Radiology

ORIGINAL RESEARCH • BREAST IMAGING

Radiomic Phenotypes of Mammographic Parenchymal Complexity: Toward Augmenting Breast Density in Breast Cancer Risk Assessment

Despina Kontos, PhD • Stacey J. Winham, PhD • Andrew Oustimov, MPH • Lauren Pantalone, BS • Meng-Kang Hsieh, MS • Aimilia Gastounioti, PhD • Dana H. Whaley, MD • Carrie B. Hruska, PhD • Karla Kerlikowske, MD • Kathleen Brandt, MD • Emily F. Conant, MD • Celine M. Vachon, PhD

From the Department of Radiology, University of Pennsylvania, Perelman School of Medicine & Hospital of the University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104 (D.K., A.O., L.P., M.K.H., A.G., E.F.C.); Department of Health Sciences Research (S.J.W., C.M.V.) and Department of Radiology (D.H.W., C.B.H., K.B.), Mayo Clinic, Rochester, Minn (S.J.W., D.H.W., C.B.H., K.B., C.M.V.); and Departments of Medicine and Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, Calif (K.K.). Received January 22, 2018; revision requested March 19; revision received August 24; accepted August 31. **Address correspondence to** D.K. (e-mail: *Despina.Kontos@uphs.upenn.edu*).

Supported by National Cancer Institute (1R01CA207084-01A1, 2R01CA161749-05) and a National Institutes of Health specialized center cooperative agreement (5U54CA163313-05).

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

See also the editorial by Pinker in this issue.

Radiology 2019; 290:41–49 • https://doi.org/10.1148/radiol.2018180179 • Content code: BR

Purpose: To identify phenotypes of mammographic parenchymal complexity by using radiomic features and to evaluate their associations with breast density and other breast cancer risk factors.

Materials and Methods: Computerized image analysis was used to quantify breast density and extract parenchymal texture features in a cross-sectional sample of women screened with digital mammography from September 1, 2012, to February 28, 2013 (n = 2029; age range, 35–75 years; mean age, 55.9 years). Unsupervised clustering was applied to identify and reproduce phenotypes of parenchymal complexity in separate training (n = 1339) and test sets (n = 690). Differences across phenotypes by age, body mass index, breast density, and estimated breast cancer risk were assessed by using Fisher exact, χ^2, and Kruskal-Wallis tests. Conditional logistic regression was used to evaluate preliminary associations between the detected phenotypes and breast cancer in an independent case-control sample (76 women diagnosed with breast cancer and 158 control participants) matched on age.

Results: Unsupervised clustering in the screening sample identified four phenotypes with increasing parenchymal complexity that were reproducible between training and test sets (P = .001). Breast density was not strongly correlated with phenotype category (R^2 = 0.24 for linear trend). The low- to intermediate-complexity phenotype (prevalence, 390 of 2029 [19%]) had the lowest proportion of dense breasts (eight of 390 [2.1%]), whereas similar proportions were observed across other phenotypes (from 140 of 291 [48.1%] in the high-complexity phenotype to 275 of 511 [53.8%] in the low-complexity phenotype). In the independent case-control sample, phenotypes showed a significant association with breast cancer (P = .001), resulting in higher discriminatory capacity when added to a model with breast density and body mass index (area under the curve, 0.84 vs 0.80; P = .03 for comparison).

Conclusion: Radiomic phenotypes capture mammographic parenchymal complexity beyond conventional breast density measures and established breast cancer risk factors.

Online supplemental material is available for this article.

Wolfe in 1976 suggested that patterns of breast parenchymal complexity, formed by the x-ray attenuation of fatty, fibroglandular, and stromal tissues (1), are associated with breast cancer risk (2). Breast density ratings, based on the extent of mammographic density, are routinely used clinically to characterize the breast parenchyma. High breast density has been associated with greater risk of breast cancer (3–5). Additionally, breast density has been associated with masking of cancers leading to interval cancers (6) in mammographic screening.

Density measures aim to capture the relative amount of fibroglandular tissue in the breast (7); however, they are increasingly considered to be coarse measures, being limited in fully capturing the complexity of the breast parenchymal pattern (8). This has motivated research toward complementing quantitative density measures with more granular characterization of parenchymal complexity and their association to breast cancer risk and detection.

Early studies with *BRCA1* and *BRCA2* (*BRCA1/2*) carriers have shown that computerized measures of mammographic parenchymal texture from the retroareolar breast region can distinguish *BRCA1/2* carriers from low-risk women (9,10). Recent studies of case-control samples from screening populations have also shown that parenchymal texture features (either from the retroareolar region or the entire breast area) are significantly associated with breast cancer independent of breast density (11–15). Nevertheless, to our knowledge, no studies to date have attempted to define distinct imaging phenotypes that reflect intrinsic complexity of the breast parenchymal tissue.

Abbreviations

BI-RADS = Breast Imaging Reporting and Data System, BMI = body mass index, LIBRA = Laboratory for Individualized Breast Radiodensity Assessment, PD = percent density

Summary

Radiomic phenotypes reflect intrinsic properties of mammographic parenchymal complexity beyond breast density and have an independent association with breast cancer.

Implications for Patient Care

- Radiomic phenotypes can assess mammographic parenchymal complexity and may provide additional information for risk assessment beyond breast density.
- Radiomic phenotypes may ultimately augment breast cancer risk prediction models.

The goal of our study was to identify phenotypes of mammographic parenchymal complexity by using radiomic features and to evaluate their associations with breast density and other breast cancer risk factors. We hypothesized that by capturing various patterns or complexities of the breast parenchyma in a more refined and quantitative manner, we can provide information beyond mammographic density and established risk factors in association to breast cancer.

Materials and Methods

Requirement of obtaining consent for our Health Insurance Portability and Accountability Act–compliant retrospective study was waived by our institutional review board.

Study Population

Cross-sectional sample of screening population for phenotype identification.—Our cross-sectional sample included all white women consecutively presenting for routine screening during a 6-month period, including any breast cancers identified at the time of screening or within 1 year of screening (n = 2241). Women were imaged per U.S. Food and Drug Administration–approved protocol consisting of full-field digital mammography and breast tomosynthesis in both mediolateral oblique and craniocaudal views (Selenia Dimensions; Hologic, Bedford, Mass). For the purposes of our study, raw (ie, "For Processing") digital mammography images were analyzed. As part of routine screening, women were given a risk factor questionnaire used to estimate breast cancer risk with the National Cancer Institute's Breast Cancer Risk Assessment Tool (16) and the Breast Cancer Surveillance Consortium Risk Calculator (17). High risk was defined as 5-year risk greater than 1.67%. Risk factor variables included age, race, age at menarche, parity, age at first live birth, first-degree relatives with breast cancer, number of prior benign biopsies, body mass index (BMI), and American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) breast density score (18).

Case-control sample for evaluating associations to breast cancer.—As a preliminary evaluation, we used an independent, previously published, case-control sample (14,19) to examine associations between breast cancer and the phenotypes identified from our cross-sectional screening sample. Briefly, women diagnosed with primary invasive breast cancer were retrospectively ascertained from a study previously completed at our institution designed to determine the value of multimodality imaging for breast cancer screening, detection, and staging. The inclusion criteria of the study recruited women at high risk for breast cancer because of *BRCA1/2* mutation, greater than or equal to 25% estimated lifetime risk by using the Gail or Claus models, history of lobular carcinoma in situ or atypical hyperplasia, history of chest wall radiation before puberty, or recently diagnosed contralateral breast cancer; women with a mammographically detected suspicious finding (BI-RADS $\geq$4) following screening and/or diagnostic evaluation of women with suspicious palpable mass directed to biopsy; and women with newly diagnosed breast cancer presenting for staging. Control participants were randomly selected asymptomatic women who underwent breast cancer screening with digital mammography at our institution over the same period. To be consistent with the screening sample used for phenotype identification, only the subset of white women was included in our analysis, resulting in a convenience sample of 76 women diagnosed with breast cancer (cases) and 158 women with no breast cancer diagnosis (control participants) (228 of 424 [54%] of total sample). All women had raw digital mammograms available (Senographe 2000D and DS; GE Healthcare, Little Chalfont, United Kingdom). Whereas our prior studies evaluated individual measures of breast density and texture (14,19), here we used this sample to perform independent evaluation of the radiomic breast complexity phenotypes, identified in our larger screening sample, in association to breast cancer.

Breast density estimation.—The previously validated, publicly available, and fully automated Laboratory for Individualized Breast Radiodensity Assessment (LIBRA, version 1.0.3) software (19,20) was used to quantitatively measure breast percent density (PD) (Fig E1 [online]). LIBRA has been shown to have good agreement with the established semiautomated Cumulus method and exhibit significant associations to breast cancer (19). For our study, both views were analyzed with LIBRA for each breast and PD measures were averaged to derive per-woman estimates (see Appendix E1 [online] for details). Assessment of breast density per the fourth edition of the BI-RADS Atlas (18) was abstracted from clinical reports as *(a)* fatty, *(b)* scattered fibroglandular densities, *(c)* heterogeneously dense, and *(d)* extremely dense. Women were categorized as having nondense versus dense breast density by grouping categories *a*–*b* versus *c*–*d*, respectively.

Parenchymal complexity analysis.—Parenchymal complexity was quantified by using custom-developed fully automated software (Matlab, version 9.4; Mathworks, Natick, Mass) (14) that uses a lattice-based strategy (Fig E1 [online]) to extract a range of mammographic texture features, including four main types: histogram, co-occurrence, run-length, and structural. Briefly, gray-

level histogram features are first-order statistics describing the distribution of gray-level intensities (21). Co-occurrence features consider the spatial relationships of pixel intensities in different directions and are based on the gray-level co-occurrence matrix that encodes the relative frequency of neighboring intensity values (22). Run-length features capture the coarseness of texture in specified directions by measuring strings of consecutive pixels (ie, runs) that have the same gray-level intensity along specific linear orientations (23). Finally, structural features reflect the architectural composition of the parenchyma by characterizing the directionality of flowlike structures and intensity variations between neighboring pixels (24,25). Measures from each lattice point, the number of which depends on breast size (and therefore can be different for each breast), were averaged over each mammographic view and for both breasts to create a per-woman measure for each feature. This resulted in a feature vector of 29 features characterizing parenchymal complexity for each woman (see Appendix E1 [online] for details).

Phenotype identification in the cross-sectional screening sample.—We performed unsupervised hierarchical clustering on the woman-specific feature vectors, similar in concept to the approach originally used for the discovery of the intrinsic molecular subtypes for breast cancer (26). Prior to clustering, each feature was standardized in the entire sample by using *z* score normalization (27), and the sign for each feature was chosen such that negative values corresponded to low complexity and positive values corresponded to high complexity for consistent interpretation across features. We first visually inspected the histograms of the *z*-scored features. To ensure that the analyzed features would be informative for clustering, we excluded features with extremely low variation (interquartile range <1). Additionally, to ensure that the analyzed features did not deviate significantly from a normal distribution, we excluded those with extreme skewness. Five features satisfied both criteria and were excluded. This resulted in 24 of 29 available features, which were subsequently fed into the clustering algorithm (see Appendix E1 [online] for details). To assess phenotype reproducibility, we randomly divided our screening sample into separate training ($n = 1339$) and test ($n = 690$) sets. After clusters were determined to be reproducible in the training and test sets, the sets were combined for further analysis. A complexity score was derived for each woman by averaging their corresponding normalized feature vector values.

Phenotype classification in the case-control sample.—Parenchymal texture analysis was performed with the same approach as in our screening sample. Features in the case-control sample were standardized independently and women were classified into phenotypes based on their proximity to the cluster centroid of the phenotypes identified in the cross-sectional screening sample, based on the minimum Euclidean distance of the corresponding feature vectors.

Statistical Analysis

For our cross-sectional screening sample, differences in age, BMI, breast density, estimates of breast cancer risk, and screening outcomes across parenchymal phenotypes were assessed with the Fisher exact, χ^2, and Kruskal-Wallis tests for discrete and continuous covariates, respectively. Missing BMI data were imputed by using the Markov Chain Monte Carlo method for full imputation on the basis of age and continuous breast density measures. Five imputations were carried out to produce five imputed values for each observation, which were then averaged to produce the final imputed values for analysis (27). Associations between the continuous breast PD estimates and corresponding complexity scores were also evaluated with linear regression. For our case-control sample, differences in age, BMI, and PD between cases and control participants were assessed with Wilcoxon signed-rank test. Because of small numbers, women were grouped into two BMI categories (low BMI, <25 kg/m^2 and high BMI, ≥ 25 kg/m^2) and complete-case analysis was used for the adjusted models without imputation for missing BMI data (31 of 228, 14%). Associations between phenotypes and breast cancer were assessed with logistic regression, both unadjusted and adjusted for covariates. Phenotypes were coded as a categorical variable with the most prevalent phenotype used as the reference. Associations were described with the odds ratio, and phenotype contribution to the model was assessed with the likelihood ratio test (28). Discriminatory performance was estimated with the area under the curve of

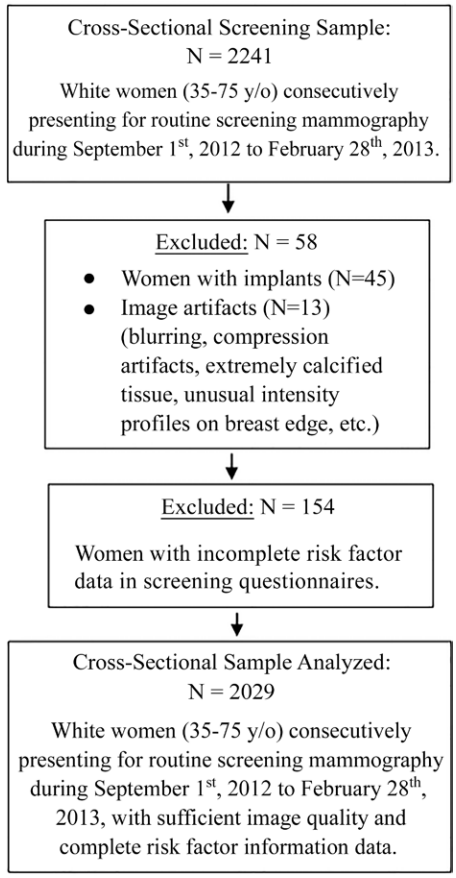


Figure 1: Flowchart shows inclusion and exclusion criteria for cross-sectional screening sample analyzed in study for parenchymal complexity phenotype identification.

Table 1: Characteristics of the Screening Sample Used for Phenotype Identification

Characteristic	Entire Cohort (n = 2029)	Training Sample (n = 1339)	Testing Sample (n = 690)	P Value*
Age (y)	55.9 ± 9.4	55.8 ± 9.3	56.0 ± 9.5	.56
Body mass index (kg/m^2)	26.4 ± 6.0	26.4 ± 6.1	26.4 ± 5.9	.48
Missing data†	164 (8.1)	116 (8.7)	48 (7.0)	…
Gail model 5-y risk (%)	1.8 ± 1.1	1.8 ± 1.1	1.8 ± 1.1	.50
Gail model lifetime risk (%)	10.8 ± 4.9	10.9 ± 4.9	10.8 ± 4.9	.99
BCSC model 5-y risk (%)	1.67 ± 0.9	1.67 ± 0.9	1.69 ± 0.9	.68
BI-RADS density†				.07
Fatty	150 (7.4)	108 (8.1)	42 (6.1)	
Scattered fibroglandular densities	1027 (50.6)	665 (49.7)	362 (52.5)	
Heterogeneously dense	800 (39.4)	525 (39.2)	275 (39.9)	
Extremely dense	52 (2.6)	41 (3.1)	11 (1.6)	
Callback†	171 (8.4)	111 (8.3)	60 (8.7)	.76
Referred to biopsy†	45 (2.2)	28 (2.1)	17 (2.5)	.59
Detected cancers†	18 (0.9)	12 (0.9)	6 (0.9)	.95
Percent density	17.5 ± 10.6	17.7 ± 10.8	17.2 ± 10.4	.41

Note.—Unless otherwise specified, data are means ± standard deviations. BCSC = Breast Cancer Surveillance Consortium, BI-RADS = Breast Imaging Reporting and Data System.

* P values for comparison of sample characteristics across training and test samples were calculated with Wilcoxon rank-sum test for continuous variables and Pearson χ^2 test for categorical variables.

† Data are numbers, with percentages in parentheses.

the receiver operating characteristic and DeLong test was used to evaluate gain in performance. All statistical analyses were performed with the R software (version 3.1.1; R Foundation for Statistical Computing, Vienna, Austria).

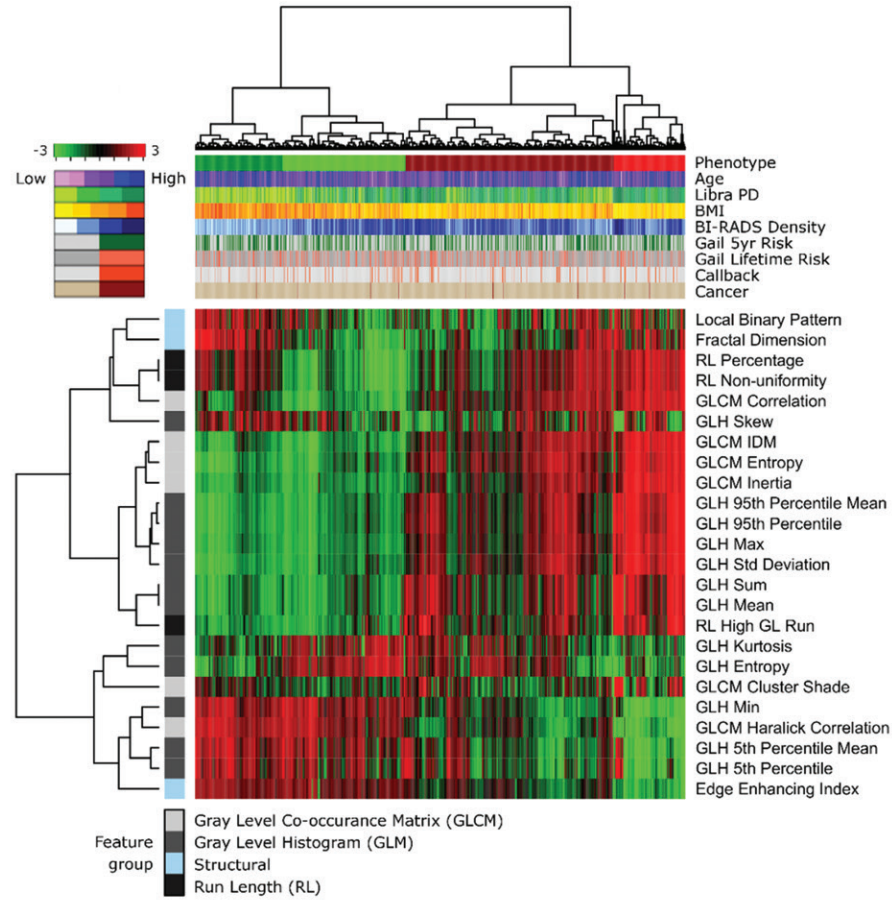
Results

Phenotype Identification (Cross-sectional Screening Sample)

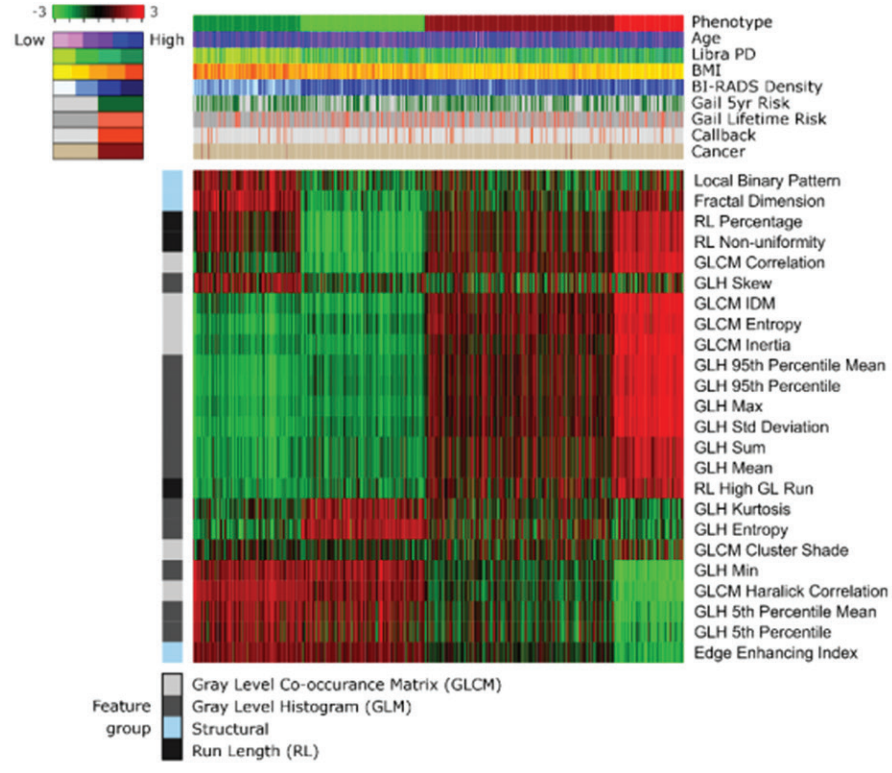
Women with implants (n = 45) or other image artifacts (n = 13) (including blurring, compression artifacts, extremely calcified tissue, unusual intensity profiles on the breast edge, etc) known to influence the application of computerized algorithms were excluded (n = 58). From the remaining sample (n = 2183), only women with complete breast cancer risk factor data required for the Gail and Breast Cancer Surveillance Consortium risk models were included in our study sample (n = 2029) (Fig 1). Demographics, risk factor characteristics, and screening outcomes were not significantly different between the training and the test sets (Table 1). Unsupervised clustering identified four distinct parenchymal phenotypes (Fig 2a), ranked based on their complexity score, as follows: low (light green, 511 of 2029 [25.2%]), low to intermediate (dark green, 390 of 2029 [19.2%]), intermediate to high (dark red, 837 of 2029 [41.3%]), and high complexity (bright red, 291 of 2029 [14.3%]). Phenotypes were reproducible in the separate training and test sets (P = .001) (Fig 2b). When examining the full sample (Fig E2 [online]), differences for age, BMI, and breast PD were observed across phenotypes (P = .001) (Fig 3). Women with low to intermediate complexity had the highest age (median, 59.3 years) and BMI (31.3 kg/m^2) but the lowest average breast PD of 8%, whereas women with the lowest complexity were the youngest (median, 47.3 years). Women with the highest complexity had the lowest BMI (21.1 kg/m^2) and highest breast PD of 23% (P = .001) (Figs 3, 4). Overall, breast PD was not strongly correlated to the parenchymal phenotypes (R^2 = 0.24 for linear association) (Fig E3 [online]).

The proportion of women with dense breasts also differed among phenotypes (Table 2, Fig E4 [online]); particularly for the low- to intermediate-complexity cluster compared with the other phenotypes (P = .001). The low- to intermediate-complexity cluster had the lowest proportion of women with dense breasts (eight of 390 [2.1%]), whereas similar proportions were observed across the other phenotypes (from 140 of 291 [48.1%] to 275 of 511 [53.8%]). BI-RADS density also had a low correlation with the breast complexity phenotypes (Table 2) (Spearman ρ, 0.11; P = .001).

The proportions of women with high risk as estimated by using the Gail and the Breast Cancer Surveillance Consortium risk models also differed across phenotypes. The low- to intermediate-complexity phenotype had the highest proportion of women at high 5-year risk estimated with the Gail model (186 of 390 [47.7%]; P = .03), but the lowest proportion of women with high risk as estimated with the Breast Cancer Surveillance Consortium model (138 of 390 [35.3%]; P = .001), which includes breast density (Figs E4, E5 [online]). The intermediate- to high- complexity phenotype had the highest proportion of women at high 5-year risk estimated with the Breast Cancer Surveillance Consortium model (413 of 837 [49.3%]; P = .001), compared with the low-complexity and low- to intermediate-complexity phenotypes. Finally, there was no difference across phenotypes for the total of 18 cancers that were detected. ($P > .05$) (Table E1 [online]).



a.



b.

Figure 2: Heat map generated by unsupervised hierarchical clustering of extracted radiomic features, applied to separate **(a)** training and **(b)** test sets of screening sample for phenotype identification, shows cluster membership (top color bar) with complexity-based color scheme (ie, light green = low parenchymal complexity, dark green = low to intermediate parenchymal complexity, dark red = intermediate to high parenchymal complexity, bright red = high parenchymal complexity). Each column in heat map represents a woman and each row represents a specific radiomic feature, with standardized feature values ordered according to parenchymal complexity (ie, low complexity = −3 standard deviations depicted by green, high complexity = 3 standard deviations depicted by red in color bar on top left). Dendrogram at top represents grouping of women in distinct phenotypes, whereas dendrogram on left represents groupings of extracted features with similar information. Corresponding color maps represent distributions of demographics, risk factor data, and screening outcomes for each phenotype (top) and categories of extracted features (left).

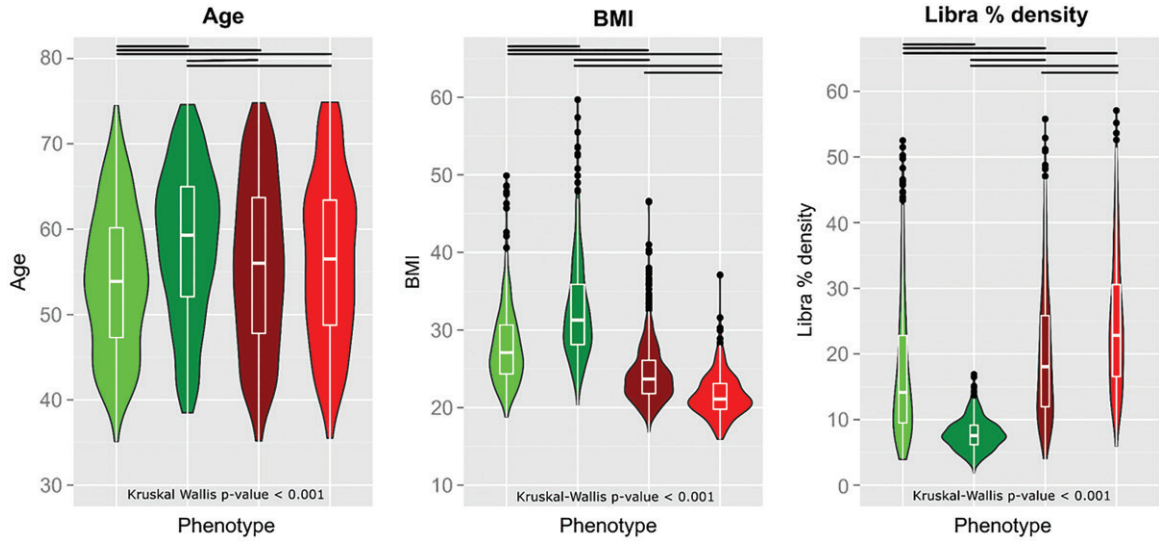


Figure 3: Phenotype-specific violin plots show distributions from left to right of age (in years), body mass index (BMI) (in kg/m^2), and percentage of density across different parenchymal complexity phenotypes. Plots are arranged in order of increasing parenchymal complexity (from light green to bright red), and solid bars at top indicate significant pairwise differences in corresponding distributions of age, BMI, and percentage of density for different phenotypes. LIBRA = Laboratory for Individualized Breast Radiodensity Assessment.

Associations to Breast Cancer (Case-Control Sample)

Women diagnosed with breast cancer had a higher proportion of low-complexity and low- tointermediate- complexity phenotypes, but also higher BMI and higher breast PD (Table 3). Radiomics phenotypes were significantly associated with breast cancer, both in unadjusted (Table 4) and in models accounting for breast PD and BMI (Table 4). The likelihood ratio test revealed model improvement (P = .001) with the inclusion of the complexity phenotypes having an area under the curve of 0.84 (95% confidence interval: 0.78, 0.89), compared with a baseline model including only breast PD and BMI with area under the curve of 0.80 (95% confidence interval: 0.74, 0.87) (P = .03 for the comparison).

Discussion

In our study, we showed that radiomic phenotypes reflect intrinsic properties of mammographic parenchymal complexity beyond breast density and have an independent association to breast cancer. Unlike mammographic density, texture features can capture the spatial distribution of the more subtle and localized complexity of the parenchymal pattern. We identified four phenotypes that represent different degrees of parenchymal complexity and were differentially associated with age, BMI, breast density, and other established risk factors. The low- to intermediate-complexity phenotype had the lowest proportion of dense breasts (eight of 390 [2.1%]), whereas similar proportions were observed across other phenotypes (from 140 of 291 [48.1%] to 275 of 511 [53.8%]). These phenotypes were also associated with breast cancer, even when accounting for breast density and BMI, resulting in higher discriminatory capacity (area under the curve, 0.84 vs 0.80; P = .03 for comparison), suggesting their potential to provide complementary information for improving estimation of breast cancer risk.

Different from prior studies, which relied on sampling primarily the retroareolar breast region (9,10) or single feature measurements from the entire breast (11–15), our study uses a lattice-based strategy to extract texture features from multiple regions covering the entire breast to better capture the heterogeneity of the parenchymal pattern. In addition, most previous studies have evaluated parenchymal texture analysis in supervised classification (9–15). Here we expand on prior work (14) to identify radiomic phenotypes generated from unsupervised hierarchical clustering to gain insight into the variation of potentially intrinsic mammographic parenchymal patterns on a population basis. By reducing the high-dimensional radiomic features into concrete and robust phenotype groups, our approach is principled and could ultimately alleviate potential overfitting in risk prediction models, thereby augmenting their ability to generalize in larger populations.

Both qualitative and quantitative breast density measures varied across complexity phenotypes. This is expected because breast density is an inherent component of the parenchymal pattern. Nevertheless, density was different primarily for the low- to intermediate-complexity phenotype but similar across the other phenotype clusters. Interestingly, the lowest complexity phenotype had the highest proportion of women with high breast density, whereas the low- to intermediate-complexity phenotype had few women with high breast density. This observation further supports the complementary aspects of the parenchymal pattern captured with the complexity phenotypes compared with density. For example, women with very dense breasts could have low overall complexity because their entire breast is predominantly dense, therefore rendering a rather homogeneous parenchymal pattern, whereas women with scattered fibroglandular densities may have a more complex parenchymal pattern due to the presence of higher degree of inherent heterogeneity in their parenchymal tissue.

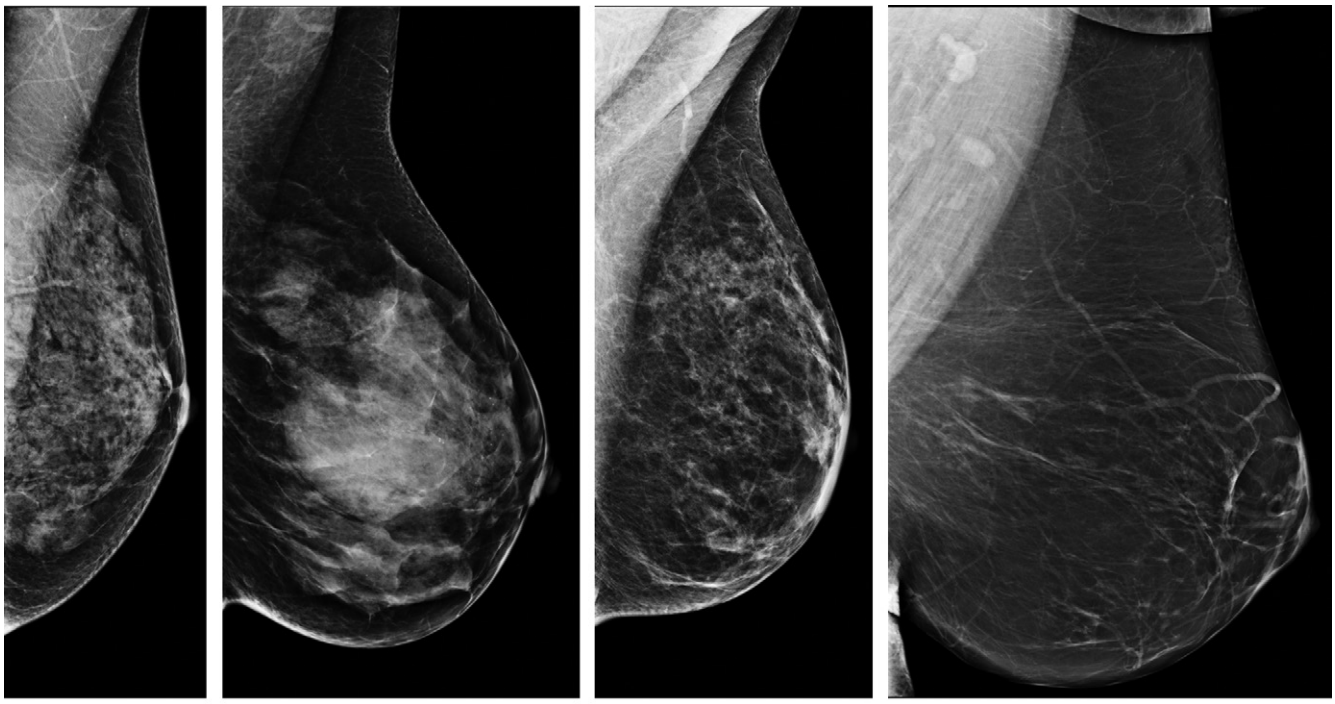


Figure 4: Images show examples of negative digital screening mammograms of women with **(a)** high density and high complexity (percent density [PD], 42%; complexity score [CS], 0.6), **(b)** high density and low complexity (PD, 43%; CS, −0.6), **(c)** low density and high complexity (PD, 14%; CS, 0.7), and **(d)** low density and low complexity (PD, 8%; CS, −0.7), demonstrating differences in breast density versus breast complexity. Corresponding age and body mass index for these women were (**a**) 54 years and 19.5 kg/m^2, (**b**) 46 years and 28.7 kg/m^2, (**c**) 57 years and 17.9 kg/m^2, and (**d**) 52 years and 48.6 kg/m^2, respectively.

Table 2: Proportion of Women with Dense Breasts (BI-RADS Density Categories *c* and *d*) in Screening Sample across the Identified Parenchymal Complexity Phenotypes

	Parenchymal Complexity Phenotypes				
Density Category	Light Green	Dark Green	Dark Red	Bright Red	Total
Low density*	236	382	408	151	1177
High density†	275	8	429	140	852
High density (%)	53.8	2.1	51.3	48.1	42
Total‡	511 (25.2)	390 (19.2)	837 (41.3)	291 (14.3)	2029

Note.—Test for linear association between membership in phenotypes of increasing parenchymal complexity and Breast Imaging Reporting and Data System (BI-RADS) density shows weak association between breast density and parenchymal complexity phenotype clusters (Spearman ρ correlation for linear trend, 0.11; 95% confidence interval: 0.07, 0.15; P = .001). Light green = low complexity, dark green = low to intermediate complexity, dark red = intermediate to high complexity, bright red = high complexity.

* Indicates BI-RADS categories *a–b.*

† Indicates BI-RADS categories *c–d.*

‡ Data in parentheses are percentages.

Important limitations of our study must also be noted. As a first step, our analyses were restricted to white women to avoid any potentially unknown feature differences due to ethnicity. In addition, images were analyzed by using a previously validated fixed feature set for texture analysis. Further, the extracted features were averaged within and across breasts to generate a parsimonious feature set given our relatively small sample, albeit limiting the information extracted. We also did not have information on menopause status and use of postmenopausal hormone therapy, which could affect breast density and the extracted parenchymal texture features. In addition, our case-control set was a relatively small convenience sample, with limited power to detect associations of the phenotypes with breast cancer. Our screening sample was also not sufficiently powered to detect significant associations with screening outcomes. Finally, we did not have available tissue or sufficient numbers of corresponding biopsies or tumor subtypes in these populations to allow for biologic interpretation of the observed phenotypes, which will be the subject of future investigation. We expect these parenchymal phenotypes to provide greater insight to the established histologic associations with breast density, including potentially the lack of or partial involution and increased breast density (29,30). In prior studies, we have seen large percentages of women with high density and partial involution (30), who may represent women in the low-complexity phenotype. Given the intriguing preliminary evidence, further work will seek to use larger, multiethnic cohorts with more comprehensive

feature sets and complementary approaches—such as deep learning—to independently validate these phenotypes in association to breast cancer risk and screening outcomes for both digital mammography and breast tomosynthesis.

The identification of women at high risk for breast cancer is becoming increasingly important in guiding personalized screening and prevention. Use of routine breast screening is an optimal setting for identifying image-based risk factors to improve this assessment. In fact, studies have shown a potential to improve risk assessment models by including breast density measures (31). We envision a breast cancer risk assessment clinic in which fully automated risk assessment tools that include quantitative imaging data can be combined with education on risk assessment and the effectiveness of screening mammography. Such data-driven tools will aide in shared decision making surrounding tailored screening, surveillance, and preventative interventions. Although important progress has been made, the gains in performance reported to date in risk model discriminatory capacity have been modest at best (31). Toward this end, our study defined distinct phenotypes of mammographic parenchymal complexity by using a quantitative radiomic feature analysis approach. Overall, our data suggests that these phenotypes capture intrinsic parenchymal complexity characteristics that could supplement breast density and other established risk factors to ultimately improve estimation of breast cancer risk.

Table 3: Characteristics for the Independent Case-Control Sample

Characteristic	Overall (n = 234)	Control (n = 158)	Case (n = 76)	P Value*
Phenotype				.001
Green	61 (26.1)	28 (17.7)	33 (43.4)	
Dark green	44 (18.8)	29 (18.4)	15 (19.7)	
Dark red	95 (40.6)	71 (44.9)	24 (31.6)	
Red	34 (14.5)	30 (19.0)	4 (5.3)	
Age (y)†	53.9 ± 7.9	53.8 ± 7.8	54.0 ± 8.2	.86
Age category (y)				.93
40–44	38 (16.2)	26 (16.5)	12 (15.8)	
45–49	35 (15.0)	23 (14.6)	12 (15.8)	
50–54	44 (18.8)	29 (18.4)	15 (19.7)	
55–59	67 (28.6)	48 (30.4)	19 (25.0)	
60–64	26 (11.1)	18 (11.4)	8 (10.5)	
65–69	18 (7.7)	10 (6.3)	8 (10.5)	
70–74	6 (2.6)	4 (2.5)	2 (2.6)	
BMI (kg/m^2)†	25.5 ± 5.0	23.9 ± 3.5	29.1 ± 6.0	.001
BMI category				
Missing data	31 (13)	…	…	…
<18.5 kg/m^2	5 (2.5)	4 (2.8)	1 (1.6)	.001
18.5–24 kg/m^2	108 (53.2)	94 (66.7)	14 (22.6)	.001
25–29 kg/m^2	58 (28.6)	36 (25.5)	22 (35.5)	.001
≥30 kg/m^2	32 (15.8)	7 (5.0)	25 (40.3)	.001
Percent density†	30.6 ± 14.9	29.1 ± 14.2	33.8 ± 15.7	.03

Note.—Unless otherwise specified, data are numbers, with percentages in parentheses. BMI = body mass index.

* P values for assessment of significant differences across case-control groups calculated with χ^2 and Fisher exact test for categorical variables, and Student t test for continuous variables.

† Data are means ± standard deviations.

Table 4: Association between Parenchymal Complexity Phenotype and Breast Cancer Diagnosis (Logistic Regression), Unadjusted and Adjusted for Standardized Percent Density and BMI

Variable	Unadjusted Odds Ratio	P Value*	Adjusted Odds Ratio†	P Value*
Phenotype		.001		.001
Green	3.5 (1.8, 6.9)		2.5 (1.1, 6.0)	
Dark green	1.5 (0.7, 3.3)		2.3 (0.8, 6.7)	
Dark red	Reference		Reference	
Red	0.4 (0.1, 1.2)		0.1 (0.0, 0.5)	
Percent density per standard deviation	1.3 (1.0, 1.8)	.027	2.7 (1.7, 4.4)	.001

Note.—Data in parentheses are 95% confidence intervals. Likelihood ratio test for presence of interaction between phenotype and percent density is not statistically significant (P = .363). Likelihood ratio test for presence of interaction between phenotype and body mass index (BMI) is not statistically significant (P = .335).

* P values are for the Type-3 hypothesis test, testing for overall association between the factor of interest and case-control status.

† Cases: n = 62. Control participants: n = 141.

Author contributions: Guarantor of integrity of entire study, D.K.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, D.K., A.G., K.K., E.F.C., C.M.V.; clinical studies, D.H.W., E.F.C.; experimental studies, D.K., A.O., M.K.H., A.G.; statistical analysis, D.K., S.J.W., A.O., M.K.H., C.M.V.; and manuscript editing, D.K., S.J.W., A.O., M.K.H., A.G., D.H.W., C.B.H., K.K., K.B., E.F.C., C.M.V.

Disclosures of Conflicts of Interest: **D.K.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed no relevant relationships. Other relationships: holds patent no. US8634610B2. **S.J.W.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: has grants/grants pending with the National Institutes of Health. Other relationships: disclosed no relevant relationships. **A.O.** disclosed no relevant relationships. **L.P.** disclosed no relevant relationships. **M.K.H.** disclosed no relevant relationships. **A.G.** disclosed no relevant relationships. **D.H.W.** disclosed no relevant relationships. **C.B.H.**

Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: is inventor of patents licensed to Gamma Medica (acquired by CMR Naviscan); received royalties for technologies licensed to Gamma Medica. Other relationships: disclosed no relevant relationships. **K.K.** disclosed no relevant relationships. **K.B.** disclosed no relevant relationships. **E.F.C.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: holds current grants and grants pending with Hologic and iCAD. Other relationships: disclosed no relevant relationships. **C.M.V.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: has grants/grants pending with Grail; institution received payment from Grail for travel/accommodations/meeting expenses unrelated to activities listed. Other relationships: disclosed no relevant relationships.

References

1. Kaufhold J, Thomas JA, Eberhard JW, Galbo CE, Trotter DE. A calibration approach to glandular tissue composition estimation in digital mammography. Med Phys 2002;29(8):1867–1880.
2. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. AJR Am J Roentgenol 1976;126(6):1130–1137.
3. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15(6):1159–1169.
4. 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.
5. Brandt KR, Scott CG, Ma L, et al. Comparison of clinical and automated breast density measurements: implications for risk prediction and supplemental screening. Radiology 2016;279(3):710–719.
6. Kerlikowske K, Zhu W, Tosteson AN, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. Ann Intern Med 2015;162(10):673–681.
7. Yaffe MJ. Mammographic density: measurement of mammographic density. Breast Cancer Res 2008;10(3):209.
8. Gastounioti A, Conant EF, Kontos D. Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. Breast Cancer Res 2016;18(1):91.
9. Huo Z, Giger ML, Olopade OI, et al. Computerized analysis of digitized mammograms of BRCA1 and BRCA2 gene mutation carriers. Radiology 2002;225(2):519–526.
10. Gierach GL, Li H, Loud JT, et al. Relationships between computer-extracted mammographic texture pattern features and BRCA1/2 mutation status: a cross-sectional study. Breast Cancer Res 2014;16(4):424.
11. Manduca A, Carston MJ, Heine JJ, et al. Texture features from mammographic images and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2009;18(3):837–845.
12. Häberle L, Wagner F, Fasching PA, et al. Characterizing mammographic images by using generic texture features. Breast Cancer Res 2012;14(2):R59.
13. Wei J, Chan HP, Wu YT, et al. Association of computerized mammographic parenchymal pattern measure with breast cancer risk: a pilot case-control study. Radiology 2011;260(1):42–49.
14. Zheng Y, Keller BM, Ray S, et al. Parenchymal texture analysis in digital mammography: a fully automated pipeline for breast cancer risk assessment. Med Phys 2015;42(7):4149–4160.
15. Heine JJ, Scott CG, Sellers TA, et al. A novel automated mammographic density measure and breast cancer risk. J Natl Cancer Inst 2012;104(13):1028–1037.
16. National Cancer Institute. Breast Cancer Risk Assessment Tool. http://www.cancer.gov/bcrisktool/. Accessed August 21, 2018.
17. Breast Cancer Surveillance Consortium. Breast Cancer Surveillance Consortium Risk Calculator. https://tools.bcsc-scc.org/BC5yearRisk/intro.htm. Accessed August 21, 2018.
18. D'Orsi CJ, Bassett LW, Berg WA, et al. BI-RADS: Mammography. In: D'Orsi CJ, Mendelson EB, Ikeda DM, et al, eds. Breast Imaging Reporting and Data System: ACR BI-RADS–Breast Imaging Atlas. 4th ed. Reston, Va: American College of Radiology, 2003.
19. Keller BM, Chen J, Daye D, Conant EF, Kontos D. Preliminary evaluation of the publicly available Laboratory for Breast Radiodensity Assessment (LIBRA) software tool: comparison of fully automated area and volumetric density measures in a case-control study with digital mammography. Breast Cancer Res 2015;17(1):117.
20. Laboratory for Individualized Breast Radiodensity Assessment (LIBRA). https://www.cbica.upenn.edu/sbia/software/LIBRA/. Accessed August 21, 2018.
21. Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern 1989;19(5):1264–1274.
22. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610–621.
23. Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run lengths for texture analysis. Pattern Recognit Lett 1990;11(6):415–419.
24. Li H, Giger ML, Olopade OI, Lan L. Fractal analysis of mammographic parenchymal patterns in breast cancer risk assessment. Acad Radiol 2007;14(5):513–521.
25. Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002;24(7):971–987.
26. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406(6797):747–752.
27. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer, 2003.
28. Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York, NY: Wiley, 2000.
29. Ghosh K, Vierkant RA, Frank RD, et al. Association between mammographic breast density and histologic features of benign breast disease. Breast Cancer Res 2017;19(1):134.
30. Ghosh K, Hartmann LC, Reynolds C, et al. Association between mammographic density and age-related lobular involution of the breast. J Clin Oncol 2010;28(13):2207–2212.
31. Cintolo-Gonzalez JA, Braun D, Blackford AL, et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. Breast Cancer Res Treat 2017;164(2):263–284 [Published correction appears in Breast Cancer Res Treat 2017;164(3):745.] https://doi.org/10.1007/s10549-017-4247-z.
32. Keller BM, Oustimov A, Wang Y, et al. Parenchymal texture analysis in digital mammography: robust texture feature identification and equivalence across devices. J Med Imaging (Bellingham) 2015;2(2):024501.
33. Ward JH Jr. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 1963;58(301):236–244.
34. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn 2003;52(1-2):91–118.
35. Liu Y, Hayes DN, Nobel A, Marron JS. Statistical significance of clustering for high-dimension, low–sample size data. J Am Stat Assoc 2008;103(483):1281–1293.