Archives Of Breast Cancer

Association between Radiological and First-Order Statistical Features of the Mammogram, and the Tumor Phenotype in Breast Cancer Patients

Hoda Mahdavi*[a,b] (iD), Vahid Kaveh[c] (iD), Fatemeh Karami[d] (iD), Mohammad Amin Rahimi[e] (iD)

[a]*Radiation Oncology Department, School of Medicine, Iran University of Medical Sciences, Tehran, Iran*

[b]*Firoozgar Clinical Research Development Center, Iran University of Medical Sciences, Tehran, Iran*

[c]*Hematology and Oncology Department, School of Medicine, Iran University of Medical Sciences, Tehran, Iran*

[d]*Shafa Radiology Center, Isfahan, Iran*

[e]*School of Medicine, Iran University of Medical Sciences, Tehran, Iran*

ARTICLE INFO　　　ABSTRACT

Background: Breast cancer is among the most prevalent cancers which can effectively be screened by mammography. The spatial distribution of grey levels of the mammogram known as first-order statistical features (FOSFs) contain higher dimensional data which describe the breast composition. We aim to test these basic measures to differentiate density categories in breast cancer patients, and use them as covariates to investigate the relationship between radiologic and pathologic features of the tumor.

Methods: Data from 85 breast cancer patients, their BI-RADS breast density category (a to d), percentage density (PD), and FOSFs of the mammogram including median, mean, interquartile range, kurtosis, maximum, minimum, standard deviation, skewness, and energy, were extracted. The tumor grade and the percentage of Ki67, ER, PR, and Her2 status were recorded. A linear discriminant analysis, and a support vector machine (SVM) were used to discriminate each density category from others. Then, the relation between the variables was investigated using ANCOVA and regression analysis.

Results: Density categories a and d were classified by SVM with high accuracy. The key feature of a and d were interquartile range and maximum intensities, respectively. Reported tumor margins were related to Her2 overexpression and PR positivity. Spiculated tumor margin predicted the percentage of PR expression, with a cumulative odds ratio of 7.85 (CI 2.5- 24.78), when adjusted for age, area of breast, density, and FOSFs (P=0.0004).

Conclusion: The findings of this study suggest that FOSFs can be incorporated in computer-aided systems to adjust for differences in breast composition and to refine risk profiles.

*Address for correspondence:
Hoda Mahdavi, MD,
School of Medicine, Iran University of Medical Sciences, Tehran, Iran/Firoozgar Hospital, Beh-Afarin St., Karimkhane-Zand Blvd., Tehran, Iran
Email: Mahdavi.h@iums.ac.ir

INTRODUCTION

Breast cancer is among the most prevalent cancers of adult women of any age, but hopefully it can be screened or diagnosed early in the majority of incidents. The importance of early diagnosis has led to several screening guidelines, with mammography

as the leading procedure.[1] Reportedly, high mammographic density, due to the high proportion of epithelial-stromal tissue and complex structure of the glands, is an independent risk factor for breast cancer incidence with an anticipated relative risk ranged from 1.8 to 6.0.[2-5] The concealing effect of the dense background causing reduced sensitivity for detection of precancerous and cancerous lesions in mammography makes the owner prone to cancers appearing at screening intervals.[5-7] Since the higher incidence of cancer does not attenuate by time or repeated screening, another possibility is the higher volume of ducts or a more suitable microenvironment for tumor growth where the cancer originates, causing higher cancer incidence.[8-11] The current 5th edition of the Breast Imaging-Reporting and Data System (BI-RADS) criteria for categorizing breast density has replaced estimates of dense area percentage with descriptions about the possibility of obscured lesions[12], which may reflect information related to different breast compositions.

Several computer algorithms have been designed to extract texture features and convert them into higher-dimensional data. For instance, Support Vector Machine (SVM) is a type of automated binary machine learning algorithm that has been used to distinguish and classify radiological patterns of mammograms. Their potential clinical implication is to improve accuracy, reduce intra and inter observer variability, and support decision-making by the radiologists.[13,14] The first-order statistical features (FOSFs) of the intensity histogram of mammograms may be considered as radiomic features which describe the distribution of intensity voxels by basic metrics. FOSFs include parameters which provide information beyond the point of human visual perception, and is possibly related to cancer incidence.[15-19] Due to lower accuracy of the mammogram and higher susceptibility to cancer incidence in dense breasts, updated screening recommendations endorse early risk assessment and supplemental screening if the risk is considered higher-than-average.[20]

While basic mammographic features are readily extractable, their potential correlation with pathological or radiological findings offers opportunities for enhancing computer-aided diagnostic and prognostic models. This study aims to investigate the relation between objective (measured) and subjective (radiologist's interpretation) assessments of breast density, investigate the relationship between mammographic FOSFs and density, and explore the potential of FOSFs as covariates in modeling the association between tumor characteristics and pathology.

METHODS

Study design and participants

In this cross-sectional study, all women who underwent diagnostic mammography for evaluation of their breast mass or abnormality and had a diagnosis of primary breast cancer at Shafa Imaging Center, Isfahan, Iran, between 2016 and 2019, were the target population included in our study. Patients with histories of past breast cancer, recurrence, surgery or radiation therapy on the affected breast, and those with unavailable digital mammograms were excluded from the study.

The general approach to suspicious lesions was as follows: those with BI-RADS scores of 4 or 5 in their mammography report, and also in some cases of BI-RADS 0 or 3 with a high suspicion of malignancy as determined by the radiologist based on history, clinical, or subsequent ultrasonographic data, were candidates for core needle biopsy. None of the patients had prior diagnosis of breast cancer, so the radiologist was not aware of the pathology results when reporting the images. Biopsies were undertaken by the coauthor radiologist via standard 14 G × 100 mm core needle for breast biopsy (Medax, San Possidonio, Italy) with ultrasonic guidance.

Data collection

In accordance with the ethical standards of the ethics committee on human experimentation a verbal informed consent was performed from patients to use the information from the paper and electronic documents of the center. The demographic, mammographic, and pathologic data were collected, but missing data was allowed. Information related to family history of breast cancer, age at menarche, age at first pregnancy, number of pregnancies, and history of oral contraceptive pills (OCP) intake or hormone replacement therapy (HRT) were asked over the phone and recorded. Exclusively, pathology reports from the academic pathologist in charge of the core needle biopsy which was performed by the co-author radiologist, were obtained from patients' files. The required information included histology, grade, and hormone receptor positivity and percentage of immunohistochemistry (IHC) positive staining of the cancerous lesions. The ER, PR, Her2, Ki67 kit used for this means was rabbit anti-human monoclonal antibody (Master diagnostica, Granada, Spain). ER and PR positivity were described as at least 1% stain. We used low (5-80%) and high (90% and above) positivity to describe hormone positive cases. Relying exclusively on IHC results, Her2 positivity was defined as protein overexpression (score 3+), Her2 equivocal or negative stains were grouped as otherwise. In addition, the four molecular subtypes of breast cancer were categorized as luminal A: hormone

receptor positive with Ki 67 less than 14%, luminal B: Ki 67 of at least 14%, Her2 enriched: hormone receptor negative, Her2 overexpression, and triple negative: all three receptors negative.

Mammograms

All patients had diagnostic full field digital mammograms with cranial caudal (CC) and mediolateral oblique (MLO) views for each breast for the diagnosis of the suspicious lesion before biopsy using Hologic, Selenia mammography system with similar settings. In order to eliminate interobserver variability, all images were interpreted by the co-author radiologist who has a ten-year sub-specialty expertise in breast image reporting. The required variables included mass density, compared to an equivalent volume of fibro-glandular tissue, and tumor margins, including micro lobulated, indistinct, and spiculated. The four descriptors for breast composition according to the 5th edition of the BI-RADS based or the American College of Radiology criteria [12] are as follows: a: almost entirely fatty, b: scattered areas of fibro glandular density, c: heterogeneously dense, and d: extremely dense.

Image variables

By means of the Cancer Imaging Phenomics Toolkit v.1.8.1 platform, all of the four mammography views were analyzed by the automated Laboratory for Individualized Breast Radiodensity Assessment algorithm (LIBRA). The algorithm detects the breast parenchyma outline and then segments absolutely dense or white stromal-epithelial clusters, and then it normalizes the ratio of the dense tissue area to the total breast area which results in breast percent density (PD). [21,22] The texture feature pipeline was also executed for texture analyses. We selected the FOSFs that included mean, the average intensity of voxels; median, the mid intensity of the voxels; variance, the variation of the voxels on histogram; skewness, the measure of histogram asymmetry; kurtosis, the weight of tails of the histogram; interquartile range, the range between the 25th and 75th percentile, and energy, the sum of squared intensities of the histogram. The formula is described in https://cbica.github.io/CaPTk/tr_FeatureExtraction.html. The area of the breast was used as a surrogate for breast size.

Statistical analysis

Descriptive statistics were reported as frequency (percentage) for categorical variables and mean (SD) for quantitative variables. Analyses were performed by Spearman's correlation coefficient to analyze the relationship between continued quantitative variables.

In order to remove redundant variables, we executed a one-way analysis of covariance test which served as a filter for feature selection by determining their differences between a to d categories of breast density, and post hoc to understand pairwise differences. Statistics and Machine Learning Toolbox of MATLAB R2019a (Math Works, MA, USA) software was used for a one versus others approach of BI-RADS density classification by selected FOFSs and PD. The discriminating accuracy for classification of the linear discriminant analysis (LDA), and the support vector machine (SVM) was tested. The classifier parameters were heuristically tuned and validated using the k-fold cross validation technique (k=10). Then, ranking the key features by class separability criteria was done. Finally, in order to explore if the percentage of hormone receptor positivity is predicted by tumor radiological features, ordinal logistic regressions were performed, and the cumulative odds value was calculated as e^β. IBM SPSS Statistics (v. 16, IBM Corp., Armonk, NY, USA) was used for statistical analyses. In all statistical tests, a P-value less than 0.05 was considered statistically significant.

RESULTS

In this study, data from 85 patients were analyzed who all had unifocal and unilateral lesions of in situ or invasive carcinoma. The age range of patients was 36 to 79 years. Ninety percent of patients had a history of child birth, with an average of 3 children, the age at first pregnancy ranged from 15 to 37. A family history of breast cancer in first- or second-degree relatives was present in 21%. A history of regular intake of OCP was reported in 46%. None of the menopausal patents had a history of HRT. Mammogram data for tumor size were available for only 20 patients. Sixteen patients were T1, two were T2, and two others were T3 according to tumor size. Other demographic information is listed in Table 1. There was statistically significant variation between categories of breast density groups in the means of PD after adjusting for age and breast area, determined by one-way ANCOVA. Cases of different density categories were not significantly different for the kurtosis, mean, standard deviation, and minimum measure of their normalized intensity histograms. Other histogram FOSFs including median, interquartile range, maximum, skewness, and energy, were significantly different in breast density categories (Table 2).

Radiological features and demographics

The radiological features of the study population include the information reported by the radiologist which is shown in Table 1.

Table 1. patient characteristics

General (years)	M (SD)
Age	53.25 (10.5)
Menarche	12.82 (1.9)
First pregnancy	20.31 (3.99)
Histology	Number (%)
Invasive ductal carcinoma	71 (83.5)
Mucinous carcinoma	4 (4.7)
Invasive lobular carcinoma	3 (3.5)
Adenoid Cystic Carcinoma	1 (1.2)
Adenocarcinoma of the breast	4 (4.7)
Grade of cancerous lesions	
1	15 (19.2)
2	35 (44.9)
3	28 (35.9)
Immunohistochemistry	
ER	
high positive (90% ≤)	44 (51.2)
low positive (5-80%)	19 (22.1)
70-80%	13 (15.1)
40-60%	1 (1.2)
5-30%	5 (5.8)
negative	9 (10.5)
PR	
positive high (90% ≤)	37 (43)
positive low (5-80%)	22 (24.6)
70-80%	13 (15.1)
40-60%	5 (5.8)
5-30%	4 (4.7)
negative	16 (18.6)
Her2	
Overexpression (3+)	11 (14.7)
Otherwise	64 (85.3)
Molecular subtypes	
Luminal A	41 (55.4)
Luminal B (Her2 negative)	14 (19.1)
Luminal B (Her2 positive)	7 (9.6)
Her2 enriched	3 (4.1)
Triple negative	8 (10.8)
Radiologic	
Breast density	
A	7 (8.2)
B	26 (30.6)
C	28 (32.9)
D	23 (27.1)
Tumor density	
Isodense or low	26 (33.8)
high	51 (66.2)
Tumor margin	
Ill-defined or indistinct	24 (30.8)
Micro-lobulated	10 (12.8)
Spiculated	44 (56.4)

The mean age (SD) in density categories of A, B, C and D were 66 (6.9), 58 (9.0), 50 (9.9), and 46 (6.5) years, respectively. Among the demographic variables, there were significant negative relations between PD and age (r_s= -.052, P<0.001), and between PD and breast mean pixel area (r_s= -0.6, P<0.00001). A significant relationship between the

age variable and PD, age and several mammography histogram variables including the median (P<0.001), interquartile range (P<0.001), kurtosis (P<0.001), maximum (P<0.001), minimum (P<0.001), standard deviation (P=0.001), skewness (P<0.001), and energy (P=0.003) was also evident. In addition, a significant difference was present between the PD and number of pregnancies (P=0.04), but this was not significant when controlled for age. No relationship was observed between density and first pregnancy age, menarche, family history of breast cancer, or OCP consumption. Also, the relationship between tumor margin and age or other demographic variables was not statistically significant.

Pathology and demographics
In the studied population, the percentage of ER and PR decreased with age, but this was not statistically significant. The mean age in Her2 overexpressed cases was lower than the mean age in otherwise Her2 state, F (1, 73) = 5.7, P = 0.02, 46 years vs. 54 years. No significant relationship was detected between tumor subtype and family history of breast cancer.

Radiological features and breast density
Correlations between the PD and FOSFs are shown in Table 4. Subsequently, the PD and selected FOSFs from the ACNOVA (Table 2), were used to discriminate each BI-RADS density category from others. The classification accuracy of the LDA and SVM classifiers are presented in Table 3 which demonstrates the highest (> 90%) accuracy was via SVM and a category. The rank feature function showed that the most important FOSFs for class discrimination was interquartile range for a and c, but PD and energy came second for each accordingly. The key feature of d was Maximum, and PD for b, both followed by skewness. No statistically significant relationship was found between the categories of breast density and tumor density or margin (P=0.49, P=0.08).

Radiological features and pathology
ANCOVA test with age and mammographic area as covariates showed that molecular subtype, ER status, or Her2 overexpression of the tumor were not differentiated by PD. The tumor subtypes were not related to the tumor margin, but a significant relationship between Her2 overexpression and tumor margins of micro-lobulated, ill-defined, and spiculated margins existed, $\chi^2(2)$= 9.86, P=0.007. A nominal logistic regression model for tumor margin prediction of hormone receptor expression (negative, low positive, high positive) was statistically significant for tumor margin expressing PR,

indicating that if the margin is spiculated, the probability of highly positive (90% and above) expression compared to PR negative is of OR= 6.25, (CI 1.23—31.84), P=0.03.

Table 2. One-way analysis of covariance of breast density categories on first-order statistical features and percentage density of the mammogram

| | variables | mammographic breast density categories | | | | F (3, 75) | *p* | *post hoc* |
		a	b	c	d			
adjusted means†	percentage density	17.57	14.14	18.4	28.06	7.38	< 0.001	b < d**, c < d**
	interquartile range	0.35	0.22	0.36	1.22	7.09	< 0.001	b < c**, b < d**, c < d**
	median	-0.42	-0.36	-0.43	-0.57	5.09	0.003	b > d**, c > d*
	energy	1579	2176	2078	2029	4.94	0.003	a < b**, a < c*
	maximum	2.63	3.11	2.7	2.03	5.16	0.03	b > d**, c > d**
	skewness	1.83	2.18	1.86	1.19	5.73	0.001	b > d**, c > d*

*P<0.05, **P<0.01, † age and area of the breast as covariates.

Table 3. Classification accuracy to discriminate each BI-RADS category from other groups based on first order statistical features.

| | Classification categories versus others | | | |
	a	b	c	d
Linear discriminant analysis	53.01	66.27	50.6	81.93
Support vector machine	91.57	72.29	67.47	83.13

An ordinal regression model of margin spiculation predicting a five-level low to high ranks of hormone receptor percentage, adjusted for age, breast area, density percentage, and FOSFs, showed that the proportional odds model for PR had a negative effect β= -2.06 which was statistically significant according to Wald test with P = 0.0004 (Table 4).

When comparing non-spiculated mass to spiculated mass, the cumulative odds ratio was 0.13,

which means that there is an OR equal to 7.85 (CI 2.5-24.78) higher chance to have PR positive tumors with higher receptor expression. The similar regression was not significant for ER positive percentage ranks (P= 0.21).

DISCUSSION

The findings of this study showed some associations between demographic, radiological, and pathological features of the breast composition and the tumor. PD significantly correlated with most FOSFs of the mammogram, and FOSFs of histograms used by the SVM algorithm, besides PD, could discriminate a or d BI-RADS density groups from others with an acceptable accuracy. Also, the spiculated border of the tumor could predict higher PR percentage positivity when the images were adjusted for age, area of the breasts, and FOSFs.

Table 4. Spearman's correlations of the features of the mammographic histogram

| | | Density | first-order statistical features | | | | | | | | |
		percent	Interquartile range	Median	Energy	Kurtosis	Maximum	Mean	Minimum	Skewness	Standard Deviation
Density	percent	1	.748**	-.910**	-.625**	-.977**	-.979**	.090	.953**	-.974**	-.678**
first-order statistical features	Interquartile range		1	-.705**	-.478**	-.754**	-.753**	.067	.770**	-.753**	-.415**
	Median			1	.549**	.927**	.924**	-.098	-.944**	.921**	.666**
	Energy				1	.618**	.621**	.152	-.606**	.645**	.237*
	Kurtosis					1	.999**	-.149	-.972**	.996**	.670**
	Maximum						1	-.127	-.973**	.996**	.675**
	Mean							1	.128	-.109	-.014
	Minimum								1	-.968**	-.687**
	Skewness									1	.637**
	Standard Deviation										1

*P<0.05, **P<0.01

The relationship between age and PD or FOSFs suggests an effect of breast texture change by age. As it has been studied extensively[11,21], breast composition is affected by factors such as age and body mass index (BMI), hormone intervention, growth factors, and nutrition. Although genetic factors may explain the majority of the etiology of breast density, only few genes have been detected so far, whose proliferative activity is linked to breast dense composition and carcinogenesis as well. The acquired factors are estimated to explain 20-30% of etiology of breast density and some are shared with cancer etiology,[11]

Our findings did not suggest any link between hormone receptor positivity of cancer and dense breasts. Previous studies have inconsistently indicated that high breast density seems to increase the risk of all breast cancer subtypes[16,23], but several studies indicate a stronger link between hormone positive breast cancers and breast density. A number of normal breast epithelial cells express PR and ER, which are involved in growth and proliferation. Blocking hormone receptors can reduce both cancer incidence and breast density. Moreover, there is evidence that at least some microenvironmental factors such as collagen type 1, small leucine-rich proteoglycans, high expression of DNA damage response genes, and downregulation of CD36 are etiologies shared for breast density and cancer[24] many of which are more or less thought to be mediated by hormone receptors.[8-10]

Patients who posed Her2 overexpressed tumors in our study were younger. Accordingly, some epidemiologic studies have shown that Her2 positive tumors, like other aggressive histologies, have a higher prevalence in young women in comparison with older patients[25,26], but evidence is limited, in particular, across various ethnicities. Another finding of this study was that Her2 overexpression had a moderate effect on the tumor margin, and tumor spiculation could predict the higher percentage of PR expression. Manifestations of different molecular subtypes of breast tumors have been radiologically described. Our findings from the logistic regression were in agreement with those studies which associate margin spiculation with positivity of hormone receptor markers. Similar studies have evaluated this relation in which different cut-off values of IHC percentage were used to define the positivity of hormone receptors.[27-31] Luminal type tumors are likely to form desmoplastic interactions with the surrounding stroma that cause the stretch of the Cooper's ligaments causing margin spiculation.[27,32] Conversely, the frequency of spiculated margins is less in Her2 overexpressed[32], or triple-negative tumors.[28] Still Her2 positive tumors usually present with spiculated margins, while triple negative tumors are more likely to present with round borders resembling benign masses.[33] Abundant PR is extremely rare when ER is negative, so most PR positive tumors express ER as well. PR is a predictive marker and absence of PR receptors even in the presence of ER is detriment to survival.[34] Most invasive carcinomas express PRA isoform dominancy which is responsible for invasiveness via cross-signaling with ER.[35,36] Therefore, linkage of tumor features and PR can provide insights into the prognosis of the patient.

We assume that FOSFs extracted by the algorithms are reproducible and operator-independent and the data of the combined histograms from all four mammographic views primarily represents the composition of the uncancerous breast tissue.[11] Mammograms with different texture in the normal population may show different susceptibility to cancer incidence. A study on screening mammograms indicated that a more uniform parenchyma and a higher density percentage were each independently proportional to the incidence of cancer. In a study, an unsupervised clustering could accurately stratify mammograms into four phenotypes based on radiomic complexity scores and density, in which subjects who were diagnosed with breast cancer had higher low-complexity and low- to intermediate-complexity phenotypes and those with very dense breasts were mostly in the minimum complexity group, so complexity was not directly related to density.[15,18] Some studies attempted to find specific prognostic models to predict the incidence of ER positive cancer subtypes, which have been linked to breast dense composition.[16,17] A case-control study indicated that a variance measure of mammographic histograms could act as a possible indicator of breast density with a stronger relationship to breast cancer incidence than the calibrated 'percent glandular'.[19] Also, a larger study indicated that manual and automated percentage of breast tissue density as well as variance are independently related to breast cancer.[37] The evidence may reflect carcinogenesis is proceeded by changes of breast composition parameters.

Except for mean, minimum, and kurtosis which were supposed to be almost uniform in the standardized mammographic histograms, mean values of inter quartile range, skewness, median, maximum, and energy of the images were diverse across the categories of breast density. These variables were not accounted in the standard calculations of breast density but were representative of texture variability. The FOSFs besides PD could almost accurately discriminate BIRADS density groups, especially for categories a and d by the SVM

algorithm. SVM accuracy of over 80-90% for discrimination of categories a and d of breast density indicates that the FOSFs besides PD are generally informative and good at discriminating each of these categories from others. Although these features exhibited strong correlations (Table 5), the rank feature function demonstrated that their contributions to distinguishing between density categories are not equivalent.

For instance, energy, a variable not so important for discriminating categories, correlated negatively with PD, and had a significant negative correlation with the minimum intensity as well, and a significant positive correlation with maximum and standard deviation. The sum of squared pixel raw intensities was used to calculate energy, while PD was calculated from the dense extreme end of the histogram. In other words, most of the energy reflects the fibro-glandular tissue in the mid-range of the histogram (Figure 1).

Our study used the most basic features none of which accounted for local variation of the image. This may explain the low accuracy of SVM classification for discriminating categories b or c. Using more complex features, this function was able to accurately classify mammograms according to their breast density using gray level co-occurrence matrices texture features.[38] One of the limitations of the current study was the retrospective nature of the data, small sample size, and relying on existing documented findings.

Table 5. Ordered logistic regression of margin spiculation predicting progesterone receptor percentage

| Variables	margin spiculation		
	Estimate	SE	P
Threshold of PR percentage			
negative	348.1	162.42	0.03
5-30%	348.5	162.43	0.03
40-60%	348.9	162.44	0.03
70-80%	349.9	162.47	0.03
≥90%			
breast area	0.01	0.003	0.06
age	0.01	0.03	0.77
density percent	0.32	0.23	0.16
interquartile range	0.49	0.60	0.41
median	78.06	15.27	0.00
energy	-0.01	0.003	0.07
kurtosis	2.16	1.09	0.05
maximum	-49.50	12.49	0.00
mean	-52.21	62.02	0.40
skewness	13.04	7.50	0.08
minimum	-22.83	0.000	-
standard deviation	378.55	168.52	0.03
margin spiculation	-2.06	0.58	0.0004
non-margin spiculation	0		
Model chi-square	28.77		0.007
Nagelkerke	0.365		
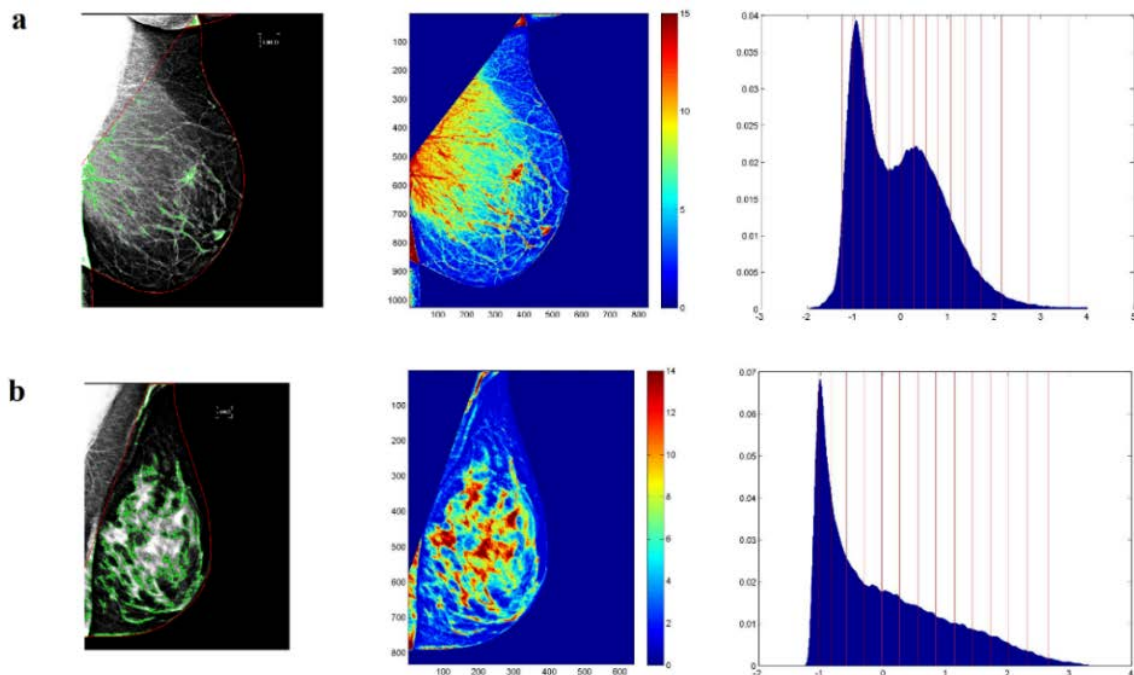
Figure 1. Examples of left mediolateral oblique mammograms of two sample cases. Left-to-right images show automated edge detection and segmented dense clusters used for percentage density (PD) calculation, color wash density values, and normalized gray level histograms of all four mammogram views of two patients. a) A Sixty-year old woman with mucinous carcinoma of the contralateral breast, grade 1, ER 90%, PR 90%, Her2 negative, Ki 67 10%, PD of 7.3%, energy of 34177, breast density category B, b) A thirty-eight year old woman with invasive ductal carcinoma of the contralateral breast, grade 3. ER 90%, PR and Her2 negative, Ki67 25%, PD of 49.0%, energy of 17582, breast density category D.

Given the distribution of pathologic subtypes, studies with higher power in specific ethnic populations are required to enable analyses of subgroups. Despite high concordance of hormone markers in needle and surgical samples[39], future studies may assess the reevaluation of negative hormone receptor stains or weak Her2 positive cases. The relationship between the radiological and morphological basis of the breast and tumor and PR positivity can be the subject of future epidemiologic studies.

CONCLUSION

The linkage between findings of imaging and pathology, although speculative at this stage, could potentially have implications to understand the pathogenesis of breast cancer. Moreover, it may inform future studies, particularly, about dense breasts in which the masking effect and higher risk of cancer may need enhancing filtered images or customized prediction models.

ETHICAL CONSIDERATIONS

The study protocol was reviewed and approved by the ethical review board of Iran University of Medical Sciences, with an ethics code number of IR.IUMS.FMD.REC.1400.112.

FUNDING

CONFLICT OF INTEREST

The authors declare no confilict of interest for this article.

ACKNOWLEDGEMENTS

REFERENCES

1. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Akinyemiju TF, Al Lami FH, Alam T, Alizadeh-Navaei R, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2016: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 2018;4:1553–1568. doi: 10.1001/jamaoncol.2018.2706.

2. Brentnall AR, Cuzick J, Buist DSM, Bowles EJA. Long-term Accuracy of Breast Cancer Risk Assessment Combining Classic Risk Factors and Breast Density. *JAMA Oncol* 2018;4:e180174. doi: 10.1001/jamaoncol.2018.0174.

3. Kerlikowske K, Ma L, Scott CG, Mahmoudzadeh AP, Jensen MR, Sprague BL, et al. Combining quantitative and qualitative breast density measures to assess breast cancer risk. Breast Cancer Res. 2017 Aug 22;19(1):97. doi: 10.1186/s13058-017-0887-5.

4. Destounis SV, Santacroce A, Arieno A. Update on breast density, risk estimation, and supplemental screening. *American Journal of Roentgenology*. 2020;214:296-305. doi: 10.2214/AJR.19.21994.

5. Ghosh K, Brandt KR, Sellers TA, Reynolds C, Scott CG, Maloney SD, Carston MJ, Pankratz VS, Vachon CM. Association of mammographic density with the pathology of subsequent breast cancer among postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2008 Apr;17(4):872-9. doi: 10.1158/1055-9965.EPI-07-0559.

6. 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;17:117. doi: 10.1186/s13058-015-0626-8.

7. Boyd N, Berman H, Zhu J, Martin LJ, Yaffe MJ, Chavez S, et al. The origins of breast cancer associated with mammographic density: a testable biological hypothesis. *Breast Cancer Res* 2018;20:17. doi: 10.1186/s13058-018-0941-y.

8. Evans DGR, Harkness EF, Brentnall AR, van Veen EM, Astley SM, Byers H, et al. Breast cancer pathology and stage are better predicted by risk stratification models that include mammographic density and common genetic variants. *Breast Cancer Res Treat* 2019;176:141–148. doi: 10.1007/s10549-019-05210-2.

9. van der Waal D, Verbeek ALM, Broeders MJM. Breast density and breast cancer-specific survival by detection mode. *BMC Cancer* 2018;18:386. doi: 10.1186/s12885-018-4316-7.

10. Edwards BL, Atkins KA, Stukenborg GJ, Novicoff WM, Larson KN, Cohn WF, et al. The Association of Mammographic Density and Molecular Breast Cancer Subtype. *Cancer Epidemiol Biomarkers Prev* 2017;26:1487–1492. doi: 10.1158/1055-9965.EPI-16-0881.

11. Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ, et al. Mammographic breast density as an intermediate phenotype for breast cancer. *lancet oncol* 2005;6:798-808. doi: 10.1016/S1470-2045(05)70390-9.

12. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES, et al. Breast imaging reporting and data system, Reston, VA: *American College of Radiology* 2013: 39-48. doi: 10.1118/1.4824319.

13. Liu J, Lei J, Ou Y, Zhao Y, Tuo X, Zhang B, et al. Mammography diagnosis of breast cancer screening through machine learning: a systematic review and

meta-analysis. *Clinical and Experimental Medicine*. 2023 Oct;23(6):2341-56. doi: 10.1007/s10238-022-00895-0.

14. Kayode AA, Akande NO, Adegun AA, Adebiyi MO. An automated mammogram classification system using modified support vector machine. *Medical Devices: Evidence and Research*. 2019 Aug 12:275-84. doi: 10.2147/MDER.S206973.

15. Pinker K. Beyond breast density: radiomic phenotypes enhance assessment of breast cancer risk. *Radiology* 2019;290:50. doi: 10.1148/radiol.2018182296.

16. 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;18:1-2. doi: 10.1186/s13058-016-0755-8.

17. Keller BM, Chen J, Conant EF, et al. Breast density and parenchymal texture measures as potential risk factors for estrogen-receptor positive breast cancer. *Proc SPIE Int Soc Opt Eng* 2014;27;9035:9035. doi: 10.1117/12.2043710.

18. Kontos D, Winham SJ, Oustimov A, Pantalone L, Hsieh MK, Gastounioti A, et al. Radiomic phenotypes of mammographic parenchymal complexity: toward augmenting breast density in breast cancer risk assessment. *Radiology*. 2019;290:41-9. doi: 10.1148/radiol.2018180179.

19. Heine JJ, Cao K, Rollison DE. Calibrated measures for breast density estimation. *Acad Radiol* 2011;18:547-555.

20. Monticciolo DL, Newell MS, Moy L, Lee CS, Destounis SV. Breast cancer screening for women at higher-than-average risk: updated recommendations from the ACR. *Journal of the American College of Radiology*. 2023 Sep 1;20(9):902-14. doi: 10.1016/j.jacr.2023.04.002.

21. Bell RJ. Mammographic density and breast cancer screening. *Climacteric* 2020;23:460-465. doi: 10.1080/13697137.2020.1785418.

22. Gastounioti A, Kasi CD, Scott CG, Brandt KR, Jensen MR, Hruska CB, et al. Evaluation of LIBRA software for fully automated mammographic density assessment in breast cancer risk prediction. *Radiology*. 2020;296:24-31. doi: 10.1148/radiol.2020192509.

23. Conroy SM, Pagano I, Kolonel LN, et al. Mammographic density and hormone receptor expression in breast cancer: The Multiethnic Cohort Study. *Cancer Epidemiol* 2011;35:448–452. doi:10.1016/j.canep.2010.11.011.

24. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. *Breast cancer* 2018;25:259-67. doi: 10.1007/s12282-018-0857-5.

25. Kwan ML, Kushi LH, Weltzien E, Maring B, Kutner SE, Fulton RS, et al. Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. *Breast Cancer Research*. 2009 Jun;11:1-3. doi: 10.1186/bcr2261.

26. Cronin KA, Harlan LC, Dodd KW, Abrams JS, Ballard-Barbash R. Population-based estimate of the prevalence of HER-2 positive breast cancer tumors for early stage patients in the US. *Cancer investigation*. 2010 Oct 1;28(9):963-8. doi: 10.3109/07357907.2010.496759.

27. Sturesdotter L, Sandsveden M, Johnson K, Larsson AM, Zackrisson S, Sartor H. Mammographic tumour appearance is related to clinicopathological factors and surrogate molecular breast cancer subtype. Sci Rep. 2020 Nov 30;10(1):20814. doi: 10.1038/s41598-020-77053-7.

28. Kojima Y, Tsunoda H. Mammography and ultrasound features of triple-negative breast cancer. *Breast Cancer* 2011;18:146-151. doi: 10.1007/s12282-010-0223-8.

29. Tamaki K, Ishida T, Miyashita M, et al. Correlation between mammographic findings and corresponding histopathology: potential predictors for biological characteristics of breast diseases. *Cancer Sci* 2011;102:2179-2185. doi: 10.1111/j.1349-7006.2011.02088.x.

30. Tamaki K, Ishida T, Miyashita M, Amari M, Ohuchi N, Tamaki N, et al. Mammography and morphobiologic characteristics of human breast cancer. *Tumori* 1993;79:422-426. doi: 10.1177/030089169307900611.

31. Jiang L, Ma T, Moran MS, Kong X, Li X, Haffty BG, et al. Mammographic features are associated with clinicopathological characteristics in invasive breast cancer. *Anticancer Res* 2011;31:2327-2334.

32. Liu S, Wu XD, Xu WJ, et al. Is there a correlation between the presence of a spiculated mass on mammogram and luminal a subtype breast cancer? *Korean J Radiol* 2016;17:846-852. doi: 10.3348/kjr.2016.17.6.846.

33. Johnson KS, Conant EF, Soo MS. Molecular subtypes of breast cancer: a review for breast radiologists. *Journal of Breast Imaging*. 2021 Jan 1;3(1):12-24. doi: 10.1093/jbi/wbaa110.

34. Purdie CA, Quinlan P, Jordan LB, Ashfield A, Ogston S, Dewar JA, et al. Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study. *Br J cancer* 2014;110:565-72. doi: 10.1038/bjc.2013.756.

35. McFall T, McKnight B, Rosati R, Kim S, Huang Y, Viola-Villegas N, et al. Progesterone receptor A promotes invasiveness and metastasis of luminal breast cancer by suppressing regulation of critical microRNAs by estrogen. *Journal of Biological Chemistry*. 2018;293:1163-77. doi: 10.1074/jbc.M117.812438.

36. Lamb CA, Fabris VT, Jacobsen BM, Molinolo A, Lanari C. Biological and clinical impact of imbalanced progesterone receptor isoform ratios in breast cancer. *Endoc relat cancer* 2018;25:R605-24. doi: 10.1530/ERC-18-0179.

37. Warner ET, Rice MS, Zeleznik OA, Fowler EE, Murthy D, Vachon CM, et al. Automated percent mammographic density, mammographic texture variation, and risk of breast cancer: a nested case-control study. NPJ Breast Cancer. 2021 May 31;7(1):68. doi: 10.1038/s41523-021-00272-2.

38. Sansone M, Fusco R, Grassi F, Gatta G, Belfiore MP, Angelone F, et al. Machine learning approaches with textural features to calculate breast density on mammography. *Current Oncology*. 2023 Jan 7;30(1):839-53. doi: 10.3390/curroncol30010064.

Li S, Yang X, Zhang Y, Fan L, Zhang F, Chen L, et al. Assessment accuracy of core needle biopsy for hormone receptors in breast cancer: a meta-analysis. *Breast cancer res treat* 2012;135:325-34. doi: 10.1007/s10549-012-2063-z.