Analysis of Parenchymal Texture with Digital Breast Tomosynthesis: Comparison with Digital Mammography and Implications for Cancer Risk Assessment

Despina Kontos, PhD
Lynda C. Ikejimba, BS
Predrag R. Bakic, PhD
Andrea B. Troxel, ScD
Emily F. Conant, MD
Andrew D. A. Maidment, PhD

Purpose:
To correlate the parenchymal texture features at digital breast tomosynthesis (DBT) and digital mammography with breast percent density (PD), an established breast cancer risk factor, in a screening population of women.

Materials and Methods:
This HIPAA-compliant study was approved by the institutional review board. Bilateral DBT images and digital mammograms from 71 women (mean age, 54 years; age range, 34–75 years) with negative or benign findings at screening mammography were retrospectively collected from a separate institutional review board–approved DBT screening trial (performed from July 2007 to March 2008) in which all women had given written informed consent. Parenchymal texture features of skewness, coarseness, contrast, energy, homogeneity, and fractal dimension were computed from the retroareolar region. Principal component analysis (PCA) was applied to obtain orthogonal texture components. Mammographic PD was estimated with software. Correlation analysis and multiple linear regression with generalized estimating equations were performed to determine the association between texture features and breast PD. Regression was adjusted for age to determine the independent association of texture to breast PD when age was also considered as a predictor variable.

Results:
Texture feature correlations to breast PD were stronger with DBT than with digital mammography. Statistically significant correlations ($P < .001$) were observed for contrast ($r = 0.48$), energy ($r = -0.47$), and homogeneity ($r = -0.36$) at DBT and for contrast ($r = 0.26$), energy ($r = -0.26$), and homogeneity ($r = -0.33$) at digital mammography. Multiple linear regression analysis of PCA texture components as predictors of PD also demonstrated significantly stronger associations with DBT. The association was strongest when age was also considered as a predictor of PD ($R^2 = 0.41$ for DBT and 0.28 for digital mammography; $P < .001$).

Conclusion:
Parenchymal texture features are more strongly correlated to breast PD in DBT than in digital mammography. The authors’ long-term hypothesis is that parenchymal texture analysis with DBT will result in quantitative imaging biomarkers that can improve the estimation of breast cancer risk.

†From the Department of Radiology, University of Pennsylvania Health System, 1 Silverstein Bldg HUP, 3400 Spruce St, Philadelphia PA 19104-4206 (D.K., L.C.I., P.R.B., E.F.C., A.D.A.M.); and Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pa (A.B.T.). Received May 27, 2010; revision requested July 13; revision received December 23; accepted January 18, 2011; final version accepted May 17. Supported by a research fellow grant from the Radiological Society of North America (RF0707) and a postdoctoral fellowship from the Susan G. Komen for the Cure Foundation (KG080756). Address correspondence to D.K. (e-mail: Despina.Kontos@uphs.upenn.edu).

© RSNA, 2011

The assessment of breast cancer risk has become increasingly important for establishing screening recommendations (1–4) and forming preventive strategies (5,6). Growing evidence suggests that mammographic breast density, in addition to affecting the sensitivity of mammography by obscuring breast lesions (7), is a strong independent risk factor for breast cancer (8–12); in fact, it is the strongest known attributable risk factor after age (13). Several studies have shown that the relative risk of breast cancer increases with increasing mammographic density (9–12) and that breast density can be used to improve risk stratification at the individual level by improving the discriminatory accuracy of breast cancer risk assessment models (14–17). Boyd et al (10,18) have established risk stratification levels based on mammographic breast percent density (PD), as follows: 0% < PD < 10%, relative risk = 1.2; 10% ≤ PD < 25%, relative risk = 2.2; 25% ≤ PD < 50%, relative risk = 2.4; 50% ≤ PD < 75%, relative risk = 3.4; and 75% ≤ PD < 100%, relative risk = 5.3.

Although useful for estimating the risk of breast cancer, the methods currently used to assess breast density lack automation and are highly subjective to inter- and intrarater variability (19,20). Byng et al (21,22) showed that mammographic texture analysis can be used to automate the analysis of breast density patterns. Manduca et al (23) demonstrated that mammographic texture features have similar odds ratios to those of breast PD and can help predict breast cancer risk at a similar magnitude. Studies with digitized mammograms have also shown that mammographic texture features can help differentiate women who are carriers of BRCA1/2 from those at low risk for breast cancer (24–28).

Mammography, however, is limited in the performance of parenchymal texture analysis. Because mammography is a projection imaging technique, texture features reflect an admixture of overlapping tissues, including fibroglandular, fatty, and superficial skin tissue layers (29). Because the risk of breast cancer based on breast texture and/or density is mainly associated with the properties of fibroglandular tissue (ie, dense tissue), the texture projected from the skin or the subcutaneous fat could be regarded as “noise” and potentially affect the information content of the computed texture features.

Digital breast tomosynthesis (DBT) is an emerging radiographic imaging modality in which tomographic breast images are reconstructed from multiple low-dose projections acquired at different angles of the x-ray tube (30–34). DBT alleviates the effect of tissue superposition, offering the ability to analyze the fibroglandular tissue texture selectively (Fig 1). We performed this study to correlate parenchymal texture features seen at DBT and digital mammography with breast PD in a screening population of women. Our hypothesis was that the superior parenchymal tissue visualization with DBT will provide texture features that are more informative in characterizing breast parenchymal patterns and ultimately result in more accurate measures with which to estimate breast cancer risk.

### Materials and Methods

#### Imaging

This study was in compliance with the Health Insurance Portability and Accountability Act and received institutional review board approval. Anonymized images from DBT and digital mammography were retrospectively collected from a separate institutional review board–approved DBT screening trial that was previously completed in our department (performed from July 2007 to March 2008). Women who...
Mammographic Density Estimation

Breast PD was estimated by a medical physicist (P.R.B, with 10 years of experience in x-ray imaging physics, breast anatomy, and breast density estimation) (35,36) from the postprocessed (Premium View) digital mammograms by using Cumulus (version 4.0; University of Toronto, Toronto, Ontario, Canada), a widely validated software used to estimate mammographic density (10,37–39). Although Cumulus was originally developed for use with digitized screen-film mammograms, recent studies have also applied Cumulus to digital mammograms to estimate cancer risk (36,40,41). We have previously reported high intra- and interreader agreement of clinical radiologists when using Cumulus to estimate breast PD from postprocessed digital mammograms (42). In addition, the difference in PD obtained...
from raw versus postprocessed digital mammograms is relatively small (<1%) compared with the validated Boyd risk categories, indicating that image postprocessing is unlikely to have a substantial effect on density-based breast cancer risk stratification in clinical practice (43). With use of Cumulus, the image background and pectoral muscle region are excluded and gray level intensity thresholds are manually defined to outline the fibroglandular tissue regions within the breast (38). Breast PD is then estimated as the percentage of the total breast area occluded by the segmented fibroglandular tissue as follows: \( \frac{A_p}{A_b} \times 100\% \), where \( A_p \) and \( A_b \) represent the area of dense tissue and the total breast area, respectively.

Parenchymal Texture Analysis

Texture analysis was performed on all digital mammograms and DBT images within manually selected retroareolar regions of interest (ROIs). The placement and physical dimensions of the ROIs were chosen on the basis of previous studies (24–28), which indicated that the retroareolar breast region can provide the most discriminative texture features for differentiating women at high risk for cancer from those at low risk. The ROIs were placed in the central breast region immediately behind the nipple, which is usually texturally dense owing to the prominence of the underlying ductal network, and the placement was such that the subcutaneous fat region immediately under the skin was excluded. The ROIs used for DBT were 2.5 cm\(^2\) (116 × 116 × 26 voxels). The ROIs used for digital mammography were placed in identical spatial locations to those on corresponding DBT images and were 2.5 cm\(^2\) (256 × 256 pixels).

Parenchymal texture descriptors of skewness, coarseness, contrast, energy, homogeneity, and fractal dimension were computed by using established computerized methods (24–28). These descriptors have been previously used to characterize breast parenchymal patterns and have been shown to be predictive of the risk for developing breast cancer (21–28). Briefly, skewness reflects the asymmetry of the gray-level pixel value distribution and has been used to assess parenchymal density (26,38). Coarseness reflects the local granularity (ie, roughness) in image texture and is based on the computation of the neighborhood gray tone difference matrix (26,44). Contrast, energy, and homogeneity, as originally proposed by Haralick et al (45), require the computation of the gray-level spatial co-occurrence matrix. Contrast quantifies variation in image intensity, providing a measure of gray-level contrast between neighboring pixels over the entire image. Energy is a measure of texture uniformity of the gray-level spatial distribution. Homogeneity reflects the heterogeneity of the texture pattern and decreases with contrast. Fractal dimension indicates the measure of self-similarity in the texture pattern and the overall texture roughness at different scales (28,46). Details of the mathematical notations and the computation of these texture features have been previously published (26,28), including in our preliminary study (33), and are provided in Appendix E1 (online).

For the DBT ROIs, a feature was computed for each of the six texture descriptors (ie, skewness, coarseness, contrast, energy, homogeneity, and fractal dimension) from each of the 26 tomographic sections of the three-dimensional ROI, resulting in a texture feature vector of 26 \( \times 6 = 156 \). The average value over the 26 tomographic ROI sections was used as the representative feature, providing an estimate of the mean global texture within the ROI for each of the six main texture descriptors considered in our experiments. For the digital mammography ROIs, the equivalent single-valued texture feature was computed from the corresponding two-dimensional ROIs (Fig 2). Gray-level quantization was implemented to optimize the computation of the co-occurrence statistics, as previously defined in the literature (23,24,47), to cover a wide range of spatial scales and feature strengths relative to noise. Briefly, the initial range \( N_g \) of gray-level values in each ROI was linearly reduced by uniformly dividing the gray-level range by a constant. Details of the gray-level quantization method are provided in Appendix E1 (online). For the experiments reported herein, a representative value (\( N_g/16 \)) was selected.

Statistical Analysis

Correlation analysis and univariate linear regression were performed to determine the association between the individual texture features and the corresponding breast PD estimates (Matlab R2010b; Mathworks, Natick, Mass). To investigate associations between texture feature combinations and breast PD, principal component analysis (PCA) was performed to construct orthogonal texture components and remove multicollinearity effects (48). Texture features were normalized before PCA by using the z score (49). Multiple linear regression with a generalized estimating equation (50,51) was applied to model the association between the orthogonal PCA texture components and breast PD. The generalized estimating equation was implemented by using the robust covariance matrix (52) to adjust for bilateral correlations of the PCA texture components in the regression analysis. Age was also evaluated as a predictor of breast PD both independently and as an additional predictor in the multivariate models. The \( R^2 \) and adjusted \( R^2 \) metrics were used to assess goodness of fit. A \( P \) value threshold of .01 was used to determine statistically significant differences. The Akaike information criterion (53) was used as a measure of goodness of fit to compare the performance between the different models. Bonferroni correction was considered to adjust for multiple comparisons in statistical significance across the different models (\( P = .01/5 = .002 \) with Bonferroni adjustment).

Results

The computed breast PD varied from 0.07% to 83.11%, with a mean of 38.54%. For both digital mammography and DBT, the strongest correlations were observed for the Haralick co-occurrence texture features—which were stronger with DBT than with digital mammography (Table 1). Statistically significant correlations (\( P \leq .001 \)) were
observed for contrast ($r = -0.48$), energy ($r = -0.47$), and homogeneity ($r = -0.36$) with DBT and for contrast ($r = 0.26$), energy ($r = 0.26$), and homogeneity ($r = -0.33$) with digital mammography. Texture features at DBT showed stronger linear trends with the computed breast PD values compared with those at digital mammography, as evidenced by the steeper slopes and higher $R^2$ estimates (Fig 3). Compared with the texture features at digital mammography, texture features at DBT had a stronger dependency on the different breast PD

<table>
<thead>
<tr>
<th>Feature</th>
<th>Digital Mammography</th>
<th>DBT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$ Value</td>
<td>$P$ Value</td>
</tr>
<tr>
<td>Skewness</td>
<td>$-0.21$</td>
<td>.01</td>
</tr>
<tr>
<td>Coarseness</td>
<td>$-0.02$</td>
<td>.81</td>
</tr>
<tr>
<td>Contrast</td>
<td>0.26</td>
<td>.001</td>
</tr>
<tr>
<td>Energy</td>
<td>$-0.26$</td>
<td>.001</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>$-0.33$</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Fractal dimension</td>
<td>0.14</td>
<td>.09</td>
</tr>
</tbody>
</table>

* Data are Pearson correlation coefficients.
categories, as evidenced by the steeper linear trends and higher $R^2$ estimates (Fig 4).

When we investigated the association between combinations of texture features and breast PD, the DBT multiple linear regression models demonstrated stronger associations between the PCA texture components and breast PD than did the digital mammography models, as evidenced by the lower mean square error of the residual fits (mean square error for digital mammography: 354; mean square error for DBT: 298.16), the higher $R^2$ estimates ($R^2 = 0.28$ for digital mammography and 0.38 for DBT), and the significant $P$ values ($P = .21$ for digital mammography and $<.001$ for DBT).

The results of the different multiple regression models correlating PCA texture components and breast PD for digital mammography and DBT are shown in Table 2 (details about the model coefficients are provided in Appendix E1, online). The model including only the digital mammography PCA texture components as PD predictors was not statistically significant ($R^2 = 0.28; P = 0.21$, analysis of variance [ANOVA]). When age was added as an additional predictor of breast PD, the regression model was statistically significant ($R^2 = 0.28$, $P = .002$, ANOVA) and the fifth and sixth PCA texture components were identified as the most significant predictors ($P = .02$ for the fifth PCA component; $P < .001$ for the sixth PCA component) in addition to age ($P = .005$).

Both of the corresponding regression models for DBT were statistically significant ($R^2 = 0.38$ for the model including only PCA texture components as predictors, $R^2 = 0.41$ for the model with PCA components and age as predictors; $P < .001$, ANOVA). In the model including only the PCA texture components as predictors of breast PD, the first two components were identified as the most significant ($P = .002$ for the first component, $P < .001$ for the second...
component). When age was included as an additional predictor of breast PD, the first two PCA texture components were again identified as the most significant ($P = .003$ for the first component, $P < .001$ for the second component) in addition to age ($P < .001$), and the observed predictive association was stronger.

For comparison, regression was also performed with age as the only predictor of breast PD ($P < .001$, ANOVA). As shown in Table 2, the observed associations between combinations of the PCA texture features and age as predictors

---

**Table 2**

<table>
<thead>
<tr>
<th>Model*</th>
<th>Mean Square Error</th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>$P$ Value</th>
<th>AIC†</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM-I</td>
<td>354.00</td>
<td>0.53</td>
<td>0.28</td>
<td>0.19</td>
<td>.21</td>
<td>847.44</td>
</tr>
<tr>
<td>DM-II</td>
<td>325.82</td>
<td>0.53</td>
<td>0.28</td>
<td>0.25</td>
<td>.002</td>
<td>837.66</td>
</tr>
<tr>
<td>DBT-I</td>
<td>298.16</td>
<td>0.61</td>
<td>0.38</td>
<td>0.32</td>
<td>&lt;.001</td>
<td>823.96</td>
</tr>
<tr>
<td>DBT-II</td>
<td>269.08</td>
<td>0.64</td>
<td>0.41</td>
<td>0.38</td>
<td>&lt;.001</td>
<td>810.49</td>
</tr>
<tr>
<td>AGE</td>
<td>395.55</td>
<td>0.30</td>
<td>0.09</td>
<td>0.09</td>
<td>&lt;.001</td>
<td>853.20</td>
</tr>
</tbody>
</table>

* DM-I = digital mammography model with only PCA texture components as predictors; DM-II = digital mammography model with PCA texture components and age as predictors; DBT-I = DBT model with only PCA texture components as predictors; DBT-II = DBT model with PCA texture components and age as predictors; AGE = model with age as the only predictor.

† AIC = Akaike information criterion.
of breast PD were stronger than the corresponding association when age alone was considered as the only predictor of breast PD ($R^2 = 0.09$) for both digital mammography and DBT. The DBT model incorporating PCA texture components and age as predictor variables of breast PD achieved the best performance based on the Akaike information criterion (Table 2, lowest Akaike information criterion = 810.49). Note that even when Bonferroni correction was considered at a significance level of $P = .01$ for the five models compared, all models, except the one with only PCA texture components as predictors, retained statistical significance ($P = .002$ with Bonferroni adjustment).

Breast PD estimates obtained with the DBT regression models were more strongly associated with the different categories of breast PD, as evidenced by the steeper slopes of the linear regression fits, the higher $R^2$ values, and the statistically significant $P$ values ($P < .001$, Fig 5). The observed associations with both digital mammography and DBT models are again stronger when age was added as an additional predictor of the breast PD category of risk.

**Discussion**

Our results demonstrate that parenchymal texture features at DBT are more strongly correlated with breast PD than
are those at digital mammography, suggesting that DBT image texture analysis can provide more informative imaging features with which to characterize parenchymal patterns. Our results from this screening population of women are in agreement with data published from our previous pilot study with a diagnostic population of women at high risk for breast cancer (35). Given that breast cancer risk assessment is likely to have value primarily in the screening setting for prevention and early detection, this confirmatory study with a screening population of women investigates the potential impact of our previously reported techniques in the actual clinical population of interest.

In addition, in this study we built on our previous investigation by performing multivariate analysis to determine the combined effect of the parenchymal texture components, rather than just the individual components reported previously, in association with parenchymal density patterns. We also considered the effect of age both in combination with the extracted image texture features and alone (ie, univariate model) as a predictor of density (ie, risk). We believe that this multivariate analysis will pave the way for an integrated modeling approach that will incorporate demographics, familiar risk factors, and novel imaging markers into an augmented breast cancer risk assessment model for the general population.

We attribute the stronger linear association of the DBT texture features with breast PD to the tomographic separation of the parenchymal tissue layers in DBT. By alleviating the effect of tissue superposition in digital mammography, DBT offers the ability to analyze the fibro glandular tissue texture selectively, therefore providing features that are more strongly associated with the underlying fibroglandular (ie, dense) tissue patterns.

Starting with the pioneering work of Wolfe in 1976 (54), several studies have demonstrated an association between breast parenchymal pattern complexity and breast cancer risk (10–12,55,56). In our study, breast PD was used as a short-term surrogate of breast cancer risk. Although significant associations were detected between the extracted parenchymal texture features and breast PD, the observed correlations indicate that the variability in breast PD does not appear to be fully accounted for by the texture feature components, as evidenced by the $R^2$ values of the corresponding regression models ($R^2 = 0.28$ with the digital mammography model when only PCA texture components were used as predictors, $R^2 = 0.38$ with use the DBT model when only PCA texture components were used as predictors).

Currently, there is no standard lexicon that reflects parenchymal complexity completely; in mammography, breast density is the only image-based descriptor in the standard clinical Breast Imaging Reporting and Data System lexicon (57). Breast PD is a measure of the amount of fibroglandular tissue in the breast. Parenchymal texture features characterize the complexity of the parenchymal tissue composition, as reflected by the underlying epithelial and stromal breast tissue components (58,59), by quantifying properties of the distributions of the gray-level values in the image. Therefore, these textural features capture different information than that provided by breast PD with respect to the complexity of the parenchymal pattern and could ultimately be combined with breast density measures to provide more comprehensive imaging descriptors of breast tissue composition and complexity.

The research question yet to be fully investigated is to what extent parenchymal texture descriptors can complement breast density measures in breast cancer risk estimation. Most studies published to date addressing this question have used digitized screen-film mammograms for parenchymal texture analysis (23,55,56). Manduca et al (23) showed that parenchymal texture features can help predict breast cancer risk with similar odds ratios to that of breast density and that these features retain statistical significance even when breast density is included in a risk prediction model. Studies with BRCA1/2 carriers have shown that, despite observing the same distribution of breast density in BRCA1/2 carriers and the general population (60), mammographic texture features can help differentiate the patterns in BRCA1/2 carriers from those in women at low risk for cancer (24–28). A prospective clinical trial with digital mammography, DBT, and long-term patient follow-up will enable estimation of the relative risk of breast cancer and corresponding odds ratios associated with parenchymal texture in digital mammography versus DBT and will also allow the investigation of the attributable predictive value of various parenchymal texture descriptors when breast density measures are also considered in risk estimation. In our study, the observed associations between parenchymal texture features and breast PD were stronger when age was added as a predictor of PD ($R^2 = 0.38$ with the DBT model when only PCA texture components were used as predictors, $R^2 = 0.41$ with DBT model when PCA components and age were used as predictors, $P < .001$), indicating the need to consider additional patient risk factors to fully predict risk.

Our conclusions are limited by the use of a specific DBT prototype imaging system. Considering that image gray levels in DBT may be affected by factors such as the image acquisition geometry and the reconstruction algorithm (33,61), future work must investigate approaches to standardize the computed texture features across imaging platforms and processing algorithms. Preliminary studies with simulated phantoms have shown that although the angular range of the x-ray tube rotation and the number of source projections in DBT can affect the distributions of the computed texture features, the individual values are highly correlated and appear to follow specific nonlinear trends (62). Therefore, standardization of computerized descriptors should be feasible through investigation and, therefore, ultimately within reach of wide-scale dissemination. The concurrence of results, however, between our current study and the previous pilot study (35) despite the use of different imaging systems further validates the
predictive value of our proposed imaging markers.

Future approaches may also include the investigation of areas of the breast other than the retroareolar region (eg, the entire breast region or the tissue regions specifically attributed to fibroglandular tissue) and the evaluation of a more comprehensive feature set of texture descriptors. In our study, considering the potential bias of overfitting due to the size of our dataset and issues with multiple comparisons, we chose to average the texture features over the DBT tomographic sections to obtain an estimate of the mean global texture within the ROI. We realize, however, that by estimating the mean value, additional information about the richness of the feature set, including higher moments and variance, could be lost. Larger clinical studies in the future will allow a more thorough investigation of feature extraction and feature merging algorithms in direct association with the clinical outcome of interest (ie, follow-up for cancer status).

There is increasing interest in the augmentation of breast cancer risk assessment models with measures of breast density to improve their discriminatory accuracy (14–17). Parenchymal texture analysis in DBT could provide a wide range of quantitative parenchymal pattern descriptors that can be implemented by automated computerized algorithms. The rapidly emerging technical improvements in DBT image quality and reconstruction (63–72), the relatively low cost, and the potential for superior clinical performance (73–79) will determine the ultimate role of DBT in clinical practice. Recent data suggest a substantial benefit from DBT in the reduction of recall rates and the number of subsequent unnecessary biopsies in women in the screening setting (74,80–82), indicating a potential future role for DBT in breast cancer screening.

The reproducibility and automation of DBT parenchymal texture descriptors could ultimately accelerate the incorporation of imaging biomarkers in clinical breast cancer risk assessment. A refined cancer risk estimator incorporating such imaging biomarkers of risk could be of great clinical advantage in reaching the goal of customizing screening recommendations on the basis of personal risk levels, tailoring individual treatments, and forming preventive strategies—especially for women at a higher risk of breast cancer.

Acknowledgments: We thank Martin Yaffe, MD, for providing the Camnaha software. We also thank Johnny Kuo, MD, for developing the BSNA Medical Imaging Resource Center (MIRC) image database archive.

Disclosures of Potential Conflicts of Interest: D.K. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: received payment from Hologic for lectures including service on speakers bureaus. Other relationships: none to disclose. L.G.I. No potential conflicts of interest to disclose. P.R.B. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: institution received money from Real-Time Tomography for grants/grants pending. Other relationships: none to disclose. A.R.T. No potential conflicts of interest to disclose. E.E.C. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: received a consulting fee from Hologic; received royalties from Elsevier. Other relationships: none to disclose. A.D.A.M. Financial activities related to the present article: none to disclose. Financial activities not related to the present article: received money for consultancy from Real-Time Tomography. Other relationships: none to disclose.

References
19. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA.


