least two screening mammograms (DM, 82.4%; DBT, 81.7%). US for supplemental screening was more common with DBT (3.1%) than DM (1.4%). A total of 23.6% of mammograms were in women determined as having an elevated risk status (DM, 23.6%; DBT, 23.5%). Where race was known, 68.1% of mammograms in the DM group and 77.6% of those in the DBT group were in White women. The most common breast tissue density (when known) was scattered fibroglandular densities (DM, 46.4%; DBT, 48.3%). Patient characteristics among those with at least 3 or 6 months of follow-up data, required for biopsy rate and CDR outcomes, were comparable with those of the overall cohort (Tables S1, S2).

Recall Rate

We observed an RR of 9.4% (95% CI: 9.4, 9.4) among 2 528 063 screening mammograms (Table 2; Figures 2, 3). The RR among examinations in women with and without prior screening mammograms was 7.9% (95% CI: 7.9, 8.0) and 16.1% (95% CI: 16.0, 16.2), respectively. Examinations in women with heterogeneously or extremely dense breast tissue had RRs of 11.4% (95% CI: 11.3, 11.4) and 10.0% (95% CI: 9.8, 10.1), respectively. Examinations in women with almost entirely fatty breasts had an RR of 5.6% (95% CI: 5.5, 5.7), and those with scattered fibroglandular densities had an RR of 8.4% (95% CI: 8.3, 8.4). RR was lower with older patient age, with 13.6% (95% CI: 13.5, 13.7) among women aged 40–44 years and 6.9% (95% CI: 6.8, 7.1) among those aged 75–79 years.

We observed a lower crude RR in the DBT group (8.9% [95% CI: 8.9, 9.0]) than in the DM group (10.3% [95% CI: 10.3, 10.4]; $P < .001$), a trend seen across demographic and clinical subgroups, with the exception of those with extremely dense breast tissue, where the crude RR was higher in the DBT group (10.1% [95% CI: 9.9, 10.3]) than in the DM group (9.7% [95% CI: 9.5, 10.0]). After adjustment for potential confounders, DBT was associated with lower RR (OR, 0.92 [95% CI: 0.91, 0.93]; $P < .001$) (Fig 4).

Cancer Detection Rate

A total of 9714 cancers were detected (DM, 3421; DBT, 6293) among 1 948 098 screening mammograms with at least 6 months of postindex follow-up (DM, 766 587; DBT, 1 181 511) (Table 3). The overall CDR was 5.0 of 1000 screening mammograms, which increased with patient age

Table 1: Characteristics of the Cohort, Overall and by Index Screening Modality (DM vs DBT)

Characteristic	Overall (*n* = 2 528 063)	DM (*n* = 834 336)	DBT (*n* = 1 693 727)
Age (y)			
Mean and SD	57 ± 10	57 ± 10	57 ± 10
Median and IQR	57 (49–65)	57 (49–65)	57 (49–66)
Age category			
40–44 years	325 598 (12.9)	103 585 (12.4)	222 013 (13.1)
45–49 years	358 883 (14.2)	118 725 (14.2)	240 158 (14.2)
50–54 years	384 229 (15.2)	132 935 (15.9)	251 294 (14.8)
55–59 years	399 564 (15.8)	136 059 (16.3)	263 505 (15.6)
60–64 years	366 384 (14.5)	124 507 (14.9)	241 877 (14.3)
65–69 years	322 822 (12.8)	99 861 (12.0)	222 961 (13.2)
70–74 years	235 571 (9.3)	74 075 (8.9)	161 496 (9.5)
75–79 years	135 012 (5.3)	44 589 (5.3)	90 423 (5.3)
Race			
Asian	160 132 (7.7)	72 636 (10.5)	87 496 (6.3)
Black	286 884 (13.8)	121 765 (17.7)	165 119 (11.8)
Other*	85 640 (4.1)	25 357 (3.7)	60 283 (4.3)
White	1 552 653 (74.5)	468 768 (68.1)	1 083 885 (77.6)
Unknown	442 754	145 810	296 944
Ethnicity			
Hispanic	146 787 (8.3)	48 332 (8.9)	98 455 (8.0)
Not Hispanic	1 630 845 (91.7)	496 003 (91.1)	1 134 842 (92.0)
Unknown	750 431	290 001	460 430
Menopause status			
Postmenopause	1 506 351 (80.8)	492 556 (80.1)	1 013 795 (81.1)
Premenopause	358 448 (19.2)	122 538 (19.9)	235 910 (18.9)
Unknown	663 264	219 242	444 022
Elevated risk status	595 721 (23.6)	197 106 (23.6)	398 615 (23.5)
Breast density			
Almost entirely fatty (A)	215 131 (8.7)	76 360 (9.3)	138 771 (8.3)
Scattered fibroglandular densities (B)	1 184 485 (47.7)	380 795 (46.4)	803 690 (48.3)
Heterogeneously dense (C)	934 293 (37.6)	308 265 (37.6)	626 028 (37.6)
Extremely dense (D)	149 428 (6.0)	55 057 (6.7)	94 371 (5.7)
Unknown	44 726	13 859	30 867
Initial screening BI-RADS category			
0	236 764 (9.4)	85 958 (10.3)	150 806 (8.9)
1	1 377 085 (54.5)	410 436 (49.2)	966 649 (57.1)
2	912 367 (36.1)	337 372 (40.4)	574 995 (33.9)
3	1263 (0.0)	377 (0.0)	886 (0.1)
4	512 (0.0)	160 (0.0)	352 (0.0)
5	72 (0.0)	33 (0.0)	39 (0.0)
Supplemental screening			
US	63 827 (2.5)	11 886 (1.4)	51 941 (3.1)
MRI	8492 (0.3)	2502 (0.3)	5990 (0.4)
None	2 455 744 (97.1)	819 948 (98.3)	1 635 796 (96.6)
At least 2 screening examinations	2 071 741 (81.9)	687 759 (82.4)	1 383 982 (81.7)
Screening interval			
≤24 months	1 445 925 (84.5)	412 238 (82.2)	1 033 687 (85.5)
>24 months	265 316 (15.5)	89 557 (17.8)	175 759 (14.5)
Unknown	816 822	332 541	484 281

Note.—Unless otherwise specified, data are numbers of screening mammograms, with percentages in parentheses. All percentages are calculated based on the number of screening mammograms in patients with nonmissing data. BI-RADS = Breast Imaging Reporting and Data System, DBT = digital breast tomosynthesis, DM = digital mammography.

* "Other" includes individuals who identified as American Indian or Alaska Native, multiracial, Native Hawaiian, other Pacific Islander, or other categories not specified within the data source.

Table 2: Recall Rates, Overall and by Select Patient Characteristics

Characteristic	Overall			
	No. of Screening Mammograms	Recall Rate	Recall Rate for DM	Recall Rate for DBT
All screening mammograms	2 528 063	9.4 (9.4, 9.4)	10.3 (10.3, 10.4)	8.9 (8.9, 9.0)
Screening mammogram count				
1	456 322	16.1 (16.0, 16.2)	17.8 (17.6, 18.0)	15.3 (15.2, 15.4)
2 or more	2 071 741	7.9 (7.9, 8.0)	8.7 (8.7, 8.8)	7.5 (7.5, 7.5)
Screening interval				
≤24 months	1 445 925	7.0 (6.9, 7.0)	7.6 (7.5, 7.6)	6.7 (6.7, 6.8)
>24 months	265 316	10.2 (10.1, 10.3)	11.1 (10.9, 11.3)	9.7 (9.6, 9.8)
Unknown	816 822	13.4 (13.3, 13.5)	13.5 (13.4, 13.7)	13.3 (13.2, 13.4)
Age category				
40–44 years	325 598	13.6 (13.5, 13.7)	15.4 (15.1, 15.6)	12.8 (12.6, 12.9)
45–49 years	358 883	11.6 (11.4, 11.7)	12.8 (12.6, 13.0)	10.9 (10.8, 11.1)
50–54 years	384 229	9.9 (9.8, 10.0)	10.6 (10.4, 10.8)	9.5 (9.4, 9.6)
55–59 years	399 564	8.3 (8.2, 8.4)	9.0 (8.8, 9.1)	8.0 (7.9, 8.1)
60–64 years	366 384	7.9 (7.8, 8.0)	8.8 (8.7, 9.0)	7.4 (7.3, 7.5)
65–69 years	322 822	7.7 (7.6, 7.8)	8.6 (8.4, 8.8)	7.3 (7.2, 7.5)
70–74 years	235 571	7.2 (7.1, 7.3)	7.9 (7.7, 8.1)	6.9 (6.8, 7.0)
75–79 years	135 012	6.9 (6.8, 7.1)	7.4 (7.2, 7.6)	6.7 (6.5, 6.9)
Race				
Asian	160 132	10.2 (10.0, 10.3)	10.2 (9.9, 10.4)	10.2 (10.0, 10.4)
Black	286 884	9.7 (9.6, 9.8)	11.2 (11.0, 11.4)	8.5 (8.4, 8.7)
Other*	85 640	10.3 (10.1, 10.5)	11.1 (10.7, 11.5)	10.0 (9.8, 10.2)
White	1 552 653	8.9 (8.9, 9.0)	9.6 (9.5, 9.7)	8.7 (8.6, 8.7)
Unknown	442 754	10.3 (10.2, 10.4)	11.9 (11.7, 12.0)	9.5 (9.4, 9.6)
Short-term risk status†				
Normal risk	1 932 342	9.9 (9.9, 10.0)	11.0 (10.9, 11.0)	9.4 (9.4, 9.5)
Elevated risk	595 721	7.6 (7.6, 7.7)	8.2 (8.1, 8.4)	7.4 (7.3, 7.4)
Breast density				
Almost entirely fatty (A)	215 131	5.6 (5.5, 5.7)	6.3 (6.2, 6.5)	5.2 (5.1, 5.3)
Scattered fibroglandular densities (B)	1 184 485	8.4 (8.3, 8.4)	9.4 (9.3, 9.5)	7.9 (7.8, 8.0)
Heterogeneously dense (C)	934 293	11.4 (11.3, 11.4)	12.5 (12.4, 12.7)	10.8 (10.7, 10.9)
Extremely dense (D)	149 428	10.0 (9.8, 10.1)	9.7 (9.5, 10.0)	10.1 (9.9, 10.3)
Unknown	44 726	10.3 (10.0, 10.5)	10.2 (9.7, 10.7)	10.3 (10.0, 10.6)
History of breast implants				
Yes	70 304	7.6 (7.4, 7.8)	8.3 (7.9, 8.7)	7.3 (7.1, 7.6)
No	2 457 759	9.4 (9.4, 9.5)	10.4 (10.3, 10.4)	9.0 (8.9, 9.0)
Supplemental screening				
US	63 827	11.9 (11.7, 12.2)	12.5 (11.9, 13.1)	11.8 (11.5, 12.0)
MRI	8492	6.6 (6.1, 7.1)	8.1 (7.1, 9.3)	5.9 (5.4, 6.6)
None	2 455 744	9.3 (9.3, 9.4)	10.3 (10.2, 10.4)	8.8 (8.8, 8.9)

Note.—The recall rate (per 1000 screening mammograms) was calculated as the proportion of screening mammograms resulting in recalls at digital mammography (DM) (n = 834 336) and digital breast tomosynthesis (DBT) (n = 1 693 727). Data in parentheses are 95% CIs.

* "Other" includes individuals who identified as American Indian or Alaska Native, multiracial, Native Hawaiian, other Pacific Islander, or other categories not specified within the data source.

† Short-term risk status was determined with use of the Gail model 5-year risk score, where elevated risk was defined as a score of 1.66 or higher.

from 2.2 of 1000 (95% CI: 2.1, 2.4) among women aged 40–44 years to 8.5 of 1000 (95% CI: 8.0, 9.1) among women aged 75–79 years. The overall crude CDR was higher in the DBT group (5.3 of 1000 [95% CI: 5.2, 5.5]) than the DM group (4.5 of 1000 [95% CI: 4.3, 4.6]; P < .001), a finding persistent across subgroups including breast tissue density and race, where Asian women who underwent DBT had the highest CDR (5.7 of 1000 [95% CI: 5.1, 6.4]). Among examinations in women with only one observed screening examination, the crude CDR was higher for DM than for DBT (17.9 of 1000 [95% CI: 17.2, 18.6] vs 16.0 of 1000 [95% CI: 15.4, 16.6]; P < .001). After adjustment, DBT was

significantly associated with a higher CDR than DM (OR, 1.24 [95% CI: 1.19, 1.30]; $P < .001$) (Fig 4).

Positive Predictive Value of Recall

A total of 186 949 screening mammograms (DM, 79 483; DBT, 107 466) led to recall, and the overall PPV1 was 5.2% (95% CI: 5.1, 5.3) (Table 4). By modality, the crude PPV1 was higher for DBT (5.9% [95% CI: 5.7, 6.0]) than for DM (4.3% [95% CI: 4.2, 4.5]; $P < .001$), a pattern that was observed across all characteristics with the exception of those with only one prior screening mammogram. Among those with a single screening examination, crude PPV1 was lower for DBT (9.7% [95% CI: 9.4, 10.1]) than for DM (10.0% [95% CI: 9.6, 10.4]). After adjustment, DBT remained associated with higher PPV1 (OR, 1.33 [95% CI: 1.27, 1.40]; $P < .001$) (Fig 4).

Biopsy Rate

The biopsy rate among 2 092 346 screening mammograms (DM, 779 716; DBT, 1 312 630) in women who had at least 3 months of follow-up was 16.4 of 1000 screening mammograms (95% CI: 16.2, 16.6) (Table S3). By modality, the crude biopsy rate was 17.6 of 1000 (95% CI: 17.4, 17.8) for DBT and 14.5 of 1000 (95% CI: 14.2, 14.7) for DM ($P < .001$). Crude biopsy rates in the DBT group were consistently higher than those in the DM group, except for women with only one observed screening mammogram, where the biopsy rate was higher for the DM group (38.4 of 1000 [95% CI: 37.4, 39.4]) than the DBT group (36.6 of 1000 [95% CI: 35.8, 37.4]). When women were stratified by race, examinations in Asian women in the DBT group had the highest crude biopsy rate at 21.8 of 1000 (95% CI: 20.7, 22.9), compared with 18.0 of 1000 (95% CI: 17.7, 18.3) for White women and 15.9 of 1000 (95% CI: 15.2, 16.6) for Black women. The association between DBT and biopsy rate was statistically significant in adjusted analyses (adjusted OR, 1.33 [95% CI: 1.30, 1.37]; $P < .001$) (Fig 4).

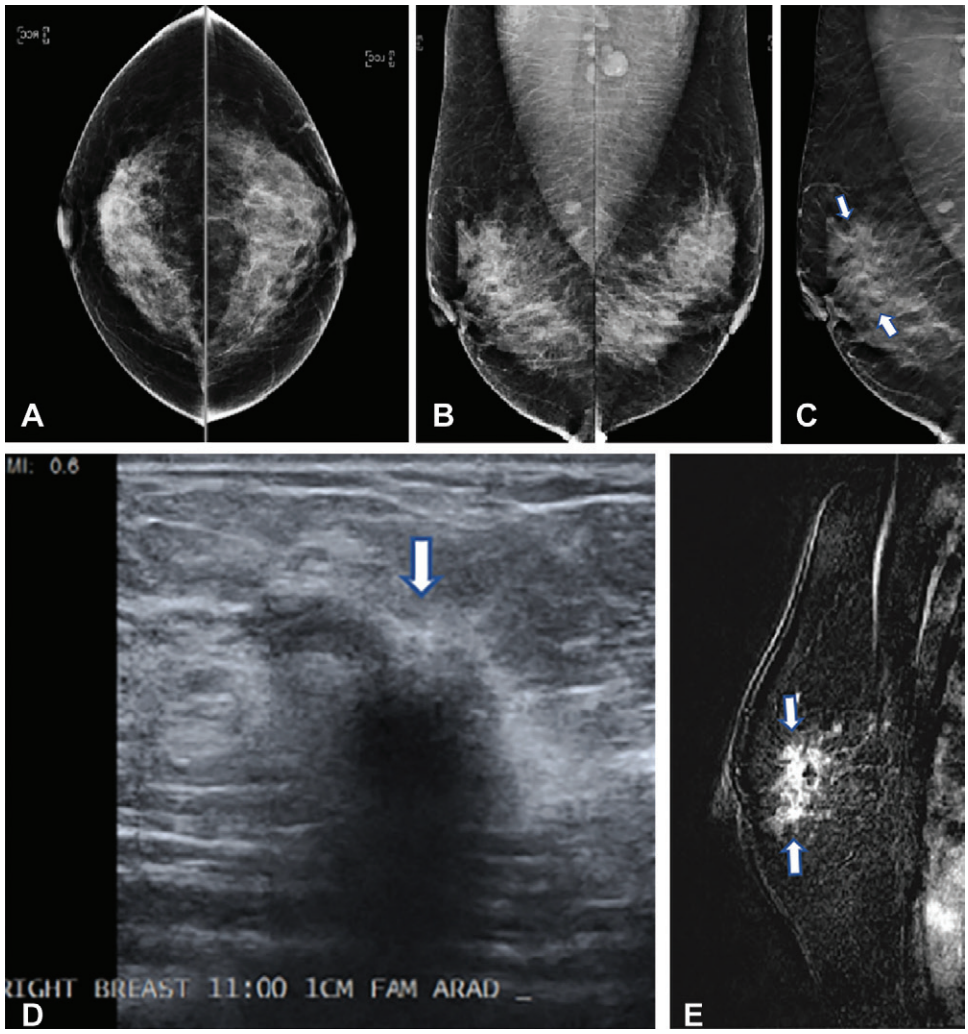

Figure 2: Images in a 42-year-old woman who presented for routine screening. **(A)** Two-dimensional craniocaudal digital mammogram and **(B)** two-dimensional mediolateral digital mammogram show heterogeneously dense breast tissue with no abnormality. **(C)** Mediolateral digital breast tomosynthesis image shows subtle architectural distortion (arrows) extending superiorly from the nipple level. **(D)** Target US image of the right breast shows a highly suspicious, hypoechoic, irregular mass (arrow). US-guided core biopsy yielded invasive ductal carcinoma with extensive ductal carcinoma in situ (T2N0M0; estrogen receptor–positive, progesterone receptor–positive, and human epidermal growth factor receptor 2–negative). **(E)** Right-breast MRI scan acquired to evaluate the extent of disease shows a clip artifact in the superior aspect of the irregular enhancing mass (arrows). No other suspicious lesions were seen.

Positive Predictive Value of Biopsy

The overall PPV3 among 34 357 screening mammograms (DM, 11 271; DBT, 23 086) that led to biopsies during the 3-month follow-up period was 29.5% (95% CI: 29.1, 30.0) (Table S4). By modality, the crude PPV3 was 29.3% (95% CI: 28.7, 29.9) and 30.0% (95% CI: 29.2, 30.9) for DBT and DM, respectively ($P = .16$). No evidence of a difference was found for PPV3 when stratified by subgroups, except among White women (32.3 [95% CI: 31.1, 33.4] for DM vs 29.6 [95% CI: 28.9, 30.4] for DBT) and Asian women undergoing DM (29.5 [95% CI: 26.7, 32.4] vs 24.4 [95% CI: 22.2, 26.7]), as well as among women with elevated risk (39.7 [95% CI: 37.9, 41.6] vs 32.1 [95% CI: 30.9, 33.4]). The null association between DBT and PPV3 remained after adjustment (OR, 0.95 [95% CI: 0.90, 1.00]; $P = .16$) (Fig 4).

Discussion

This retrospective cohort study included 2 528 063 screening mammograms conducted in 1 100 447 women between 2014 and 2020. Screening mammogram outcomes were compared between digital breast tomosynthesis (DBT) and digital mammography (DM). In crude analyses, DBT was associated with reduced recall rate (10.3% for DM vs 8.9% for DBT), increased cancer detection rate (4.5 of 1000 screening mammograms vs 5.3 of 1000), positive predictive value of recall (4.3% vs 5.9%), and biopsy rate (14.5 of 1000 screening mammograms vs 17.6 of 1000). After adjustment for potential confounders, these associations remained.

Our study contributes to the existing literature by describing the use and outcomes of DBT and DM among a large cohort of women across the United States. Our results were consistent with those of prior studies reporting reduced RRs and increased CDRs, PPV1, and biopsy rates with DBT compared with DM (5,6,8), with a few key differences. For example, in an adjusted analysis of over 1.5 million screening mammograms, Lowry et al (11) found that the recall and CDR benefits associated with DBT were most pronounced among women undergoing their baseline mammogram. In our study, those in the DM group with a single mammogram had higher CDRs, PPV1, and biopsy rates. However, because the presence of a single screening mammogram in our analysis may represent the only observed mammogram in the data rather than a true baseline mammogram, results among women with only one screening mammogram cannot be directly compared with those from studies that identified baseline mammograms.

The DM and DBT groups were comparable in our study, with few imbalances, including the higher proportions of screening mammograms from postmenopausal and White women in the DBT group. While the focus of this analysis was not to evaluate access to or differences in DBT screening performance based on race or socioeconomic factors, this has been previously described (25), and our adjusted findings for key screening outcomes remained consistent with only slight shifts in point estimates and 95% CIs from the crude analyses.

It is possible that residual confounding may contribute to at least some of the observed differences between DM and DBT, including upward bias or positive confounding from unmeasured or imperfectly measured factors associated with both the choice of modality and the risk of screening outcomes. For example, data on race, ethnicity, and breast density were missing for some patients in the study, and if this missingness was not random across screening modality, it could account for at least some of the observed differences in screening outcomes. Additionally, our analysis was not able to adjust for site or practice-level variables like operator and reader experience,

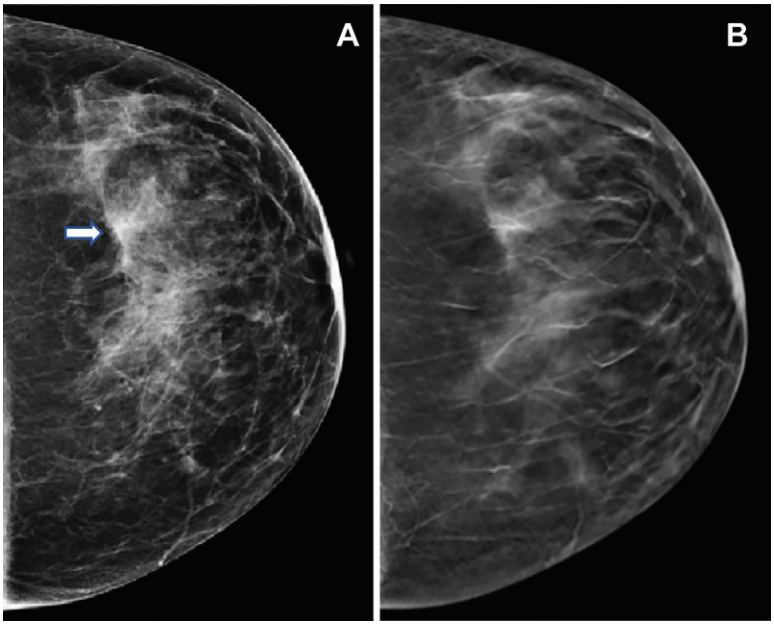

Figure 3: Images in a 47-year-old woman who presented for routine screening. **(A)** Craniocaudal digital mammogram shows scattered fibroglandular densities. On the left digital mammographic craniocaudal view laterally, an asymmetry (arrow) is present. **(B)** Left craniocaudal digital breast tomosynthesis image shows no suspicious lesion but rather a superimposition of normal fibroglandular and ligamentous structures.

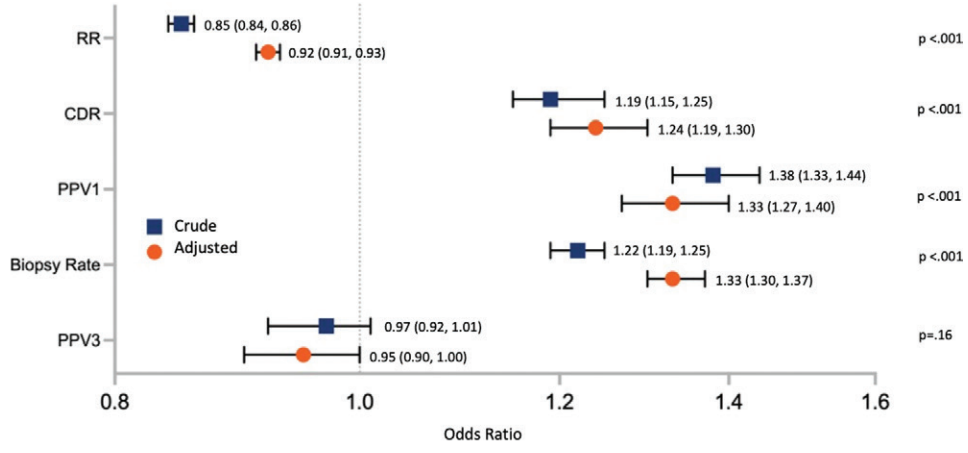

Figure 4: Crude and adjusted associations between screening modality and breast cancer screening outcomes: comparison of digital breast tomosynthesis (DBT) with digital mammography (DM). Forest plot shows the crude and adjusted odds ratios (ORs), with 95% CIs in parentheses and vertical bars showing the range of 95% CIs, for the primary outcomes of recall rate (RR), cancer detection rate (CDR), positive predictive value of recall (PPV1), biopsy rate, and positive predictive value of biopsy (PPV3), comparing DBT with DM. ORs were adjusted for age category (40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, and 75–79 years), breast density, health system, and index year of examination.

Table 3: CDRs, Overall and by Select Patient Characteristics

Characteristic	Overall			
	No. of Screening Mammograms	CDR	CDR for DM	CDR for DBT
All screening mammograms	1 948 098	5.0 (4.9, 5.1)	4.5 (4.3, 4.6)	5.3 (5.2, 5.5)
Screening mammogram count				
1	323 983	16.8 (16.3, 17.2)	17.9 (17.2, 18.6)	16.0 (15.4, 16.6)
2 or more	1 624 115	2.6 (2.6, 2.7)	1.6 (1.5, 1.7)	3.3 (3.2, 3.4)
Screening interval				
≤24 months	1 077 898	4.1 (4.0, 4.3)	3.6 (3.4, 3.8)	4.4 (4.3, 4.6)
>24 months	185 809	6.5 (6.1, 6.9)	6.3 (5.8, 6.9)	6.6 (6.1, 7.1)
Unknown	684 391	5.9 (5.7, 6.1)	5.1 (4.8, 5.3)	6.7 (6.4, 6.9)
Age category				
40–44 years	255 194	2.2 (2.1, 2.4)	1.8 (1.5, 2.0)	2.5 (2.3, 2.8)
45–49 years	277 706	3.4 (3.2, 3.6)	3.1 (2.8, 3.5)	3.6 (3.3, 3.9)
50–54 years	299 461	4.2 (4.0, 4.5)	3.7 (3.4, 4.0)	4.6 (4.3, 4.9)
55–59 years	309 666	4.6 (4.3, 4.8)	4.2 (3.8, 4.6)	4.8 (4.5, 5.1)
60–64 years	280 640	5.6 (5.3, 5.9)	5.2 (4.8, 5.6)	5.9 (5.5, 6.3)
65–69 years	245 847	6.9 (6.5, 7.2)	6.0 (5.5, 6.5)	7.4 (7.0, 7.8)
70–74 years	177 196	7.8 (7.4, 8.2)	7.3 (6.6, 7.9)	8.1 (7.6, 8.7)
75–79 years	102 388	8.5 (8.0, 9.1)	7.7 (6.9, 8.6)	9.0 (8.3, 9.8)
Race				
Asian	123 372	4.9 (4.5, 5.3)	4.2 (3.7, 4.7)	5.7 (5.1, 6.4)
Black	203 412	5.3 (5.0, 5.6)	5.1 (4.7, 5.5)	5.6 (5.1, 6.1)
Other*	56 961	4.2 (3.7, 4.7)	3.1 (2.4, 3.9)	4.9 (4.2, 5.7)
White	1 179 274	5.2 (5.1, 5.3)	4.8 (4.6, 5.0)	5.5 (5.3, 5.6)
Unknown	385 079	4.3 (4.1, 4.5)	3.3 (3.0, 3.6)	4.8 (4.5, 5.1)
Short-term risk status†				
Normal risk	1 489 247	4.7 (4.6, 4.8)	4.0 (3.8, 4.2)	5.2 (5.0, 5.3)
Elevated risk	458 851	5.9 (5.7, 6.1)	6.0 (5.6, 6.3)	5.9 (5.6, 6.1)
Breast density				
Almost entirely fatty (A)	163 707	4.2 (3.9, 4.5)	4.0 (3.5, 4.5)	4.4 (4.0, 4.8)
Scattered fibroglandular densities (B)	902 874	5.0 (4.9, 5.2)	4.6 (4.4, 4.9)	5.3 (5.1, 5.5)
Heterogeneously dense (C)	721 756	5.3 (5.1, 5.5)	4.6 (4.3, 4.8)	5.8 (5.6, 6.0)
Extremely dense (D)	120 672	3.8 (3.5, 4.2)	3.1 (2.6, 3.6)	4.4 (3.9, 4.9)
Unknown	39 089	4.9 (0.3, 74.4)	5.4 (4.3, 6.8)	4.7 (3.9, 5.6)
History of breast implants				
Yes	53 305	2.3 (1.9, 2.7)	2.2 (1.5, 3.1)	2.3 (1.9, 2.8)
No	1 894 793	5.1 (5.0, 5.2)	4.5 (4.4, 4.7)	5.4 (5.3, 5.6)
Supplemental screening				
US	41 001	NA	NA	NA
MRI	7349	NA	NA	NA
None	1 899 748	5.1 (5.0, 5.2)	4.5 (4.4, 4.7)	5.5 (5.4, 5.6)

Note.—The cancer detection rate (CDR) (per 1000 screening mammograms) was calculated for examinations in women with at least 6 months of follow-up as the number of screening-detected cancers identified within 6 months of index at digital mammography (DM) (766 587 examinations) and digital breast tomosynthesis (DBT) (1 181 ,511 examinations). Data in parentheses are 95% CIs. NA = not applicable.

* "Other" includes individuals who identified as American Indian or Alaska Native, multiracial, Native Hawaiian, other Pacific Islander, or other categories not specified within the data source.

† Short-term risk status was determined with use of the Gail model 5-year risk score, where elevated risk was defined as a score of 1.66 or higher.

which are known to influence the screening performance of DBT (26–28). However, the robustness of our results to sensitivity analyses suggests that the strength of any confounding would have to be relatively large to completely account for the associations we observed. The E-value is a measure related to evidence for causality representing the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and outcome to fully explain a treatment-outcome association, conditional on the

Table 4: PPV1 for Cancer Screening, Overall and by Select Patient Characteristics

Characteristic	Overall			
	No. of Examinations Resulting in Recall	PPV1 (%)	PPV1 for DM (%)	PPV1 for DBT (%)
All screening mammograms	186 949	5.2 (5.1, 5.3)	4.3 (4.2, 4.5)	5.9 (5.7, 6.0)
Screening mammogram count				
1	55 086	9.9 (9.6, 10.1)	10.0 (9.6, 10.4)	9.7 (9.4, 10.1)
2 or more	131 863	3.2 (3.2, 3.3)	1.9 (1.8, 2.0)	4.3 (4.1, 4.4)
Screening interval				
≤24 months	76 130	5.8 (5.7, 6.0)	4.7 (4.4, 4.9)	6.5 (6.3, 6.8)
>24 months	19 217	6.3 (5.9, 6.6)	5.7 (5.2, 6.2)	6.8 (6.3, 7.3)
Unknown	91 602	4.4 (4.3, 4.6)	3.8 (3.6, 4.0)	5.0 (4.8, 5.2)
Age category				
40–44 years	35 168	1.6 (1.5, 1.8)	1.1 (1.0, 1.3)	2.0 (1.8, 2.2)
45–49 years	32 774	2.9 (2.7, 3.1)	2.4 (2.2, 2.7)	3.2 (3.0, 3.5)
50–54 years	30 170	4.2 (4.0, 4.4)	3.5 (3.2, 3.8)	4.8 (4.5, 5.1)
55–59 years	26 328	5.4 (5.1, 5.6)	4.6 (4.3, 5.0)	5.9 (5.6, 6.3)
60–64 years	22 633	6.9 (6.6, 7.3)	5.8 (5.4, 6.3)	7.8 (7.4, 8.3)
65–69 years	19 571	8.6 (8.2, 9.0)	6.9 (6.3, 7.5)	9.8 (9.3, 10.3)
70–74 years	13 127	10.5 (10.0, 11.1)	9.2 (8.4, 10.0)	11.5 (10.8, 12.2)
75–79 years	7178	12.1 (11.4, 12.9)	10.4 (9.3, 11.5)	13.4 (12.4, 14.5)
Race				
Asian	12 901	4.7 (4.3, 5.0)	4.1 (3.7, 4.6)	5.2 (4.7, 5.8)
Black	20 915	5.2 (4.9, 5.5)	4.5 (4.2, 4.9)	6.1 (5.6, 6.7)
Other*	6095	3.9 (3.4, 4.4)	2.8 (2.2, 3.5)	4.7 (4.0, 5.4)
White	106 826	5.8 (5.6, 5.9)	4.9 (4.7, 5.2)	6.3 (6.1, 6.5)
Unknown	40 212	4.1 (3.9, 4.3)	2.7 (2.5, 3.0)	5.0 (4.7, 5.3)
Short-term risk status†				
Normal risk	151 473	4.6 (4.5, 4.7)	3.6 (3.5, 3.8)	5.4 (5.2, 5.5)
Elevated risk	35 476	7.6 (7.4, 7.9)	7.2 (6.8, 7.6)	7.9 (7.6, 8.3)
Breast density				
Almost entirely fatty (A)	9606	7.2 (6.7, 7.7)	6.2 (5.5, 6.9)	8.0 (7.3, 8.7)
Scattered fibroglandular densities (B)	77 601	5.8 (5.7, 6.0)	4.9 (4.7, 5.1)	6.6 (6.3, 6.8)
Heterogeneously dense (C)	83 819	4.6 (4.4, 4.7)	3.6 (3.5, 3.8)	5.3 (5.1, 5.5)
Extremely dense (D)	12 046	3.8 (3.5, 4.2)	3.2 (2.7, 3.7)	4.3 (3.8, 4.8)
Unknown	3877	5.0 (4.3, 5.7)	5.4 (4.3, 6.7)	4.7 (4.0, 5.6)
History of breast implants				
Yes	4177	2.9 (2.4, 3.4)	2.6 (1.8, 3.7)	3.0 (2.4, 3.7)
No	182 772	5.2 (5.1, 5.4)	4.3 (4.2, 4.5)	5.9 (5.8, 6.1)
Supplemental screening				
US	6134	NA	NA	NA
MRI	450	NA	NA	NA
None	180 365	5.4 (5.3, 5.5)	4.4 (4.2, 4.5)	6.1 (6.0, 6.3)

Note.—Positive predictive value of recall (PPV1) is reported as the proportion of examinations in women diagnosed with screening-detected breast cancer within 6 months of recall at digital mammography (DM) (79 483 examinations) and digital breast tomosynthesis (DBT) (107 466 examinations). Data in parentheses are 95% CIs. NA = not applicable.

* "Other" includes individuals who identified as American Indian or Alaska Native, multiracial, Native Hawaiian, other Pacific Islander, or other categories not specified within the data source.

† Short-term risk status was determined with use of the Gail model 5-year risk score, where elevated risk was defined as a score of 1.66 or higher.

measured covariates (29). For this study, E-values ranging from approximately 1.3 to 1.8 were calculated for the primary outcomes of RR, CDR, and biopsy rate, while associations reported in the literature (30) are more modest in magnitude than those resulting from the E-value sensitivity analysis; thus, our findings are not likely to be fully accounted for by confounding.

Our study had several limitations. First, as this is an observational study, cautious interpretation of associations is warranted given the potential for residual and unmeasured

confounding. Nonetheless, as discussed earlier, the primary comparison groups of interest were relatively similar with respect to baseline characteristics and were robust even after multivariable adjustment.

Second, this study used secondary data collection, and therefore, information not documented at the time of data retrieval was not available. Data on some baseline characteristics (breast density, race, ethnicity, menopause status, screening interval) were missing for substantial proportions (1.7%–32.2%) of screening mammograms, and while supplemental screening data were extracted for this analysis, these practices vary by site, and the reasons for using US or MRI were not collected. As such, some imaging examinations may have been misclassified as supplemental screening. There may also be cases where breast cancers were identified by means of incidental findings at other imaging (eg, PET), and these imaging procedures were not captured. Additionally, the identification of a breast cancer diagnosis relied on institutional cancer registry data. Because the potential lag in case reporting, some cancer cases may not have been reported at the time of data retrieval, and screening mammograms conducted before an observed lag were excluded. However, overestimation of the CDR is possible because false-negative results may have been included. As this would be a rare event with either imaging modality, it would not be expected to appreciably bias the results.

Third, while this study allowed exploration of the comparison of DBT versus DM in key patient subgroups of interest, such as those stratified by race and breast density, these analyses were hypothesis-generating in nature and should not be overinterpreted given the risks inherent to multiple-subgroup analysis that is exploratory.

In summary, data from a multi-institutional, United States–based database of over 2.5 million screening mammograms found consistent improvement in screening outcomes with digital breast tomosynthesis (DBT). This data set may be used in future analyses that require large data sets, such as in the comparison of the detection rates of advanced breast cancers or in subgroup analysis of women who underwent DBT or digital mammography.

Acknowledgments: The authors acknowledge Alex Sardina, MD, for his contributions to development of the study protocol and Janis Breeze, MPH, for her contributions to the writing of this manuscript.

Author contributions: Guarantors of integrity of entire study, **C.R.P., A.R.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **E.F.C., J.K.P., N.A.**; clinical studies, **E.F.C.**; statistical analysis, **B.C.S., S.Y.L., Y.J., L.A.S.S., J.K.P.**; and manuscript editing, **E.F.C., M.M.T., C.R.P., S.Y.L., S.P., A.R., Y.J., L.A.S.S., J.K.P., N.A.**

Disclosures of conflicts of interest: **E.F.C.** Grants to institution from iCAD; consulting fees from iCAD and Hologic; payment for lectures from Hologic, MedScape, and Aunt Minnie; support for travel to meetings from the RSNA; unpaid leadership role in the Society of Breast Imaging. **M.M.T.** No relevant relationships. **C.R.P.** No relevant relationships. **B.C.S.** No relevant relationships. **S.Y.L.** No relevant relationships. **S.P.** Stock in Hologic. **A.R.** No relevant relationships. **Y.J.** No relevant relationships. **L.A.S.S.** No relevant relationships. **J.K.P.** No relevant relationships. **N.A.** No relevant relationships.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021;71(1):7–33. [Published correction appears in CA Cancer J Clin 2021;71(4):359.]
2. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 U.S. Preventive Services Task Force recommendation. Ann Intern Med 2016;164(4):244–255.
3. Korhonen KE, Conant EF, Cohen EA, Synnestvedt M, McDonald ES, Weinstein SP. Breast cancer conspicuity on simultaneously acquired digital mammographic images versus digital breast tomosynthesis images. Radiology 2019;292(1):69–76.
4. Østerås BH, Martinsen ACT, Gullien R, Skaane P. Digital mammography versus breast tomosynthesis: impact of breast density on diagnostic performance in population-based screening. Radiology 2019;293(1):60–68.
5. Alsheik N, Blount L, Qiong Q, et al. Outcomes by race in breast cancer screening with digital breast tomosynthesis versus digital mammography. J Am Coll Radiol 2021;18(7):906–918.
6. Alsheik NH, Dabbous F, Pohlman SK, et al. Comparison of resource utilization and clinical outcomes following screening with digital breast tomosynthesis versus digital mammography: findings from a learning health system. Acad Radiol 2019;26(5):597–605.
7. Conant EF, Zuckerman SP, McDonald ES, et al. Five consecutive years of screening with digital breast tomosynthesis: outcomes by screening year and round. Radiology 2020;295(2):285–293.
8. Conant EF, Barlow WE, Herschorn SD, et al. Association of digital breast tomosynthesis vs digital mammography with cancer detection and recall rates by age and breast density. JAMA Oncol 2019;5(5):635–642.
9. McDonald ES, Oustimov A, Weinstein SP, Synnestvedt MB, Schnall M, Conant EF. Effectiveness of digital breast tomosynthesis compared with digital mammography: outcomes analysis from 3 years of breast cancer screening. JAMA Oncol 2016;2(6):737–743.
10. McDonald ES, McCarthy AM, Akhtar AL, Synnestvedt MB, Schnall M, Conant EF. Baseline screening mammography: performance of full-field digital mammography versus digital breast tomosynthesis. AJR Am J Roentgenol 2015;205(5):1143–1148.
11. Lowry KP, Coley RY, Miglioretti DL, 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(7):e2011792.
12. Conant EF, Beaber EF, Sprague BL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast Cancer Res Treat 2016;156(1):109–116.
13. Libesman S, Zackrisson S, Hofvind S, 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;22(5):e647–e654.
14. Rafferty EA, Park JM, Philpotts LE, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. Radiology 2013;266(1):104–113.
15. Richman IB, Long JB, Hoag JR, et al. Comparative effectiveness of digital breast tomosynthesis for breast cancer screening among women 40-64 years old. J Natl Cancer Inst 2021;113(11):1515–1522.
16. Siu AL; U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2016;164(4):279–296. [Published correction appears in Ann Intern Med 2016;164(6):448.]
17. Monticciolo DL, Newell MS, Hendrick RE, et al. Breast cancer screening for average-risk women: recommendations from the ACR Commission on Breast Imaging. J Am Coll Radiol 2017;14(9):1137–1143.
18. American Cancer Society Recommendations for the Early Detection of Breast Cancer. https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html. Accessed April 16, 2022.
19. Rafferty EA, Rose SL, Miller DP, et al. Effect of age on breast cancer screening using tomosynthesis in combination with digital mammography. Breast Cancer Res Treat 2017;164(3):659–666.
20. Phi XA, Tagliafico A, Houssami N, Greuter MJW, de Bock GH. Digital breast tomosynthesis for breast cancer screening and diagnosis in women with dense breasts – a systematic review and meta-analysis. BMC Cancer 2018;18(1):380.

21. Hovda T, Brandal SHB, Sebuødegård S, et al. Screening outcome for consecutive examinations with digital breast tomosynthesis versus standard digital mammography in a population-based screening program. Eur Radiol 2019;29(12):6991–6999.
22. Houssami N, Zackrisson S, Blazek K, et al. Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus mammography population screening. Eur J Cancer 2021;148:14–23.
23. Kerlikowske K, Su YR, Sprague BL, et al. Association of screening with digital breast tomosynthesis vs digital mammography with risk of interval invasive and advanced breast cancer. JAMA 2022;327(22):2220–2230.
24. Winter AM, Kazmi S, Hardy AK, Bennett DL. Comparison of interval breast cancers with 2D digital mammography versus 3D digital breast tomosynthesis in a large community-based practice. Breast J 2020;26(10):1953–1959.
25. Lee CI, Zhu W, Onega T, 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(2):e2037546.
26. Sprague BL, Coley RY, Kerlikowske K, et al. Assessment of radiologist performance in breast cancer screening using digital breast tomosynthesis vs digital mammography. JAMA Netw Open 2020;3(3):e201759.
27. Bernstein MH, Baird GL, Lourenco AP. Digital Breast tomosynthesis and digital mammography recall and false-positive rates by time of day and reader experience. Radiology 2022;303(1):63–68.
28. Tucker L, Gilbert FJ, Astley SM, et al. Does reader performance with digital breast tomosynthesis vary according to experience with two-dimensional mammography? Radiology 2017;283(2):371–380.
29. Mathur MB, Ding P, Riddell CA, VanderWeele TJ. Web site and R package for computing e-values. Epidemiology 2018;29(5):e45–e47.
30. Castells X, Molins E, Macià F. Cumulative false positive recall rate and association with participant related factors in a population based breast cancer screening programme. J Epidemiol Community Health 2006;60(4):316–321.