



OPEN

Breast density analysis of digital breast tomosynthesis

John Heine^{1✉}, Erin E. E. Fowler¹, R. Jared Weinfurter², Emma Hume¹ & Shelley S. Tworoger¹

Mammography shifted to digital breast tomosynthesis (DBT) in the US. An automated percentage of breast density (PD) technique designed for two-dimensional (2D) applications was evaluated with DBT using several breast cancer risk prediction measures: normalized-volumetric; dense volume; applied to the volume slices and averaged (slice-mean); and applied to synthetic 2D images. Volumetric measures were derived theoretically. PD was modeled as a function of compressed breast thickness (CBT). The mean and standard deviation of the pixel values were investigated. A matched case-control (CC) study ($n = 426$ pairs) was evaluated. Odds ratios (ORs) were estimated with 95% confidence intervals. ORs were significant for PD: identical for volumetric and slice-mean measures [OR = 1.43 (1.18, 1.72)] and [OR = 1.44 (1.18, 1.75)] for synthetic images. A 2nd degree polynomial (concave-down) was used to model PD as a function of CBT: location of the maximum PD value was similar across CCs, occurring at $0.41 \times$ CBT, and PD was significant [OR = 1.47 (1.21, 1.78)]. The means from the volume and synthetic images were also significant [ORs ~ 1.31 (1.09, 1.57)]. An alternative standardized 2D synthetic image was constructed, where each pixel value represents the percentage of breast density above its location. Several measures were significant and an alternative method for constructing a standardized 2D synthetic image was produced.

Breast density is a significant and well accepted breast cancer risk factor assessed from mammograms^{1–4}. Areas of increased breast density (i.e., the degree of bright tissue) correspond to tissue with greater x-ray attenuation, as observed in mammograms used for clinical purposes. Breast density is one factor among others that could be considered in the development of breast cancer risk prediction models for clinical purposes. These models could be used for developing personalized healthcare strategies, such as setting risk-modulated screening intervals or imaging modality choice, providing the accuracy permits^{5,6}.

In the current clinical environment, there are several breast cancer risk models used for specific purposes^{4,7} for example: the Gail model⁸ is used to advise on chemoprevention for reducing risk; the Tyrer-Cuzick⁹, BRCAPRO^{10,11} and Claus¹² models are useful for determining if supplemental imaging with magnetic resonance might be beneficial⁴. The Breast Cancer Screening Consortium (BCSC) model^{13,14} may be useful for determining if women with dense breasts require supplemental screening¹³. These models do not use the same set of risk factors and are useful for different subpopulations¹⁵. For example, the BRACPRO model is used for predicting the risk of carrying a genetic mutation, and the Claus model is based on a family history of breast cancer. It is worth noting the Tyrer-Cuzick and BCSC models use breast density. Recently, the American College of Radiology (ACR) provided recommendations for supplemental imaging based on risk in conjunction with breast density¹⁶. It is also worth noting at present for screening in many high-income countries, risk is based primarily on age¹⁷. Risk prediction methods that use some type of image derived information (simple modeling through deep learning) show that texture may be an important factor, yet it is not commonly used in practice^{4,18,19}.

There are many methods under investigation for measuring both breast density and more generally texture^{19–22}. The percentage of breast density measure (PD) has been studied for many years and has repeatedly shown to be significantly associated with breast cancer risk^{23,24}. PD requires determining a threshold in a 2D mammogram. All pixels above this threshold are labeled as dense or otherwise labeled as non-dense creating a binary image. The final measure is derived from normalizing the dense area by the total breast area, presented as percentage. The ACR Breast Imaging Reporting & Data System (BI-RADS) four state ordinal breast composition classification²⁵ has also been used as a risk measure and shows consistent risk prediction capability across studies and time¹. These measures capture the volumetric tissue characteristics projected onto a plane. Newly derived image markers are often compared to PD as it has been considered as the de facto benchmark standard.

¹Cancer Epidemiology Department, Moffitt Cancer Center and Research Institute, 12902 Bruce B. Downs Blvd, Tampa, FL 33612, USA. ²Diagnostic Imaging and Interventional Radiology, Moffitt Cancer Center and Research Institute, 12902 Bruce B. Downs Blvd, Tampa, FL 33612, USA. ✉email: john.heine@moffitt.org

For example, recent work shows that PD produces a risk measure marginally superior to a commercially available volumetric breast density product when studying 2D full field digital mammography (FFDM) images²⁶.

Breast screening recently has largely transitioned from FFDM to digital breast tomosynthesis (DBT) in the US. DBT provides a *three-dimensional* (3D) rendering of the breast via stacked 2D images (slices) derived from 2D projection images acquired over a limited angular range. Clinical DBT images result from heavy processing. There is little published work with measures derived from clinical DBT images for risk factor purposes at this time^{19,27}. It is reasonable to assume that a more precise measure of breast density would result from analyzing volumetric images in comparison with conventional 2D mammograms. Accordingly, recent work in comparing an automated volumetric measure from DBT with measures applied to 2D mammograms shows improvements in risk prediction capability²⁸. Moreover, recent modeling using DBT data, incorporating various images features, illustrates it is possible to guide image care¹⁸.

We previously developed an automated PD approach (PD_a) that was evaluated in studies with digitized film and FFDM images^{29–31}. In this report, we will apply this method to DBT data³² and evaluate its merits for risk prediction. A matched case–control study was investigated with women that have DBT volume images (2D slices) and synthetic 2D mammograms (referred to C-View images in the mammography technology applicable to this report). There are three main study objectives. First, we investigated an algorithm modification used recently when studying relatively low-resolution digitized film mammograms³¹ and evaluated it with DBT without training or additional testing (i.e., a blind evaluation). Secondly, we derived different breast density measures from DBT volume data and compared these measures with PD determined from the C-View images in their risk prediction capabilities. Thirdly, PD was modeled as a function of compressed breast thickness and investigated.

Materials and methods

Study data

Study data was obtained from women participating in one of two studies collected under the same Institutional Review Board (IRB) of the University of South Florida, Tampa, FL (08/13/2007) selection protocol. Data was collected both retrospectively on a waiver for informed consent and prospectively with informed consent, both approved by the IRB referenced above. We investigated a matched case–control study (n = 426 pairs). Selection criteria and study population were described previously^{33,34}. Briefly, cases were women with first time pathologically verified unilateral breast cancer from two sources: (1) women attending the breast clinics at Moffitt Cancer Center (MCC) diagnosed with breast cancer, and (2) attendees of surrounding area clinics sent to MCC for either breast cancer treatment or diagnostic workup and found to have breast cancer. Controls were attendees of this center without history of breast cancer, verified by a two-year follow-up. Controls were individually matched to cases on these criteria: age (± 2 years), hormone replacement therapy (HRT) usage and current duration (never users or not current users in the prior 2 years, current user ± 2 years duration), screening history (any prior screening with time since last screening < 30 months, any prior screening > 30 months before baseline, no prior screening), and mammography unit (described below). Cranial caudal orientation mammograms were used for this study to reduce pectoral interference in the automated analyses. The unaffected breast image(s) were used for cases (all study images were acquired before treatment) and the matching lateral breast image for controls. Population characteristics are shown in Table 1.

Mammograms were acquired with Hologic Dimensions DBT units (Hologic, Inc., Bedford, MA): *volume* images (in-plane 100 μm average pixel spacing, 1 mm slice thickness, with 10-bit pixel dynamic range); and synthetic (i.e., C-View as named by this manufacturer) 2D images (100 μm average pixel spacing with 10-bit pixel dynamic range). The in-plane pixel spacing varies for DBT *volume* slices and C-View 2D images simultaneously and in tandem for each acquisition, ranging from approximately 85 μm to 106 μm (about 100 μm average) for our data. The number of slices in each DBT volume image is approximately one slice per mm of compressed breast thickness. It is important to note that DBT slices and the C-View images result from heavy processing unlike raw (*for processing*) FFDM images. Because this work involves creating an alternative 2D synthetic image, we use the manufacturer's C-View nomenclature when referring to the respective synthetic 2D images used for clinical purposes and synthetic when referring to the experimental images produced in this report to avoid confusion.

Automated PD detection algorithm

Our automated PD detection mechanism (referred to here as, PD_a) has been under investigation for many years, necessitated by both changing mammography technologies and implementing algorithm improvements based on our image-understanding^{29,30,35–37}. The methodology operates by analyzing signal dependent noise locally in 2D images²⁹. We refer to this approach as a *detection algorithm* due to the way it makes systematic (probabilistic) local decisions to identify dense tissue. PD_a is a two-stage detection algorithm that first determines a global reference variance for adipose tissue (over the entire breast area) in a high pass wavelet filtered version of the image. In the first detection stage, this reference variance is used for making statistical comparisons with local variances determined from a $n \times n$ box ($n = 4$) scanned systematically across the breast area. Local regions that deviate significantly from this reference are labeled as dense or otherwise non-dense by default. In the second detection stage, the reference adipose variance is refined by estimating it from the non-dense areas identified in the first stage. The localized comparisons are then repeated with the refined reference resulting in the PD_a labeled image. Each detection stage requires a threshold defined by a significance level selected a priori from a Chi-square distribution.

In previous work, we noted applying a non-linear transform to the images before performing the density detection process had a beneficial impact on the algorithm's output^{29,36,38}. Through experimentation with 2D *for presentation* images from General Electric Senographe 2000D (Milwaukee, WI) [180 case–control pairs] and Hologic Selenia units [320 case–control pairs], we found that a 0.1 significance level was robust for both detection

Measure or characteristic	p value	Case n	Case mean, SD, or relative frequency	Control n	Control mean, SD, or relative frequency	Total n	Total mean, SD, or relative frequency
Age*	0.7787	426	58.2 (11.6)	426	58.2 (11.6)	852	58.2 (11.6)
BMI	<0.0001	426	29.6 (6.9)	424	27.7 (6.3)	850	28.7 (6.7)
Screening group*	N/A						
Group 1		262	61.50%	262	61.50%	524	61.50%
Group 2		87	20.42%	87	20.42%	174	20.42%
Group 3		77	18.08%	77	18.08%	154	18.08%
HRT usage*	N/A						
Current		36	8.45%	36	8.45%	72	8.45%
Not currently		390	91.55%	390	91.55%	780	91.55%
Race							
Caucasian	0.5394	340	79.81%	348	81.69%	688	80.75%
African American	1.0000	47	11.03%	46	10.80%	93	10.92%
Asian	0.5716	16	3.76%	12	2.82%	28	3.28%
Other	1.0000	12	2.82%	12	2.82%	24	2.82%
More than one	0.0703	7	1.64%	1	0.23%	8	0.94%
Unknown	0.5488	4	0.94%	7	1.64%	11	1.29%
Ethnicity							
Non-Hispanic		369	86.62%	323	75.82%	692	81.22%
Hispanic		57	13.38%	98	23.00%	155	18.19%
Unknown		0	0.00%	5	1.17%	5	0.59%
Menopausal status							
Pre-menopausal	1.0000	114	26.70%	111	26.06%	225	26.41%
Post-menopausal	0.7877	311	73.13%	311	73.00%	622	73.00%
Unknown	0.3750	1	0.17%	4	0.94%	5	0.59%

Table 1. Population characteristics: participant numbers (n) are broken down by case–control status and totals. Where applicable, distribution means and standard deviations (SDs) are provided. Matching factors are indicated by asterisks.

stages [i.e., determined from images that were processed in some way after the acquisition process unlike raw (*for processing* images)] by first multiplying mammograms with random noise (zero-mean unit variance normally distributed) before applying the high pass wavelet filter (i.e., the same *trick* used for the low-resolution film analysis). This extra step appears to boost the signal dependent noise signal in the high pass wavelet image; understanding this mechanism is a separate investigation underway. DBT data has undergone heavy processing and has relatively greater pixel spacing (i.e., reduced resolution) than Hologic 2D FFDM images used in this study. Therefore, we applied this noise multiplication modification to all DBT images investigated in this report and performed the detection with the preset detection parameters. As part of *this blind investigation* (verification), detection thresholds (both stages) were set using a significance level = 0.1 with 15 degrees of freedom (i.e., $n^2 - 1$) without analyzing DBT data with the modified PD_a methodology a priori. Because there is difficulty in determining ground truth for breast density, we will use the statistical significance of the OR findings as the objective endpoint benchmarks.

Breast density measurement modeling with DBT

Multiple breast density measures can be derived from PD measurements when applied to DBT images. At the local region within a slice (or any image), we assume that the density detection process is a surrogate for approximating the probability that the region has the x-ray attenuation coefficient of dense breast tissue, and the dense tissue acceptance (detection) is based on this probability (i.e., referred to as the probability conjecture below).

For the volume derivations, we let a given DBT volume image have N slices with an isotropic thickness, t, measured in mm and pixel area given by A measured in mm^2 . The i^{th} slice has n_i pixels in the breast area with d_i pixels labeled as dense. The expression for the breast volume (BV), required in the derivations, is given by

$$BV = t \times A \sum_{i=1}^N n_i = c \times N \times \langle n_i \rangle, \quad (1)$$

where the brackets indicate the average (expectation) operator, $\langle n_i \rangle$ is the average number of pixels over the slice breast areas, and $c = t \times A$. As a measure, the total dense tissue volume (D_v) within the breast volume is given by

$$D_v = t \times A \sum_{i=1}^N d_i = c \times N \times \langle d_i \rangle, \quad (2)$$

where $\langle d_i \rangle$ is average of PD taken over the N slices. Using the BV and D_v expressions, the volumetric PD measure (PD_{vol}) is given by

$$PD_{vol} = 100\% \times \frac{D_v}{BV} = 100\% \times \frac{\langle d_i \rangle}{\langle n_i \rangle}. \quad (3)$$

Another metric can be derived by averaging PD over the N slices (PD_m) giving

$$PD_m = \frac{100\%}{N} \times \sum_{i=1}^N \frac{d_i}{n_i}. \quad (4)$$

Equation (4) is an indication of why the normalization for PD may be important, as follows. Making the approximation that the breast area in each slice is constant for a given woman with n pixels gives

$$PD_m \approx \frac{100\%}{N} \times \frac{\sum_{i=1}^N d_i}{n} = 100\% \times \frac{N \times \langle d_i \rangle}{Nn} = 100\% \times \frac{\langle d_i \rangle}{\langle n_i \rangle}, \quad (5)$$

where $n = \langle n_i \rangle$. When the breast areas are similar in each slice, $PD_{vol} \sim PD_m$; this approximation will be evaluated. PD from the labeled slices can be projected (summed) onto a plane giving a (coarse) 2D image with pixel values ranging between zero and N . The summation of pixel values within this image gives the projected total (PT) expressed as

$$PT = N \times \langle d_i \rangle. \quad (6)$$

If we assume the traditional PD thresholding in 2D captures the dense pixel proportion above a given location within the compressed breast, the total number of dense pixels in a 2D labeled image is approximately the normalized projected total (NPT) determined by dividing Eq. (6) by N giving

$$NPT = \langle d_i \rangle. \quad (7)$$

Given there are n pixels within the breast area, the standardized 2D measure is approximated by

$$PD \approx 100\% \times \frac{NPT}{n} = 100\% \times \frac{\langle d_i \rangle}{n}, \quad (8)$$

which is just Eq. (3) or Eq. (5) relabeled. The projected image normalized by N and multiplied by 100% produces a standardized synthetic image [defined as $s(x,y)$], where each pixel represents the percentage of dense tissue volume above its location (using a parallel beam approximation); we note, these interpretations follow from the probability conjecture defined above. In these derivations, image parameters (A and t) were not relevant except for D_v . When assessing D_v across women, t is constant while A will vary and must be accounted for in the metric.

We investigated odd ratios (ORs) produced when applying PD_a to the C-View images (PD_{syn}) and when producing PD_{vol} and PD_m . ORs produced by analyzing PD from the isolated central DBT slice were also analyzed (i.e., exploring the possibility that PD from one slice may be representative of the volume). As further comparators, we evaluated the mean and standard deviation of the pixel values within the DBT volume without PD_a processing, referred to as m_{vol} and v_{vol} . Likewise, we used the mean from the C-view image pixel values (m_{syn}) as another comparator.

To study breast density characteristics throughout the DBT volume, we used an empirically driven second-degree polynomial model expressed as

$$PD = a + b \times P + c \times P^2, \quad (9)$$

where PD is from each slice, P is the normalized independent slice number variable ranging from [1, 100] measured as a unitless proportion ranging from the slice at the breast support surface ($P = 1$) to the furthest slice from the support surface ($P = 100$), and (a, b, c) are parameters to be determined with curve fitting analysis. The slice distance from the breast support surface can be recovered approximately by $\frac{P}{100} \times (\text{compressed breast thickness})$. The convention used for increasing P follows that of the clinical volume rendering from the image header files. This normalization for distance puts the fit parameters on the same scale over all participants. We investigated parameter distributions (empirical normalized histograms) and made comparisons across case-control status. The normalized distance for the maximum PD quantity can be derived by setting the derivative of Eq. (9) to zero giving

$$P = -\frac{b}{2 \times c}, \quad (10)$$

which was investigated.

Statistical analyses

Image measure associations with breast cancer were evaluated using conditional logistic regression while controlling for body mass index (BMI) and ethnicity. Unadjusted models are included in the tables for completeness. We used ORs with 95% confidence intervals (CIs) as the primary metric to evaluate and compare breast cancer associations between the various image measures defined above. ORs were estimated for continuous measures in per standard deviation (SD) increment. We note, an OR derived from a case-control study from a given population

is often used as an approximation for relative risk for the same population, providing the disease incidence is small. Although building predictor models is not the purpose of this report, for completeness the area under the receiver operator characteristic curve (Az) was calculated with 95% CIs to summarize the discriminatory ability for each model. Both ORs and Azs are presented with CIs parenthetically. The matching design in this case–control study was implemented specifically to isolate the associations of image measures with breast cancer risk and generally precludes developing risk prediction models, which require detailed information regarding selection probabilities. Pearson correlation coefficient (R) was used to show the linear relations between select breast density measures and BMI.

Ethics and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations. All experimental procedures were approved by the Institutional Review Board (IRB) of the University of South Florida, Tampa, FL under protocol #Ame13_104715. Mammography data was collected both retrospectively on a waiver for informed consent and prospectively with informed consent both approved by the IRB referenced above.

Results

Population characteristics

Table 1 shows the characteristics of the case–control participants. As expected, matching variables (age, screening, HRT) were similar across case–control groups. Similarly, neither race (Caucasian, African American, or Asian) or menopausal status varied significantly by status. Menopausal status did not vary significantly as it is likely a surrogate for age. In contrast, both ethnicity (Hispanic, non-Hispanic) and BMI (larger for cases) varied significantly. The BMI finding is expected, as increased risk is associated with increased BMI³⁹. The difference in ethnicity is due to the shifting demographics at our clinics.

Illustrations and related analyses

Several illustrations are used to show image data and the detection algorithm's output. Figure 1 (top row) shows images from the two participants selected at random from left to right: (a) C-View image for illustration-1; (b) central DBT slice for illustration-1; (c) C-View image for illustration-2; and (d) central DBT slice for illustration-2. Illustration-1 has 89 μm pixel spacing and illustration-2 has 107 μm pixel spacing. The largest rectangle

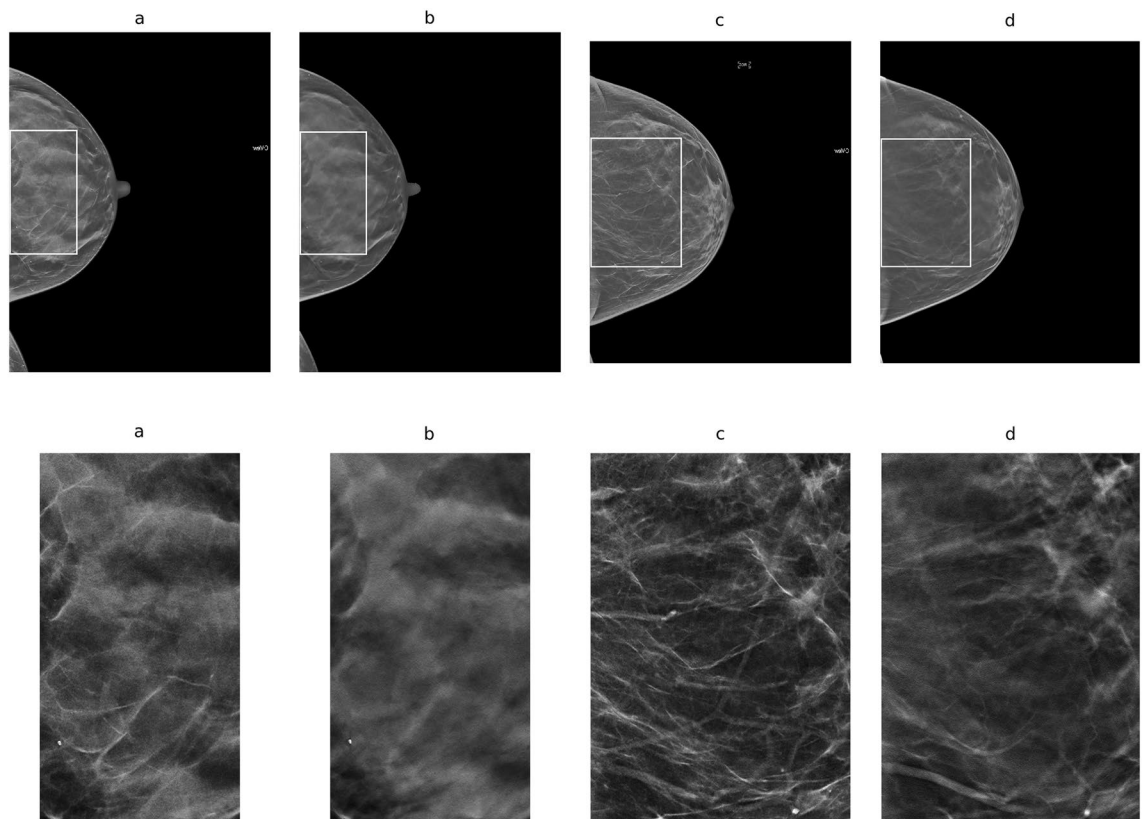


Figure 1. Image Illustrations: in the top row (a) C-View image for illustration-1; (b) respective central slice image; (c) C-View image for illustration-2; and (d) respective central slice image. Outlines in the top-row images are the largest rectangles that fit within the breast areas. These regions are shown in the second row with the same ordering for illustration purposes. Illustration-1 has 89 μm pixel spacing and illustration-2 has 107 μm pixel spacing.

that fits within the breast area³⁴ is outlined in each image. These regions are used for improved viewing purposes and are shown in the second row of Fig. 1. Although the breast structure in the volume slices track that of the C-View images, it appears less resolved. The density-detected images are shown in Fig. 2. These show the density labeling in the C-view images also tracks that of the labeling in the central volume slices. Figure 3 shows the synthetic $s(x,y)$, images produced by the PD slice projections for these illustrations (from their labeled volumes). These appear as processed images (*for presentation*) but with level contrast as they are 8-bit images with limited dynamic range. These images illustrate another technique for creating synthesized 2D images in comparison with the C-View type images.

Measurement modeling

Equations (5)–(8) show PD_{vol} and PD_m should be equivalent under the breast area similarity approximation. Figure 4 shows the scatter plots of these two measures (points) and the fitted regression line (solid red) with $R \approx 1.0$, slope ≈ 1.002 , and intercept ≈ -0.0201 indicating the two measures are essentially *identical*. These findings show that both Eq. (8) and the interpretation of the synthesized images shown in Fig. 3 are valid. DBT slice modeling using Eq. (9) is shown in Fig. 5. There is also a cluster of PD quantities similar to the maximum in close slice proximity in both examples. Figure 6 shows the histograms for the Eq. (9) coefficients separated by case–control status. Averaging like coefficients for cases and controls gave: $(a, b, c)_{mean} \approx (21.1, 0.06, -0.008)$ and $(a, b, c)_{mean} \approx (20.8, 0.07, -0.0008)$, respectively. Applying a t-test across groups showed only the intercept (i.e., a) varied marginally (p value ≈ 0.046), as expected. Using parameter-means, the position with the greatest PD finding from Eq. (10) was approximately $P \approx 42$ for either group. Empirically the mean maximum was $P \approx 40.7$ with a standard error $\approx \pm 0.28$, showing close agreement with the model. Considering these findings, we investigated the maximum PD value as another measure.

Breast cancer risk associations

Breast cancer associations are shown in Table 2. Both PD_{syn} [OR = 1.44 (1.18, 1.75)] and PD_{vol} [OR = 1.43 (1.18, 1.72)] were significant. ORs from PD_m were identical to PD_{vol} . PD from the central slice was significant [OR = 1.42 (1.17, 1.72)] and similar to the largest PD quantity [OR = 1.47 (1.21, 1.78)]. The mean of DBT volume pixels (m_{vol}) was also significant [OR = 1.31 (1.09, 1.57)] and similar to m_{syn} [OR = 1.29 (1.10,

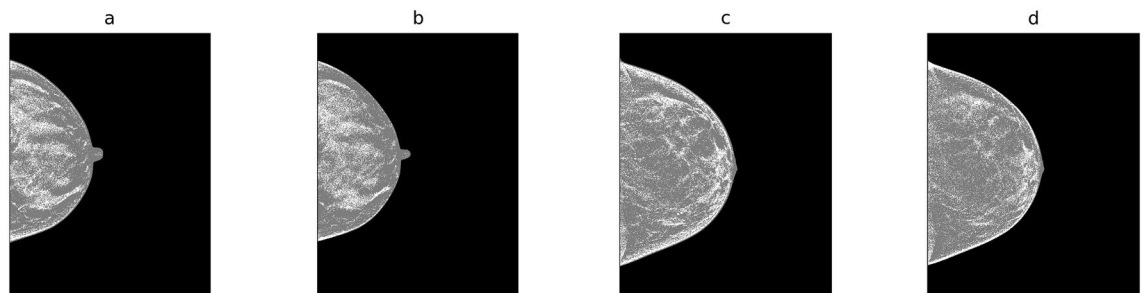


Figure 2. Breast Density Detection: this shows the density detection for the illustrations: (a) C-view, illustration-1; (b) respective central slice image; (c) C-view, illustration-2; and (d) respective central slice.

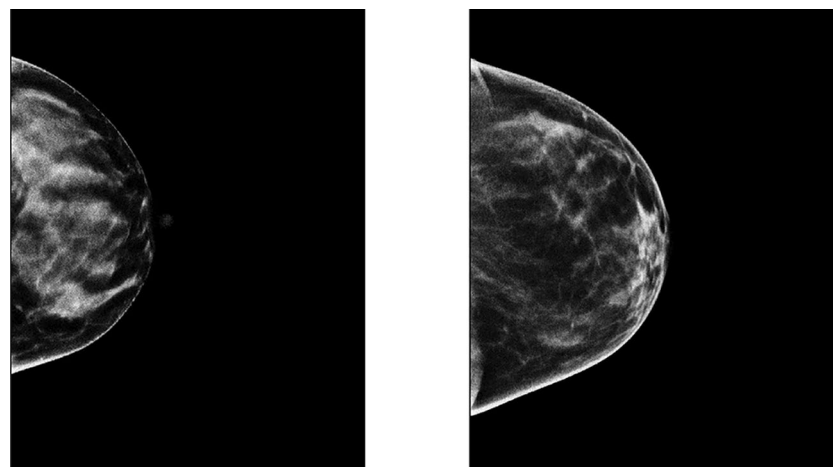


Figure 3. Projected Standardized Synthetic Breast Density Images: these show the standardized, $s(x,y)$, images for illustration-1 (left) and illustration-2 (right) resulting from Eq. (8). Pixel values represent the percentage of dense tissue in the breast volume above their locations.

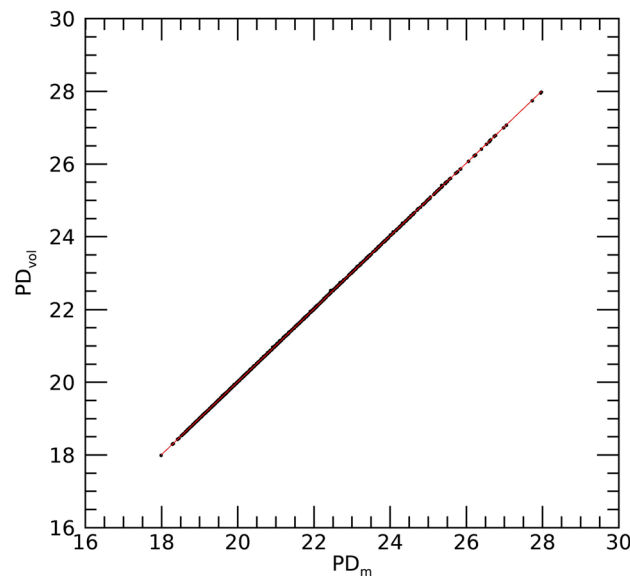


Figure 4. Regression analysis with PD_{vol} as a function of PD_m . This shows the scatter plot between the two measures (points) and regression line (solid red). The analysis gave: slope ≈ 1.002 , intercept ≈ -0.0201 , and linear correlation ≈ 1.0 . The [mean, standard deviations] for the distributions were [21.60, 1.98] for PD_{vol} and [21.58, 1.98] for PD_m .

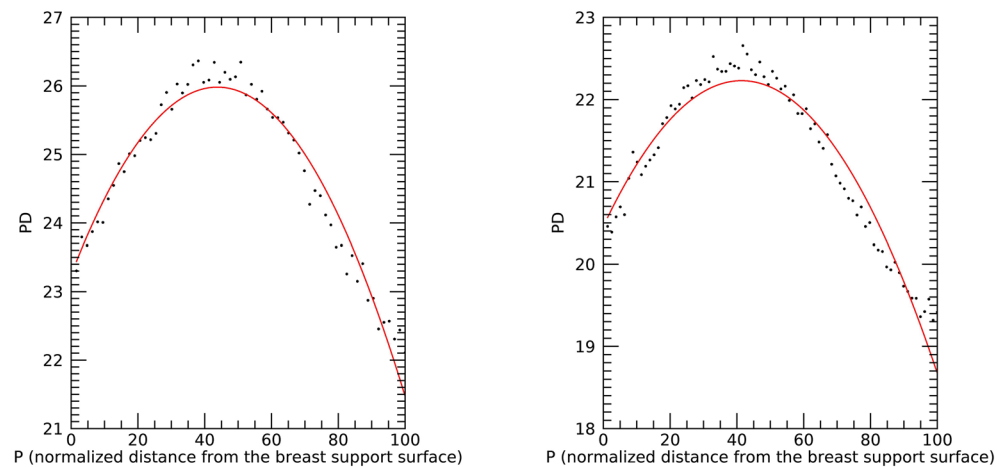


Figure 5. PD slice Profiles: this shows PD (y-axis) by slice number using the normalized distance from the breast support surface (P on the x-axis) for illustration-1 (left) and illustration-2 (right). PD values per slice number (points) were fitted with a second-degree polynomial (solid). The slice distance increases as the distance increases from the breast support surface (left side of each plot).

1.52)]. Neither the total dense volume (D_v) or the standard deviation within the volume (v_{vol}) produced significant associations. We also compared PD from $P = 1$ (closest to breast support surface) and $P = 100$ (compression paddle) [not shown]; these provided significant breast cancer associations that were similar to either m_{syn} or m_{vol} . When comparing the unadjusted to adjusted models in Table 2, the ORs for the BD measurements shifted considerably. Therefore, we investigated the correlation between BMI and the three main findings by first removing BMI outliers giving: $R = [PD_{syn}, PD_{vol}, PD_m] = [-0.43, -0.33, -0.33]$. Although this range of correlation is weak, it explains the confounding influence of BMI on BD measurements; the risk of breast cancer increases as either BD or BMI increases while these two factors move in opposition.

The OR findings above were similar for PD_{vol} and PD_{syn} . For risk factor purposes, this implies analyzing the C-View images (i.e., the breast volume structure projected onto a plane with heavy processing) is not subordinate to analyzing the volume images. To study these measures further, we investigated their relationship with linear regression. Figure 7 shows the scatter plot (points) and regression analysis (solid line): slope ≈ 0.92 , intercept ≈ -1.4 , and $R \approx 0.93$. Because the slope is close to unity, the intercept is not far from $PD_{vol} = 0$, and the strong positive correlation, these measures are approximately on the same scale and similar, although the plot does show

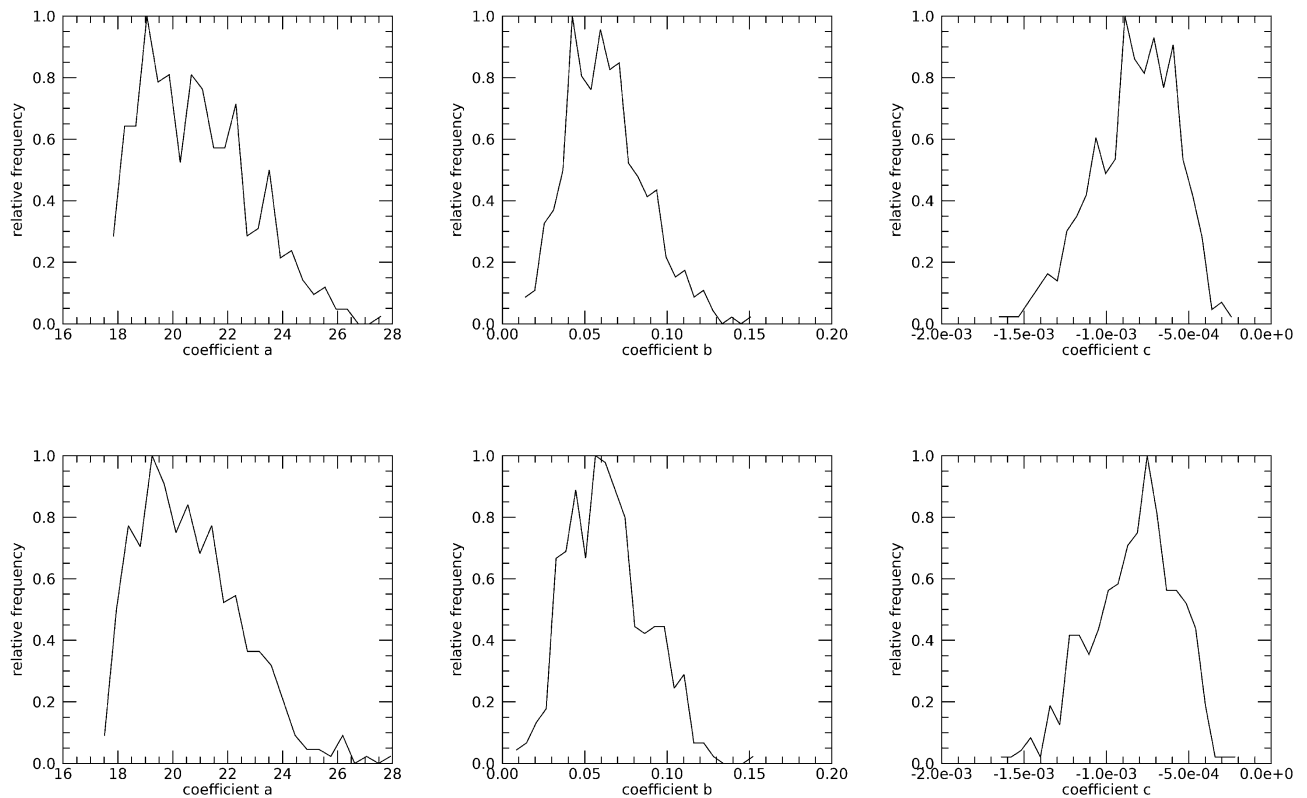


Figure 6. DBT Volume Slice 2nd Degree Polynomial Coefficient Histograms: these show the normalized histograms for the fit-coefficients, (a,b,c), separated by case (top-row) and control status (bottom-row).

nonlinear variation. Although the variation between these two measures increases as the respective measures increase, these regression findings assist in explaining the OR similarities found above.

Discussion

The study demonstrated the validity of the density detection algorithm's capability of translating to DBT by producing significant ORs. The volumetric measure was equivalent to the average PD values taken across the DBT slice images and agreed with the derivation showing these measures are approximately equivalent. Three other findings were notable as well: (1) the DBT slice with the largest PD finding was offset considerably from the central slice; (2) PD from the central slice, from the slice with maximum PD finding, or from the C-View image provided ORs similar to those from volumetric measure, and (3) the mean of the pixel values from the DBT volume slices or from the C-View images produced significant ORs without applying the detection processing. Where applicable, reference will be made to each of these three other notable findings in the following paragraphs with their respective finding number.

The PD slice profiling analysis is another novel aspect of this report. The related plots (Fig. 5) showed clusters of points (around 5–10 slices) with values close to the maximum PD value occurring around the curvature crest indicating why these isolated slices (finding number 2) provided similar ORs. We believe this is the first study to represent PD in this slice profile (derived from images that represent a volumetric rendering of the breast from x-ray technology). Other work that compared various PD-type measures (using FFDM) with a commercial volumetric breast density product that operates on 2D raw mammograms did not find large differences in ORs across the measures⁴⁰. Our work agrees with these findings.

Breast cancer ORs between the volume and the synthetic images were almost identical; this indicates there is no benefit derived from analyzing the volume directly, admitting the C-View image is derived from the volume. This finding applies to our method specifically but agrees with volumetric measures derived from 2D FFDM as follows. Comparisons with other techniques are often not exact due to study design variations such as sample size and model differences. Likewise, there is not an accepted convention for the standard deviation increment in the image measurement, which is distribution dependent for each image measure. However, the ORs for PD found in the report parallel results in other work to varying degrees: (1) agrees with those determined in these reports^{24,38}; (2) are similar to those determined with volumetric measures⁴¹ (derived from conventional 2D mammograms); and (3) and are marginally less than a volumetric measure applied to DBT²⁸, and it is worth noting that this DBT approach first operates on the 2D projection images, then applies machine learning to the reconstructed volume, and as evaluated had relatively few cancer observations. Finding number 3 above follows intuition as larger pixel values represent elevated levels of dense breast tissue. In the past, we have found the variation in conventional 2D mammograms provided significant ORs^{41–43}, which did not hold in this study for the volume images. In this

n = 426 (pair)	SD	Unadjusted OR (95% CI) and Az (95% CI)	BMI and ethnicity adjusted OR (95% CI) and Az (95% CI)
PD _{syn}			
OR	0.08	1.08 (0.92, 1.26)	1.44 (1.18, 1.75)
Az		0.51 (0.47, 0.56)	0.65 (0.61, 0.70)
D _v			
OR	0.5513	1.23 (1.07, 1.42)	1.00 (0.82, 1.22)
Az		0.55 (0.51, 0.60)	0.63 (0.58, 0.67)
PD _{vol}			
OR	0.0902	1.14 (0.97, 1.33)	1.43 (1.18, 1.72)
Az		0.53 (0.48, 0.57)	0.64 (0.60, 0.69)
PD _m			
OR	0.0902	1.14 (0.97, 1.33)	1.43 (1.18, 1.72)
Az		0.52 (0.48, 0.57)	0.65 (0.60, 0.69)
m _{vol}			
OR	0.00878	1.26 (1.06, 1.50)	1.31 (1.09, 1.57)
Az		0.55 (0.50, 0.59)	0.63 (0.59, 0.68)
v _{vol}			
OR	0.5582	1.13 (0.89, 1.42)	1.19 (0.93, 1.53)
Az		0.51 (0.46, 0.56)	0.64 (0.60, 0.69)
m _{syn}			
OR	1.0168	1.23 (1.05, 1.43)	1.29 (1.10, 1.52)
Az		0.56 (0.51, 0.60)	0.65 (0.60, 0.69)
PD (central slice)			
OR	0.0923	1.11 (0.95, 1.30)	1.42 (1.17, 1.72)
Az		0.52 (0.48, 0.57)	0.66 (0.61, 0.70)
PD (maximum slice)			
OR	0.0932	1.13 (0.97, 1.33)	1.47 (1.21, 1.78)
Az		0.53 (0.48, 0.57)	0.66 (0.61, 0.70)

Table 2. Conditional Logistic Regression Modeling Results: this gives odd ratios (ORs) with 95% confidence intervals (CIs) parenthetically for each model; image measures were log-transformed. Standard deviations (SDs) are provided for log-transformed distributions. The area under the receiver operating characteristic curve (Az) is given for each model with CIs parenthetically. From top to bottom: PD_{syn} is from C-View image analysis; D_v is the dense volume (mm³); PD_v is normalized volumetric PD; PD_m is the slice-mean; m_{vol} is the mean of the pixel values within the DBT volume (no processing); v_{vol} is the standard deviation of the pixel values within the DBT volume (no processing); and m_{syn} is the mean of the C-view image pixel values (no processing); PD from the central slice in the DBT volume at P ≈ 50; and PD (maximum slice) determined from the DBT slice with the largest PD finding.

current study, D_v was not a significant risk factor but is the critical factor in the other measures. A significant OR was produced when normalizing D_v by the total breast volume, which supports the probability conjecture. Thus, the study provides insight into the nature of the traditional PD measurement (applied to 2D mammograms) and produced a related prescription for constructing a standardized synthesized 2D image.

Our findings can be summarized into two areas of investigation: (1) PD measurement development and validation; and (2) the nature of the anatomical volumetric distribution of PD (ORs and slice modeling), which may have biological importance (unknown at this time). We believe both areas represent new findings. As for measurement development, there are no universally accepted measures of breast density for 2D (let alone DBT) for multivariate risk prediction models in general. However, there are trends in this direction clinically. The standard measure for breast density in the US used for clinical reporting is the BI-RADS ordinal composition classification provided by the attending radiologist, originally developed for masking, or indicating when mammography may be ineffective. This measure is also used for risk prediction in both the BCSC^{13,14} and Tyrer-Cuzick models. The Tyrer-Cuzick model (and other models including the BCSC and Claus models) is also available in a widely used commercial mammography reporting software product (<https://magview.com/risk-assessment/>) to identify high-risk women within the radiology workflow. For DBT, the ACR recommends making the BI-RADS tissue assessments from either the synthetic 2D images (i.e., C-View images in this study) or the accompanying 2D FFDM images (supplement to BI-RADS lexicon 5th edition, 2013). It is also worth noting, in this 5th edition the tissue composition categories changed by dropping the quantitative component of the reporting due to reproducibility problems with volumetric measurements⁴⁴. At this point, is not clear if a conventional measurement of BD, such as PD (studied here) translated to DBT, or if more involved methods derived from artificial intelligence⁴⁵ will provide benefits if incorporated into risk prediction models used for clinical purposes beyond that provided by the BI-RADS measure because it is one measurement among many factors. Mammography

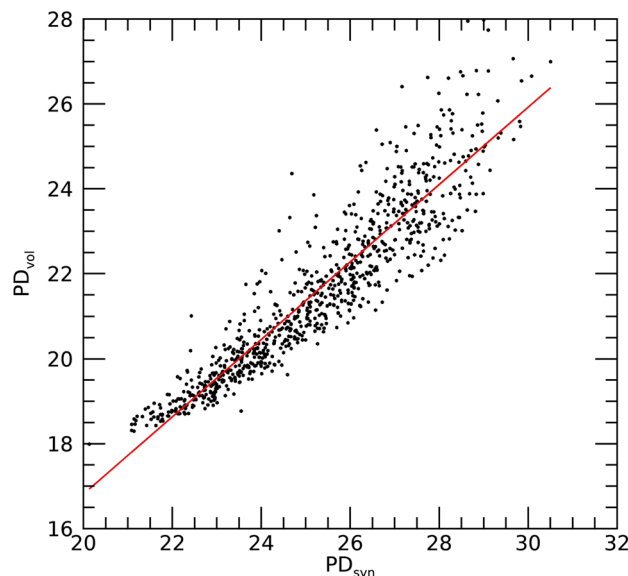


Figure 7. Regression analysis with PD_{vol} as a function of PD_{syn} : this shows the scatter plot between the two measures (points) and regression line (red-solid): slope ≈ 0.92 and standard error ≈ 0.01 , intercept ≈ -1.4 , and $R \approx 0.93$. The [mean, standard deviations] for the distributions were [21.60, 1.98] for PD_{vol} and [25.25, 2.01] for PD_{syn} .

imaging technology shifts can occur rapidly (discussed below) and are naturally ahead of the automated breast density measurement advancements. We could also posit the possibility that the available information within a mammogram of any type related to risk is somewhat limited and measurement reproducibility over longer timeframes is a critical measurement attribute for clinical applications.

There are several comments worth noting about this study. Absolute ground truth for volumetric breast density for an arbitrary breast was not known. Ideally, comparing our findings with breast phantoms designed for DBT with known mixtures of adipose and fibro-glandular tissue would be beneficial. In 2D (FFDM raw images), pixel values approximate the attenuated signal. Unfortunately, this characteristic is not preserved in the volume slices or 2D synthetic images due to the processing required for their construction. In any event, such phantoms would require realistic breast tissue spatial distributions for our approach to operate optimally due to its localized detection characteristic; this would preclude using more uniform type phantoms. To the best of our knowledge, the development of realistic anthropomorphic breast phantoms for DBT is a challenging problem and open area of research^{46,47}. We have shown agreement ($R > 0.7$) between our PD measure and a calibrated phantom approach with FFDM images, where pixel values were mapped to standardized values³⁶. We also found similar correlations when making comparisons with the operated-assisted PD method applied to both digitized film³⁰ and FFDM²⁹ images. Although not on the same scale, the monotonic relationship that our measure has with these other measurements is likely an essential attribute responsible for its risk prediction replication characteristic. Additionally, we analyzed a hospital-based population, where matching was used to account for case-control differences. Both the OR findings and summaries from Table 1 indicate this did not materially influence the outcomes. PD derivations are general and apply to any like metric, whereas the findings in this report apply specifically to our automated approach. We only analyzed cranial caudal mammograms, indicating we may have missed density information from the mediolateral views. Although the results indicated that 2D and 3D measures from PD were similar, the study design establishes a template that could be used for investigating other measures such as texture. Study images were from one type of DBT technology. It is worth noting, DBT technology is also shifting. For instance, the manufacturer of the units used for this study is now offering artificial intelligence enhanced images for DBT, smaller pixel spacing, and interleaved slice spacing (increased). The noise field multiplication modification analyzed here offers potential to apply to images derived from evolving DBT advances. The results from this study will require verification in other populations and DBT technologies as well.

Data availability

Mammography data can be obtained upon request to the corresponding author (JH, john.heine@moffitt.org).

Received: 24 August 2023; Accepted: 19 October 2023

Published online: 31 October 2023

References

1. Bodewes, F. T. H., van Asselt, A. A., Dorrius, M. D., Greuter, M. J. W. & de Bock, G. H. Mammographic breast density and the risk of breast cancer: A systematic review and meta-analysis. *Breast* **66**, 62–68. <https://doi.org/10.1016/j.breast.2022.09.007> (2022).

2. Boyd, N. F. *et al.* Breast tissue composition and susceptibility to breast cancer. *J. Natl. Cancer Inst.* **102**, 1224–1237. <https://doi.org/10.1093/jnci/djq239> (2010).
3. Bertrand, K. A. *et al.* Mammographic density and risk of breast cancer by age and tumor characteristics. *Breast Cancer Res.* **15**, R104 (2013).
4. Butler, R. Invited commentary: Breast cancer risk assessment and screening strategies-what's new?. *Radiographics* **40**, 937–940. <https://doi.org/10.1148/rg.2020190218> (2020).
5. Roman, M. *et al.* Personalized breast cancer screening strategies: A systematic review and quality assessment. *PLoS One* **14**, e0226352 (2019).
6. Louro, J. *et al.* A systematic review and quality assessment of individualised breast cancer risk prediction models. *Br. J. Cancer* **121**, 76–85 (2019).
7. DBI. *DenseBreast-info*. <https://densebreast-info.org/>.
8. Costantino, J. P. *et al.* Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J. Natl. Cancer Inst.* **91**, 1541–1548 (1999).
9. Brentnall, A. R. & Cuzick, J. Risk models for breast cancer and their validation. *Stat. Sci. Rev. J. Inst. Math. Stat.* **35**, 14 (2020).
10. Mazzola, E., Blackford, A., Parmigiani, G. & Biswas, S. Recent enhancements to the genetic risk prediction model BRCAPRO. *Cancer Inform.* **14**, CIN. S17292 (2015).
11. Parmigiani, G., Berry, D. A. & Aguilar, O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am. J. Hum. Genet.* **62**, 145–158 (1998).
12. Claus, E. B., Risch, N. & Thompson, W. D. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* **73**, 643–651 (1994).
13. Tice, J. A. *et al.* Validation of the breast cancer surveillance consortium model of breast cancer risk. *Breast Cancer Res. Treat.* **175**, 519–523 (2019).
14. Vachon, C. M. *et al.* The contributions of breast density and common genetic variation to breast cancer risk. *J. Natl. Cancer Inst.* **107**, dju397 (2015).
15. Gail, M. H. Vol. 112 433–435 (Oxford University Press, 2020).
16. Weinstein, S. P. *et al.* ACR appropriateness criteria(R) supplemental breast cancer screening based on breast density. *J. Am. Coll. Radiol.* **18**, S456–S473. <https://doi.org/10.1016/j.jacr.2021.09.002> (2021).
17. Pashayan, N. *et al.* Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat. Rev. Clin. Oncol.* **17**, 687–705. <https://doi.org/10.1038/s41571-020-0388-9> (2020).
18. Eriksson, M. *et al.* A risk model for digital breast tomosynthesis to predict breast cancer and guide clinical care. *Sci. Transl. Med.* **14**, eabn3971. <https://doi.org/10.1126/scitranslmed.abn3971> (2022).
19. Mendes, J. & Matela, N. Breast cancer risk assessment: A review on mammography-based approaches. *J. Imaging* **7**, 98 (2021).
20. Gastouniotti, A., Conant, E. F. & Kontos, D. Beyond breast density: A review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. *Breast Cancer Res.* **18**, 91 (2016).
21. Kontos, D. *et al.* Radiomic phenotypes of mammographic parenchymal complexity: Toward augmenting breast density in breast cancer risk assessment. *Radiology* **290**, 41–49. <https://doi.org/10.1148/radiol.2018180179> (2019).
22. Chen, J.-H., Gulsen, G. & Su, M.-Y. Imaging breast density: Established and emerging modalities. *Transl. Oncol.* **8**, 435–445 (2015).
23. Boyd, N. F., Martin, L. J., Yaffe, M. J. & Minkin, S. Mammographic density and breast cancer risk: Current understanding and future prospects. *Breast Cancer Res.* **13**, 223. <https://doi.org/10.1186/bcr2942> (2011).
24. Pettersson, A. *et al.* Mammographic density phenotypes and risk of breast cancer: A meta-analysis. *J. Natl. Cancer Inst.* <https://doi.org/10.1093/jnci/dju078> (2014).
25. D'Orsi, C. J. *et al.* ACR BI-RADS Atlas: Breast Imaging Reporting and Data System; Mammography, Ultrasound, Magnetic Resonance Imaging, Follow-up and Outcome Monitoring. Data Dictionary (ACR, American College of Radiology, 2013).
26. Jeffers, A. M. *et al.* Breast cancer risk and mammographic density assessed with semiautomated and fully automated methods and BI-RADS. *Radiology* **282**, 348–355. <https://doi.org/10.1148/radiol.2016152062> (2017).
27. Geras, K. J., Mann, R. M. & Moy, L. Artificial intelligence for mammography and digital breast tomosynthesis: Current concepts and future perspectives. *Radiology* **293**, 246–259 (2019).
28. Gastouniotti, A. *et al.* Fully automated volumetric breast density estimation from digital breast tomosynthesis. *Radiology* **301**, 561–568. <https://doi.org/10.1148/radiol.2021210190> (2021).
29. Fowler, E. E., Vachon, C. M., Scott, C. G., Sellers, T. A. & Heine, J. J. Automated percentage of breast density measurements for full-field digital mammography applications. *Acad. Radiol.* **21**, 958–970. <https://doi.org/10.1016/j.acra.2014.04.006> (2014).
30. Heine, J. J. *et al.* An automated approach for estimation of breast density. *Cancer Epidemiol. Biomark. Prev.* **17**, 3090–3097 (2008).
31. Warner, E. T. *et al.* Automated percent mammographic density, mammographic texture variation, and risk of breast cancer: A nested case-control study. *NPJ Breast Cancer* **7**, 68. <https://doi.org/10.1038/s41523-021-00272-2> (2021).
32. Heine, J., Fowler, E. E., Weinfurter, R. J., Tworoger, S. & Hume, E. Breast density analysis of digital breast tomosynthesis. *bioRxiv*, 2023.2002.2010.527911 (2023).
33. Fowler, E. E. *et al.* Spatial correlation and breast cancer risk. *Biomed. Phys. Eng. Express* **5**, 045007. <https://doi.org/10.1088/2057-1976/ab1dad> (2019).
34. Fowler, E. E. *et al.* Generalized breast density metrics. *Phys. Med. Biol.* <https://doi.org/10.1088/1361-6560/aaf307> (2019).
35. Heine, J. J., Deans, S. R., Velthuisen, R. P. & Clarke, L. P. On the statistical nature of mammograms. *Med. Phys.* **26**, 2254–2265 (1999).
36. Heine, J. J. & Kaufhold, J. in *IWDM 2002: 6th International Workshop on Digital Mammography, June 22–25, 2002* (ed Peitgen, H.-O.) 544–546 (Springer, 2002).
37. Heine, J. J. & Velthuisen, R. P. A statistical methodology for mammographic density detection. *Med. Phys.* **27**, 2644–2651 (2000).
38. Manduca, A. *et al.* Texture features from mammographic images and risk of breast cancer. *Cancer Epidemiol. Biomark. Prev.* **18**, 837–845. <https://doi.org/10.1158/1055-9965.EPI-08-0631> (2009).
39. Ji, P. *et al.* The burden and trends of breast cancer from 1990 to 2017 at the global, regional, and national levels: Results from the global burden of disease study 2017. *Front. Oncol.* **10**, 650. <https://doi.org/10.3389/fonc.2020.00650> (2020).
40. Gastouniotti, A. *et al.* Evaluation of LIBRA software for fully automated mammographic density assessment in breast cancer risk prediction. *Radiology* **296**, 24–31. <https://doi.org/10.1148/radiol.2020192509> (2020).
41. Heine, J. *et al.* Mammographic variation measures, breast density, and breast cancer risk. *AJR Am. J. Roentgenol.* **217**, 326–335. <https://doi.org/10.2214/AJR.20.22794> (2021).
42. Heine, J. J., Fowler, E. E. & Flowers, C. I. Full field digital mammography and breast density: Comparison of calibrated and noncalibrated measurements. *Acad. Radiol.* **18**, 1430–1436. <https://doi.org/10.1186/1475-925X-12-114> (2011).
43. Heine, J. J. *et al.* A novel automated mammographic density measure and breast cancer risk. *J. Natl. Cancer Inst.* **104**, 1028–1037. <https://doi.org/10.1093/jnci/djs254> (2012).
44. Spak, D. A., Plaxco, J., Santiago, L., Dryden, M. & Dogan, B. BI-RADS® fifth edition: A summary of changes. *Diagn. Interv. Imaging* **98**, 179–190 (2017).
45. Gastouniotti, A., Desai, S., Ahluwalia, V. S., Conant, E. F. & Kontos, D. Artificial intelligence in mammographic phenotyping of breast cancer risk: A narrative review. *Breast Cancer Res.* **24**, 1–12 (2022).

46. Varallo, A. *et al.* Fabrication of 3D printed patient-derived anthropomorphic breast phantoms for mammography and digital breast tomosynthesis: Imaging assessment with clinical X-ray spectra. *Physica Medica* **98**, 88–97 (2022).
47. Sarno, A. *et al.* Physical and digital phantoms for 2D and 3D x-ray breast imaging: Review on the state-of-the-art and future prospects. *Radiat. Phys. Chem.* **204**, 110715 (2022).

Author contributions

J.H. is the corresponding author, conceived the plan and methods; E.F. is a coauthor, developed the computer code and assisted in the plan and methods development; R.J.W. is a coauthor, E.H. is a coauthor, and S.T. is a coauthor.

Funding

This work was supported by the National Institutes of Health Grants R01CA166269 and U01CA200464.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023