Letters

RESEARCH LETTER

Breast Cancer Screening Using Tomosynthesis and Digital Mammography in Dense and Nondense Breasts

Breast density is associated with reduced mammographic sensitivity and specificity. Additionally, increased tumor size and worsened prognosis are associated with increased breast density.[1,2] Dense breast tissue may also represent an independent risk factor for breast cancer.[3] Currently, 24 states have laws mandating that women be notified of the implications of breast density, thereby encouraging discussions between patients and physicians regarding the need for supplemental screening.[4] However, which, if any, additional modalities should be recommended for women with dense breasts is not known.

Related article page 1786

Using data from our previous multicenter study,[5] we evaluated differential screening performance of digital mammography combined with tomosynthesis compared with digital mammography alone as a function of breast density.

Methods | The protocol was approved by institutional review boards of participating institutions with a waiver of informed consent. Screening performance metrics from 13 US institutions were reported for 12 months using digital mammography alone (beginning March 2011 to October 2012) and from the date of introduction of tomosynthesis until December 31, 2012 (range, 3-22 months).

Subgroups included the 4 breast density categories used for clinical reporting. Almost entirely fat and scattered fibroglandular densities were considered nondense tissue patterns, whereas heterogeneously dense and extremely dense were considered dense tissue patterns.

Overall and invasive cancer detection rates and recall rate with and without tomosynthesis were analyzed in patients with both nondense and dense breasts. Positive predictive value for recall was calculated. Exploratory analyses were conducted for all 4 density categories. Additive models were used to estimate rates as previously described (adjusting for screening method and site).[5] An additional multivariable model including all subgroup effects was fit to determine age-adjusted density effect. Adjusted rates and 95% confidence intervals were calculated based on fitted models using SAS (SAS Institute), version 9.3. All tests were 2-sided and a *P* value less than .05 was considered statistically significant. Because data on interval cancers were not available, complete assessment of sensitivity and specificity could not be done.

Results | Of 452 320 examinations, 278 906 were digital mammography alone and 173 414 digital mammography plus tomosynthesis; 2157 cancers were diagnosed. The **Table** summarizes results of primary (dense vs nondense) and exploratory (breast density categories) analyses showing model-adjusted rates. Recall rates per 1000 screens in nondense breasts decreased from 90 to 79 (difference, −12 [95% CI, −14 to −9]; $P < .001$); and in dense breasts from 127 to 109 (difference, −18 [95% CI, −21 to −15]; $P < .001$) with tomosynthesis. Positive predictive value of recalls increased in both nondense and dense breasts. Cancer detection rates also increased in both groups. Invasive cancer detection rate per 1000 screens in nondense breasts increased from 3.0 to 4.0 (difference, 0.9 [95% CI, 0.4 to 1.5]; $P < .001$) and in dense breasts from 2.9 to 4.2 (difference, 1.4 [95% CI, 0.9-1.9]; $P < .001$) with tomosynthesis.

For subgroups of breast density, improvements in rates were greatest for women with scattered fibroglandular densities and heterogeneously dense breasts. Differences were mostly not significant for almost entirely fat and extremely dense subgroups.

The **Figure** depicts density effect adjusted for age, consistent with the increased cancer detection and reduced recall rates after implementation of tomosynthesis not being solely attributable to confounding by age but possibly independently associated with improved screening performance.

Discussion | Addition of tomosynthesis to digital mammography for screening was associated with an increase in cancer detection rate and a reduction in recall rate for women with both dense and nondense breast tissue. These combined gains were largest for women with heterogeneously dense breasts, potentially addressing limitations in cancer detection seen with digital mammography alone in this group, but were not significant in women with extremely dense breasts.

Limitations of this study include its retrospective design, collection of data at the population level rather than the patient level, and insufficient follow-up to determine if increased invasive cancer detection improved clinical outcomes. For women classified as having dense breast tissue, most have heterogeneously dense breasts, mandating caution in drawing conclusions regarding the performance of tomosynthesis for the small proportion of women with extremely dense breasts.

Elizabeth A. Rafferty, MD
Melissa A. Durand, MD
Emily F. Conant, MD
Debra Somers Copit, MD
Sarah M. Friedewald, MD
Donna M. Plecha, MD
Dave P. Miller, MS

Author Affiliations: Department of Radiology, Massachusetts General Hospital, Boston (Rafferty); now with L&M Radiology, West Acton, Massachusetts (Rafferty); Yale University School of Medicine, New Haven, Connecticut (Durand); Department of Radiology, University of Pennsylvania Perelman School of Medicine, Philadelphia (Conant); Albert Einstein Healthcare Network,

Table. Model-Adjusted Rates and Positive Predictive Values for Screening Examinations vs Breast Density Among US Women

			Nondense Tissue Patterns		Dense Tissue Patterns	
Breast Density	Nondense	Dense	Almost Entirely Fat	Scattered Fibroglandular Densities	Heterogeneously Dense	Extremely Dense
No. of screens (%)						
Digital mammography	146 910 (52.7)	131 996 (47.3)	24 467 (8.8)	122 443 (43.9)	113 290 (40.6)	18 706 (6.7)
Digital mammography + tomosynthesis	89 171 (51.4)	84 243 (48.6)	15 319 (8.8)	73 852 (42.6)	72 481 (41.8)	11 762 (6.8)
Model-Adjusted Rates[a]						
Recalls per 1000 screens, estimate (95% CI) [No. of screens][b]						
Digital mammography	90 (75 to 106) [12 845]	127 (107 to 147) [16 582]	57 (44 to 70) [1297]	97 (81 to 114) [11 548]	128 (107 to 149) [14 484]	114 (94 to 133) [2098]
Digital mammography + tomosynthesis	79 (63 to 94) [6955]	109 (89 to 129) [9030]	55 (41 to 68) [909]	84 (68 to 101) [6046]	110 (90 to 131) [7852]	98 (78 to 118) [1178]
Difference (95% CI)	−12 (−14 to −9)	−18 (−21 to −15)	−2 (−8 to 3)	−13 (−16 to −10)	−18 (−21 to −15)	−16 (−23 to −8)
P value	<.001	<.001	.34	<.001	<.001	<.001
Cancers per 1000 screens, estimate (95% CI) [No. of screens][c]						
Digital mammography	4.2 (3.7 to 4.7) [610]	4.5 (4.0 to 4.9) [597]	3.2 (2.4 to 4.0) [77]	4.4 (3.7 to 5.0) [533]	4.5 (3.9 to 5.1) [528]	3.8 (2.6 to 4.9) [69]
Digital mammography + tomosynthesis	5.1 (4.5 to 5.8) [455]	5.8 (5.3 to 6.4) [495]	4.2 (3.2 to 5.2) [64]	5.3 (4.6 to 6.1) [391]	6.1 (5.4 to 6.8) [450]	3.9 (2.6 to 5.2) [45]
Difference (95% CI)	1.0 (0.4 to 1.5)	1.4 (0.8 to 2.0)	1.0 (−0.2 to 2.3)	1.0 (0.3 to 1.6)	1.6 (0.9 to 2.3)	0.1 (−1.3 to 1.6)
P value	.001	<.001	.10	.004	<.001	.88
Invasive cancers per 1000 screens, estimate (95% CI) [No. of screens][d]						
Digital mammography	3.0 (2.6 to 3.5) [439]	2.9 (2.5 to 3.2) [376]	2.3 (1.6 to 3.1) [55]	3.2 (2.6 to 3.8) [384]	3.0 (2.6 to 3.4) [340]	1.9 (1.3 to 2.6) [36]
Digital mammography + tomosynthesis	4.0 (3.4 to 4.5) [351]	4.2 (3.8 to 4.6) [356]	3.5 (2.5 to 4.4) [52]	4.1 (3.5 to 4.8) [299]	4.5 (4.0 to 5.0) [326]	2.6 (1.7 to 3.4) [30]
Difference (95% CI)	0.9 (0.4 to 1.5)	1.4 (0.9 to 1.9)	1.1 (0.0 to 2.2)	0.9 (0.4 to 1.5)	1.5 (1.0 to 2.1)	0.6 (−0.5 to 1.7)
P value	<.001	<.001	.046	.001	<.001	.25
Positive predictive value for recall, % (95% CI)[e]						
Digital mammography	5.1 (3.9 to 6.2)	3.8 (2.9 to 4.7)	6.2 (4.2 to 8.2)	4.9 (3.7 to 6.0)	3.8 (2.8 to 4.9)	3.7 (2.2 to 5.1)
Digital mammography + tomosynthesis	7.1 (5.9 to 8.4)	5.7 (4.7 to 6.6)	8.4 (6.0 to 10.7)	6.9 (5.7 to 8.2)	5.9 (4.9 to 7.0)	4.3 (2.7 to 5.9)
Difference (95% CI)	2.1 (1.4 to 2.8)	1.9 (1.4 to 2.4)	2.1 (−0.1 to 4.4)	2.0 (1.3 to 2.8)	2.1 (1.5 to 2.7)	0.6 (−0.7 to 1.9)
P value	<.001	<.001	.07	<.001	<.001	.38

[a] Model estimates were used to estimate rates with screening method (digital mammography and digital mammography + tomosynthesis) as a fixed effect and site as a random effect. Additive models used SAS PROC MIXED (SAS Institute), version 9.3.

[b] Recall rate (proportion of screening examinations requiring additional imaging based on screening examination result).

[c] Cancer detection rate (proportion of screening examinations with screen-detected breast cancer).

[d] Invasive cancer detection rate (proportion of screening examinations with screen-detected invasive breast cancer).

[e] Positive predictive value for recall (proportion of recalls after screening subsequently diagnosed with breast cancer).

Philadelphia, Pennsylvania (Copit); Caldwell Breast Center, Advocate Lutheran General Hospital, Park Ridge, Illinois (Friedewald); now with Lynn Sage Comprehensive Breast Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Friedewald); University Hospitals Case Medical Center, Cleveland, Ohio (Plecha); Genomic Health, Redwood City, California (Miller).

Corresponding Author: Elizabeth A. Rafferty, MD, L&M Radiology, PO Box 615, West Acton, MA 01720 (erafferty@lmradiology.com).

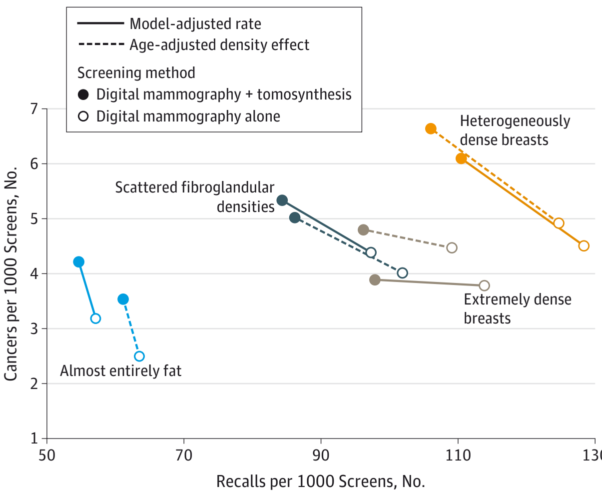
Author Contributions: Dr Rafferty had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: All authors.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: All authors.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Miller.
Obtained funding: Conant, Friedewald, Plecha.
Administrative, technical, or material support: Rafferty, Conant, Friedewald.
Study supervision: Rafferty, Conant, Friedewald, Miller.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding and Support: This study was funded by a research grant from Hologic (Drs Rafferty, Durand, Copit, Friedewald, and Plecha).

Role of the Funder/Sponsor: Principal investigators from the 13 participating institutions along with Hologic jointly designed the study. Hologic provided logistical and technical support for the collection of data. The study contracts and data sharing agreements specified that Hologic had the right to review all publications prior to submission, but could not mandate any revision of the manuscript or prevent submission for publication. Each institution verified the accuracy of their performance data, which was then transferred to an independent statistical group (ICON) for analysis. The results reported in this article are the result of the independent analysis by Faith Beery and David Pasta (both from ICON).

Figure. Combined Change in Recall and Cancer Detection Rates for Digital Mammography vs Digital Mammography Plus Tomosynthesis for Each Breast Density Category



The model-adjusted rate was adjusted for screening method and site. The density effect was adjusted for age to account for the potential confounding effect of age on breast density.

Additional Contributions: We thank the following site principal investigators: Stephen L. Rose, MD (TOPS Comprehensive Breast Center); Linda N. Greer, MD (John C. Lincoln Breast Health and Research Center); Mary K. Hayes, MD (Radiology Associates of Hollywood); Ingrid L. Ott, MD (Washington Radiology Associates); Kara L. Carlson, MD (Evergreen Breast Health Center); Thomas M. Cink, MD (Edith Sanford Breast Health Institute); Lora D. Barke, DO (Invision Sally Jobe Breast Center). We also acknowledge Anne Marie McCarthy, PhD (Massachusetts General Hospital), for data analysis and David Pasta, MS, and Faith Beery, MS (both from ICON), for statistical analysis of the results. Drs Rose, Greer, Hayes, Ott, Carlson, Cink, and Barke received a research grant for their contributions in this study from Hologic. All other contributors did not receive compensation for their contribution aside from their salaries.

1. Yaghjyan L, Colditz GA, Collins LC, et al. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. *J Natl Cancer Inst*. 2011;103(15):1179-1189.

2. Bertrand KA, Tamimi RM, Scott CG, et al. Mammographic density and risk of breast cancer by age and tumor characteristics. *Breast Cancer Res*. 2013;15(6):R104.

3. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007;356(3):227-236.

4. Slanetz PJ, Freer PE, Birdwell RL. Breast-density legislation—practical considerations. *N Engl J Med*. 2015;372(7):593-595.

5. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA*. 2014;311(24):2499-2507.

Content, Readability, and Understandability of Dense Breast Notifications by State

Along with their screening mammogram results, women in nearly half of US states also receive notifications of breast density, a result of legislation intended to assist in making personalized decisions about further action. Dense breasts can mask cancer on mammography (masking bias), and are an independent cancer risk factor, but evidence does not yet indicate whether or what supplemental screening is appropriate. Rather, risk stratification is proposed to determine who may benefit from supplemental screening (eg, magnetic resonance imaging for women at high risk).[1,2]

Related article page 1784

The text of dense breast notifications (DBNs) may affect women's ability to understand their message. We examined DBN characteristics across states to inform future policy.

Methods | We reviewed the laws requiring DBNs for states with legislation effective as of January 1, 2016 (except Delaware, whose legislation language was not sufficiently detailed to analyze DBN content). In most states, the legislation specified the exact language for DBNs. We compared the content, readability, and understandability of DBNs across states. We noted the mandates and required recipients stated in the laws and whether the DBNs addressed masking bias, density as a cancer risk factor, and supplemental screening. We measured readability using the Flesch-Kincaid reading grade level in MS Word (range: theoretical lower bound, −3.4; no upper bound) and the Dale-Chall readability grade score (range, ≤4 to ≥16).[3] Understandability was assessed using the Patient Education Materials Assessment Tool (PEMAT; range, 1% to 100%).[4] We obtained the proportion of adults in each state lacking basic prose literacy skills from available statistics,[5] comparing DBN Flesch-Kincaid readability with state population literacy level.

Results | Twenty-four states require DBNs as of January 1, 2016; we analyzed all but Delaware. Most states (n = 19, 83%) mandate specific language (**Table**); 4 states (17%) only mandate required components. Seven states (30%) require a generic DBN for every woman receiving a screening mammogram, whereas all others only require notification to those with dense findings. All DBNs mention masking bias, 17 (74%) mention the association with increased cancer risk, and 15 (65%) mention supplemental screening as an option, advising women to consult their physician. Of 15 DBNs requiring mention of supplemental screening, 6 (40%) inform women that they might benefit from such screening; 4 mention specific modalities.

Flesch-Kincaid readability levels ranged from grades 7 to 19.4 (mean, 11.1), most exceeding the recommended readability level (grades 7-8); about 20% of the population reads below a grade 5 level.[5] Dale-Chall readability grade scoring[3] produced slightly higher scores overall (grade range: 9-10 to 13-15). All DBNs scored poorly on understandability (PEMAT; range, 11%-33%). There was widespread discordance between states' DBN readability and corresponding basic literacy levels (**Figure**). Only 3 states' DBN readability level was at the grade 8 level or below; some of the highest readability levels occurred in states with the lowest literacy levels.

Discussion | We found wide variation in 23 states' DBN content, with most having readability at the high school level or above, poor understandability, and discontinuity with states' average literacy. Such problems may create uncer-