

Radiomics in Predicting Recurrence for Patients with Locally Advanced Breast Cancer using Quantitative Ultrasound

Archya Dasgupta

Sunnybrook Health Science Centre

Divya Bhardwaj

Sunnybrook Health Science Centre

Daniel DiCenzo

Sunnybrook Health Science Centre

Kashuf Fatima

Sunnybrook Health Science Centre

Laurentioscar Osapoetra

Sunnybrook Health Science Centre

Karina Quaioit

Sunnybrook Health Science Centre

Murtuza Saifuddin

Sunnybrook Health Science Centre

Stephen Brade

Sunnybrook Health Science Centre

Maureen Trudeau

Sunnybrook Health Science Centre

Sonal Gandhi

Sunnybrook Health Science Centre

Andrea Eisen

Sunnybrook Health Science Centre

Frances Wright

Sunnybrook Health Science Centre

Nicole Look-Hong

Sunnybrook Health Science Centre

Ali Sadeghi-Naini

York University

William Tran

Sunnybrook Health Science Centre

Belinda Curpen

Sunnybrook Health Science Centre

Michael Kolios

Ryerson University

Lakshmanan Sannachi

Sunnybrook Health Science Centre

Gregory Czarnota (✉ gregoryczarnota.submissions@gmail.com)

Sunnybrook Health Science Centre

Research Article

Keywords: Radiomics, Breast cancer, Quantitative ultrasound, Recurrence, Neoadjuvant chemotherapy, Texture analysis, Texture derivatives, Machine learning.

Posted Date: December 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-127490/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Oncotarget on December 7th, 2021. See the published version at <https://doi.org/10.18632/oncotarget.28139>.

Abstract

Background The purpose of the study was to investigate the role of pre-treatment quantitative ultrasound (QUS)-radiomics in predicting recurrence for patients with locally advanced breast cancer (LABC).

Methods A prospective study was conducted with patients with LABC (n=83). Primary tumours were scanned using a clinical ultrasound device before starting treatment. Ninety-five imaging features were extracted-spectral features, texture, and texture-derivatives. Patients were determined to have recurrence or no recurrence based on clinical outcomes. Machine learning classifiers with k-nearest neighbour (KNN) and support vector machine (SVM) were evaluated for model development using a maximum of 3 features and leave-one-out cross-validation.

Results With a median follow up of 69 months (range 7-118 months), 28 patients had disease recurrence (local or distant). The best classification results were obtained using an SVM classifier with a sensitivity, specificity, accuracy and area under curve of 71%, 87%, 82%, and 0.76, respectively. Using the SVM model for the predicted non-recurrence and recurrence groups, the estimated 5-year recurrence-free survival was 83% and 54% (p=0.003), and the predicted 5-year overall survival was 85% and 74% (p=0.083), respectively.

Conclusion A QUS-radiomics model using higher-order texture derivatives can predict patients with LABC at higher risk of disease recurrence before starting treatment.

Introduction

Breast cancer is one of the most common cancer in women and accountable for the leading cause of death¹. Locally advanced breast cancer (LABC) is seen in approximately 10–30% of patients and is associated with a poor prognosis compared to early breast cancer (EBC)². LABC encompasses advanced primary disease with or without metastatic involvement of regional lymph nodes. The 5-year survival for patients with LABC can vary between 50–80%, depending upon several factors including clinical characteristics and molecular features like the expression of estrogen (ER), progesterone (PR) receptors, or human epidermal growth factor receptor 2 (HER 2) expression³⁻⁵. Treatment typically involves a multimodality approach, including surgery, systemic therapy (chemotherapy, targeted therapy, endocrine therapy), and radiotherapy (RT). Translational and clinical research are investigating different strategies to improve outcomes and developing up-front biomarkers to identify patients at higher risk of disease recurrence. Several genetic tests are available in predicting tumour aggressiveness in parallel to what is classically ascertained through tissue assessment, with their role well established for early breast cancer⁶⁻⁹. Such tests have found utility in guiding clinicians to decide the role of treatment intensification (particularly chemotherapy) in patients predicted to harbour relatively higher risk disease. However, the application of such biomarkers is limited in LABC, where clinical outcomes could be potentially improved given the higher risk of relapse.

Imaging in oncology has a well-established role in diagnosis, staging, response assessment, and surveillance. There has been a paradigm shift in recent years with the introduction of artificial intelligence in medicine, with the promising role of imaging to be used as a noninvasive biomarker in understanding tumour biology^{10,11}. More popularly known as “radiomics,” advanced imaging analysis has generated promise in disease stratification and predicting clinical outcomes. Several imaging modalities like mammography (MMG), ultrasonography (USG), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) have proved utility in the field of radiomics for breast cancer^{12,13}. Quantitative ultrasound (QUS) is similar to conventional USG with the advantage of capturing and analysis of raw radiofrequency (RF) data, which can better characterize tissue microstructure^{14,15}. The basis for QUS is its ability to detect the microstructural elastic properties, which are different between benign or malignant tissues or also between various grades of tumour. The commonly used spectral features include mid-band fit (MBF), spectral slope (SS), spectral intercept (SI), spacing among scatterers (SAS), acoustic scatterer diameter (ASD), average acoustic-scatterer concentration (AAC), and attenuation coefficient estimate (ACE). Texture analysis from spectral images using gray level co-occurrence matrix (GLCM) can extract second-order imaging features like contrast (CON), correlation (COR), energy (ENE), and homogeneity (HOM), which can provide insights into different aspects of tumour heterogeneity. Studies have demonstrated the clinical efficacy of QUS in predicting response to neoadjuvant chemotherapy (NAC) in LABC^{16–20}, and in patients with head-neck malignancies treated with radiotherapy²¹.

In this study, we investigated the role of QUS obtained before the start of treatment in predicting the risk of tumour recurrence in patients with LABC. The imaging features were obtained from the QUS imaging, which included spectral parameters, texture of spectral parameters (QUS-Tex¹), and second-order texture analysis of QUS-Tex¹ features (QUS-Tex¹-Tex²). Model development was done using *k*-nearest neighbours (KNN) and support vector machines-radial basis function (SVM). To the best of our knowledge, this is the first study of QUS-radiomics to predict the recurrence groups in patients with LABC.

Results

Clinical Features

A total of 83 patients were included in the final analysis. The median follow up was 69 months (range 7-118 months) for all patients and 74 months (range 49-118 months) for patients without any evidence of disease recurrence. The total number of patients with recurrence was 28, whereas 55 were free from any recurrence until the last follow up. The distribution of various features between the two patient groups (recurrence versus non-recurrence) is summarized in **Table 1**. The most common histological type was invasive ductal carcinoma, found in 92% of patients. The majority of the patients had hormone-positive disease with ER+ and PR+ status in 58% and 52% of patients, respectively. Her2 expression was observed in 35% of patients.

Survival Outcomes

The 3 and 5-year recurrence-free survivals (RFS) for the entire cohort were 77% and 68%, respectively. The median time to recurrence was 24 months (range 4-82 months). Out of all recurrences, more than 80% occurred in the initial 4 years. The predominant pattern of initial recurrence was distant metastasis (DM) in 22 patients, followed by local relapse in 7, and regional nodes in 6 (**Supplementary Figure 1**). The common sites of DM were bone, lung, and liver in 61%, 54%, and 46% of patients, respectively. The 3 and 5-year overall survival (OS) in the entire group was 89% and 79%, respectively.

Feature Analysis and Classifier Performances

The representative B-mode, QUS parameter, texture and texture-derivative parametric maps for one patient, each with and without any recurrence, are presented in **Figure 1**. Two QUS- Tex^1 features had a significantly different distribution between the two groups- SAS-COR ($p=0.025$), ASD-ENE ($p=0.026$) (**Table 2**). Another three QUS- Tex^1 - Tex^2 features exhibited significant differences-ASD-COR-CON ($p=0.042$), SI-COR-CON ($p=0.033$), and SI-COR-HOM ($p=0.049$). The scatter plots indicating the distributions of these features between the two groups are presented in **Figure 2**. The scatter plots for all the 95 features have been included in **Supplementary Figure 2**.

The classifier performances using KNN and SVM are summarized in **Table 3**. The best result was obtained with the SVM classifier (selecting from all 95 features) with a sensitivity, specificity, accuracy, and AUC of 71%, 87%, 82%, and 0.76, respectively. Using the KNN model and again selecting from all features, the sensitivity, specificity, accuracy, and area under curve (AUC) were 84%, 68%, 76%, 0.78, respectively. The inclusion of third-order features (QUS- Tex^1 - Tex^2) further improved the diagnostic performances remarkably for the KNN classifier from 70% to 76%, while for SVM, it had only a small performance increment, changing from 80% to 82%. The corresponding ROC plots for the models are presented in **Figures 3A and 3B** and the representative bar diagram in **Figures 3C and 3D**.

Finally, the model-based predicted groups were evaluated to investigate the impact on recurrence-free survival (RFS) and OS. The SVM model performed best in segregating the two groups with predicted 5-year RFS being 83% (predicted non-recurrence) versus 54% (predicted recurrence), with a p -value of 0.003. The 5-year OS for the predicted non-recurrence and recurrence groups using the SVM classifier was 85% and 74%, respectively ($p=0.08$). The estimated RFS and OS plots using the SVM model have been shown in **Figure 4**.

Discussion

Breast cancer represents a heterogeneous disease entity with survival dependant upon several clinical, biological, and treatment-related factors. *In-situ* and early breast cancer lie on one end of the spectrum with excellent survival rates, whereas metastatic breast cancer represents a disease with dismal prognosis²²⁻²⁴. Locally advanced breast cancer is associated with an intermediate prognosis, with disease recurrence more commonly encountered than earlier breast cancer. In recent decades, the identification of specific molecular pathways and the availability of systemic and targetted agents have

helped to improve the outcomes to a certain extent. Still, there is an unmet need for the development of biomarkers, which can further help refine existing risk-stratification and pave the way towards precision and personalized medicine. The study presented here presents a novel strategy of using imaging like QUS and artificial intelligence-based tools to predict patients with a higher risk of recurrence before the initiation of any treatments.

Several genetic markers have established their role in stratifying risk recurrence in patients with EBC⁶⁻⁸. A 21-gene recurrence score had been used to guide treatment, particularly in EBC, showing improvement of clinical outcomes with consideration of adjuvant systemic therapy and treatment de-escalation in low-risk patients^{25,26}. However, there is limited literature related to such genetic markers with clinical application in patients with LABC. Molecular profile has been shown to influence the recurrence risk in LABC, with ER/PR+/HER2- tumours having better outcomes as compared to triple-negative cancers^{27,28}. Circulating microRNA and DNA methylation have been indicated to correlate with the clinical outcomes in patients with LABC^{29,30}. The introduction of radiomics has shown promise of developing noninvasive biomarkers in risk stratification and response monitoring. Several radiomic studies have been adopted in breast cancer patients using different imaging modalities^{12,13}. Radiomic analysis of MRI scans has been correlated with the established genetic tools in prognostication of patients with breast cancer³¹. In a study including 294 patients with breast cancer, Park *et al.* demonstrated an MRI-based radiomic signature correlated with disease-free survival³². QUS-Radiomics obtained before starting treatment and early during NAC have also been shown to predict patients' final pathological response with LABC^{16,17}. Ha *et al.* demonstrated the utility of radiomic features obtained from PET/CT to be associated with treatment response and prognosis in patients with LABC³³.

Ultrasound is a widely available portable imaging modality with rapid scan acquisition and significantly lower costs than other imaging modalities. Conventional B-mode US is commonly used in screening and diagnostic evaluation^{34,35}, with morphological features demonstrating correlation with breast cancer biology to some extent^{36,37}. As compared to conventional B-mode imaging, QUS provides more detailed information as obtained from the unprocessed radiofrequency data and is machine-independent, thus less influenced by technical variations or subjective interpretations. QUS spectral parameters have been demonstrated to be related to tissue microstructural elastic properties. In preclinical studies, QUS has been shown to detect cell death associated with various treatment modalities effectively^{14,38}. Further clinical studies have demonstrated the efficacy of pre-chemotherapy QUS-radiomics in predicting responses to NAC with an accuracy of 88%¹⁶. In the current study, the pre-treatment QUS-radiomics parameters could be used to identify patients developing disease recurrence with an accuracy of 82% using an SVM-based model. The SAS (scatterer spacing) parameter was one of the selected features in the SVM model, suggesting a differential microstructural organization and architecture between the two risk groups with its influence on biological behaviour. In a previous study, the SAS and associated textural features were shown to help differentiate grade I from grade II/ III breast cancer, with tumour grade known to reflect cancer differentiation and aggressiveness³⁹. The other two features selected in the SVM model

were ASD-CON and MBF-COR (texture features). The ASD parameter reflects the microstructural size, while MBF represents scatterer size, shape, number, organization, and elastic properties. The texture features are particularly helpful in characterizing tissue heterogeneity. Intratumoral heterogeneity is a well-known entity in breast cancer, leading to the evolution of therapeutic resistance and the development of disease recurrence^{40,41}. It is, therefore, reasonable to assume during presentation that high-risk and low-risk tumours harbour diverse levels of intratumoral heterogeneity, which can be detected by advanced imaging analysis methods. QUS probes into tumour structure and is influenced by organization at the cellular level. In this regard, it is important to note that third-order imaging features or texture-derivatives led to significant improvement here of the KNN model from 70–76% (modest improvement for SVM). It is possible the higher-order imaging features can further detect the heterogeneity at a deeper level helping in the refinement of the radiomics model.

The 5-year RFS of 68% and 5-year OS of 79% in our study is comparable to previous reports^{3–5}. As noted earlier, in EBC, genetic-based recurrence scores had initially shown the way towards the development of risk-adapted treatment protocols^{26,42}. In the neoadjuvant setting, response-guided chemotherapy⁴³ or the addition of bevacizumab to docetaxel and trastuzumab⁴⁴ has been demonstrated to improve disease-free survival and response rates. It is crucial to identify in advance patients with an aggressive disease having a higher risk of recurrence, which can help prognosticate and design appropriate treatment intensification strategies. For instance, high-risk LABC patients could be investigated for additional systemic treatment options, including maintenance therapies, given the higher risk of distant metastases. Similarly, some patients clinically presenting with LABC can have biologically relatively indolent tumours, and careful de-escalation strategies can be explored in such cohorts avoiding or minimizing treatment-related toxicities.

Limitations And Future Directions

The study presented here has a relatively small number of patients and has been expanded to continue in a larger population. It is possible with a higher number of patients, advanced strategies like deep learning can be used to improve classification performance and the reliability of generated radiomics models. Although we had a relatively longer follow up (median follow up > 6 years in patients without recurrence), small groups can exhibit late recurrence with a possible switch of the output groups. Based on the use of QUS-radiomics to predict the response to NAC, a randomized trial is currently underway to study the effect of adaptive chemotherapy (clinicaltrials.gov.in identifier NCT04050228).

Methods And Materials

Patient Selection and Treatment

This prospective, observational study was approved by the Sunnybrook Health Sciences Centre Research Ethics Committee and registered (clinicaltrials.gov identifier [NCT00437879](https://clinicaltrials.gov/ct2/show/study/NCT00437879)). The research was conducted following good clinical practice and as per the declaration of Helsinki. Written consent was obtained for

all study participants. Patients with a diagnosis of LABC were deemed eligible for the study. All the patients were treated with neoadjuvant chemotherapy (NAC) followed by surgery in the form of breast-conserving surgery or mastectomy with sentinel lymph node biopsy or axillary dissection (according to clinical standards and patient preferences). Additional subsequent adjuvant treatment was carried out with radiation and systemic therapies (endocrine therapy, targeted therapy) according to standard institutional practice. Patients were seen in the clinic every 3-6 months in the first 2 years, and after that every 6-12 months or as indicated clinically. Recurrence was confirmed by clinical investigations, medical imaging and supplemented with tissue diagnosis where available and appropriate, based on the decision of the treating physicians. Patients without any evidence of recurrent disease who had a minimum follow up of 4 years or recurrence (any time point within 4 years) were included in this analysis. Patients with recurrence suggesting a new breast primary (e.g. recurrence involving different quadrant or contralateral breast) or second different primary malignancy (e.g. lung primary) were excluded from the analysis, as they were likely to be associated with different biological etiology.

Acquisition of Ultrasound Data

All patients were scanned using a clinical ultrasound device Sonix RP (Analogic Medical Corp., Vancouver), with a linear array transducer having a central frequency of 7 MHz (bandwidth 4-9 MHz). For digital RF data acquisition, a sampling frequency of 40 MHz was used with a 16-bit resolution. Scans were obtained before the initiation of any cancer-directed treatment. For imaging, the transducer was focused towards the centre of the tumour and was scanned at regular intervals of 1 cm. The primary tumour was contoured manually and designated as the region of interest (ROI).

Image Analysis

RF Data corresponding to the ROI was analyzed using a sliding window analysis with an overlap of 94% in the axial and lateral directions to generate sub-regions of interest (sub-ROI). The corresponding dimensions of the sub-ROIs were 2 mm x 2 mm. The RF data obtained from each sub-ROI was subjected to a fast Fourier transform (FFT) and normalized using a tissue-mimicking phantom to obtain power spectra. Individual QUS parameters were determined throughout the entire tumour-MBF, SS, SI, SAS, ASD, and AAC. An ACE parameter was used in the spectral correction and served as an independent feature as well. The QUS parameters from all the sub-ROIs were determined, and the final values from the entire ROI served as first-order imaging features. Further details related to data processing are described in previous publications³⁹.

For the determination of second-order features, first, a parametric-coded map was generated for each of the QUS parameters (except ACE), accounting for individual values within the sub-ROIs, called QUS parametric maps. Texture analysis was carried out subsequently using GLCM, which computes the relation of the index pixel with the neighbouring ones at four different angular relations, 0°, 45°, 90°, and 135°. Four GLCM features were analyzed-CON, COR, ENE, and HOM. These led to a total of 24 QUS-Tex¹ features (from 6 QUS parametric maps, no parametric map generated for ACE).

In order to obtain the third-order imaging parameters, 16 QUS- Tex^1 features were used (the texture maps of SS and SAS were not used). Firstly, a colour-coded map was regenerated from the values obtained from the texture values corresponding to the sub-ROIs. In a similar manner, the 4 GLCM texture analysis was repeated, which led to 64 QUS- Tex^1 - Tex^2 features. The study proceeded with a total of 95 features (7 QUS, 24 QUS- Tex^1 , 64 QUS- Tex^1 - Tex^2). Each of these features was averaged from all the tumour slices obtained from a single patient, and the mean value was used for final analysis.

Statistical Analysis and Machine Learning Classifiers

The patients were labelled using binary classes depending upon the final clinical outcomes, including recurrence (R) and non-recurrence (NR). The distribution of each feature between the two groups was tested using a Shapiro-Wilk test. Unpaired t-tests and Mann-Whitney tests were conducted for normally distributed data and non-parametric data, respectively. A p-value of <0.05 was considered statistically significant. Two machine learning classifiers were used for model development using forward feature selection algorithms- KNN and SVM. A maximum of three features was used for classification to avoid overfitting of data into the model. As the number of individuals was unevenly distributed between the two groups (recurrence versus non-recurrence), the data was balanced divided into several subsets prior to the application of the machine learning classifiers. Subsets were selected randomly from the entire group selecting proportionately equal number of patients with and without disease recurrence. Leave-one-out cross-validation was used to test the reliability and obtain the confusion matrix for the classifiers. The data were analyzed independently using first-order and second-order features (QUS+ QUS- Tex^1) and then incorporating both using all 95 features (QUS+ QUS- Tex^1 + QUS- Tex^1 - Tex^2) to evaluate the impact of higher-order features on classification performances. The AUC was obtained from receiver operating characteristics (ROC) analyses. Also, other indices like sensitivity, specificity, accuracy were obtained and compared between the outputs from the different classifiers. For processes of tumour segmentation, data processing, data extraction, and machine learning classification MATLAB R2011B (Mathworks, USA) were used. The Kaplan Meier product-limit method was used for survival analysis using the Statistical Package for the Social Sciences (SPSS V21, IBM Corporation, Armonk, New York, USA). The date of starting NAC was considered as the baseline for survival analysis. The final performance of the classifier models was tested using log-rank tests (comparison of survival rates between predicted recurrence and predicted non-recurrence groups).

Conclusion

QUS-Radiomic features obtained before the start of treatment can predict the risk of disease recurrence with reasonable accuracy. The incorporation of higher-order imaging features in the form of texture derivatives leads to the improvement of the classifier performances. The noninvasive imaging biomarker can lead to future strategies in the prognostication of patients with LABC and pave the way towards personalized medicine.

Declarations

Acknowledgments: We would like to thank all the patients for their participation in the study. Our sincere gratitude to the Terry Fox Foundation and the Canadian Institutes of Health Research for funding support. We express our regard to the physicians and other health care staff for their support in patient care.

Author contributions:

Study concept: AD, GJC

Acquisition of data, interpretation of data, writing-original draft and revision, writing-final approval: All authors

Project administration, funding acquisition: GJC

Competing Interest Statement: None of the authors have any conflict of interest to declare.

Data sharing: Data will be shared upon request to the corresponding author and the institutional ethics committee according to the policy of Sunnybrook Health Sciences Centre, Toronto.

References

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394–424 (2018).
2. Newman, L. A. Epidemiology of Locally Advanced Breast Cancer. *Seminars in Radiation Oncology* **19**, 195–203 (2009).
3. Huang, E. H. *et al.* Postmastectomy Radiation Improves Local-Regional Control and Survival for Selected Patients With Locally Advanced Breast Cancer Treated With Neoadjuvant Chemotherapy and Mastectomy. *JCO* **22**, 4691–4699 (2004).
4. Colleoni, M. *et al.* Annual Hazard Rates of Recurrence for Breast Cancer During 24 Years of Follow-Up: Results From the International Breast Cancer Study Group Trials I to V. *J. Clin. Oncol.* **34**, 927–935 (2016).
5. Wang, M. *et al.* Evaluation of the prognostic stage in the 8th edition of the American Joint Committee on Cancer in locally advanced breast cancer: An analysis based on SEER 18 database. *Breast* **37**, 56–63 (2018).
6. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
7. van de Vijver, M. J. *et al.* A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. *N Engl J Med* **347**, 1999–2009 (2002).
8. Paik, S. *et al.* A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *New England Journal of Medicine* **351**, 2817–2826 (2004).

9. Sotiriou, C. & Pusztai, L. Gene-Expression Signatures in Breast Cancer. *New England Journal of Medicine* **360**, 790–800 (2009).
10. Lambin, P. *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* **48**, 441–446 (2012).
11. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563–577 (2015).
12. Valdora, F., Houssami, N., Rossi, F., Calabrese, M. & Tagliafico, A. S. Rapid review: radiomics and breast cancer. *Breast Cancer Res Treat* **169**, 217–229 (2018).
13. Tagliafico, A. S. *et al.* Overview of radiomics in breast cancer diagnosis and prognostication. *Breast* **49**, 74–80 (2020).
14. Czarnota, G. J. *et al.* Ultrasound imaging of apoptosis: high-resolution noninvasive monitoring of programmed cell death in vitro, in situ and in vivo. *British Journal of Cancer* **81**, 520–527 (1999).
15. Kolios, M. C., Czarnota, G. J., Lee, M., Hunt, J. W. & Sherar, M. D. Ultrasonic spectral parameter characterization of apoptosis. *Ultrasound in Medicine & Biology* **28**, 589–597 (2002).
16. Tadayyon, H. *et al.* A priori Prediction of Neoadjuvant Chemotherapy Response and Survival in Breast Cancer Patients using Quantitative Ultrasound. *Scientific Reports* **7**, 45733 (2017).
17. Sannachi, L. *et al.* Response monitoring of breast cancer patients receiving neoadjuvant chemotherapy using quantitative ultrasound, texture, and molecular features. *PLoS ONE* **13**, e0189634 (2018).
18. DiCenzo, D. *et al.* Quantitative ultrasound radiomics in predicting response to neoadjuvant chemotherapy in patients with locally advanced breast cancer: Results from multi-institutional study. *Cancer Med* **9**, 5798–5806 (2020).
19. Quiaoit, K. *et al.* Quantitative ultrasound radiomics for therapy response monitoring in patients with locally advanced breast cancer: Multi-institutional study results. *PLoS One* **15**, e0236182 (2020).
20. Dasgupta, A. *et al.* Quantitative ultrasound radiomics using texture derivatives in prediction of treatment response to neo-adjuvant chemotherapy for locally advanced breast cancer. *Oncotarget* **11**, 3782–3792 (2020).
21. Tran, W. T. *et al.* Quantitative ultrasound delta-radiomics during radiotherapy for monitoring treatment responses in head and neck malignancies. *Future Sci OA* **6**, FSO624 (2020).
22. Toriola, A. T. & Colditz, G. A. Trends in breast cancer incidence and mortality in the United States: implications for prevention. *Breast Cancer Res. Treat.* **138**, 665–673 (2013).
23. Jatoi, I. & Miller, A. B. Why is breast-cancer mortality declining? *The Lancet Oncology* **4**, 251–254 (2003).
24. Waks, A. G. & Winer, E. P. Breast Cancer Treatment: A Review. *JAMA* **321**, 288–300 (2019).
25. Albain, K. S. *et al.* Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on

- chemotherapy: a retrospective analysis of a randomised trial. *The Lancet Oncology* **11**, 55–65 (2010).
26. Poorvu, P. D. *et al.* Prognostic Impact of the 21-Gene Recurrence Score Assay Among Young Women With Node-Negative and Node-Positive ER-Positive/HER2-Negative Breast Cancer. *J. Clin. Oncol.* **38**, 725–733 (2020).
 27. Huber, K. E., Carey, L. A. & Wazer, D. E. Breast cancer molecular subtypes in patients with locally advanced disease: impact on prognosis, patterns of recurrence, and response to therapy. *Semin Radiat Oncol* **19**, 204–210 (2009).
 28. Meyers, M. O. *et al.* Impact of breast cancer molecular subtypes on locoregional recurrence in patients treated with neoadjuvant chemotherapy for locally advanced breast cancer. *Ann. Surg. Oncol.* **18**, 2851–2857 (2011).
 29. Wu, X. *et al.* De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. *J Transl Med* **10**, 42 (2012).
 30. Dejeux, E. *et al.* DNA Research methylation profiling in doxorubicin treated primary locally advanced breast tumours identifies novel genes associated with survival and treatment response. *13* (2010).
 31. Li, H. *et al.* MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. *Radiology* **281**, 382–391 (2016).
 32. Park, H. *et al.* Radiomics Signature on Magnetic Resonance Imaging: Association with Disease-Free Survival in Patients with Invasive Breast Cancer. *Clin Cancer Res* **24**, 4705–4714 (2018).
 33. Ha, S., Park, S., Bang, J.-I., Kim, E.-K. & Lee, H.-Y. Metabolic Radiomics for Pretreatment 18F-FDG PET/CT to Characterize Locally Advanced Breast Cancer: Histopathologic Characteristics, Response to Neoadjuvant Chemotherapy, and Prognosis. *Sci Rep* **7**, 1556 (2017).
 34. Nothacker, M. *et al.* Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC Cancer* **9**, 335 (2009).
 35. Brem, R. F., Lenihan, M. J., Lieberman, J. & Torrente, J. Screening Breast Ultrasound: Past, Present, and Future. *American Journal of Roentgenology* **204**, 234–240 (2015).
 36. Irshad, A. *et al.* Assessing the Role of Ultrasound in Predicting the Biological Behavior of Breast Cancer. *American Journal of Roentgenology* **200**, 284–290 (2013).
 37. Au, F. W.-F., Ghai, S., Lu, F.-I., Moshonov, H. & Crystal, P. Histological Grade and Immunohistochemical Biomarkers of Breast Cancer: Correlation to Ultrasound Features. *J Ultrasound Med* **36**, 1883–1894 (2017).
 38. Vlad, R. M., Alajez, N. M., Giles, A., Kolios, M. C. & Czarnota, G. J. Quantitative ultrasound characterization of cancer radiotherapy effects in vitro. *Int. J. Radiat. Oncol. Biol. Phys.* **72**, 1236–1243 (2008).
 39. Tadayyon, H., Sadeghi-Naini, A. & Czarnota, G. J. Noninvasive characterization of locally advanced breast cancer using textural analysis of quantitative ultrasound parametric images. *Transl Oncol* **7**,

- 759–767 (2014).
40. Turashvili, G. & Brogi, E. Tumor Heterogeneity in Breast Cancer. *Front Med (Lausanne)* **4**, (2017).
 41. Rye, I. H. *et al.* Intratumor heterogeneity defines treatment-resistant HER2 + breast tumors. *Mol Oncol* **12**, 1838–1855 (2018).
 42. Eiermann, W. *et al.* The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. *Annals of Oncology* **24**, 618–624 (2013).
 43. von Minckwitz, G. *et al.* Response-guided neoadjuvant chemotherapy for breast cancer. *J. Clin. Oncol.* **31**, 3623–3630 (2013).
 44. Coudert, B. *et al.* Use of [18F]-FDG PET to predict response to neoadjuvant trastuzumab and docetaxel in patients with HER2-positive breast cancer, and addition of bevacizumab to neoadjuvant trastuzumab and docetaxel in [18F]-FDG PET-predicted non-responders (AVATAXHER): an open-label, randomised phase 2 trial. *The Lancet Oncology* **15**, 1493–1502 (2014).

Tables

Table 1: Clinical characteristics for the two groups (recurrence versus non-recurrence)

Features		Recurrence (n=28)		Non-recurrence (n=55)	
Variables	Categories	n	%	n	%
Age	Median (Range)	50 (29-79) years		48 (31-72) years	
Menstrual status	Premenopausal	16	56	33	60
	Perimenopausal	1	4	3	6
	Postmenopausal	10	36	17	30
	Unknown	1	4	2	4
Laterality	Right	15	54	27	49
	Left	13	46	28	51
Histology	IDC	25	89	51	92
	ILC	2	7	1	2
	Others	1	4	3	6
ER Status	Negative	13	46	22	40
	Positive	15	54	33	60
PR Status	Negative	13	46	27	49
	Positive	15	54	28	51
HER2 Status	Negative	18	64	36	66
	Positive	10	36	19	34
T stage	T1	0	0	0	0
	T2	7	25	28	50
	T3	13	46	24	44
	T4	8	29	3	6
N stage	N0	5	18	15	27
	N1	16	57	34	61
	N2	4	14	3	6
	N3	3	11	3	6

Abbreviations: IDC: Invasive ductal carcinoma; ILC: Invasive lobular carcinoma; ER: Estrogen receptor; PR: Progesterone receptor; HER2: Human epidermal receptor 2.

Table 2: Features with differential distribution between the two groups with statistical significance

Parameter	Recurrence	Non-recurrence	p-value
	Mean ± SEM	Mean ± SEM	
SAS-COR	.3396±.02734	.3684±.06400	0.025
ASD-ENE	.0354±.00853	.0510±.05399	0.026
ASD-COR-CON	5.30 ± 0.95	4.85 ± 0.85	0.042
SI-COR-CON	5.0289±.75007	4.6532±.76247	0.033
SI-COR-HOM	.5484±.01998	.5585±.02659	0.049

Abbreviations: SEM: Standard error of the mean; SAS: Spacing among scatterers; ASD: Acoustic scatterer diameter; SI: Spectral slope; COR: Correlation; ENE: Energy; CON: Contrast; HOM: Homogeneity.

Table 3: Classification performance of the two machine learning classifiers with the best-selected features.

Classifier	Features	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	Selected features
KNN	QUS+ QUS- <i>Tex</i> ¹	84	54	70	0.73	ASD-COR SAS-HOM SAS-ENE
	QUS+ QUS- <i>Tex</i> ¹ + QUS- <i>Tex</i> ¹ - <i>Tex</i> ²	84	68	76	0.78	ACE AAC-CON-CON AAC-ENE-HOM
SVM	QUS+ QUS- <i>Tex</i> ¹	69	87	80	0.75	SAS ASD-CON MBF-COR
	QUS+ QUS- <i>Tex</i> ¹ + QUS- <i>Tex</i> ¹ - <i>Tex</i> ²	71	87	82	0.76	SAS ASD-CON MBF-COR

Abbreviations: KNN: k-nearest neighbour; SVM: Support vector machine; AUC: Area under curve; QUS: Quantitative ultrasound; QUS- Tex^1 : QUS-texture; QUS- Tex^1 - Tex^2 : QUS-texture derivatives.

Figures

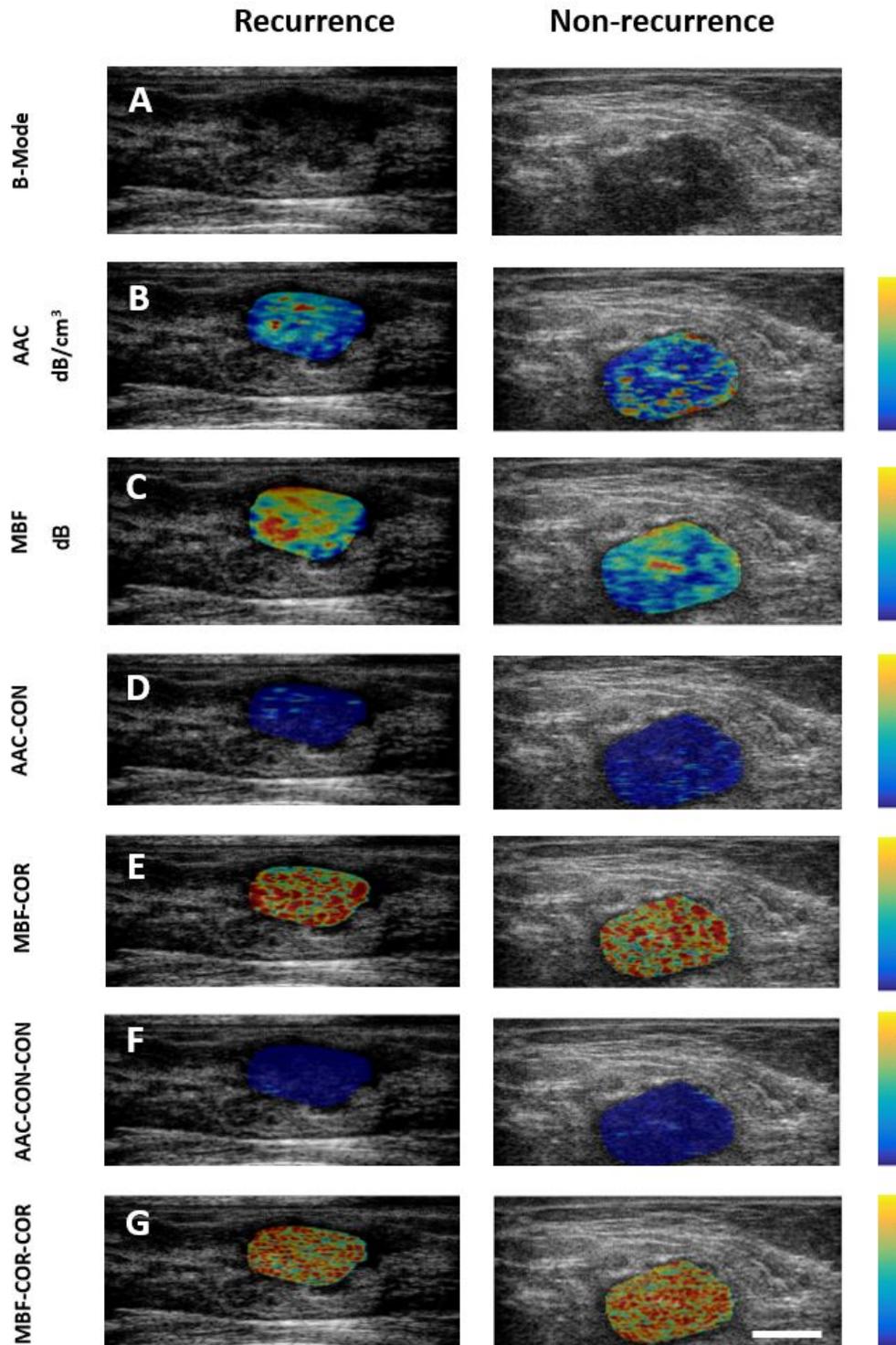


Figure 1

(A) Pre-treatment B-Mode images, QUS parametric maps ((B) AAC, (C) MBF), QUS-texture maps ((D) AAC-CON, (E) MBF-COR), and QUS-texture derivative maps ((F) AAC-CON-CON, (G) MBF-CR-COR) for one patient with recurrent disease (left panel) and one without recurrence (right panel). The colour-coded maps are generated over the region corresponding to the tumour using normalized values for individual features within the sub-ROIs. The colour scale on the right side represents the values for the individual features (B) 7 to 64 dB/cm³, (C) -21 to 21 dB, (D) 0 to 34, (E) -0.43 to 0.94, (F) 0 to 53, and (G) -0.54 to 0.94. The scale bar represents 2 cm. Abbreviations: QUS: Quantitative ultrasound; AAC: Average acoustic concentration; MBF: Mid band fit; CON: Contrast; COR: Correlation; ROI: Region of Interest.

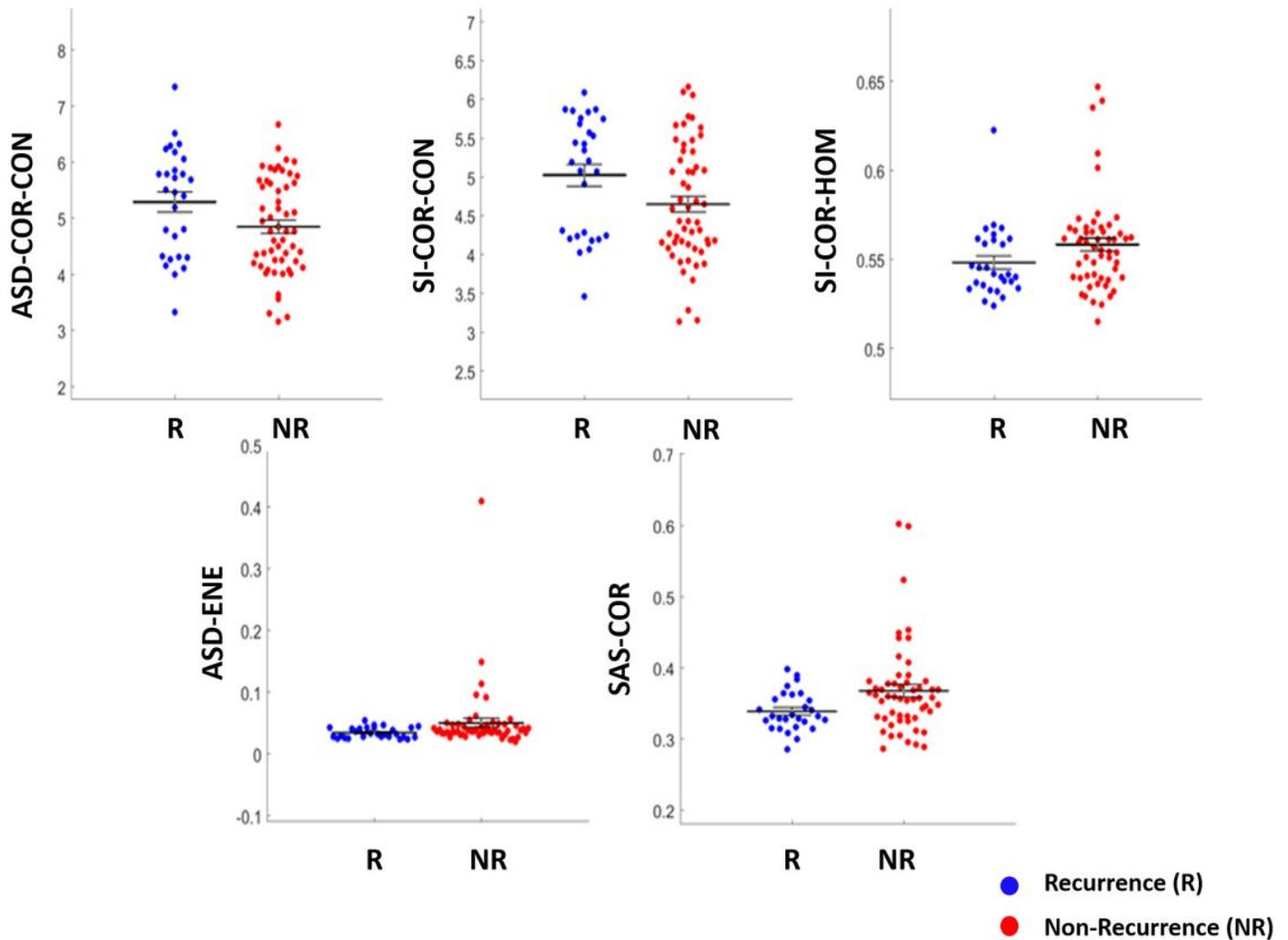


Figure 2

Scatter plots showing the features with the difference in distribution between the two groups (Recurrence vs Non-recurrence) reaching the threshold of statistical significance ($p < 0.05$). Abbreviations: SAS: Spacing among scatterers; ASD: Acoustic scatterer diameter; SI: Spectral slope; COR: Correlation; ENE: Energy; CON: Contrast; HOM: Homogeneity.

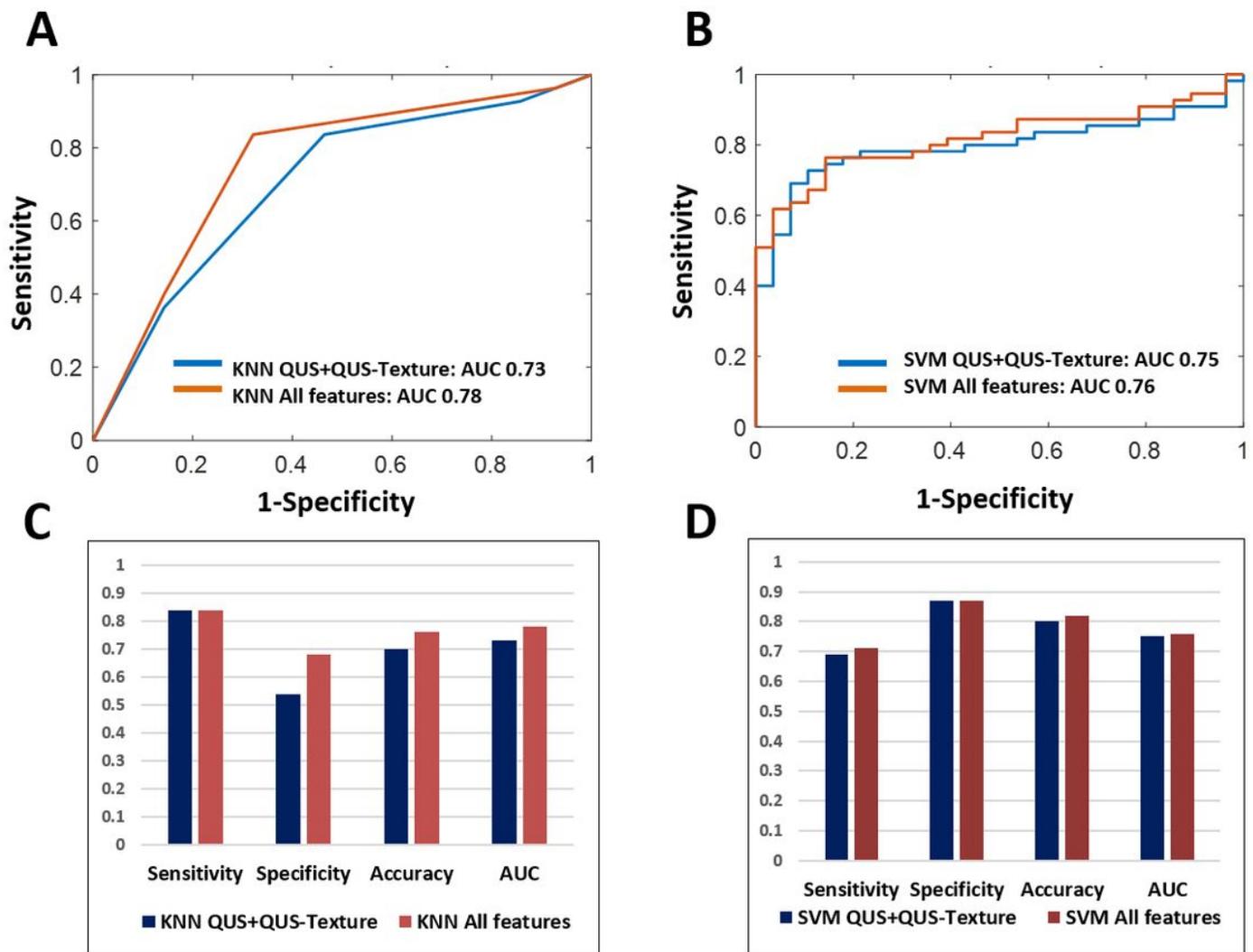
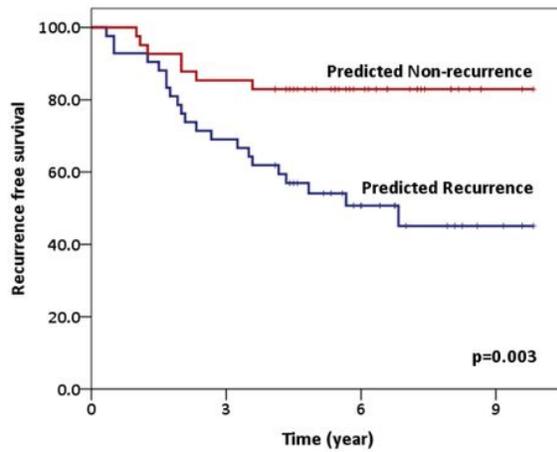


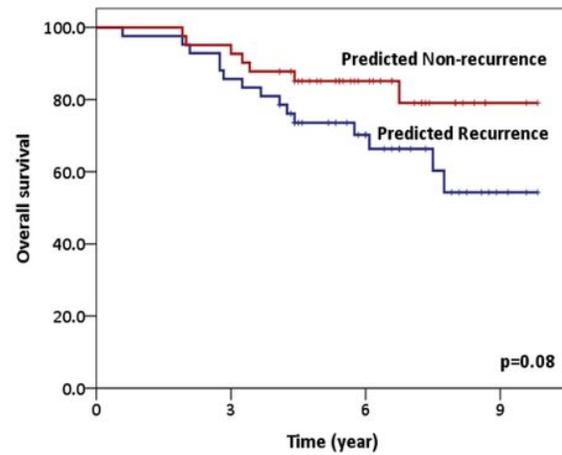
Figure 3

The classifier indices for the two machine learning classifiers. Figures 3A and 3B show the ROC curves using KNN and SVM classifiers, respectively, showing the effect of inclusion of higher-order imaging features (texture-derivatives). Figures 3C and 3D are the bar diagrams representing the diagnostic indices (sensitivity, specificity, accuracy, and AUC) for the KNN and SVM model with and without the use of texture-derivatives. Abbreviations: ROC: Receiver operating characteristics; KNN: k-nearest neighbour; SVM: Support vector machine; AUC: Area under curve;

A

Numbers at risk

Predicted Non-recurrence	41	35	18	2
Predicted recurrence	42	29	14	3

B

	41	39	19	2
	42	36	20	3

Figure 4

Predicted survival plots using support vector machine classifier predicted groups (predicted recurrence vs. predicted non-recurrence)-recurrence free survival (4A) and overall survival (OS).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfigurescombined.pdf](#)