

# **A priori prediction of breast tumour response to chemotherapy using quantitative ultrasound imaging and artificial neural networks**

**Hadi Tadayyon<sup>1,2</sup>, Mehrdad Gangeh<sup>1,2</sup>, Lakshmanan Sannachi<sup>1,2</sup>, Maureen Trudeau<sup>3</sup>, Kathleen Pritchard<sup>3</sup>, Sonal Ghandi<sup>3</sup>, Andrea Eisen<sup>3</sup>, Nicole Look-Hong<sup>4</sup>, Claire Holloway<sup>4</sup>, Frances Wright<sup>4</sup>, Eileen Rakovitch<sup>5,6</sup>, Danny Vesprini<sup>5,6</sup>, William Tyler Tran<sup>5,6</sup>, Belinda Curpen<sup>7</sup> and Gregory Czarnota<sup>1,2,3,5,6</sup>**

<sup>1</sup>Physical Sciences, Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

<sup>2</sup>Department of Medical Biophysics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

<sup>3</sup>Division of Medical Oncology, Department of Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

<sup>4</sup>Surgical Oncology, Department of Surgery, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

<sup>5</sup>Department of Radiation Oncology, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

<sup>6</sup>Department of Radiation Oncology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

<sup>7</sup>Department of Medical Imaging, Sunnybrook Health Sciences Centre, and Faculty of Medicine, University of Toronto, Toronto, ON, Canada

**Correspondence to:** Gregory Czarnota, **email:** gregory.czarnota@sunnybrook.ca

**Keywords:** quantitative ultrasound; artificial neural networks; ultrasound spectroscopy; tumour response assessment; prognostic biomarker

**Received:** November 26, 2018

**Accepted:** May 13, 2019

**Published:** June 11, 2019

**Copyright:** Tadayyon et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## **ABSTRACT**

**We demonstrate the clinical utility of combining quantitative ultrasound (QUS) imaging of the breast with an artificial neural network (ANN) classifier to predict the response of breast cancer patients to neoadjuvant chemotherapy (NAC) administration prior to the start of treatment.**

Using a 6 MHz ultrasound system, radiofrequency (RF) ultrasound data were acquired from 100 patients with biopsy-confirmed locally advanced breast cancer prior to the start of NAC. Quantitative ultrasound mean parameter intensity and texture features were computed from the tumour core and margin, and were compared to the clinical/pathological response and 5-year recurrence-free survival (RFS) of patients. A multi-parametric QUS model in conjunction with an ANN classifier predicted patient response with  $96 \pm 6\%$  accuracy, and a  $0.96 \pm 0.08$  area under the receiver operating characteristic curve (AUC), compared to  $65 \pm 10\%$  accuracy and  $0.67 \pm 0.14$  AUC achieved using a K-Nearest Neighbour (KNN) algorithm. A separate ANN model predicted patient RFS with  $85 \pm 7\%$  accuracy, and a  $0.89 \pm 0.11$  AUC, whereas the KNN methodology achieved a  $58 \pm 6\%$  accuracy and a  $0.64 \pm 0.09$  AUC.

The application of ANN for classifying patient response based on tumour QUS features performs well in terms of predicting response to chemotherapy. The findings here provide a framework for developing personalized *a priori* chemotherapy selection for patients that are candidates for NAC, potentially resulting in improved patient treatment outcomes and prognosis.

## INTRODUCTION

Neoadjuvant chemotherapy (NAC) is the primary up-front treatment modality for patients with locally advanced breast cancer (LABC). This aggressive form of cancer typically presents with tumours larger than 5 cm and extensive nodal involvement. Since the tumours are often inoperable, the goal of NAC is to reduce tumour volume. NAC may also be given to facilitate breast conserving surgery for patients who would otherwise require a mastectomy. There is a strong correlation between a pathological complete response (pCR) to NAC and cancer-free survival. Despite the availability of a wide spectrum of systemic and targeted drugs, due to genetic and epigenetic factors, most patients do not achieve pathologic complete response to NAC. Response is typically determined at the end of several months of treatment. In this light, there is growing interest in the discovery of biomarkers to predict therapy response with the aim of optimizing treatment reducing morbidity by avoiding futile treatments, and improving prognosis. For instance, diffusion-weighted MRI (DW-MRI) has been demonstrated to predict clinical response of breast tumours as early as after one cycle of chemotherapy [1]. Positron emission tomography (PET) imaging of breast cancer patients using a fluorodeoxyglucose (FDG) contrast agent has detected response-related changes in the tumour after one cycle of chemotherapy [2]. Additionally, diffuse optical imaging (DOI) studies of breast cancer have measured a significant increase in haemoglobin concentration, water content, and tissue optical index in responding patients as early as one week after the start of chemotherapy [3].

To date, the majority of studies have focused on biomarkers reflective of treatment-induced changes in functional and/or structural properties of the tumour (i.e. monitoring biomarkers). However, there is a growing shift of attention toward biomarkers reflecting inherent tumour biology (i.e. predictive biomarkers), which do not require any treatment to be administered. From a non-imaging perspective, pre-treatment levels of immunohistochemical markers, including Ki-67, HER2, and circulating nucleosomes have been linked to the likelihood of breast tumour's response to NAC [4–7]. From an imaging perspective, a growing body of research exists in the area of pre-treatment imaging biomarkers. In a recent study, diffuse optical spectroscopic (DOS) imaging of LABC patients indicated that patients with a pathologically complete response have significantly higher up-front haemoglobin concentration levels than those with pathologically incomplete response with  $p = 0.01$ ,  $AUC = 1.0$  [8]. A more recent DOS study demonstrated that changes in tissue optical index and baseline oxygen saturation levels are indicators of pCR with an  $AUC$  of 0.83 [9]. Intra-tumoural and peri-tumoural radiomic features of dynamic contrast-enhanced MRI (DCE-MRI)

of the breast have been demonstrated to be predictive of pCR prior to treatment with  $AUC$  of 0.74 [10]. In the area of PET imaging, the median progression-free survival of patients with estrogen receptor (ERe)- positive, human epithelial growth factor 2 (HER2)- negative breast tumours undergoing endocrine therapy was linked to their FDG uptake prior to treatment [11].

The use of clinical ultrasound has been established in the field of medical imaging as a cost-effective modality with high penetration depth (~7 cm) and real-time imaging capability. Furthermore, the raw ultrasound radiofrequency (RF) backscatter signal contains information about tissue microstructure, which is not resolvable in conventional ultrasound images (B-mode images). Quantitative ultrasound (QUS) techniques examine the frequency dependence of the RF signal backscattered from tissues and have been applied *in vivo* in a variety of applications to reveal information about tissue microstructure, enabling the differentiation of disease from normal tissue and the characterization of disease into its subtypes. For instance, parameters derived from the linear regression analysis of the RF power spectrum, including midband fit (MBF), spectral slope (SS), and spectral 0-MHz intercept (SI), have been used to characterize intraocular tumours and to detect prostate cancer, cardiovascular disease, and cancerous lymph nodes [12–15].

Broader frequency bandwidths further permit the estimation of advanced parameters such as average (effective) scatterer diameter (ASD) and average (effective) acoustic concentration (AAC), which are derived by fitting a scattering model to the RF data [16]. These parameters have effectively differentiated mouse carcinomas from rat fibroadenomas [17] and have demonstrated potential for use in breast tumour grading [18, 19] and diagnosis [20]. Recent pre-clinical studies have determined, using both high frequency (>20 MHz) and clinical frequency (<10MHz) ranges of ultrasound, that QUS can be used to detect and quantify tumour cell death *in vivo* in response to various treatments including photodynamic therapy, radiation therapy, chemotherapy, and anti-vascular therapy [21–24]. Furthermore, a recent pilot clinical study [25] demonstrated the effectiveness of using textural features extracted from QUS spectral images (MBF and SI) to detect breast tumour responses to neoadjuvant chemotherapy as early as one week into several-month-long chemotherapy treatments. The mean of intensity of ASD and AAC images derived from ultrasound backscatter data have also been effective at detecting treatment response in a similar clinical application [26].

Scatterer spacing, also known as spacing among scatterers (SAS), has also been investigated as a tissue characterization biomarker for tissues containing detectable periodicity in their structural organization. Previous studies have investigated the potential of SAS mainly for characterizing diffuse diseases of the liver [27–30].

For instance, in [28], the inter-scatterer-distribution (ISS) and the mean scatterer spacing (MSS) have been investigated for characterizing focal diseases of the liver using wavelet transform-based methods [28]. The MSS was considered for characterization of pathological human liver using Fourier transform-based methods [29]. The terms SAS and MSS are used interchangeably in the literature to refer to the mean scatterer spacing. More recently, SAS demonstrated discriminative power in breast tumour grading and therapy response applications [19, 31]. Motivated by those studies, we recently investigated whether pre-treatment values of QUS biomarkers can differentiate between therapy responsive and non-responsive tumours [32]. A multiparametric QUS model was developed using two regions of interest (ROIs) – the tumour core and a 5 mm margin of surrounding tissue. For each ROI, mean of intensity and texture features of QUS images were computed and incorporated into a k-nearest neighbour (KNN) classifier. Results from 56 LABC patients, indicated a response prediction accuracy of 88%, which was linked to a 5-year recurrence-free survival (RFS). However, as data becomes more complex (i.e. as data dimensionality increases), KNN performance typically deteriorates. In classification problems with high dimensionality, such as this study, an artificial neural network (ANN) classifier is a suitable choice [33]. An ANN is a nonlinear classifier that learns patterns in a data set using an interconnected network of “neurons” (elements with multiple inputs and one output) with a predefined activation rule. In the present study a previously examined cohort [32] was expanded from 56 patients to 100 patients. The data set was balanced prior to supervised learning and a more advanced model – an ANN model - was trained to predict the response and 5-year RFS of LABC patients undergoing NAC. The results were then compared with those obtained from the previously used KNN model. In addition to the conventional binary response classification (response versus non-response) done previously, a three-class grouping scheme was also investigated here. This included complete, partial, and non-response classification. Finally, whereas previous work reported Kaplan-Meier 5-year RFS curves of responding and non-responding patients, here, the 5-year RFS of patients was separately predicted using QUS-based biomarkers directly.

## RESULTS

### Patient clinical characteristics

Table 1 presents a statistical summary of patient clinical characteristics including age, tumour size, estrogen receptor (ERe) status, progesterone receptor (PRe) status, and human epithelial growth factor 2 (HER2) status. The patients are separated according to response groups. Based on the modified response (MR) scoring system and binary

classification of response described previously, of the 100 patients in the study, 83 patients responded to treatment and 17 patients did not respond to treatment. Responders had a mean age of  $50 \pm 10$  years and non-responders had a mean age of  $47 \pm 12$  years. Responders and non-responders had similar mean initial tumour sizes of  $5.6 \pm 2.7$  cm and  $5.9 \pm 2.8$  cm, respectively. The proportion of patients with ERe, PRe, and HER2 positive tumours in responder and non-responder groups are presented in Table 1. Statistical analysis using a chi-square test of independence demonstrated a significant correlation between complete response and HER2-positivity ( $p = 0.002$ ), whereas no statistically significant correlation was found between response and any of the hormone-based markers. In terms of histological subtype, the majority of the patients in both groups were diagnosed with invasive ductal carcinoma (91 % and 94 % in responder and non-responder groups, respectively), with a small number of other subtypes such as invasive lobular carcinoma and invasive mammary carcinoma, as presented in Table 1. Individual patient details are presented in Supplementary Tables 1 and 2 in Supplementary Information.

### Classification results

In this study, both two-class and three-class response grouping schemes were examined. In order to attain a sufficient number of samples for applying machine learning, the three-class response groups were combined into several two-class groups in the following manner: complete response (CR) versus (partial response (PR) + non-response (NR3)); (CR + PR) versus NR3; and PR versus NR3. Please refer to Materials and Methods for the definition of response types. Random down-sampling was performed on the majority class in order to obtain a balanced set. Table 2 presents the majority and minority class sizes, balanced set size, and the number of balanced sets obtained after down-sampling for each grouping scheme. As observed, the number of balanced sets varies between grouping schemes depending on the number of times random sampling was required to sample all patients in the database.

Figure 1 presents representative responder and non-responder patient QUS images with outlines of the core and margin ROIs. Displayed are images of B-mode (A), SS (B), SI (C), MBF (D), SAS (E), ASD (F), and AAC (G) parametric maps of the patient's tumour prior to treatment initiation. In Figure 2, the corresponding low-magnification (A) and high magnification (B) images of hematoxylin and eosin (H&E) stained sections of breast tissue specimens (excised after treatment completion and surgery) are displayed. It is evident from this figure that there are spatial variations in pixel intensities of the parametric images, highlighting the importance of texture-based features when discriminating responding tumours from non-responding ones.

**Table 1: Summary of clinical characteristics including age, initial tumour size, hormone receptor statuses, and cancer subtypes, of the studied LABC patients grouped by their clinical/pathological response to NACT**

		Responders (N = 83)	Non-responders (N = 17)
Age (yr)	Min	31	29
	Max	83	67
	Mean	50	47
	SD	11	12
Tumor size pre (cm)	Min	1	3
	Max	12	13
	Mean	6	6
	SD	3	3
ERe positive	No.	51	12
	%	61	71
PRe positive	No.	45	11
	%	54	65
HER2 positive	No.	31	4
	%	37	24
IDC	No.	76	16
	%	92	94
Other (ILC, IMC)	No.	7	1
	%	8	6

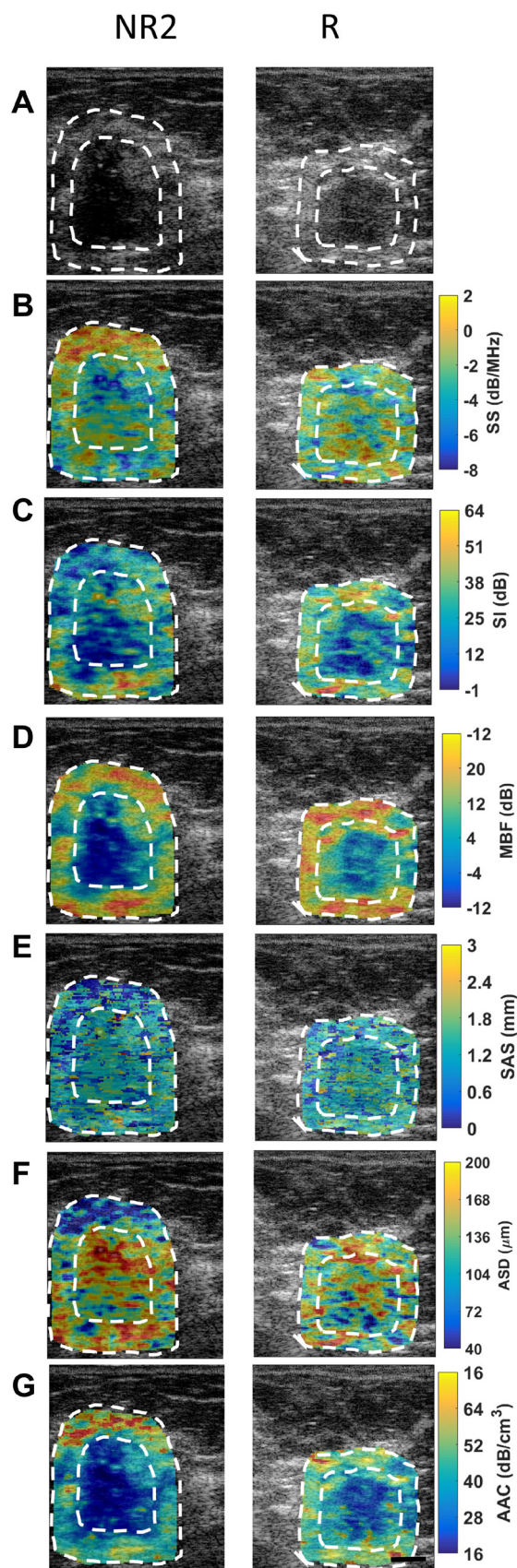
Abbreviations: ER, estrogen receptor status; PR, progesterone receptor status; HER2, human epithelial growth factor receptor 2 status; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; IMC, invasive mammary carcinoma.

**Table 2: Majority and minority class sizes, balanced set size, and number of balanced sets used for each classification type, including responder (R) vs non-responder (NR2), complete responder (CR) + partial responder (PR) vs NR3, CR vs (PR+NR3), PR vs NR3, and survivor vs non-survivor**

	Majority class	Minority class	Balanced set size	No. of balanced sets
<b>R vs NR2</b>	R (83)	NR2 (17)	34 (17+17)	21
<b>CR vs (PR+NR3)</b>	PR+NR3 (55)	CR (45)	90 (45+45)	5
<b>(CR+PR) vs NR3</b>	CR+PR (92)	NR3 (8)	16 (8+8)	41
<b>PR vs NR3</b>	PR (47)	NR3 (8)	16 (8+8)	22
<b>Survival</b>	Survived (86)	Not survived (14)	28 (14+14)	27

Figure 3 compares responding and non-responding patients through a panel of overlaid scatter plots and box plots of the top 15 QUS texture features and the top 15 QUS margin features in order of statistical significance (*t*-test or Mann–Whitney test). None of the features plotted were found to be statistically significant on their own ( $p > 0.05$ ). However, one parameter, AAC energy, was found to be marginally significant ( $p = 0.05$ ). As evident from Figure 3, none of the individual QUS features are linearly separable between the responder and non-responder groups. This highlights the need for multi-feature classification and non-linear classifiers in order to solve this complex classification problem. Table 3 presents mean classification performance metrics obtained from running the ANN model on all balanced sets (the

number of balanced sets varied from 5 to 41 depending on class distribution as reported in Table 2). Reported metrics include sensitivity, specificity, accuracy, and AUC evaluated on the test set. For conventional response (R) versus non-response (NR2) classification, mean values of sensitivity, specificity, accuracy, and AUC of  $89 \pm 9\%$ ,  $85 \pm 12\%$ ,  $87 \pm 6\%$ , and  $0.90 \pm 0.07$  were obtained, respectively. In a three-class grouping scheme, when CR and PR patients were combined into one group and were compared against the NR3 patients, a 9% higher classification accuracy was observed on average (over the samples) compared to the conventional grouping scheme. This permitted non-responder patients and patients with response (partial or complete) to be identified up-front with an accuracy of  $96 \pm 6\%$ . However, when CR



**Figure 1: QUS images from a representative non-responder (NR2) patient and a representative responder patient with outlines of core and margin ROIs. (A) B-mode images, (B) SS image, (C) SI images, (D) MBF images, (E) SAS images, (F) ASD, and (G) AAC images obtained prior to chemotherapy treatment initiation. Scale bars: 1 cm.**

patients were compared against PR+NR3 patients, the classification accuracy dropped by 8% compared to the conventional grouping scheme. Classification of PR versus NR3 patients yielded an accuracy of  $86 \pm 10\%$ . However, due to the relatively small sample size (47 PR and 8 NR3), the model has limited statistical power compared to the other classification types.

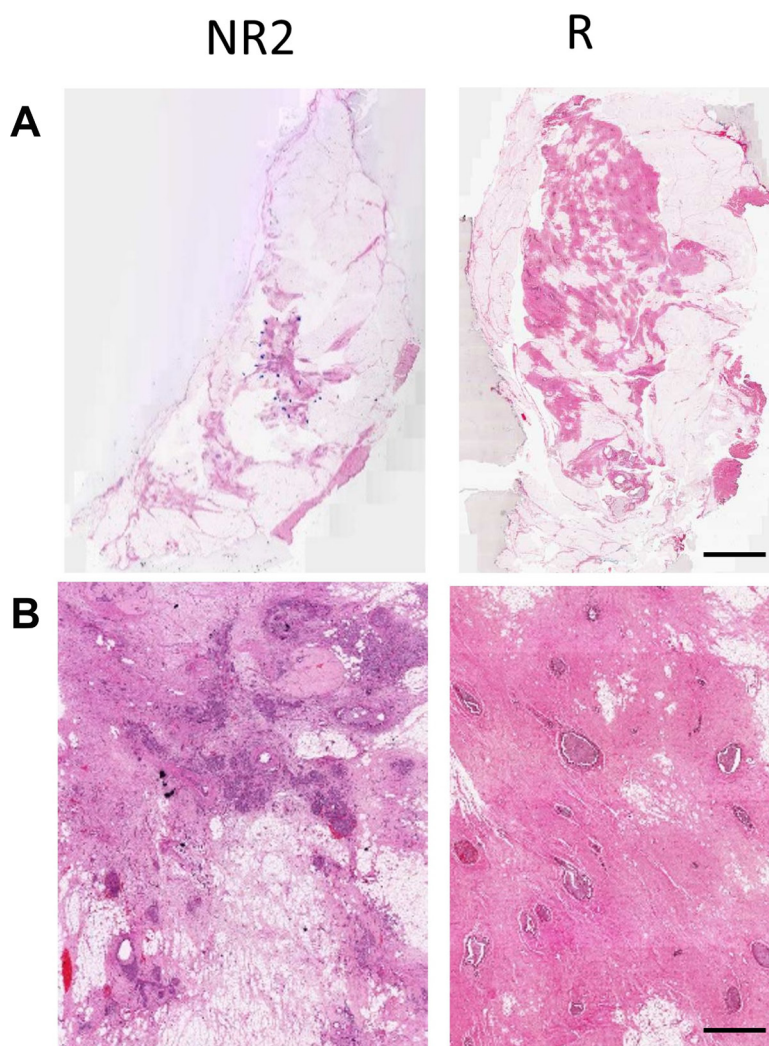
Classification performance for survival was also evaluated. The classification performance measures for classifying 5-year survivors versus patients with recurrence were similar to those for the conventional response classification (sensitivity, specificity, accuracy, and AUC of  $89 \pm 8\%$ ,  $84 \pm 11\%$ ,  $85 \pm 7\%$ , and  $0.89 \pm 0.11$ , respectively).

Figure 4 compares the AUCs obtained using the ANN and KNN classifiers for predicting two-class and three-class responses and survival of patients. In classification tasks, the ANN classifier outperformed the KNN classifier. Table 4 presents, for each grouping scheme, the five QUS and hormone features selected

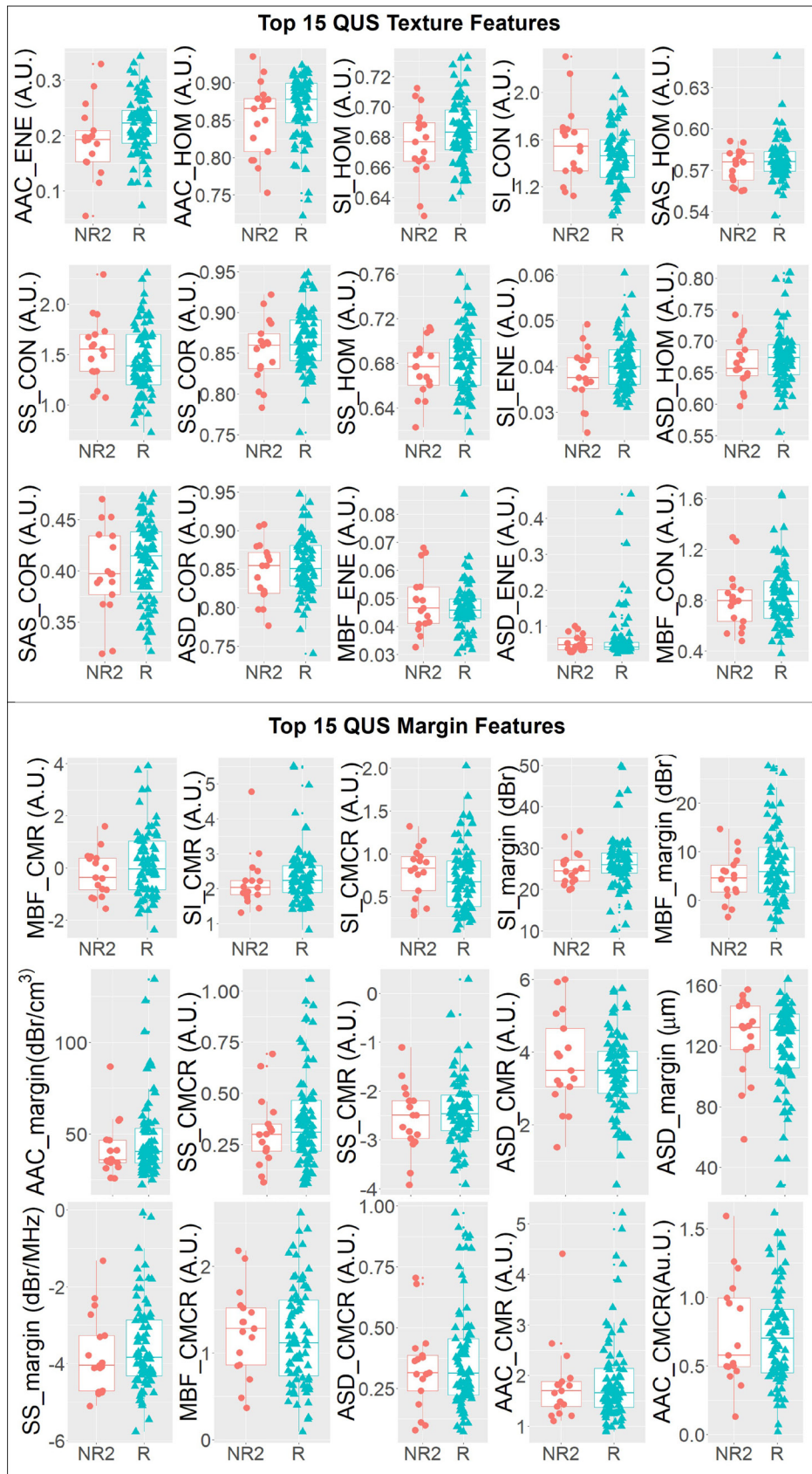
by the sequential forward feature selection method that yielded the highest AUC. As evident, QUS texture features contributed prominently to the response prediction models in all grouping schemes. Hormone features did not contribute to the binary classification (conventional response prediction and survival prediction), whereas the opposite was true when patients were grouped based on their three-category response criteria: for CR vs (PR+NR3) classification, ERe and PRe contributed to the prediction; and for PR vs NR3 prediction, ERe contributed to the prediction.

## DISCUSSION

In this study, the statistical features of QUS images combined with an artificial neural network classifier were demonstrated, for the first time, to be effective in the pre-treatment prediction of response and 5-year recurrence-free survival of LABC patients receiving neoadjuvant chemotherapy. Both the conventional clinical response



**Figure 2:** (A) H & E stained histology images of the excised breast specimen after resection. (B) High-magnification images. Scale bars: H & E low magnification – 1 cm, H & E high magnification – 100  $\mu\text{m}$ .



**Figure 3: One-dimensional scatter plots and overlaid boxplots of the top 15 QUS texture features and top 15 QUS margin features comparing responder (R) and non-responder (NR2) groups. The features are plotted in order of statistical significance (smallest  $p$ -value to largest  $p$ -value) from left to right, top to bottom in a raster fashion.**

**Table 3: Comparison of classification performances by ANN for different types of patient classification**

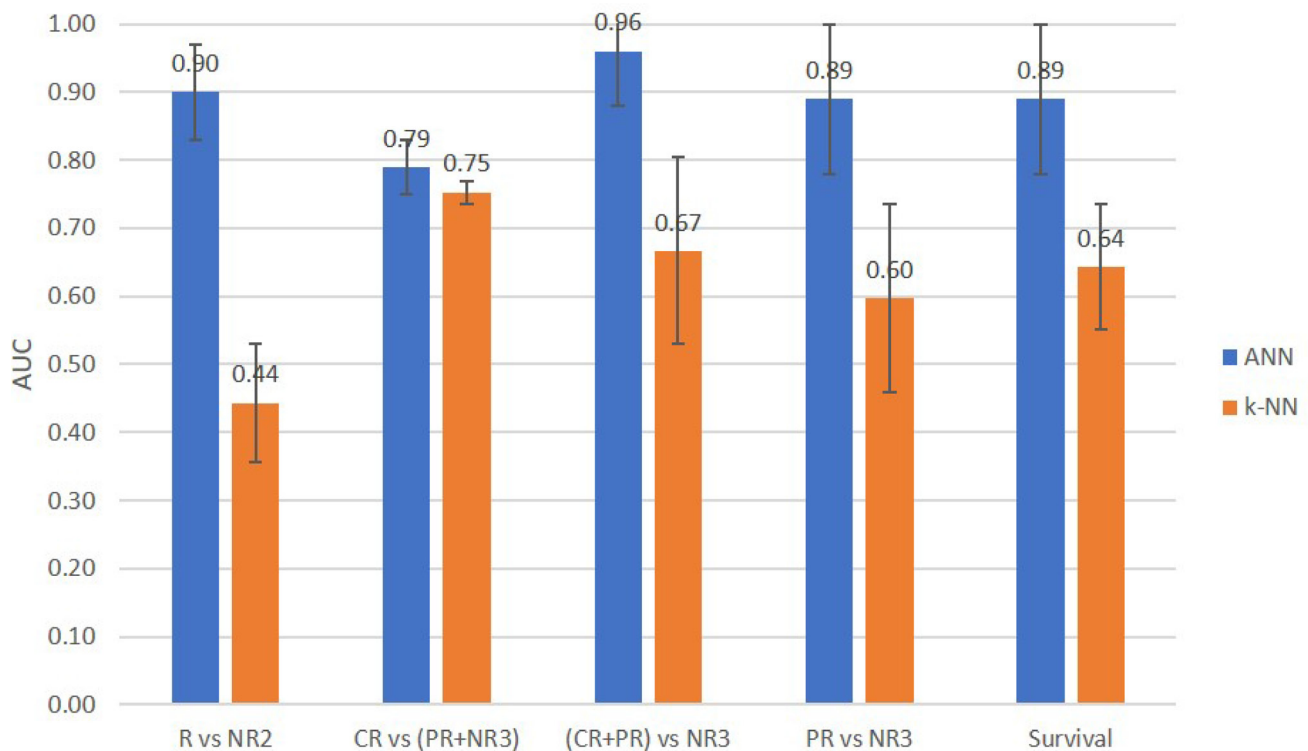
		Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
<b>R vs NR2</b>	Mean	89	85	87	0.90
	SD	9	12	6	0.07
<b>CR vs (PR+NR3)</b>	Mean	83	75	79	0.79
	SD	3	1	1	0.04
<b>(CR+PR) vs NR3</b>	Mean	93	98	96	0.96
	SD	9	6	6	0.08
<b>PR vs NR3</b>	Mean	88	84	86	0.89
	SD	12	16	10	0.11
<b>Survival</b>	Mean	89	84	85	0.89
	SD	8	11	7	0.11

Reported values are mean and standard deviation (SD) values obtained by averaging the results over the subsets.

and recurrence-free survival were predicted with high accuracies (87 % and 85 % on average, respectively). Importantly, the best results were obtained when differentiating patients with no response versus those with some response ((CR+PR) versus NR3) with an accuracy of 96% on average (93 % sensitivity, 98% specificity and 0.96 AUC). The classification results were validated with patient modified response scores determined using post-surgical pathology data. The method proposed here can be incorporated, as a pre-treatment screening step, in

the clinical workflow of LABC patients. This step would provide insight into the effectiveness of a given treatment regimen and allow the personalization of treatment. If it is known up-front that a patient will not respond to a particular chemotherapy, other agents or treatments can be selected instead of embarking on a several-month course of ineffective chemotherapy.

In this study, the ANN provided the best classification results. Results obtained using a KNN classifier were worse but were limited to 5 input



**Figure 4: Comparison of prediction performance AUCs of the ANN and KNN classifiers for two-class and three-class response and survival prediction tasks.**



**Table 4: The five best QUS + molecular features obtained by the ANN classifier for different types of classification**

	Best Features				
<b>R vs NR2</b>	AAC <sub>ENE_IN</sub>	AAC <sub>HOM_IN</sub>	MBF <sub>CMR</sub>	SI <sub>HOM_IN</sub>	SAS <sub>ENE_IN</sub>
<b>CR vs (PR+NR3)</b>	ER	SAS <sub>ENE_IN</sub>	SAS <sub>CON_IN</sub>	PR	MBF <sub>HOM_IN</sub>
<b>(CR+PR) vs NR3</b>	SS <sub>ENE_IN</sub>	SAS <sub>HOM_IN</sub>	AAC <sub>ENE_IN</sub>	ASD <sub>HOM_IN</sub>	ASD <sub>COR_IN</sub>
<b>PR vs NR3</b>	SAS <sub>HOM_IN</sub>	SAS <sub>ENE_IN</sub>	SAS <sub>CON_IN</sub>	SI <sub>COR_IN</sub>	ER
<b>Survival</b>	AAC <sub>CON_IN</sub>	ASD <sub>CON_IN</sub>	ASD <sub>CORE_IN</sub>	ASD <sub>MARGIN</sub>	SI <sub>ENE_IN</sub>

Reported are the five QUS + molecular features that were determined by way of cross-validation to yield the highest AUC.

parameters to avoid overfitting, whereas our previous work used more than 10. The high accuracy attained here is important for such methods to be used clinically. This can potentially lead to an improvement in patient quality of life as well as substantial savings in time, costs, and resources for both the patient and the health care provider.

The results demonstrated that the gray-level co-occurrence matrix (GLCM)-based texture features contributed to both response prediction models (conventional and three-class). The sensitivity of QUS texture features to therapy responsiveness are likely linked to the heterogeneous nature of tumour response to chemotherapy at the early stages. This theory has been suggested in previous studies examining GLCM-based [25] and local binary pattern-based [34] QUS texture analyses of LABC tumours undergoing chemotherapy. Aside from treatment response characterization, a previous LABC tumour characterization study demonstrated that QUS texture features provide a strong discrimination between low grade and medium-to-high grade tumours. [19], suggesting a link between QUS texture features and tumour heterogeneity.

Our response prediction results highlighted the sensitivity of the QUS feature set, identified by the ANN classifier, to the labels used in the training data set (Table 4). This may be due, in part, to the small number of non-responding patients ( $N = 17$ ) compared to responding patients ( $N = 83$ ). As data from new non-responding patients is collected in the future, the inter-patient variations in QUS features will be more effectively accounted for through machine learning and a more robust set of QUS features will be identified. The partial correlations between QUS features here are acknowledged. In parameters calculated through linear regression of the RF power spectrum, SS is related to the size of diffuse scatterers, SI is related to the acoustic concentration, and MBF is related to SS and SI. Among parameters using the Gaussian form factor model, ASD is an estimate of scatterer size and AAC is an estimate of acoustic concentration. However, due to the difference in the underlying models and assumptions, ASD and SS are partially correlated, and (MBF, SI) and AAC are partially correlated. In terms of ASD versus SAS, ASD characterizes the size of diffuse scatterers whereas SAS measures the spacing between both regular and

diffuse scatterers. In a study characterizing diffuse liver disease, SAS measurements have been correlated with the distribution of collagen fibers [35]. In breast studies, ASD measurements have been correlated with the size of cells [17, 18]. Thus, it is plausible for SAS measurements in this study to be correlated with collagen fibers and cells, whereas ASD measurements are correlated with the distribution of cells.

There is mounting evidence suggesting that molecular subtypes (i.e. hormone receptor expressions) of tumours play an important role in developed or inherent drug resistance [36]. The fact that ERe and PRe contributed to the CR vs (PR+NR3) differentiation and that ER status was determined as a contributing parameter to PR vs NR3 differentiation confirmed the importance of tumour hormone receptor expression as a predictive marker. This has also been suggested in previous studies [6, 7]. Here, a hybrid model consisting of image-based and molecular-based markers was constructed employing an ANN classifier, which yielded similar accuracy to that of our previous work [32]. The current study includes two improvements: the patient cohort investigated is nearly double the size of cohort in the previous study, and data imbalance correction was made prior to classification through random sub-sampling. Furthermore, in this study, a survival predictor model was developed using an ANN classifier. In the previous study [32], retrospective survival analysis was performed (Kaplan-Meier survival curves), providing predictive insight into patient survival. In addition, various three-way classifiers resulted in better results for what is an important clinical indicator-identifying patients *a priori* who will have no response to chemotherapy.

The method developed here may also be combined with monitoring of cell-death responses using quantitative ultrasound. Sadeghi-Naini *et al* [25] have recently used similar approaches to monitor treatment response based on cell death detection using quantitative ultrasound. They indicated that QUS markers of response to NAC capture microstructural changes in the tumour induced by anticancer drugs, which correlate very well with long-term outcomes. Thus, it is not surprising that such markers could provide insight into the likelihood of response prior to starting treatment.

Previous studies have investigated methods for *a priori* prediction of treatment response. Tran *et al.* [37] recently demonstrated the utility of diffuse optical spectroscopy imaging, particularly the homogeneity texture feature of oxygenated haemoglobin concentration within the breast in predicting breast tumour response to NAC with an accuracy of 88%. However, that study was limited to a smaller analysis of 37 patients, nearly a third of the size of the cohort used here. Furthermore, uncertainties in tumour delineation arose due to the relatively low resolution of DOS.

Molecular markers have also been used to predict breast cancer recurrence. A 21-gene reverse transcriptase-polymerase chain reaction (RT-PCR) assay, or Oncotype DX [38], has been used to grade a recurrence risk of breast cancer in patients with lymph node negative, estrogen receptor-positive breast cancer. The recurrence score was found to be predictive of whether or not a patient would benefit from adjuvant chemotherapy. However, for now that technique applies only to the aforementioned subgroup of breast cancer patients, whereas the QUS method proposed here applies to all LABC patients. Furthermore, our method is potentially extendable to early breast cancer patients receiving up-front chemotherapy or adjuvant chemotherapy, regardless of their lymph node or hormone receptor statuses.

Drug resistance of cancer cells to chemotherapy can be inherent or developed through exposure to the drug [36]. A large body of research has established multidrug resistance (MDR) transporter proteins as one of the key mechanisms of cancer cell resistance to chemotherapy drugs [36], particularly anthracyclines and taxanes. Thus, as a future investigation, correlating QUS properties of a tumour with its MDR biomarkers may shed light on the mechanism by which QUS could detect inherent MDR in a tumour. As mentioned previously, Ki-67 is also an important pre-treatment biomarker of tumour responsiveness [4]. As Ki-67 is a cell proliferation biomarker that is present in the active phases of the cell cycle (G1, S, G2, and mitosis), it is a marker of cellular and glandular morphology. QUS-based tissue characterization works by discriminating tissues based on differences in their microstructure. In the 1-10 MHz range of frequencies used in clinical applications, ultrasound is sensitive to scatterers in the range of 20-500  $\mu\text{m}$  in diameter [16]. Thus, it is plausible that ultrasound is sensitive to differences in the glandular morphology of tumours, which ultimately determines the likelihood of chemotherapy response. Most likely cellular changes associated with malignancy have an effect at one level, and as tumours become more aggressive the organization of cells becomes more and more deranged at the ductal level and then at the glandular level.

Due to the highly heterogeneous nature of tumours, particularly those of the breast, the prediction of their response to NAC requires advanced machine learning

algorithms that can effectively learn a non-linear pattern from data and build a strong classifier from several weak classifiers (QUS & molecular features). One of the most popular machine learning techniques with this capability is artificial neural networks. Recently, artificial neural networks have gained interest in oncology through successful applications in the detection of breast cancer in mammography images (AUC of 0.82) [39] and in the detection and localization of cancer metastasis in whole-slide pathology images of lymph nodes (AUC above 0.97) [40].

In the study here, an image-based model including textural features and tumour/periphery analyses was proposed for predicting response to NAC and survival of LABC patients. The sonographic analyses here can be thought of as generating “sonomic” biomarkers of response prediction akin to genomic biomarkers for predictive or prognostic assays but derived through ultrasound analyses as opposed to genetic analyses. Pre-treatment image-based biomarker surrogates of response stand to personalize health care by minimizing drug toxicity and maximizing chances of long-term survival. The technology can be incorporated into existing commercial clinical ultrasound imaging systems capable of RF data acquisition and potentially extended to other cancer types.

## MATERIALS AND METHODS

This prospective study was reviewed and approved by the institution’s research ethics board. After obtaining informed consent, ultrasound RF data were acquired from 100 patients with biopsy-confirmed LABC prior to start of their NAC. Data acquisition was performed by an experienced sonographer using a Sonix RP system (Ultrasonix, Vancouver, Canada) equipped with a 6 MHz linear array transducer (L14-5/60W) with a digital sampling rate of 40 MHz. The focus was set at the midline of the tumour using electronic beam focusing, and the imaging depth ranged from 4 to 6 cm, depending on tumour size and location. Images were acquired at 5 mm intervals over the tumour volume.

### Patient clinical characteristics

Patient data including age, initial tumour size (measured by imaging), ERe status, PRe status, and HER2 status were recorded. The clinical/pathological tumour response of each patient to treatment was determined at the end of their treatment using a modified response (MR) grading system which was based on RECIST [41] and histological [42] criteria. The MR score was defined as follows: MR Score 1: no diminishment in tumour size (cNR); MR2: up to 30% diminishment in tumour size (cNR); MR 3: between an estimated 30% and 90% reduction in tumour size (cPR); MR 4: a diminishment of more than 90% in tumour size (almost

**Table 5: QUS image-based features that were computed per patient, identified by ROI (core or margin) and image metric (mean, texture, CMR, or CMCR)**

	Core ROI	Margin ROI	Core vs. Margin	no. of features
Mean	MBF, SS, SI, SAS, ASD, AAC, ACE	MBF, SS, SI, SAS, ASD, AAC		13
Texture (CON, COR, ENE, HOM)	MBF, SS, SI, SAS, ASD, AAC,			24
CMR			MBF, SS, SI, SAS, ASD, AAC,	6
CMCR			MBF, SS, SI, SAS, ASD, AAC,	6
<b>TOTAL</b>				49

A total of 49 features were computed per patient.

pCR); MR 5: no evident tumour and no malignant cells identifiable in sections from the site of the tumour; only vascular fibroelastotic stroma remains, often containing macrophages; however, ductal carcinoma *in situ* may be present (pCR).

Both binary and three-class classifications were investigated. In the binary scenario, a patient with an MR score of 3-5 was deemed to be a responder (R) and a patient with an MR score of 1-2 was deemed to be a non-responder (NR2). In the three-class scenario, a patient with MR grade of 4-5 was deemed to be a complete responder (CR), 2-3 a partial responder (PR), and 1 a non-responder (NR3). The number proceeding NR (i.e. NR2 or NR3) differentiates the non-responders in the two-class and three-class grouping schemes. All patients received anthracycline/taxane-based treatment lasting several months. Each patient received a treatment regimen according to their disease type, stage, and hormone receptor expressions. Details about the specific types of treatments administered to individual patients are provided in Supplementary Table 1 in Supplementary Information. Recurrence-free survival was determined based on a 5-year follow up timeframe, during which the patient was free of any local or distant cancer recurrence.

### QUS feature evaluation

QUS analysis was carried out using the dual ROI method published previously [32]. In each B-mode breast ultrasound image, two separate ROIs consisting of 1) tumour core and 2) a rim of surrounding tissue of 5 mm thickness were manually contoured. This process was repeated on 4-7 image planes across the tumour. All images were para-sagittal. All images were non-overlapping. For each ROI, QUS features were computed within sliding RF windows that were  $2 \times 2$  mm in size and had 94% overlap in axial and lateral directions to produce a parametric map. From each parametric map, mean of intensity, texture, and image quality metrics were extracted and averaged across the image planes and subsequently used as features

(inputs) for the ANN-based patient response classifier. The QUS parametric maps that were generated from the raw ultrasound data were: MBF, SS, SI, ASD, AAC, SAS, and attenuation coefficient estimate (ACE). In order to characterize structural patterns in the parametric maps, a gray-level co-occurrence matrix (GLCM)-based texture analysis was performed on the newly obtained parametric images as described in Tadayyon *et al.* [19]. This method was originally developed by Haralick *et al.* [43]. The texture features that were extracted from the parametric maps were contrast (CON), correlation (COR), energy (ENE), and homogeneity (HOM). Additionally, two image quality metrics were extracted from the parametric maps that compared the statistical properties of the core ROI to those of the margin ROI: core-to-margin ratio (CMR), and core-to-margin contrast ratio (CMCR) as per the method described in [32]. Table 5 presents, in detail, the QUS-based features that were included in the ANN-based response classifier model based on ROI location and image metric. In addition to above-mentioned QUS features, tumour receptor expression statuses including PRe, ERe, an HER2 were investigated in the analysis. In total, 52 features were investigated as potential predictors of response. Details about the QUS analysis are provided in Supplementary Information.

### Response and recurrence-free survival classification

Prior to applying a classification rule, a data balancing step was performed by way of down-sampling to account for the smaller sample size of non-responding patients (MR=1-2, N=17) compared to responding patients (MR=3-5, N=83). In this step, random samples (with replacement) were drawn from the majority class QUS data (responding patients) with a size equal to that of the minority class (non-responding patients). This was repeated as many times as required to sample all patients in the majority class. Sequential forward feature selection (SFFS) [44] with *p*-value initialization was applied to the balanced dataset

to determine the optimal feature set for classification. This involved sorting the features based on their *p*-values of significance (from smallest to largest) obtained from an unpaired two-sample *t*-test or Mann–Whitney test, starting with the first feature as the initial feature set and adding or discarding features using the SFFS method until all features were evaluated. The maximum feature size was set to 5 in order to avoid overfitting due to the high dimensionality of the data set. The ANN classifier was configured as a single hidden layer model. In each balanced set, the data was randomly split into 70% training, 15% validation, and 15% test sets. In the training phase, hyper-parameter tuning was performed on the hidden layer size (1–10 nodes) using the training set to train the network and the validation set to evaluate it based on AUC. The test set was used to evaluate the generalization error of the network after fixing its hyper-parameters. The process was repeated 10 times on 10 bootstrapped, train-validate-test sets in order to account for variations in the ANN output due to random sampling. For each balanced set, a verification step was conducted to ensure that there was at least one sample from each class in each of the training, validation, and testing sets. For each balanced set, classifier performance metrics including sensitivity, specificity, accuracy, and AUC were measured on the test sets, which were obtained by averaging the values over the 10 bootstrap samples. For comparison, a KNN model was trained and tested in the same manner. Sensitivity was defined as the ratio of the number of true positives (responders) to the total number of positives in the test set. Specificity was defined as the ratio of the number of true negatives (non-responders) to the total number of negatives in the test set. Accuracy was defined as the ratio of the number of correctly classified patients to the total number of patients in the test set. All values are reported in percentages.

### Author contributions

HT, MG, LS, and GJC developed the methodology. HT, MG, LS, and GJC designed the experiment and analyzed the data. MT, KP, SG, AE, NLH, CH, FW, ER, DV, WTT, BC, and GJC acquired and analyzed the data. HT and GJC wrote and revised the manuscript.

### ACKNOWLEDGMENTS

Data were collected and available at the Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Canada. M.J.G held the Natural Sciences and Engineering Research Council of Canada Post-doctoral Fellowship. G.J.C. holds a University of Toronto James and Mary Davie Chair in Breast Cancer Imaging and Ablation.

### CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

### FUNDING

Funding for this project was provided by the Terry Fox Foundation, the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

### REFERENCES

1. Sharma U, Danishad KK, Seenu V, Jagannathan NR. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. *NMR Biomed.* 2009; 22:104–13. <https://doi.org/10.1002/nbm.1245>. [PubMed]
2. Schelling M, Avril N, Nāhrig J, Kuhn W, Römer W, Sattler D, et al. Positron emission tomography using [<sup>18</sup>F] Fluorodeoxyglucose for monitoring primary chemotherapy in breast cancer. *J Clin Oncol.* 2000; 18:1689–95. <https://doi.org/10.1200/JCO.2000.18.8.1689>. [PubMed]
3. Falou O, Soliman H, Sadeghi-Naini A, Iradji S, Lemon-Wong S, Zubovits J, Spayne J, Dent R, Trudeau M, Boileau JF, Wright FC, Yaffe MJ, Czarnota GJ. Diffuse optical spectroscopy evaluation of treatment response in women with locally advanced breast cancer receiving neoadjuvant chemotherapy. *Transl Oncol.* 2012; 5:238–46. <https://doi.org/10.1593/tlo.11346>. [PubMed]
4. Chang J, Ormerod M, Powles TJ, Allred DC, Ashley SE, Dowsett M. Apoptosis and proliferation as predictors of chemotherapy response in patients with breast carcinoma. *Cancer.* 2000; 89:2145–52. [https://doi.org/10.1002/1097-0142\(20001201\)89:11<2145::AID-CNCR1>3.0.CO;2-S](https://doi.org/10.1002/1097-0142(20001201)89:11<2145::AID-CNCR1>3.0.CO;2-S). [PubMed]
5. Stoetzer OJ, Fersching DM, Salat C, Steinkohl O, Gabka CJ, Hamann U, Braun M, Feller AM, Heinemann V, Siegele B, Nagel D, Holdenrieder S. Prediction of response to neoadjuvant chemotherapy in breast cancer patients by circulating apoptotic biomarkers nucleosomes, DNase, cytokeratin-18 fragments and survivin. *Cancer Lett.* 2013; 336:140–48. <https://doi.org/10.1016/j.canlet.2013.04.013>. [PubMed]
6. Lehner J, Stötzer OJ, Fersching D, Nagel D, Holdenrieder S. Circulating plasma DNA and DNA integrity in breast cancer patients undergoing neoadjuvant chemotherapy. *Clin Chim Acta.* 2013; 425:206–11. <https://doi.org/10.1016/j.cca.2013.07.027>. [PubMed]
7. Andre F, Mazouni C, Liedtke C, Kau SW, Frye D, Green M, Gonzalez-Angulo AM, Symmans WF, Hortobagyi GN, Pusztai L. HER2 expression and efficacy of preoperative paclitaxel/FAC chemotherapy in breast cancer. *Breast Cancer Res Treat.* 2008; 108:183–90. <https://doi.org/10.1007/s10549-007-9594-8>. [PubMed]
8. Jiang S, Pogue BW, Kaufman PA, Gui J, Jermyn M, Frazee TE, Poplack SP, DiFlorio-Alexander R, Wells WA, Paulsen KD. Predicting breast tumor response

- to neoadjuvant chemotherapy with diffuse optical spectroscopic tomography prior to treatment. *Clin Cancer Res.* 2014; 20:6006–15. <https://doi.org/10.1158/1078-0432.CCR-14-1415>. [PubMed]
9. Tromberg BJ, Zhang Z, Leproux A, O'Sullivan TD, Cerussi AE, Carpenter PM, Mehta RS, Roblyer D, Yang W, Paulsen KD, Pogue BW, Jiang S, Kaufman PA, et al, and ACRIN 6691 investigators. Predicting Responses to Neoadjuvant Chemotherapy in Breast Cancer: ACRIN 6691 Trial of Diffuse Optical Spectroscopic Imaging. *Cancer Res.* 2016; 76:5933–44. <https://doi.org/10.1158/0008-5472.CAN-16-0346>. [PubMed]
  10. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, Plecha D, Madabhushi A. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res.* 2017; 19:57. <https://doi.org/10.1186/s13058-017-0846-1>. [PubMed]
  11. Kurland BF, Peterson LM, Lee JH, Schubert EK, Currin ER, Link JM, Krohn KA, Mankoff DA, Linden HM. Estrogen Receptor Binding (18F-FES PET) and Glycolytic Activity (18F-FDG PET) Predict Progression-Free Survival on Endocrine Therapy in Patients with ER+ Breast Cancer. *Clin Cancer Res.* 2017; 23:407–15. <https://doi.org/10.1158/1078-0432.CCR-16-0362>. [PubMed]
  12. Coleman DJ, Lizzi FL, Silverman RH, Helson L, Torpey JH, Rondeau MJ. A model for acoustic characterization of intraocular tumors. *Invest Ophthalmol Vis Sci.* 1985; 26:545–50. [PubMed]
  13. Feleppa EJ, Kalisz A, Sokil-Melgar JB, Lizzi FL, Liu T, Rosado AL, Shao MC, Fair WR, Wang Y, Cookson MS, Reuter VE, Heston WD. Typing of prostate tissue by ultrasonic spectrum analysis. *IEEE Trans Ultrason Ferroelectr Freq Control.* 1996; 43:609–19. <https://doi.org/10.1109/58.503779>.
  14. Yang M, Krueger TM, Miller JG, Holland MR. Characterization of anisotropic myocardial backscatter using spectral slope, intercept and midband fit parameters. *Ultrason Imaging.* 2007; 29:122–34. <https://doi.org/10.1177/016173460702900204>. [PubMed]
  15. Mamou J, Coron A, Oelze ML, Saegusa-Beecroft E, Hata M, Lee P, Machi J, Yanagihara E, Laugier P, Feleppa EJ. Three-dimensional high-frequency backscatter and envelope quantification of cancerous human lymph nodes. *Ultrasound Med Biol.* 2011; 37:345–57. <https://doi.org/10.1016/j.ultrasmedbio.2010.11.020>. [PubMed]
  16. Insana MF, Wagner RF, Brown DG, Hall TJ. Describing small-scale structure in random media using pulse-echo ultrasound. *J Acoust Soc Am.* 1990; 87:179–92. <https://doi.org/10.1121/1.399283>. [PubMed]
  17. Oelze ML, O'Brien WD Jr, Blue JP, Zachary JF. Differentiation and characterization of rat mammary fibroadenomas and 4T1 mouse carcinomas using quantitative ultrasound imaging. *IEEE Trans Med Imaging.* 2004; 23:764–71. <https://doi.org/10.1109/TMI.2004.826953>. [PubMed]
  18. Tadayyon H, Sadeghi-Naini A, Wirtzfeld L, Wright FC, Czarnota G. Quantitative ultrasound characterization of locally advanced breast cancer by estimation of its scatterer properties. *Med Phys.* 2014; 41:012903. <https://doi.org/10.1118/1.4852875>. [PubMed]
  19. Tadayyon H, Sadeghi-Naini A, Czarnota GJ. Noninvasive characterization of locally advanced breast cancer using textural analysis of quantitative ultrasound parametric images. *Transl Oncol.* 2014; 7:759–67. <https://doi.org/10.1016/j.tranon.2014.10.007>. [PubMed]
  20. Sadeghi-Naini A, Suraweera H, Tran WT, Hadizad F, Bruni G, Rastegar RF, Curpen B, Czarnota GJ. Breast-Lesion Characterization using Textural Features of Quantitative Ultrasound Parametric Maps. *Sci Rep.* 2017; 7:13638. <https://doi.org/10.1038/s41598-017-13977-x>. [PubMed]
  21. Banihashemi B, Vlad R, Debeljevic B, Giles A, Kolios MC, Czarnota GJ. Ultrasound imaging of apoptosis in tumor response: novel preclinical monitoring of photodynamic therapy effects. *Cancer Res.* 2008; 68:8590–96. <https://doi.org/10.1158/0008-5472.CAN-08-0006>. [PubMed]
  22. Vlad RM, Brand S, Giles A, Kolios MC, Czarnota GJ. Quantitative ultrasound characterization of responses to radiotherapy in cancer mouse models. *Clin Cancer Res.* 2009; 15:2067–75. <https://doi.org/10.1158/1078-0432.CCR-08-1970>. [PubMed]
  23. Sadeghi-Naini A, Falou O, Tadayyon H, Al-Mahrouki A, Tran W, Papanicolau N, Kolios MC, Czarnota GJ. Conventional frequency ultrasonic biomarkers of cancer treatment response *in vivo*. *Transl Oncol.* 2013; 6:234–43. <https://doi.org/10.1593/tlo.12385>. [PubMed]
  24. Czarnota GJ, Karshafian R, Burns PN, Wong S, Al Mahrouki A, Lee JW, Caissie A, Tran W, Kim C, Furukawa M, Wong E, Giles A. Tumor radiation response enhancement by acoustical stimulation of the vasculature. *Proc Natl Acad Sci USA.* 2012; 109:E2033–41. <https://doi.org/10.1073/pnas.1200053109>. [PubMed]
  25. Sadeghi-Naini A, Sannachi L, Pritchard K, Trudeau M, Gandhi S, Wright FC, Zubovits J, Yaffe MJ, Kolios MC, Czarnota GJ. Early prediction of therapy responses and outcomes in breast cancer patients using quantitative ultrasound spectral texture. *Oncotarget.* 2014; 5:3497–511. <https://doi.org/10.18632/oncotarget.1950>. [PubMed]
  26. Sannachi L, Tadayyon H, Sadeghi-Naini A, Tran W, Gandhi S, Wright F, Oelze M, Czarnota G. Non-invasive evaluation of breast cancer response to chemotherapy using quantitative ultrasonic backscatter parameters. *Med Image Anal.* 2015; 20:224–36. <https://doi.org/10.1016/j.media.2014.11.009>. [PubMed]
  27. Suzuki K, Hayashi N, Sasaki Y, Kono M, Kasahara A, Imai Y, Fusamoto H, Kamada T. Evaluation of structural change in diffuse liver disease with frequency domain analysis of ultrasound. *Hepatology.* 1993; 17:1041–46. <https://doi.org/10.1002/hep.1840170616>. [PubMed]

28. Abeyratne UR, Tang X. Ultrasound scatter-spacing based diagnosis of focal diseases of the liver. *Biomed Signal Process Control*. 2007; 2:9–15. <https://doi.org/10.1016/j.bspc.2007.01.001>.
29. Machado CB, Pereira WC, Meziri M, Laugier P. Characterization of *in vitro* healthy and pathological human liver tissue periodicity using backscattered ultrasound signals. *Ultrasound Med Biol*. 2006; 32:649–57. <https://doi.org/10.1016/j.ultrasmedbio.2006.01.009>. [PubMed]
30. Wear KA, Wagner RF, Insana MF, Hall TJ. Application of autoregressive spectral analysis to cepstral estimation of mean scatterer spacing. *IEEE Trans Ultrason Ferroelectr Freq Control*. 1993; 40:50–58. <https://doi.org/10.1109/58.184998>. [PubMed]
31. Tadayyon H, Sannachi L, Gangeh M, Sadeghi-Naini A, Tran W, Trudeau ME, Pritchard K, Ghandi S, Verma S, Czarnota GJ. Quantitative ultrasound assessment of breast tumor response to chemotherapy using a multi-parameter approach. *Oncotarget*. 2016; 7:45094–111. <https://doi.org/10.18632/oncotarget.8862>. [PubMed]
32. Tadayyon H, Sannachi L, Gangeh MJ, Kim C, Ghandi S, Trudeau M, Pritchard K, Tran WT, Slodkowska E, Sadeghi-Naini A, Czarnota GJ. A priori Prediction of Neoadjuvant Chemotherapy Response and Survival in Breast Cancer Patients using Quantitative Ultrasound. *Sci Rep*. 2017; 7:45733. <https://doi.org/10.1038/srep45733>. [PubMed]
33. Freeman J, Skapura D. *Neural Networks: Algorithms, Applications, and Programming Techniques*. Addison-Wesley; 1991.
34. Gangeh MJ, Tadayyon H, Sannachi L, Sadeghi-Naini A, Tran WT, Czarnota GJ. Computer Aided Theragnosis Using Quantitative Ultrasound Spectroscopy and Maximum Mean Discrepancy in Locally Advanced Breast Cancer. *IEEE Trans Med Imaging*. 2016; 35:778–90. <https://doi.org/10.1109/TMI.2015.2495246>. [PubMed]
35. Fellingham LL, Sommer FG. Ultrasonic Characterization of Tissue Structure in the *In Vivo* Human Liver and Spleen. *IEEE Trans Sonics Ultrason*. 1984; 31:418–28. <https://doi.org/10.1109/T-SU.1984.31522>.
36. Luqmani YA. Mechanisms of drug resistance in cancer chemotherapy. *Med Princ Pract*. 2005; 14:35–48. <https://doi.org/10.1159/000086183>. [PubMed]
37. Tran WT, Gangeh MJ, Sannachi L, Chin L, Watkins E, Bruni SG, Rastegar RF, Curpen B, Trudeau M, Gandhi S, Yaffe M, Slodkowska E, Childs C, et al. Predicting breast cancer response to neoadjuvant chemotherapy using pretreatment diffuse optical spectroscopic texture analysis. *Br J Cancer*. 2017; 116:1329–39. <https://doi.org/10.1038/bjc.2017.97>. [PubMed]
38. Brufsky AM. Predictive and prognostic value of the 21-gene recurrence score in hormone receptor-positive, node-positive breast cancer. *Am J Clin Oncol*. 2014; 37:404–10. <https://doi.org/10.1097/COC.000000000000086>. [PubMed]
39. Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A. Deep Learning in Mammography: Diagnostic Accuracy of a Multipurpose Image Analysis Software in the Detection of Breast Cancer. *Invest Radiol*. 2017; 52:434–40. <https://doi.org/10.1097/RLI.0000000000000358>. [PubMed]
40. Liu Y, Gadepalli K, Norouzi M, Dahl GE, Kohlberger T, Boyko A, Venugopalan S, Timofeev A, Nelson PQ, Corrado GS, Hipp JD, Peng L, Stumpe MC. Detecting Cancer Metastases on Gigapixel Pathology Images. 8 Mar 2017. Available from <https://arxiv.org/abs/1703.02442v2>.
41. Eisenhauer E, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009; 45:228–47. <https://doi.org/10.1016/j.ejca.2008.10.026>. [PubMed]
42. Ogston KN, Miller ID, Payne S, Hutcheon AW, Sarkar TK, Smith I, Schofield A, Heys SD. A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast*. 2003; 12:320–27. [https://doi.org/10.1016/S0960-9776\(03\)00106-1](https://doi.org/10.1016/S0960-9776(03)00106-1). [PubMed]
43. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Trans Syst Man Cybern*. 1973; SMC-3:610–21. <https://doi.org/10.1109/TSMC.1973.4309314>.
44. Jain A, Duin R, Mao J. Statistical Pattern Recognition: A Review. *IEEE Trans Pattern Anal Mach Intell*. 2000; 22:4–37. <https://doi.org/10.1109/34.824819>.