

RESEARCH ARTICLE

Editorial Process: Submission:08/03/2026 Acceptance:21/10/2026 Published:25/10/2026

Development of A Radiomics Signature Using CT Texture Analysis of Breast Lesions Secondary to Infiltrating Ductal Carcinoma to Predict Histologic Features: A Cross-Sectional Study

Anila Rahim¹, Muhammad Awais^{1,2*}, Mahnoor Hafeez¹, Abdul Rehman^{2,3}, Amjad Sattar¹

Abstract

Objective: Radiomics has emerged as a promising approach for the non-invasive characterization of tumor biology using quantitative imaging features. Most radiomics studies in breast cancer have focused on magnetic resonance imaging (MRI), while the potential utility of computed tomography (CT)-based texture analysis remains underexplored. This study aimed to develop radiomic signatures derived from CT texture analysis of breast lesions to predict histologic features. **Methods:** Patients with biopsy-proven infiltrating ductal carcinoma (IDC) who had undergone contrast-enhanced chest CT less than 1 month prior to biopsy at Dow University of Health Sciences during the study period (January 2023 to December 2023) were included. A region of interest (ROI) was manually drawn around the primary breast lesion, and radiomic features were extracted using LIFEx® software. An elastic-net regression model was used to select radiomic features for the development of the final radiomics signature. **Results:** Ninety-two women with biopsy-proven IDC were included (median age, 41.5 years; IQR, 32.7–51.0). Radiomic signatures demonstrated fair predictive performance for several histologic features. The model predicting high-grade tumors achieved an area under the receiver operating characteristic curve (AUROC) of 0.763, with an accuracy of 71.7%. For a high acinar score, the AUROC was 0.706, with an accuracy of 66.3%. The radiomic signature for high nuclear grade showed the best performance, with an AUROC of 0.827 and an accuracy of 77.2%. Prediction of PR expression yielded an AUROC of 0.731 and an accuracy of 70.6%. No reliable radiomic signatures could be developed for mitotic rate, estrogen receptor expression, HER2/neu status, or Ki-67 proliferation index. **Conclusion:** In this proof-of-concept study, radiomic signatures based on CT texture analysis of IDC allowed prediction of histologic grade (nuclear grade and acinar score) and PR status with fair diagnostic accuracy. Further validation of these findings in larger cohorts of women with varying breast densities and scanned using different types of CT scanners is warranted.

Keywords: Breast Ductal Carcinoma, Radiomics, Computer-Assisted Radiographic Image Interpretation

Asian Pac J Cancer Prev, 27 (10), 3829-3838

Introduction

Breast cancer remains the most common cancer among women worldwide. In the United States, 313,530 new cases of breast cancer were estimated to occur in the year 2024, leading to an estimated 42,780 deaths [1]. The mortality is estimated to be even higher for women in low- to middle-income countries where screening by mammography is not universally implemented [2]. Prognosis and treatment of breast cancer varies widely depending on the size of the tumor, histologic type, grade of differentiation, estrogen receptor (ER)- and progesterone receptor (PR)-expression, HER-2/neu status, Ki-67 proliferation index, invasion of

local tissues, number and extent of nodal involvement, and presence of distant metastases [3]. Moreover, certain antineoplastic therapies, such as hormone therapy and targeted agents, can only be administered to patients whose cancer harbors the requisite molecular or genetic alterations. As such, there is tremendous clinical value in being able to predict these molecular and genetic features by non-invasive means, such as radiomics [4].

Texture analysis is a type of radiomics technique that employs computerized extraction of numerous features from radiographic images that are otherwise undetectable by the naked eye of a human radiologist. Texture analysis captures the voxel-to-voxel heterogeneity in radiographic

¹Department of Radiology, Dow University of Health Sciences, Karachi, Sindh, Pakistan. ²Department of Radiology, Aga Khan University Hospital, Karachi, Sindh, Pakistan. ³Department of Medicine, TidalHealth Peninsula Regional, Salisbury, MD, United States of America. *For Correspondence: awais_aku@yahoo.com

densities within a single region of interest (ROI) by calculating first-order and second-order statistics [5]. Previous studies have identified radiomic signatures on magnetic resonance imaging (MRI) and computed tomography (CT) for possible diagnosis, molecular typing, response prediction, and prognostication of patients with a variety of solid tumors [6-9]. Some studies have explored the utility of radiomic features on MRI as well as ultrasonography for care of patients with breast cancer [4, 10, 11], although this approach has not been adopted in routine clinical use yet [3]. On the other hand, the role of CT scans in evaluating breast pathologies has been historically limited because of poor soft-tissue visualization and carcinogenic potential of ionizing radiation [12]. Having said this, CT-based texture analysis has shown promise in the evaluation of other solid tumors [6, 7], but whether CT texture analysis of breast cancer has any clinical value for diagnosis, subtyping or prognostication remains unclear.

Texture features of cancerous breast lesions have been explored on contrast-enhanced MRI of the breast [4, 11, 10]. However, MRI breast is less commonly used in clinical practice, especially for patients with a palpable breast lump and a high pretest probability of breast cancer on mammographic and ultrasonographic features [12]. Mammography and ultrasonography remain the standard tools for detection and diagnostic evaluation of breast lesions. While CT is not recommended as a primary modality for breast cancer screening or diagnosis, CT chest with contrast is frequently obtained as part of a staging work-up, especially in patients with lesions 5 cm or larger [12]. We hypothesized that CT texture analysis of cancerous breast lesions may allow prediction of histologic features of the primary breast tumor. The rationale for the present study was therefore to explore whether clinically indicated staging CT scans, when they demonstrate the primary tumor, can provide additional quantitative information about tumor biology. In the present study, we employed CT texture analysis of breast lesions secondary to biopsy-proven infiltrating ductal carcinoma (IDC) to predict histologic features of cancers.

Materials and Methods

Study design and setting

A retrospective cross-sectional study was performed at the department of radiology of a 1000-bed tertiary care hospital located in the city of Karachi in Pakistan. The World Bank currently designates Pakistan as a lower-middle income country. Karachi is the world's twelfth largest city with an estimated population of 20.4 million.

Ethical considerations

This retrospective study was approved by the institutional Ethics Review Committee (ERC) at Dow University of Health Sciences and conducted in line with the ethical principles laid out in the Declaration of Helsinki by the World Medical Association. Given that no actual patient contact was involved in this retrospective review and the collected data did not contain any personal identifiers, the requirement of informed consent was

waived. Access to study data was strictly restricted to study personnel using a two-factor authentication procedure.

Sampling strategy

We obtained a consecutive sample of all patients who had undergone ultrasound-guided breast biopsy at Dow University of Health Sciences from January, 2023 to December, 2023 and had histopathologic evidence of IDC of the breast. Out of these patients, only those patients were eligible for inclusion who had undergone a contrast-enhanced CT chest within a period of 1 month prior to the biopsy procedure. Additionally, patients were included only if the primary breast lesion was identifiable on contrast-enhanced CT and could be segmented reproducibly. Patients whose breast lesion was not visible or could not be confidently separated from adjacent normal tissue were excluded from radiomic analysis. We also excluded patients whose CT scans were degraded by motion artifacts or streak artifacts secondary to hardware. Additionally, patients who were undergoing biopsy for evaluation of recurrent breast cancer, had received any chemotherapy or radiation therapy prior to the CT scan, or had undergone any type of surgical procedure of the breast (such as lumpectomy, breast implant or mastoplasty) in the past were also excluded.

Data collection

Medical records of all included patients were accessed, and data was extracted using a structured pro forma. We obtained data pertaining to age, demographics, body mass index, stage of the tumor and histopathologic features of the tumor. A single consultant radiologist, with more than 5 years of experience in breast radiography, reviewed the CT scans of all included patients and performed CT textural analysis. Breast biopsy specimens were fixed in formalin, placed in standard cassettes, dehydrated using alcohol, cleared using xylene, infiltrated and embedded into paraffin wax blocks, sectioned using a microtome and placed onto slides in a standardized fashion as recommended by the guidelines from the College of American Pathologists [13]. ER- and PR-status was assessed by immunohistochemical (IHC) staining and reported as negative (<1%), low-positive (1% to 10%) and positive (>10%)—in line with updated guidelines from the College of American Pathologists [13]. *HER-2/neu* status was assessed by a combination of IHC staining and fluorescent in situ hybridization (FISH) testing. Patients were considered to be *HER-2/neu*-positive if score was +3 by IHC staining, and negative if the *HER-2/neu* score was 0 or 1+ by IHC staining. For patients with an IHC score of +2, dual-probe FISH testing was performed and results were interpreted as negative or positive depending on the concomitant results of IHC and FISH in line with guidelines from the National Comprehensive Cancer Network [3]. Grading of breast cancer was performed based on the Nottingham histologic score, which incorporated components of acinar score, nuclear grade, mitotic rate and an overall grade based on the composite of these three components [3].

Image acquisition

The GE® Optima CT660 (GE® HealthCare, Inc.; Chicago, IL, United States) CT scanner was used to acquire all contrast-enhanced chest CT scans for patients included in the study. CT scans were performed using a tube voltage of 120 kVP and radiation exposure of 200–300 mAs to generate 5 mm thick slices for evaluation. Non-ionic intravenous contrast medium i.e. iohexol (Omnipaque®; GE Healthcare; Chicago, IL, United States), at a dose of 1.0–1.5 mL/kg of actual body weight, was utilized for all contrast-enhanced chest CT scans. Automated bolus contrast injector was used for the intravenous injection of contrast and contrast-enhanced images were obtained 35 seconds post-contrast injection. The institutional Picture Archival and Communication System (PACS) was used to extract anonymized DICOM files for the contrast-enhanced CT chest scans.

Image processing and segmentation

Open-source software, LIFEx® version 7.4.0 (LifexSoft; Orsay, France) [14], was used to open DICOM files on standard diagnostic workstations equipped with two RadiForce RX340 monitors (EIZO, Inc.; Cypress, CA, United States). Axial CT slices of the breast lesion were selected for segmentation. A three-dimensional region of interest (ROI) was manually drawn over the breast lesion with special attention paid to avoid inadvertent inclusion of any other adjacent tissue. In patients with multiple lesions, the ROI was drawn only over the lesion that was actually biopsied. Only one lesion per patient was used for feature extraction to avoid problems related to data clustering.

Radiomics features extraction

Radiomics features from the ROI were extracted using LIFEx® after using ROI relative minimum-maximum resampling with number of grey levels set to 400 and without the use of any transformations or de-noising filters. We only considered those CT texture parameters (radiomic features) which were considered standardized with a strong level of consensus by the Image Biomarker Standardization Initiative (IBSI) [15]. The first-order features were calculated from the gray attenuation values histogram without regard to spatial orientation and included texture features derived from morphology, intensity, intensity histogram, intensity rim and intensity histogram rim. Second-order features took into consideration the spatial relationship of attenuation values within voxels included in the ROI and were derived from the following matrices:

- i. Gray-level co-occurrence matrix (GLCM) capturing the arrangement of pairs of voxels
- ii. Gray-level run-length matrix (GLRLM) reflecting the size of homogeneous runs for each gray level
- iii. Neighborhood gray-tone difference matrix (NGTDM) describing the difference in gray-level between one voxel and its immediate neighbors in all directions
- iv. Gray-level size-zone matrix (GLSZM) capturing the size of homogeneous zones for each gray-level in all directions

Standardized definitions and mathematical formulae

for each of the radiomic features (texture parameters) derived from these matrices can be found in the IBSI reference manual [15].

Statistical considerations

Statistical analysis was performed in R version 4.4.0 (R Foundation for Statistical Computing; Vienna, Austria) [16]. Frequencies were calculated for qualitative variables, while median and interquartile ranges (IQR) were computed for quantitative variables. In order to explore the correlation amongst various CT texture parameters, correlation plots were constructed using the corplot package for R [17]. Elastic net (penalized) regression (glmnet package [18] was used to explore the association between histologic features and CT texture parameters. The optimal lambda and alpha parameters were selected through cross-validation and iterative methods (package caret) [19]. Coefficients with non-zero values in the elastic net regression models were scrutinized manually and only coefficients with high absolute values and minimal collinearity were selected for developing the radiomic signature. A logistic regression model was created using only the CT texture parameters that were retained for the radiomic signature and predictions of histologic features were made for the entire sample. For *Ki-67* expression, a linear regression model was created (instead of a logistic regression model) in a similar fashion. The Hosmer-Lemeshow goodness-of-fit test (package ResourceSelection) was applied to the logistic regression models to assess the adequacy of predictions over the entire range of the data [20]. A p-value of less than 0.05 was considered statistically significant. The optimal cut-off point for making predictions by the logistic regression model was selected by maximizing the Youden's J statistic (cutpointr package) [21]. Receiver operating characteristic (ROC) curves were visualized using the pROC package in R [22]. Area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR), negative LR and overall accuracy were computed as performance metrics for the radiomic signature (final logistic regression model). Bootstrapping was performed to calculate 95% confidence intervals (CI) for the performance metrics.

Results

A total of 92 patients were included in the final analysis. For each patient, only a single lesion was segmented and used for extracting CT texture parameters. Biopsy results from the same lesion were used for analyzing Nottingham histologic grade, acinar score, nuclear grade, mitotic rate, ER expression, PR expression and *HER-2/neu* expression. The median age of included patients was 41.5 (IQR: 32.7–51.0) years and all were women. All included patients had IDC. Nearly all patients (n=91, 98.9%) had detectable *Ki-67* expression and the median *Ki-67* expression was 21% (IQR: 15%–60%). Median BMI of included patients was 27.6 (IQR: 23.9–34.3) kg/m². Given the very nature of our cohort, most patients had advanced disease (stage II: 22.8%, stage III: 68.5% and stage IV: 8.7%) with a median

lesion size of 63 (50–71) mm. Rest of the histologic features are summarized in Table 1.

Elastic net regression was used to select CT texture parameters most strongly associated with individual histologic features. Optimal values for alpha and lambda parameters were selected through iterative cross-validation techniques. None of the CT texture features were strongly associated with mitotic rate, ER expression and *HER-2/neu* expression, and radiomic signatures could not be developed for these histologic features. For the remaining histologic features, coefficients with non-zero values in the elastic net regression model were selected for further scrutiny. Details of these coefficients and the performance of the regression model with all of these CT texture parameters are provided in Supplementary Tables 1 to 4 and Supplementary Figure 1.

Correlation plots were constructed to explore the correlations amongst the CT texture parameters. CT texture parameters were divided into 9 groups for this purpose:

- (1). Morphology
- (2). Intensity
- (3). Intensity Histogram
- (4). Intensity Rim
- (5). Intensity Histogram Rim
- (6). Grey level co-occurrence matrix (GLCM)
- (7). Grey level run-length matrix (GLRLM)
- (8). Neighboring grey tone difference matrix (NGTDM)
- (9). Grey level size-zone matrix (GLSZM)

Numerous CT texture parameters were found to be strongly correlated with each other and texture parameters within a single group were more strongly inter-correlated than CT texture parameters within other groups. Correlation plots depicting these correlations are provided in Supplementary Figures 2 to 8. Based on the results of these correlation plots, coefficients (CT texture parameters) retained in the elastic net regression models were further scrutinized and only those coefficients were retained, which had the highest absolute values and were not correlated with each other. Using this process, a list of 6–11 CT texture parameters were retained for building logistic regression models (radiomic signatures) for predicting histologic features.

Radiomic signature for high-grade tumors

A logistic regression model was developed for predicting high-grade (Nottingham grade III) tumors from six features viz. morphology–compactness 2 (M_{C2}), intensity histogram rim–average volume (IHR_{AV}), GLCM–normalized inverse difference ($GLCM_{NID}$), NGTDM–busyness ($NGTDM_B$), GLRLM–short run-low grey level-emphasis ($GLRLM_{SRLGLE}$) and GLSZM–large zone-low grey level-emphasis ($GLSZM_{LZLGLE}$):

$$\hat{p} = \frac{1}{1 + e^{-\hat{y}}}$$

$$\hat{y} = -6.164 + (18.855)(M_{C2}) - (3.486)(IHR_{AV}) + (2.986)(GLCM_{NID}) + (55.128)(NGTDM_B) + (413.031)(GLRLM_{SRLGLE}) - (87.43)(GLSZM_{LZLGLE})$$

Where tumor is predicted to be high-grade if $p > 0.3085$.

The ROC curve for this radiomic signature for predicting high-grade tumors is shown in Figure 1, while the performance metrics are detailed in Table 2. Hosmer-Lemeshow goodness-of-fit test failed to show any evidence of misfit ($p=0.3377$). The AUC for this radiomic signature was 0.763 (95% CI: 0.684–0.853) with a sensitivity of 62.1% (95% CI: 47.1%–70.7%) and specificity of 76.2% (95% CI: 71.0%–81.3%).

Radiomic signature for high acinar score

A logistic regression model was developed for predicting tumors of high acinar score (acinar score of 3) from six features viz. intensity histogram–root mean square (IH_{RMS}), GLCM–angular second moment ($GLCM_{ASM}$), NGTDM–busyness ($NGTDM_B$), GLCM–normalized inverse difference moment ($GLCM_{NIDM}$), GLRLM–long run-low grey level-emphasis ($GLRLM_{LRLGLE}$) and GLSZM–normalized grey level non-uniformity ($GLSZM_{NGLNU}$):

$$\hat{p} = \frac{1}{1 + e^{-\hat{y}}}$$

$$\hat{y} = +127.394 - (138.245)(IH_{RMS}) + (35.400)(GLCM_{ASM}) - (160.018)(GLRLM_{LRLGLE}) - (1.561)(NGTDM_B) - (127.928)(GLCM_{NIDM}) + (15.977)(GLSZM_{NGLNU})$$

Where tumor is predicted to have high acinar score if $p > 0.5183$

The ROC curve for this radiomic signature for predicting high acinar score is shown in Figure 1, while the performance metrics are detailed in Table 2. Hosmer-Lemeshow goodness-of-fit test failed to show any evidence of misfit ($p=0.4124$). The AUC for this radiomic signature was 0.706 (95% CI: 0.597–0.810) with a sensitivity of 66.0% (95% CI: 59.6%–72.4%) and specificity of 66.7% (95% CI: 59.7%–73.6%).

Radiomic signature for high nuclear grade

A logistic regression model was developed for predicting tumors of high nuclear grade (nuclear grade of 3) from 11 features viz. morphology–compactness 2 (M_{C2}), intensity–intensity-based coefficient of variation ($IIBC_V$), NGTDM–busyness ($NGTDM_B$), GLCM–joint maximum ($GLCM_{JM}$), GLCM–difference average ($GLCM_{DA}$), GLCM–angular second moment ($GLCM_{ASM}$), GLCM–inverse variance ($GLCM_{IV}$), GLCM–dissimilarity ($GLCM_D$), GLRLM–long run-low grey level-emphasis ($GLRLM_{LRLGLE}$), GLSZM–small zone-low grey level-emphasis ($GLSZM_{SZLGLE}$) and GLSZM–large zone-low grey level-emphasis ($GLSZM_{LZLGLE}$):

$$\hat{p} = \frac{1}{1 + e^{-\hat{y}}}$$

$$\begin{aligned}\hat{y} = & -5.0705 + (10.4037)(M_{C2}) - (0.5292)(I_{BCV}) + (24.4989)(NGTDM_B) \\ & - (78.644)(GLCM_M) + (9.1194)(GLCM_{DA}) + (362.3366)(GLCM_{ASM}) \\ & - (15.687)(GLCM_{IV}) - (7.8021)(GLCM_D) + (634.1095)(GLRLM_{LRLGLE}) \\ & - (1048.5621)(GLSZM_{SZLGLE}) - (158.5207)(GLSZM_{LZLGLE})\end{aligned}$$

Where tumor is predicted to have high nuclear grade if $p > 0.5660$

The ROC curve for this radiomic signature for predicting high nuclear grade is shown in Figure 1, while the performance metrics are detailed in Table 2. Hosmer-Lemeshow goodness-of-fit test failed to show any evidence of misfit ($p=0.7433$). The AUC for this radiomic signature was 0.827 (95% CI: 0.735–0.912) with a sensitivity of 79.6% (95% CI: 74.4%–84.9%) and specificity of 73.7% (95% CI: 66.8%–80.5%).

Radiomic signature for PR expression

A logistic regression model was developed for predicting tumors with high PR expression (PR +3 by immunohistochemistry) from 7 features viz. intensity–intensity-based quartile coefficient of dispersion (I_{BQCD}), intensity rim–approximate volume (IR_{AV}), intensity histogram–root mean square (IH_{RMS}), NGTDM–coarseness ($NGTDM_C$), NGTDM–busyness ($NGTDM_B$), GLSZM–low grey-level zone emphasis ($GLSZM_{LGLZE}$) and GLSZM–small zone-low grey level-emphasis ($GLSZM_{SZLGLE}$):

$$\hat{p} = \frac{1}{1 + e^{-\hat{y}}}$$

$$\begin{aligned}\hat{y} = & -2.1005 + (1.1040)(I_{BQCD}) - (5.6908)(IR_{AV}) + (68.5873)(IH_{RMS}) \\ & + (6.7363)(NGTDM_C) + (2.7474)(NGTDM_B) + (169.6898)(GLSZM_{LGLZE}) \\ & + (113.2707)(GLSZM_{SZLGLE})\end{aligned}$$

Where tumor is predicted to have high PR expression if $p > 0.4006$

The ROC curve for this radiomic signature for predicting high PR expression is shown in Figure 1, while the performance metrics are detailed in Table 2. Hosmer-Lemeshow goodness-of-fit test failed to show any evidence of misfit ($p=0.9652$). The AUC for this radiomic signature was 0.731 (95% CI: 0.597–0.839) with a sensitivity of 61.8% (95% CI: 53.8%–69.8%) and specificity of 75.9% (95% CI: 70.5%–81.2%).

Ki-67 expression

After elastic net regression, a total of 38 CT texture parameters had non-zero values for their regression coefficients (see Supplementary Table 5). However, the overall performance of this regression model was poor ($R^2 = 0.199$). A simpler linear regression model was also constructed from just 3 radiomic features viz. morphology–compactness 2 (M_{C2}), NGTDM–busyness ($NGTDM_B$) and GLSZM–large zone-low grey level-emphasis ($GLSZM_{LZLGLE}$). The performance of this

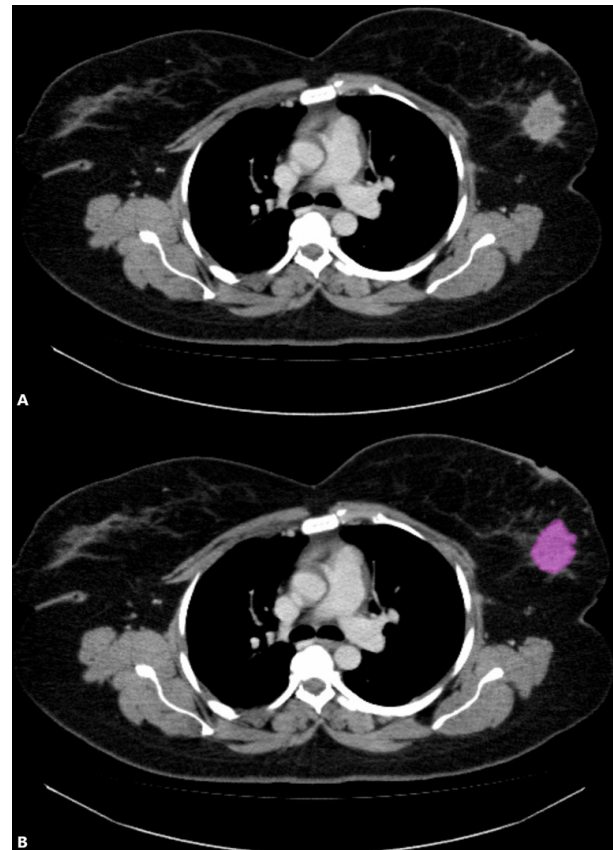


Figure 1. (A) Axial slices of contrast-enhanced computed tomography of the chest showing a visible lesion within the left breast. (B) A manual ROI (region of interest) overlay drawn over the visible lesion.

linear regression model was also poor ($R^2=0.121$). A scatterplot of the values predicted by this regression model versus the actual values of Ki-67 expression is shown in Figure 2. Therefore, an accurate radiomic signature for Ki-67 expression could not be developed.

Discussion

In this proof-of-concept study, we employed CT texture analysis of cancerous breast lesions secondary to biopsy-proven IDC among treatment-naïve women and identified radiomics signatures for predicting histologic grade (including acinar score and nuclear grade) and PR expression with fair diagnostic accuracy. The findings of our study demonstrate the feasibility of CT texture analysis of breast lesions for predicting histologic features and molecular alterations of breast cancer, which warrants further validation of this approach in larger cohorts of women with varying breast densities utilizing different CT scanners and image acquisition protocols.

Traditionally, CT imaging of the breast had been limited because of poor soft-tissue visualization of early CT scanners and carcinogenic potential of ionizing radiation. However, renewed interest in CT of the breast started with the FDA approval of cone-beam breast CT in 2015 [23], which could provide superior images to that provided by mammography alone. With the advent of radiomics, a plethora of research was conducted on the

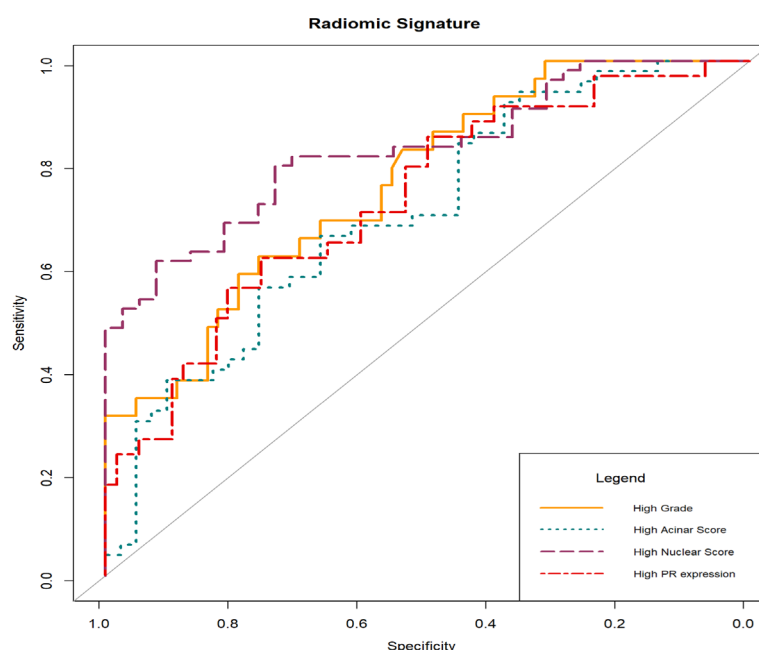


Figure 2. Receiver Operating Characteristic (ROC) Curve Depicting the Performance of CT Radiomic Signatures for Predicting High-Grade, high acinar score, high nuclear grade and positive progesterone receptor (PR) expression among breast lesions secondary to infiltrating ductal carcinoma of breast.

Table 1. Descriptive Statistics of Patients Included in the Present Study (n=92)

Parameters	n	(%)
Age, years (median [IQR])	41.5	(32.7–51.0)
Body mass index, kg/m ²	27.6	(23.9–34.3)
Lesion size, mm (median [IQR])	63	(50–71)
Clinical TNM categories		
T2	21	(22.8%)
T3	65	(70.6%)
T4	6	(6.5%)
N1	22	(23.9%)
N2	59	(64.1%)
N3	11	(12.0%)
M1	8	(8.7%)
Overall stage		
II	21	(22.8%)
III	63	(68.5%)
IV	8	(8.7%)
Histologic grade		
Low-grade (Nottingham grade I or II)	63	(68.5%)
High-grade (Nottingham grade III)	29	(31.5%)
Acinar score		
Low (1-2)	42	(45.6%)
High (3)	50	(54.3%)
Mitotic score		
Low (1-2)	80	(87.0%)
High (3)	12	(13.0%)
Nuclear grade		
Low (1-2)	38	(41.3%)
High (3)	54	(58.7%)

Table 1. Continued

Parameters	n	(%)
Estrogen receptor		
Negative (<1%)	51	(55.4%)
Weak positive (1%-10%)	0	(0.0%)
Positive (>10%)	41	(44.6%)
Progesterone receptor		
Negative (<1%)	58	(63.0%)
Weak positive (1%-10%)	0	(0.0%)
Positive (>10%)	34	(37.0%)
<i>HER-2/neu</i>		
Negative*	69	(75.0%)
Positive*	23	(25.0%)
<i>Ki-67</i> expression, % (median [IQR])	21%	(15.0%-60.0%)

IQR, Interquartile range; * *HER-2/neu* status was determined to be positive or negative based on a combination of immunohistochemical staining and fluorescent in situ hybridization results in line with guidelines from the National Comprehensive Cancer Network [3]

quantitative imaging of solid lesions and identification of radiographic biomarkers for diagnosis, prognostication and response prediction of cancers [5-7, 12, 24-28]. With respect to breast cancer, most research has focused on the utility of MRI texture analysis of breast cancer for differentiating benign from malignant lesions [29, 30], detecting the presence of nodal metastases [31], predicting *HER-2/neu* status [32], predicting *Ki-67* expression [4] and prognosticating patients' survival [33]. However, CT texture analysis of breast lesions has received limited attention over the same time period. In the present study, we employed CT texture analysis on breast lesions secondary to biopsy-proven IDC and identified radiomics signatures for prediction of histologic features and PR

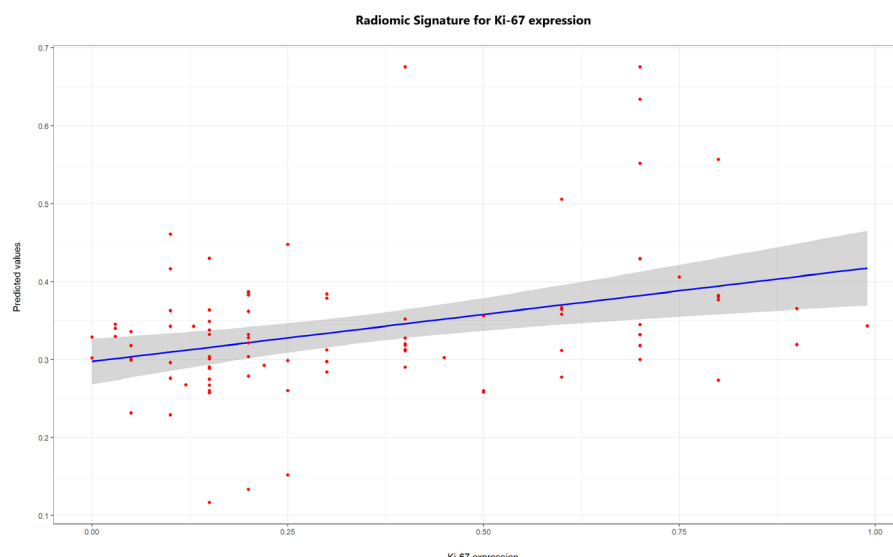


Figure 3. Scatterplot Depicting the Correlation between *Ki-67* Expression and the Predicted Value from CT Radiomic Signature (linear regression model based on CT texture parameters of breast lesions secondary to infiltrating ductal carcinoma of breast).

Table 2. Performance Metrics of Radiomic Signatures for Predicting Histologic Features of Breast Lesions Secondary to Infiltrating Ductal Carcinoma of Breast

Metric	High-grade	High acinar score	High nuclear grade	PR
Sensitivity (95% CI)	62.1% (47.1%–70.7%)	66.0% (59.6%–72.4%)	79.6% (74.7%–84.6%)	61.8% (53.8%–69.8%)
Specificity (95% CI)	76.2% (71.0%–81.3%)	66.7% (59.7%–73.6%)	73.7% (66.8%–80.5%)	75.9% (70.5%–81.2%)
PPV (95% CI)	54.5% (45.9%–63.2%)	70.2% (63.5%–76.9%)	81.1% (75.7%–86.5%)	60.0% (51.7%–68.3%)
NPV (95% CI)	81.6% (70.0%–89.4%)	62.3% (47.9%–74.8%)	71.8% (58.5%–82.1%)	77.2% (64.0%–86.6%)
Positive LR (95% CI)	2.61 (1.59–4.44)	1.98 (1.3–3.3)	3.03 (1.90–5.50)	2.56 (1.57–4.40)
Negative LR (95% CI)	0.50 (0.29–0.75)	0.51 (0.32–0.76)	0.28 (0.15–0.45)	0.50 (0.31–0.74)
Accuracy (95% CI)	71.7% (66.7%–76.7%)	66.3% (61.3%–71.3%)	77.2% (72.2%–82.1%)	70.6% (65.7%–75.6%)
AUC (95% CI)	0.763 (0.684–0.853)	0.706 (0.597–0.810)	0.827 (0.745–0.915)	0.731 (0.628–0.849)

AUC, Area under curve; CI, confidence interval; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; PR, progesterone receptor

status with 71.7% and 70.6% accuracy respectively. Our study demonstrates the feasibility and potential of CT texture analysis of breast lesions for predicting histologic features and molecular alterations of breast cancer. From a clinical standpoint, chest CT scans are frequently obtained as part of staging work-up for patients with either symptoms suggestive of metastases or those with large lesions (>5 cm) [12]. Breast MRI scans are less commonly utilized in the care of patients with breast lesions and are often reserved for equivocal cases or patients with dense breasts [12]. Therefore, if the utility of CT texture analysis of breast lesions is thoroughly validated, it can be more readily incorporated into the care of patients at the bedside as compared to MRI radiomics.

CT texture analysis has been investigated thoroughly for the evaluation of solid lesions pertaining to different organs. For instance, prior studies have validated CT texture analysis for its ability to differentiate between benign and malignant lesions [28], predict histological features [27], detect nodal involvement [26], prognosticate patients [25, 34] and predict response to treatment [24]. With respect to evaluation of breast lesions, we were

only able to identify a handful of studies employing CT texture analysis. Caballo and colleagues employed CT texture analysis of breast lesions in conjunction with machine learning models to differentiate benign from malignant breast masses [35]. Yoo et al. [36], Song et al. [37], Park et al. [38] and Moghadas-Dastjerdi et al. [39] reported on the predictive ability of CT texture analysis of breast lesions for predicting response to therapy and potential treatment failure. Our study adds to this the predictive ability of radiomics signatures derived from CT texture analysis for predicting histologic grade of IDC and PR status. An encouraging trend to note here is that all studies utilized different scanners: Phillips® Brilliance 64 or Ingenuity 128 (Yoo et al. [36] and Song et al. [37]), Phillips® IQon Spectral (Park et al. [38]), GE® LightSpeed (Moghadas-Dastjerdi et al. [39]) and GE® Optima CT660 CT scanner in our study. Moreover, the software used in this study i.e. LIFEx® version 7.4.0 was different from the software used in the aforementioned studies i.e. TexRAD® [36–38], which further attests to the applicability and generalizability of CT texture analysis.

The ability of CT texture analysis to predict histologic

grade and molecular alterations in cancerous breast lesions is plausible from a mechanistic perspective. Texture in the radiologic sense refers to the heterogeneity of radiographic densities within the individual voxels included in a ROI. On CT scans, radiographic densities are measured in Hounsfield units and variation in Hounsfield units is related to the ability of ionizing X-rays to permeate through tissues [5]. Within breast lesions, the permeability of X-rays is related to the presence of calcifications, fat content, vascularity and distortion of subcellular apparatus [12]. Therefore, the degree of architectural distortion within a lesion (presumably related to grade of cancer) as well as its attendant molecular characteristics (PR, ER or HER-2/neu) can influence radiographic texture, thereby allowing discrimination of these features through radiomics [40].

Mammography and ultrasound remain the primary modalities for detection and initial characterization of breast lesions in clinical practice, while MRI is used selectively for problem-solving and local staging. CT radiomics cannot replace biopsy, immunohistochemistry, mammography, ultrasound, or MRI; nor do we propose CT as a replacement for mammography, ultrasound, MRI, or tissue diagnosis. Rather, the purpose of the present study was to evaluate whether staging CT scans, when already obtained for clinical reasons and when the primary lesion is visible, contain quantitative radiomic information that correlates with tumor biology. The potential clinical value of our approach is that it uses CT images that have already been acquired for staging, without exposing the patient to additional radiation. If externally validated, CT radiomics may provide supplementary information about tumor aggressiveness, such as high nuclear grade or PR expression, before or alongside final histopathologic characterization. This could help multidisciplinary teams prioritize patients for expedited oncologic evaluation, identify tumors likely to have aggressive biology, and design future models that combine imaging, histopathology, and molecular markers. Thus, CT radiomics should be viewed as an opportunistic adjunct rather than a primary diagnostic pathway.

Our study comes with a number of limitations. First and foremost, our study employed CT texture analysis of breast lesions proven to be secondary to IDC by biopsy. As such, we did not investigate the ability of CT texture analysis to differentiate between benign etiologies of breast lesions versus malignant tumors. Secondly, our study was performed in the radiology suite of a single hospital located in the city of Karachi. Therefore, whether the results of this study can be extrapolated to CT scans performed using different CT scanners on women of different racial groups remains unknown. Thirdly, our cohort consisted of patients with IDC undergoing clinically indicated CT staging. Therefore, the findings from our study should not be generalized to screening populations. Moreover, our study findings are strictly applicable to patients who had visible breast lesions on CT scans, and therefore, these findings may not be generalizable to women with very early-stage breast cancer (e.g. Tis or T1a). Lastly, our study only included a modest sample of patients with a single histologic type of breast cancer i.e. IDC. Therefore,

it is unclear if the radiomics signatures developed in this study can be applied to women with other histologic types of cancer (e.g. invasive lobular carcinoma). Despite these limitations, the findings of our proof-of-concept study are important as they demonstrate the feasibility of employing CT texture analysis for prediction of histologic features and PR status of cancerous breast lesions.

In conclusion, in this proof-of-concept study, radiomics signatures based on CT texture analysis of cancerous breast lesions secondary to biopsy-proven infiltrating ductal carcinoma among treatment-naïve women allowed prediction of histologic grade (nuclear grade and acinar score) and PR status with fair diagnostic accuracy. Further validation of these findings in larger cohorts with women of varying breast densities using different types of CT scanners is warranted. While CT breast imaging cannot supplant the need for mammography or ultrasonography in routine clinical practice, CT radiomics when obtained as part of staging work-up can serve as an opportunistic adjunct for predicting histological features, prognostication and treatment planning.

Author Contribution Statement

All authors contributed substantially to the study design and conception of the study. AnR, MA and MH were involved in data acquisition, analysis and interpretation. AnR, MA and MH wrote the initial draft of the manuscript. AbR and AS reviewed and revised the manuscript for important intellectual content. All authors reviewed the final manuscript and approved the manuscript for submission.

Acknowledgements

None.

Statement of IRB approval

This research study was approved by the institutional Ethics Review Committee of the Dow University of Health Sciences (Karachi, Sindh, Pakistan).

Data sharing statement

The data on which the findings of this study are based can be requested from the corresponding author, Dr. Muhammad Awais, upon reasonable request.

Conflicts of interest

The authors of this study have no conflicts of interest to declare.

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin.* 2024;74(1):12-49. <https://doi.org/10.3322/caac.21820>.
2. Siddiqui R, Mehmood MH, Khan NA. An overview of breast cancer in pakistan. *Discover Medicine.* 2024;1(1):82. <https://doi.org/10.1007/s44337-024-00089-5>.
3. Gradishar WJ, Moran MS, Abraham J, Abramson V, Aft R, Agnese D, et al. Breast cancer, version 3.2024, nccn clinical practice guidelines in oncology. *J Natl Compr Canc Netw.* 2024;22(5):331-57. <https://doi.org/10.6004/>

- jncn.2024.0035.
4. Tabnak P, HajiEsmailPoor Z, Baradaran B, Pashazadeh F, Aghebati Maleki L. Mri-based radiomics methods for predicting *Ki-67* expression in breast cancer: A systematic review and meta-analysis. *Acad Radiol*. 2024;31(3):763-87. <https://doi.org/10.1016/j.acra.2023.10.010>.
5. Awais M, Khan N, Khan AK, Rehman A. Ct texture analysis for differentiating between peritoneal carcinomatosis and peritoneal tuberculosis: A cross-sectional study. *Abdom Radiol (NY)*. 2024;49(3):857-67. <https://doi.org/10.1007/s00261-023-04103-9>.
6. Li T, Gan T, Wang J, Long Y, Zhang K, Liao M. Application of ct radiomics in brain metastasis of lung cancer: A systematic review and meta-analysis. *Clin Imaging*. 2024;114:110275. <https://doi.org/10.1016/j.clinimag.2024.110275>.
7. Cheng DO, Khaw CR, McCabe J, Pennycuik A, Nair A, Moore DA, et al. Predicting histopathological features of aggressiveness in lung cancer using ct radiomics: A systematic review. *Clin Radiol*. 2024;79(9):681-9. <https://doi.org/10.1016/j.crad.2024.04.022>.
8. Ponsiglione A, Gambardella M, Stanzione A, Green R, Cantoni V, Nappi C, et al. Radiomics for the identification of extraprostatic extension with prostate mri: A systematic review and meta-analysis. *Eur Radiol*. 2024;34(6):3981-91. <https://doi.org/10.1007/s00330-023-10427-3>.
9. Ma D, Zhou T, Chen J, Chen J. Radiomics diagnostic performance for predicting lymph node metastasis in esophageal cancer: A systematic review and meta-analysis. *BMC Med Imaging*. 2024;24(1):144. <https://doi.org/10.1186/s12880-024-01278-5>.
10. Polat DS, Xi Y, Hulsey K, Lewis M, Dogan BE. Radiomics analysis of contrast-enhanced breast mri for optimized modelling of virtual prognostic biomarkers in breast cancer. *Eur J Breast Health*. 2024;20(2):122-8. <https://doi.org/10.4274/ejbh.galenos.2024.2023-12-12>.
11. Zheng S, Yang Z, Du G, Zhang Y, Jiang C, Xu T, et al. Discrimination between her2-overexpressing, -low-expressing, and -zero-expressing statuses in breast cancer using multiparametric mri-based radiomics. *Eur Radiol*. 2024;34(9):6132-44. <https://doi.org/10.1007/s00330-024-10641-7>.
12. Yasaka K, Sato C, Hirakawa H, Fujita N, Kurokawa M, Watanabe Y, et al. Impact of deep learning on radiologists and radiology residents in detecting breast cancer on ct: A cross-vendor test study. *Clin Radiol*. 2024;79(1):e41-e7. <https://doi.org/10.1016/j.crad.2023.09.022>.
13. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, et al. Estrogen and progesterone receptor testing in breast cancer: American society of clinical oncology/college of american pathologists guideline update. *Arch Pathol Lab Med*. 2020;144(5):545-63. <https://doi.org/10.5858/arpa.2019-0904-SA>.
14. Nioche C, Orlhac F, Boughdad S, Reuze S, Goya-Outi J, Robert C, et al. Lifex: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res*. 2018;78(16):4786-9. <https://doi.org/10.1158/0008-5472.CAN-18-0125>.
15. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295(2):328-38. <https://doi.org/10.1148/radiol.2020191145>.
16. Core Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. <https://www.R-project.org/>. Accessed 2023/08/18.
17. T W, V S. R package “corrplot”: Visualization of a correlation matrix (version 0.84). <https://github.com/taiyun/corrplot>. Accessed 5th September, 2024.
18. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1-22.
19. Kuhn M. Building predictive models in r using the caret package. *Journal of Statistical Software*. 2008;28(5):1 - 26. <https://doi.org/10.18637/jss.v028.i05>.
20. Hosmer DW, Lemeshow S. Goodness of fit tests for the multiple logistic regression model. *Commun. Stat - Theory Methods*. 1980;9(10):1043-69. <https://doi.org/10.1080/03610928008827941>.
21. Thiele C, Hirschfeld G. Cutpointr: Improved estimation and validation of optimal cutpoints in r. *arXiv preprint arXiv:200209209*. 2020.
22. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. Proc: An open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics*. 2011;12:77. <https://doi.org/10.1186/1471-2105-12-77>.
23. O’Connell AM, Marini TJ, Kawakyu-O’Connor DT. Cone-beam breast computed tomography: Time for a new paradigm in breast imaging. *J Clin Med*. 2021;10(21). <https://doi.org/10.3390/jcm10215135>.
24. Goh V, Ganeshan B, Nathan P, Juttla JK, Vinayan A, Miles KA. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: Ct texture as a predictive biomarker. *Radiology*. 2011;261(1):165-71. <https://doi.org/10.1148/radiol.11110264>.
25. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by ct texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol*. 2012;67(2):157-64. <https://doi.org/10.1016/j.crad.2011.08.012>.
26. Bayanati H, R ET, Souza CA, Sethi-Virmani V, Gupta A, Maziak D, et al. Quantitative ct texture and shape analysis: Can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? *Eur Radiol*. 2015;25(2):480-7. <https://doi.org/10.1007/s00330-014-3420-6>.
27. Liu S, Liu S, Ji C, Zheng H, Pan X, Zhang Y, et al. Application of ct texture analysis in predicting histopathological characteristics of gastric cancers. *Eur Radiol*. 2017;27(12):4951-9. <https://doi.org/10.1007/s00330-017-4881-1>.
28. Hu X, Ye W, Li Z, Chen C, Cheng S, Lv X, et al. Non-invasive evaluation for benign and malignant subcentimeter pulmonary ground-glass nodules (≤ 1 cm) based on ct texture analysis. *Br J Radiol*. 2020;93(1114):20190762. <https://doi.org/10.1259/bjr.20190762>.
29. Daimiel Naranjo I, Gibbs P, Reiner JS, Lo Gullo R, Sooknaran C, Thakur SB, et al. Radiomics and machine learning with multiparametric breast mri for improved diagnostic accuracy in breast cancer diagnosis. *Diagnostics (Basel)*. 2021;11(6). <https://doi.org/10.3390/diagnostics11060919>.
30. Daimiel Naranjo I, Gibbs P, Reiner JS, Lo Gullo R, Thakur SB, Jochelson MS, et al. Breast lesion classification with multiparametric breast mri using radiomics and machine learning: A comparison with radiologists’ performance. *Cancers (Basel)*. 2022;14(7). <https://doi.org/10.3390/cancers14071743>.
31. Yu Y, He Z, Ouyang J, Tan Y, Chen Y, Gu Y, et al. Magnetic resonance imaging radiomics predicts preoperative axillary lymph node metastasis to support surgical decisions and is associated with tumor microenvironment in invasive

- breast cancer: A machine learning, multicenter study. *EBioMedicine*. 2021;69:103460. <https://doi.org/10.1016/j.ebiom.2021.103460>.
32. Ramtohul T, Djerroudi L, Lissavalid E, Nhy C, Redon L, Ikni L, et al. Multiparametric mri and radiomics for the prediction of her2-zero, -low, and -positive breast cancers. *Radiology*. 2023;308(2):e222646. <https://doi.org/10.1148/radiol.222646>.
33. Lu D, Yan Y, Jiang M, Sun S, Jiang H, Lu Y, et al. Predictive value of radiomics-based machine learning for the disease-free survival in breast cancer: A systematic review and meta-analysis. *Front Oncol*. 2023;13:1173090. <https://doi.org/10.3389/fonc.2023.1173090>.
34. Sandrasegaran K, Lin Y, Asare-Sawiri M, Taiyini T, Tann M. Ct texture analysis of pancreatic cancer. *Eur Radiol*. 2019;29(3):1067-73. <https://doi.org/10.1007/s00330-018-5662-1>.
35. Caballo M, Hernandez AM, Lyu SH, Teuwen J, Mann RM, van Ginneken B, et al. Computer-aided diagnosis of masses in breast computed tomography imaging: Deep learning model with combined handcrafted and convolutional radiomic features. *J Med Imaging (Bellingham)*. 2021;8(2):024501. <https://doi.org/10.1117/1.JMI.8.2.024501>.
36. Yoo H, Cho KR, Song SE, Cho Y, Jung SP, Sung K. Tumor heterogeneity of breast cancer assessed with computed tomography texture analysis: Association with disease-free survival and clinicopathological prognostic factor. *Diagnostics (Basel)*. 2023;13(23). <https://doi.org/10.3390/diagnostics13233569>.
37. Song SE, Seo BK, Cho KR, Woo OH, Ganeshan B, Kim ES, et al. Prediction of inflammatory breast cancer survival outcomes using computed tomography-based texture analysis. *Front Bioeng Biotechnol*. 2021;9:695305. <https://doi.org/10.3389/fbioe.2021.695305>.
38. Park HS, Lee KS, Seo BK, Kim ES, Cho KR, Woo OH, et al. Machine learning models that integrate tumor texture and perfusion characteristics using low-dose breast computed tomography are promising for predicting histological biomarkers and treatment failure in breast cancer patients. *Cancers (Basel)*. 2021;13(23). <https://doi.org/10.3390/cancers13236013>.
39. Moghadas-Dastjerdi H, Rahman SH, Sannachi L, Wright FC, Gandhi S, Trudeau ME, et al. Prediction of chemotherapy response in breast cancer patients at pre-treatment using second derivative texture of ct images and machine learning. *Transl Oncol*. 2021;14(10):101183. <https://doi.org/10.1016/j.tranon.2021.101183>.
40. Qi YJ, Su GH, You C, Zhang X, Xiao Y, Jiang YZ, et al. Radiomics in breast cancer: Current advances and future directions. *Cell Rep Med*. 2024;5(9):101719. <https://doi.org/10.1016/j.xcrm.2024.101719>.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.