

Predicting Breast Cancer Response to Neoadjuvant Chemotherapy Using Pretreatment Diffuse Optical Spectroscopic-Texture Analysis

TRAN, William T., GANGEH, Mehrdad J., SANNACHI, Lakshmanan, CHIN, Lee, WATKINS, Elyse, BRUNI, Silvio G., RASTEGAR, Rashin Fallah, CURPEN, Belinda, TRUDEAU, Maureen, GANDHI, Sonal, YAFFE, Martin, SLODKOWSKA, Elzbieta, CHILDS, Charmaine <<http://orcid.org/0000-0002-1558-5633>>, SADEGHI-NAINI, Ali and CZARNOTA, Gregory J.

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/15429/>

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

TRAN, William T., GANGEH, Mehrdad J., SANNACHI, Lakshmanan, CHIN, Lee, WATKINS, Elyse, BRUNI, Silvio G., RASTEGAR, Rashin Fallah, CURPEN, Belinda, TRUDEAU, Maureen, GANDHI, Sonal, YAFFE, Martin, SLODKOWSKA, Elzbieta, CHILDS, Charmaine, SADEGHI-NAINI, Ali and CZARNOTA, Gregory J. (2017). Predicting Breast Cancer Response to Neoadjuvant Chemotherapy Using Pretreatment Diffuse Optical Spectroscopic-Texture Analysis. *British journal of cancer*, 116 (10), 1329-1339.

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Keywords: diffuse optical spectroscopy imaging; breast cancer; imaging biomarkers

Predicting breast cancer response to neoadjuvant chemotherapy using pretreatment diffuse optical spectroscopic texture analysis

William T Tran^{1,2}, Mehrdad J Gangeh¹, Lakshmanan Sannachi^{1,3}, Lee Chin¹, Elyse Watkins¹, Silvio G Bruni⁴, Rashin Fallah Rastegar⁴, Belinda Curpen⁴, Maureen Trudeau⁵, Sonal Gandhi⁵, Martin Yaffe⁶, Elzbieta Slodkowska⁷, Charmaine Childs², Ali Sadeghi-Naini^{1,3,6,8} and Gregory J Czarnota^{*1,3,6,8}

¹Department of Radiation Oncology, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada; ²Centre for Health and Social Care Research, Sheffield Hallam University, 32 Collegiate Crescent, Sheffield S10 2BP, UK; ³Department of Medical Biophysics, University of Toronto, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada; ⁴Department of Medical Imaging, Sunnybrook Hospital, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada; ⁵Division of Medical Oncology, Department of Medicine, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, ON M4N 3M5, Canada; ⁶Physical Sciences Platform, Sunnybrook Research Institute, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada; ⁷Department of Anatomic Pathology, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada and ⁸Department of Radiation Oncology, University of Toronto, 2075 Bayview Avenue, Toronto, Ontario M4N 3M5, Canada

Background: Diffuse optical spectroscopy (DOS) has been demonstrated capable of monitoring response to neoadjuvant chemotherapy (NAC) in locally advanced breast cancer (LABC) patients. In this study, we evaluate texture features of pretreatment DOS functional maps for predicting LABC response to NAC.

Methods: Locally advanced breast cancer patients ($n=37$) underwent DOS breast imaging before starting NAC. Breast tissue parametric maps were constructed and texture analyses were performed based on grey-level co-occurrence matrices for feature extraction. Ground truth labels as responders (R) or non-responders (NR) were assigned to patients based on Miller–Payne pathological response criteria. The capability of DOS textural features computed on volumetric tumour data before the start of treatment (i.e., 'pretreatment') to predict patient responses to NAC was evaluated using a leave-one-out validation scheme at subject level. Data were analysed using a logistic regression, naive Bayes, and k -nearest neighbour classifiers.

Results: Data indicated that textural characteristics of pretreatment DOS parametric maps can differentiate between treatment response outcomes. The HbO₂ homogeneity resulted in the highest accuracy among univariate parameters in predicting response to chemotherapy: sensitivity (%Sn) and specificity (%Sp) were 86.5% and 89.0%, respectively, and accuracy was 87.8%. The highest predictors using multivariate (binary) combination features were the Hb-contrast + HbO₂-homogeneity, which resulted in a %Sn/%Sp = 78.0/81.0% and an accuracy of 79.5%.

Conclusions: This study demonstrated that the pretreatment DOS texture features can predict breast cancer response to NAC and potentially guide treatments.

*Correspondence: Dr GJ Czarnota; E-mail: gregory.czarnota@sunnybrook.ca

Received 7 March 2017; revised 16 March 2017; accepted 17 March 2017

© 2017 Cancer Research UK. All rights reserved 0007–0920/17

Breast cancer is the second leading cause of cancer-related mortality in women, and accounts for approximately one-third of new malignancies (Desantis *et al*, 2014). Nearly 20% of cases present as locally advanced breast cancer (LABC), which are characterised as stage IIB or stage III disease; thus, having large/bulky tumours which are often > 5 cm in size and involving the lymph nodes or skin (Giordano, 2003; Whitman and Strom, 2009). Survival outcomes for LABC are poor; only 50% of patients survive beyond 5 years (Lee and Newman, 2007). Guidelines from the National Institute for Health and Care Excellence (NICE) in the United Kingdom recommend neoadjuvant chemotherapy (NAC) for clinical management to downstage tumours prior to locoregional treatment with surgery and radiotherapy (Cance *et al*, 2002; Lee and Newman, 2007; National Institute For Health And Clinical Excellence (NICE), 2009). However, variable tumour responses have been shown in patients receiving NAC and there is evidence to suggest that favourable response to NAC correlates to improved disease-free survival (DFS) (Mathew *et al*, 2009; von Minckwitz *et al*, 2012).

Measuring tumour response at early stages of NAC may help guide treatments for potentially improved DFS. Standard methods use clinical palpation or medical imaging such as magnetic resonance imaging (MRI) to measure anatomical changes (Eisenhauer *et al*, 2009). However major limitations include variability in reader expertise, image reproducibility, and the time delay between biological and anatomical changes (Eisenhauer *et al*, 2009; Roblyer *et al*, 2011). Thus, emerging studies are investigating quantitative imaging methods to evaluate tumour response based on functional and biological tumour alterations to NAC within days to weeks after treatment initiation (Jakubowski *et al*, 2004; Roblyer *et al*, 2011). Recent studies have indicated that diffuse optical spectroscopy (DOS) imaging can characterise tumour response by focusing on changes in tissue composition (Cerussi *et al*, 2007; Roblyer *et al*, 2011; Jiang *et al*, 2014). Maps of tumour physiological features, such as haemoglobin are computed from tissue optical properties that are based on near-infrared optical scattering and absorption (Cerussi *et al*, 2006). For breast tissue, significant optical absorbers include oxy-haemoglobin (HbO₂), deoxy-haemoglobin (Hb), water (H₂O), and lipids (Li) (Cerussi *et al*, 2006). Their concentrations can be estimated by measuring the absorption coefficient (μ_a) and using Beer's law equation (Cerussi *et al*, 2011). Also, tissue optical parameters such as the reduced scattering coefficient (μ'_s) can provide additional information on tissue microstructure ($\sim 0.2 \mu\text{m}$), corresponding to optical scattering effects from mitochondria and the cell nucleus (Mourant *et al*, 2000; Cerussi *et al*, 2006). In previous studies, changes in DOS markers such as HbO₂ have been correlated to tumour response within 1 week of starting NAC (Roblyer *et al*, 2011) and after several cycles of chemotherapy (Schaafsma *et al*, 2015). A study by Ueda *et al* (2012) also measured baseline tumour oxygen saturation (StO₂) prior to NAC and reported significantly higher StO₂ in tumours that demonstrated pathological complete response (pCR) vs non-pCR.

Second-order statistical analysis of DOS parametric images can also yield texture features and aid in discriminating tumour response during NAC (Sadeghi-Naini *et al*, 2015). Feature extraction methods such as those based on grey-level co-occurrence matrices (GLCM) can be applied to compute the probabilities of relative pixel intensities of images from the spatial distribution of their voxels (Haralick *et al*, 1973). Grey-level co-occurrence matrices texture features include contrast (con), correlation (cor), homogeneity (hom), and energy (ene), and are dependent on the number of grey-levels (N_g) in the image or within the region of interest (ROI). Such techniques have been applied to several modalities, such as X-ray mammography (Li *et al*, 2005), MRI (Lerski *et al*, 1993; Chen *et al*, 2007), positron-emission tomography (Chicklore *et al*, 2013), and ultrasound

(Yang *et al*, 2012) in order to discriminate and characterise tissue types (Castellano *et al*, 2004). In breast studies, GLCM analysis has been able to classify benign and malignant lesions using planar (2D) and volumetric (3D) MRI images (Gibbs and Turnbull, 2003; Chen *et al*, 2007). For X-ray mammography, GLCM analysis has been used to segment lesion borders of stellate (malignant) breast masses (Gupta and Undrill, 1995). It was also recently reported that texture-based features from quantitative ultrasound imaging can be used to classify responders and non-responders early during NAC treatment (Sadeghi-Naini *et al*, 2014). These previous findings suggested that textural features may detect the acute, heterogeneous microstructural features carried in the parametric layout (Sadeghi-Naini *et al*, 2014). The research here is built on those previous findings but focusses on the pretreatment DOS texture features of the tumour area in 37 patients with LABC.

The aims of the present study were to carry out baseline DOS texture analysis to characterise pretreatment tumour heterogeneity as a marker for NAC response. Diffuse optical spectroscopy breast maps were acquired in LABC before treatment and here, texture features were used to predict treatment response (responder vs non-responder) based on three classifier models: a logistic regression analysis, naive Bayes, and k -nearest neighbour (k -NN). The classification ground truth was based on final Miller–Payne (MP) pathologic response criteria. The results of this study demonstrate a significant difference ($P < 0.05$) in the DOS textural features between response groups for features such as the Hb-hom and StO₂-con. The results indicate that textural characteristics of pretreatment DOS parametric maps can serve as predictors of pathological response to NAC. Diffuse optical spectroscopy imaging biomarkers could potentially help guide treatment for improved treatment outcomes in LABC patients, and for tailoring personalised treatment.

MATERIALS AND METHODS

Patient treatment and response criteria. This study was approved by the institutional research ethics board at Sunnybrook Health Sciences Centre, Toronto, Canada and all patients signed a written informed consent before participating in the study. Patients ($n = 37$) were included following a biopsy-confirmed diagnosis of LABC, aged 18–85, and were recommended for NAC. As part of the patient's standard of care, tumour molecular and histological features such as: oestrogen receptor (ER) status, progesterone receptor (PR) status, and human epidermal growth factor receptor amplification (Her2/Neu) were determined during the patient's diagnostic work-up. Patients had pretreatment MRI for initial size measurements. Clinical data were recorded in the patient's electronic medical record and retrieved for this study.

Neoadjuvant chemotherapy consisted of combination anthracycline- and taxane-based therapy for 6–8 cycles. These included combination therapies of either AC-T (Adriamycin, Cyclophosphamide, Taxol) or FEC-D (Fluorouracil, Epirubicin, Cyclophosphamide, Docetaxel). For Her2/Neu-amplified tumours, patients also received Trastuzumab concurrently with taxanes. Patients were clinically assessed by physical examination for the duration of treatment by their medical oncologist. Patients were assessed for NAC response pathologically using MP criteria (Ogston *et al*, 2003; Provenzano *et al*, 2015). MP criteria is a five-point grading system where MP-5 indicates no malignant cells identified in the tumour bed after treatment; MP-4 indicates a significant disappearance of tumour cells (>90% loss of tumour cells); MP-3 indicates a 30–90% disappearance of tumour cells; and MP-2 and MP-1 indicate no change or minor loss (<30%) of tumour cells (Ogston *et al*, 2003). Patients were considered responders (R) if there was at least a 30% reduction in tumour cells under microscopy (i.e., MP

3–5) (Ogston *et al*, 2003). Patients were otherwise classified as non-responders (NR) (MP 1–2). On the basis of these criteria, 27 patients were classified as responders and 10 patients were non-responders; that is, the ratio of responders to non-responders was ~ 3 . This labelling was used as the ground truth in the classifiers analyses performed as explained in the next subsections.

Tomographic diffuse optical spectroscopy imaging. Time-domain DOS imaging was performed on the patient's breast prior to the start of NAC, using a commercially developed tomographic DOS imaging device (SoftScan, Montreal, QC, Canada). Patients were laid prone with the breast positioned inside an imaging chamber. The breast was stabilised with compression plates in the craniocaudal direction and optical compensation medium was added for optical coupling between surfaces ($\mu_a = 0.04 \text{ cm}^{-1}$; $\mu_s = 11 \text{ cm}^{-1}$ (Intes, 2005)). Near-infrared light transmission was carried out using four individually pulsed semiconductor diode lasers operating at 690, 730, 780, and 830 nm, with a pulse duration of $< 150 \text{ ps}$, 20 MHz repetition frequency, and an average power of 0.5 mW (PicoQuant, Berlin, Germany). Optical detection used five output-fibered collimators that collected photons transmitted through the breast. Fibered collimators were coupled to photo-multiplier tubes (Hamamatsu, Bridgewater, NJ, USA) and directed to a photon counting module (Becker & Hickl, Berlin, Germany). The temporal-point spread function was measured and used to compute the absorption (μ_a) and scattering (μ_s) coefficients based on photon diffusion theory in tissue (Xu and Povoski, 2007).

Parametric data included measurements of dominant DOS chromophores in breast comprising of HbO₂ and Hb (Cerussi *et al*, 2006). The absorbance spectra of these tissue components were used to determine their concentrations (C) using the Beer–Lambert law, with known molar extinction coefficients (ε) (Equation 1):

$$\mu_a = \varepsilon \times C \quad (1)$$

Other DOS parameters, such as %water, %lipid, scattering power (b), and scattering amplitude (A) were measured using the power-law fit of the scattering spectra within a given wavelength (λ) (Equation 2). This relationship is based on a Mie scattering approximation (Tromberg *et al*, 2005):

$$\mu_s(\lambda) = A \times \lambda^b \quad (2)$$

Additionally, other optical parameters such as StO₂, oxygen desaturation (St), total haemoglobin (HbT), and the tissue optical index (TOI) were calculated from the Hb, HbO₂, %water, and %lipid, and these calculations have been described elsewhere (Intes, 2005; Cerussi *et al*, 2011).

DOS texture analysis. Whole-breast tomographic DOS parametric maps were constructed with an in-plane resolution of $3 \times 3 \text{ mm}^2$ and slice thickness of 7.5 mm. The ROI was an ellipsoid contour that was manually annotated around the tumour (Review Workstation, SoftScan; Schaafsma *et al*, 2015). Tumour ROIs were also selected with reference to the patients' pretreatment MRI with breast radiologists (BC, SGB, RFR).

Grey-level co-occurrence matrices texture analysis was performed on tomographic DOS data for the entire tumour volume and averaged over multiple frames; corresponding to the DOS parameters. A bilinear interpolation was applied to compensate for differences in the spatial resolutions, thus obtaining volumetric images with isotropic voxels ($3 \times 3 \times 3 \text{ mm}^3$; Sadeghi-Naini *et al*, 2015). For the GLCM, grey-tone intensities (N_g) were quantised into 16 grey-levels from the DOS parametric maps. A symmetric GLCM was constructed based on the spatial relationship of each voxel's neighbours using a displacement vector; the magnitude of the displacement vector was one to four voxel distances. A cumulative GLCM was calculated in 13 directions (45° rotations in each adjacent direction) relative to the central voxel (Chen *et al*, 2007). The resulting co-occurrence features included: energy,

homogeneity, contrast, and correlation. These were previously defined by (Haralick *et al* (1973)):

$$\text{Energy (Ene)} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} g_{ij}^2 \quad (3)$$

where the energy (angular second moment) describes the textural uniformity of the image, $0 \leq \text{Ene} \leq 1$, and g_{ij} is the i th and j th entry of the GLCM representing a grey-tone intensity in the matrix;

$$\text{Homogeneity (Hom)} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} g_{ij} \quad (4)$$

where the homogeneity (inverse difference moment) measures the diagonal elements within the displacement vector of the GLCM and relates the similarities in grey tones between voxels;

$$\text{Contrast (Con)} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 g_{ij} \quad (5)$$

where the contrast measures the differences between the lowest and highest voxels' grey tones and finally;

$$\text{Correlation (Cor)} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)g_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (6)$$

where the correlation measures the linear dependency on neighbouring grey-tone intensities; and σ_x , σ_y are the s.d., and μ_x , μ_y are the means of the probability matrix. Therefore, a total of 40 DOS texture features was included for analysis; there were 10 DOS parameters comprised of four GLCM features for each parameter (Supplementary Figure 1). Diffuse optical spectroscopy texture features were calculated using MatLab R2011b (The MathWorks Inc., Natick, MA, USA).

Statistical analysis and classification. All statistical analyses were performed using SPSS V.22 (IBM Corp., Armonk, NY, USA) and MatLab R2011b (The MathWorks Inc.). The DOS–GLCM features were compared for significant differences between R and NR patients ($n = 37$). Additionally, a subgroup analysis was completed based on tumour molecular features and chemotherapy treatments. The subgroups included ER+ ($n = 27$), 'triple-negative' (ER−/PR−/HER2−; $n = 7$), FEC-D ($n = 16$) and AC-T ($n = 21$) chemotherapy. HER2+ patients were not considered due to the limited number of non-responders in this subgroup ($n = 2$). For the classifiers to be used, at least three subjects are needed within each response group to perform the leave-one-subject-out validation scheme; as it may happen that one subject be in the test set, one in the validation set, and therefore, one additional subject is needed in the training set.

Comparisons were first tested for normality violations using a Shapiro–Wilk test. Normally distributed parameters were tested for significance using an unpaired t -test (two-sided within the 95% confidence interval). Otherwise, a Mann–Whitney test was performed (two-sided, 95% confidence).

In this study, a logistic regression analysis, a naive Bayes classification method, and a k -NN classifier were used to analyse the predictive value of DOS–GLCM features. A naive Bayes classification method assumes that features are independent of each other within the class variable. The k -NN classifier considers a test point and analyses the neighbouring points to form a class based on the frequency and distance between points. Four performance measures including the sensitivity (%Sn), specificity (%Sp), accuracy, and area under curve (AUC) of the receiver-operating characteristic (ROC) were calculated for both univariate parameters and multivariate parameters. These measures were used to compare the performance of the three classifiers. Prior to training and test validation, the data set was randomly subsampled

into 20 subsets with replacement. Each subset had equal numbers of responders and non-responders; this method was used to account for the data imbalance between the two response groups. Since each patient was represented using 40 DOS–GLCM features, and due to a limited sample size, in order to prevent the ‘curse of dimensionality’ (Jain *et al*, 2000), a feature selection based on sequential forward selection (SFS) algorithm in a wrapper framework (Duda *et al*, 2001) was performed to find the best (in the univariate case) or at most two (in the multivariate case) features. In order to prevent the peaking phenomenon due to the curse of dimensionality (Jain *et al*, 2000), the number of features should be at most 1/10th of the number of data samples, and since in the balanced data, there were only 20 data samples, a maximum of two features were selected for multivariate analysis using the SFS algorithm. The classifiers were evaluated using a leave-one-out cross validation at subject level. At each fold, the test set (one patient) remained unseen during the feature selection, tuning, and training of a classifier. Furthermore, at each fold, a leave-one-out cross validation was performed on the training set for the purpose of feature selection and tuning a classifier parameter (such as k in k -NN). Thus, at each fold, the training set was further divided into training and validation sets. The most discriminative feature(s) and the optimal classifier parameter were selected on the training set at each fold without involving the left-out test sample. Subsequently at each fold, the classifier was trained on the whole-training set using the optimal classifier parameter and selected features, and tested on the test sample. This process was repeated on all samples (in the leave-one-out process) to evaluate the performance of the classifier.

In order to test if there was a correlation between DOS–GLCM features and tumour biology and clinical features, a multiple linear regression analysis was completed using methods previously reported for imaging biomarker analysis (Evans *et al*, 2013). The following clinical variables were considered in the model: Patient’s age, ER/PR status, Her2 status, tumour size, and pathologic response. The regression coefficient (r) was calculated between the clinical variables and DOS–GLCM features. A statistical test of significance was also performed using an ANOVA test with an alpha of 0.05.

RESULTS

Patient characteristics. This study included $n = 37$ patients with biopsy-confirmed LABC. The median age of subjects was 50 years old. The mean tumour size in the largest dimension prior to NAC was 5.4 cm for responders and 7.0 cm for non-responders. The molecular features of tumours indicated that 27 patients were ER/PR positive. A total of 12 patients were Her2/Neu positive. A total of $n_1 = 27$ patients were classified as responders and $n_2 = 10$ patients as non-responders based on the ultimate pathological data (described below). All patients received taxane- and anthracycline-based chemotherapies: 21 patients received AC-T chemotherapy and 16 patients underwent FEC-D chemotherapy. Twelve patients who were Her2/Neu positive received Trastuzumab during taxane chemotherapy. Patient characteristics are summarised in Table 1. Representative DOS parametric maps for responders and non-responders, and the pretreatment MRI are shown in Figure 1.

Tumour haemoglobin and oxygenation – texture features demonstrate significant differences between response groups. Box-and-whisker plots for DOS–GLCM haemoglobin and StO₂ features are presented in Figures 2 and 3, respectively. The deoxy-haemoglobin-homogeneity (Hb-hom) feature demonstrated a significant difference between responders and non-responders; ($P = 0.030$). The Hb-hom feature was greater in NR compared to R (mean value: 0.329 ± 0.06 (a.u.) (\pm s.d.) vs 0.282 ± 0.06 , respectively). However, other features such as the Hb-con did not demonstrate a significant difference between groups ($P = 0.066$) (Figure 2). For the HbO₂-GLCM features, the HbO₂-cor was

Table 1. Patient and clinical characteristics	
Patient and tumour characteristics	$n = 37$ (All subjects)
Patients	
Age (years)	50
Pretreatment tumour size (MRI, cm)	
Responders	5.4
Non-responders	7.0
Molecular and histological features	
ER +	27
Triple negative/basal-like	7
HER2 +	12
Invasive ductal carcinoma	36
Invasive lobular carcinoma	1
Treatment response classification	
Responders	27
Non-responders	10
Chemotherapy and targeted therapies	
AC-T	21
FEC-D	16
Trastuzumab	12
Abbreviations: AC-T = adriamycin, cyclophosphamide, taxol; ER, oestrogen receptor; FEC-D = fluorouracil, epirubicin, cyclophosphamide, docetaxel; MRI, magnetic resonance imaging. Patients included in the study were diagnosed with biopsy-confirmed locally advanced breast cancer, and received a combination of anthracycline and taxane-based chemotherapies following standard institutional guidelines.	

greater for responders ($R = 0.205 \pm 0.06$ (a.u.) compared to NR = 0.156 ± 0.05 (a.u.) (mean values), $P < 0.024$). The HbT measurements in tumours demonstrated significant differences in homogeneity (HbT-hom) ($P = 0.047$). Other features were not significantly different such as the HbO₂-con ($P = 0.058$) and HbO₂-hom ($P = 0.088$).

There were also significant differences between response groups in tumour oxygenation texture features (St-con and StO₂-con, $P < 0.05$). The St-con was significantly different between response groups ($P = 0.044$); while other features such as St-hom were close to being significantly different ($P = 0.058$). St-con measurements were greater in responders (mean value: 21.87 ± 2.70 (a.u.)) vs non-responders (mean value = 20.43 ± 1.41 (a.u.)) ($P < 0.05$). Conversely the StO₂-con parameter was greater in non-responders (mean value = 22.13 ± 2.63 (a.u.)) compared to responders (mean value = 19.87 ± 3.02 (a.u.)) ($P < 0.05$).

Diffuse optical spectroscopy texture features that were significantly different between response groups (N vs NR) were analysed using the classifier models (Table 2A). In general, naive Bayes classification performed the best among the classifier models used in this study. For the HbT-hom, naive Bayes classification resulted in a %Sn of 84%, and %Sp of 85% (AUC = 0.813), in comparison to k -NN classification, which resulted in a classification of only %Sn = 74%, %Sp = 47%, and AUC of 0.552 (Table 2A). Analysis of all DOS texture features was also performed, independent of statistical significance between groups, using the three classifiers (logistic regression analysis, the naive Bayes model, or k -NN classifier). However, Table 2B presents the best DOS texture features from all possible univariate features ($d = 40$) extracted from each classifier model. The corresponding ROC curves with AUCs are presented in Figure 4. Classification results from significant univariate texture features indicated an AUC range between 0.756 and 0.821 (Figure 4). A maximum AUC was observed for HbO₂-hom (AUC = 0.821) using a naive Bayes model. Cross validated %Sn and %Sp were 86.5%, and 89.0%, respectively, and corresponded to an accuracy of 87.8% (Table 2B).

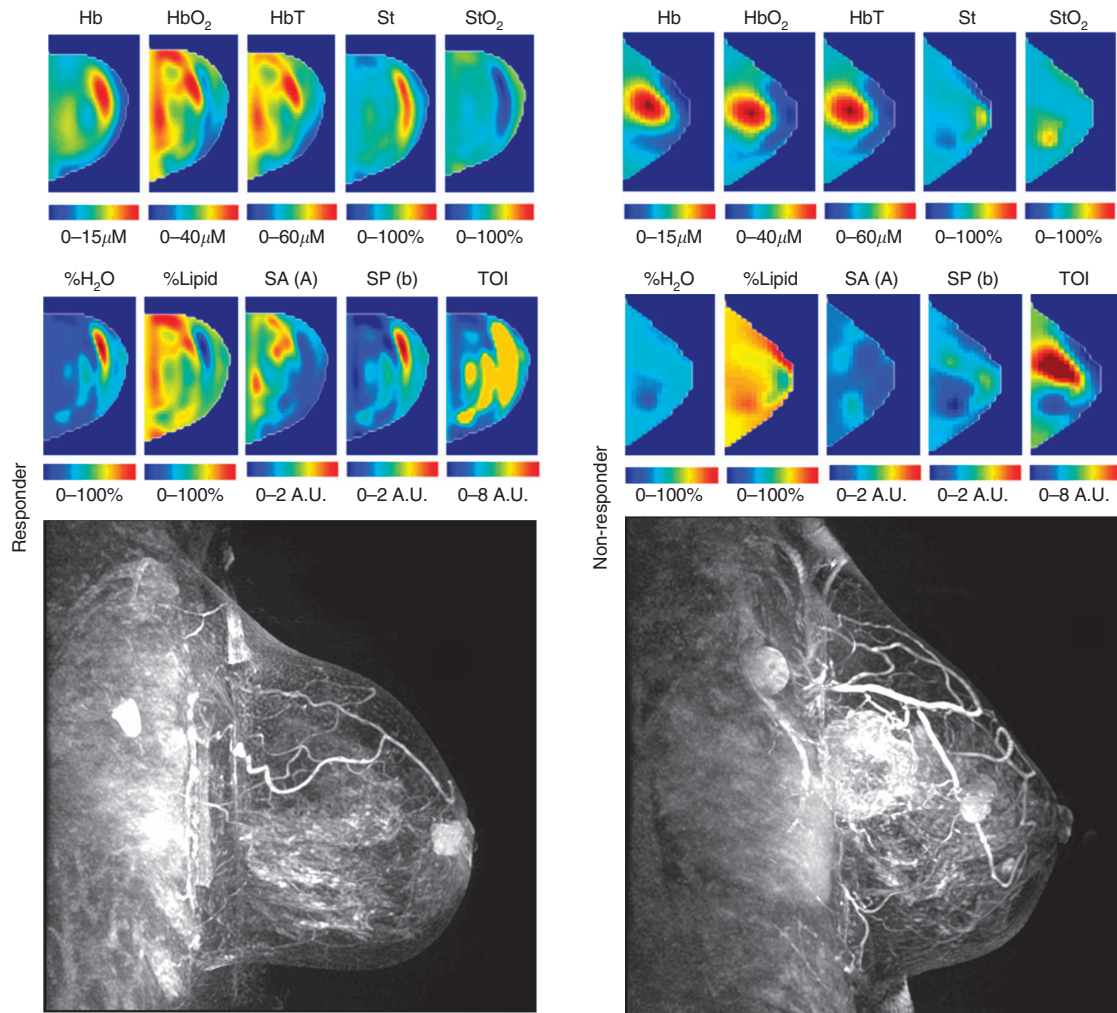


Figure 1. Representative responder vs non-responder. Representative DOS parametric maps for a responder (left column) and a non-responder (right column) are presented, and corresponding clinical contrast-enhanced magnetic resonance images of the breast. Baseline DOS images were acquired prior to starting NAC, using a tomographic diffuse optical spectroscopy device. Parametric maps were constructed volumetrically for analysis in order to calculate the GLCM texture features.

Other classifiers demonstrated a %Sn range between 70.0 and 81.0%, and a Sp of 70.0–73.0%. This corresponded to an accuracy between 70 and 77% for logistic regression and *k*-NN classifiers, respectively.

Multivariate DOS–GLCM features. Table 2C presents classification results for pairwise DOS–GLCM feature combinations. The accuracy of optimal pairwise combinations was 77.8–79.5% for classifying response groups. Using a logistic regression analysis, the combination of HbO₂-cor + Hb-hom demonstrated a %Sn of 80%, and %Sp of 78.0%. This corresponded to an AUC of 0.815, and an accuracy of 79.5%. In comparison to the naive Bayes model, the optimal pairwise combination was observed using Hb-con + HbO₂-hom, which indicated a %Sn and %Sp of 78.0% and 81.0%, respectively. The AUC for these combined parameters was 0.773, and the accuracy was 79.5% (Figure 5; Table 2C). Finally, using the *k*-NN classifier, the best pairwise combination resulted from Hb-con and HbO₂-con, which showed a %Sn and %Sp of 79.5% and 76.0%, respectively. The corresponding AUC was 0.802 and the accuracy was 77.8% (Figure 5; Table 2C).

Multiple linear regression analysis and subgroup analysis. The results of the multiple linear regression demonstrated insignificant correlations between clinical features (age, ER/PR status, Her2 status, and tumour size) and DOS–GLCM features for this patient

cohort. However, the Hb-hom, HbO₂-cor, and StO₂-con features demonstrated significant correlations to MP grading, corresponding to a regression coefficient value (*r*) of –0.358, +0.375, and –0.325, respectively (*P* < 0.05). Results of the multiple regression analysis are presented in Table 3.

Subgroup analysis showed that the HbO₂-hom feature was the best predictor in ER+ patients using a naive Bayes classifier (Table 4). For patients with triple-negative tumours, the Hb-hom was the best predictor resulting in an AUC of 0.917 (%Sn = 75.0%, 66.7%) using a *k*-NN classifier. Patients separated according to chemotherapy groups also showed variances in optimal features; FEC-D-treated patients classified into responders and non-responders with a %Sn of 100.0% and Sp of 92.3% using a logistic regression analysis for TOI-hom. Patients treated with AC-T-based chemotherapy demonstrated an AUC of 0.896 using the HbO₂-hom feature with *k*-NN classification (Table 4).

DISCUSSION AND CONCLUSION

This study examined pretreatment DOS texture analysis and machine learning techniques to statistically measure textural heterogeneity in locally advanced breast tumours. These techniques were used to predict breast cancer response to anthracycline- and

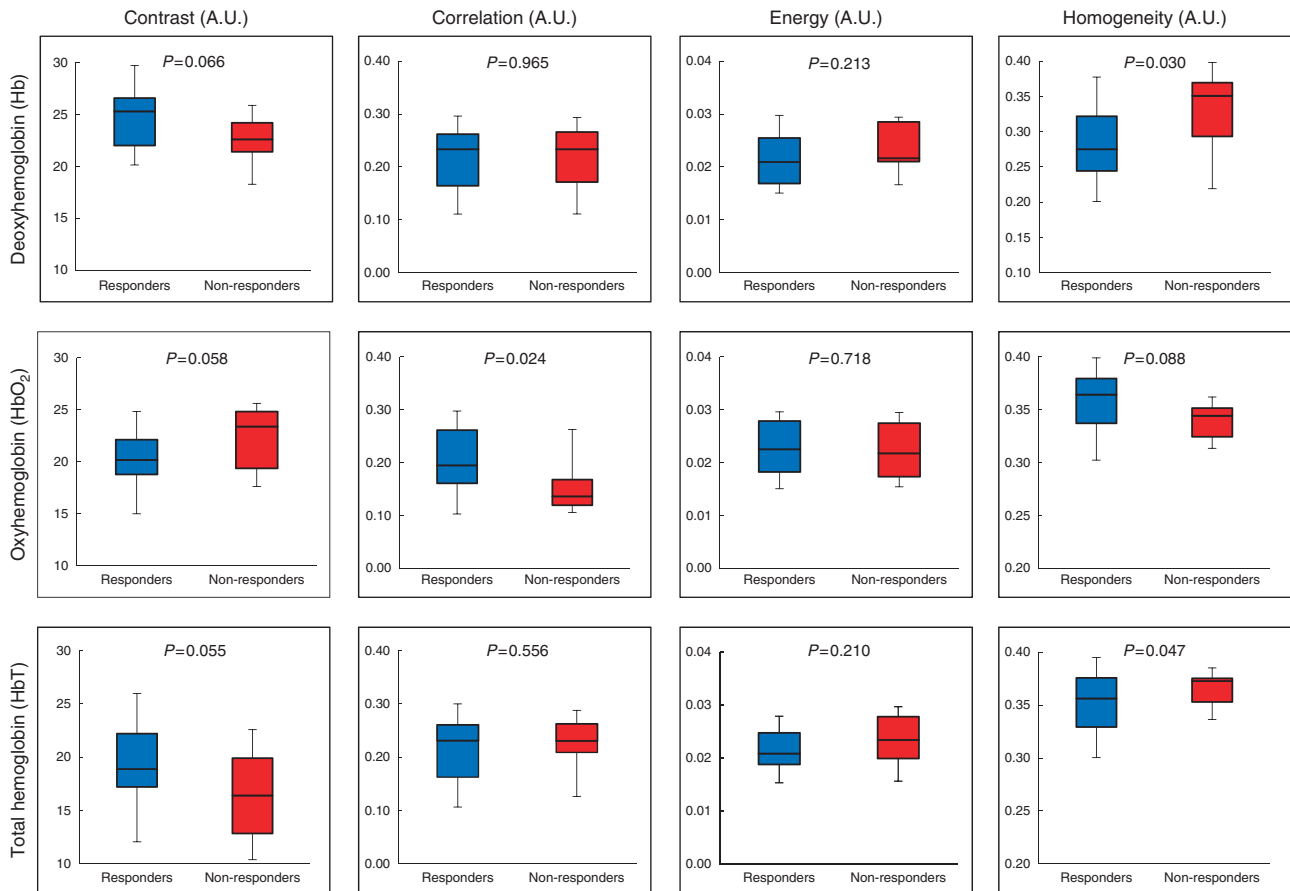


Figure 2. GLCM texture features for haemoglobin. Box-and-whisker plots showing significant differences in DOS textural markers for responders and non-responders. Haemoglobin-based features at baseline demonstrated a significant difference ($P < 0.05$) between response groups. An unpaired student *t*-test was used to test the significance for normally distributed data. *P*-values indicated.

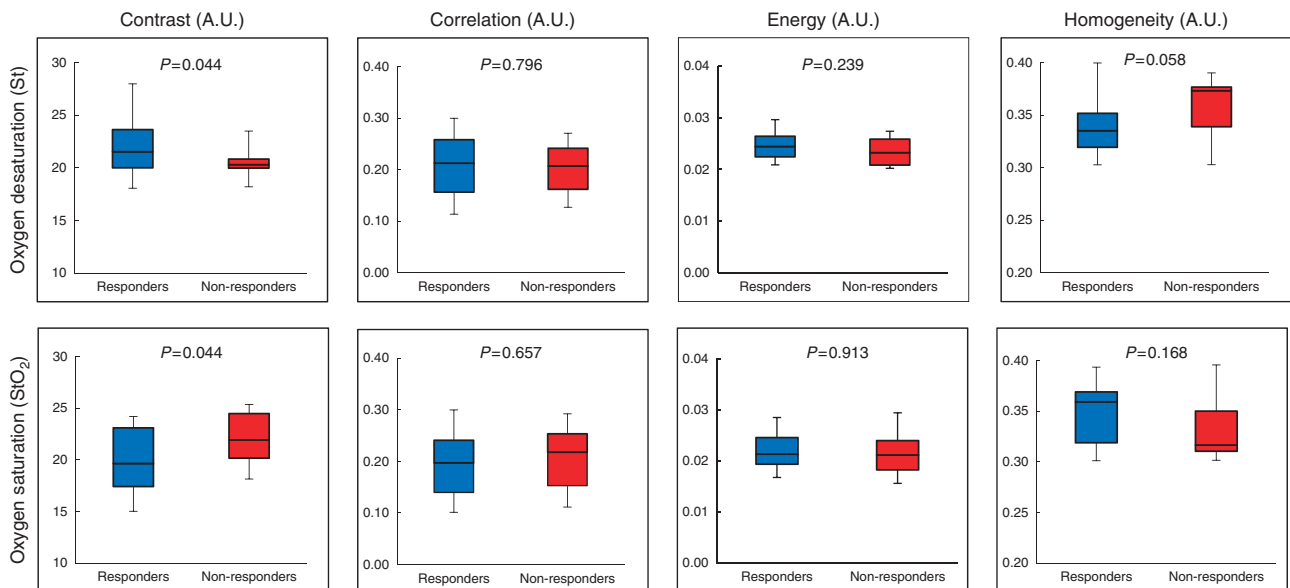


Figure 3. GLCM texture features for oxygen saturation. Box-and-whisker plots showing significant differences in DOS textural markers for responders and non-responders. Oxygen saturation parameters at baseline demonstrated a significant difference ($P < 0.05$) between response groups (unpaired student *t*-test, *P*-values indicated).

taxane-based NAC. The results demonstrate for the first time that textural heterogeneities in DOS measures of haemoglobin and oxygen content in breast tumours predict NAC response with high accuracy. Volumetric tumour analysis indicated that tumour-

specific ROI-data can yield highly sensitive and specific univariate and multivariate textural markers using various classifier models. Pretreatment DOS texture features (Hb, HbO₂, HbT, St, and StO₂) were significantly different between responders and

Table 2. Results of univariate (A, B) and multivariate analysis (C) using three classification models: logistic regression analysis, naive Bayes classifier, and k-NN

A						
Significant univariate feature	Classifier/Model	%Sn	%Sp	AUC	P-value	Statistical power (n ₂)
Hb-homogeneity	Logistic regression	60.0	60.0	0.726	0.030	71.8 (14)
	Naive Bayes	82.0	82.0	0.799		
	k-NN	61.5	67.5	0.577		
HbO ₂ -correlation	Logistic regression	70.0	70.0	0.756	0.024	78.9 (11)
	Naive Bayes	80.0	81.0	0.778		
	k-NN	66.5	74.5	0.602		
HbT-homogeneity	Logistic regression	60.0	60.0	0.657	0.047	79.9 (11)
	Naive Bayes	84.0	85.0	0.813		
	k-NN	74.0	47.0	0.552		
St-contrast	Logistic regression	60.0	63.0	0.670	0.044	73.5 (13)
	Naive Bayes	79.5	82.0	0.779		
	k-NN	70.5	64.5	0.582		
StO ₂ -contrast	Logistic regression	70.0	63.0	0.715	0.044	85.6 (enough)
	Naive Bayes	83.0	85.5	0.803		
	k-NN	70.0	66.5	0.610		
B						
Univariate features	Classifier/Model	%Sn	%Sp	%Acc		
HbO ₂ -correlation	Logistic regression	70.0	70.0	70.0		
HbO ₂ -homogeneity	Naive Bayes	86.5	89.0	87.8		
HbO ₂ -contrast	k-NN	81.0	73.0	77.0		
C						
Multivariate features	Classifier/Model	%Sn	%Sp	%Acc		
HbO ₂ -correlation + Hb-homogeneity	Logistic regression	80.0	78.0	79.5		
Hb-contrast + HbO ₂ -homogeneity	Naive Bayes	78.0	81.0	79.5		
Hb-correlation + HbO ₂ -contrast	k-NN	79.5	76.0	77.8		

Abbreviations: %Acc = accuracy; AUC = area under curve; Hb = deoxy-haemoglobin; HbO₂ = oxy-haemoglobin; HbT = total haemoglobin; k-NN = k-nearest neighbour; Sn = sensitivity; Sp = specificity; St = oxygen desaturation; StO₂ = tumour oxygen saturation. Bold values indicate best classifiers. The last column in Table 2A reports the percentage of the statistical power. The numbers inside parentheses in this column indicate the number of non-responders (n₂) required in this study to achieve a statistical power of minimum 80% in case that the number of responders (n₁) is fixed at 27.

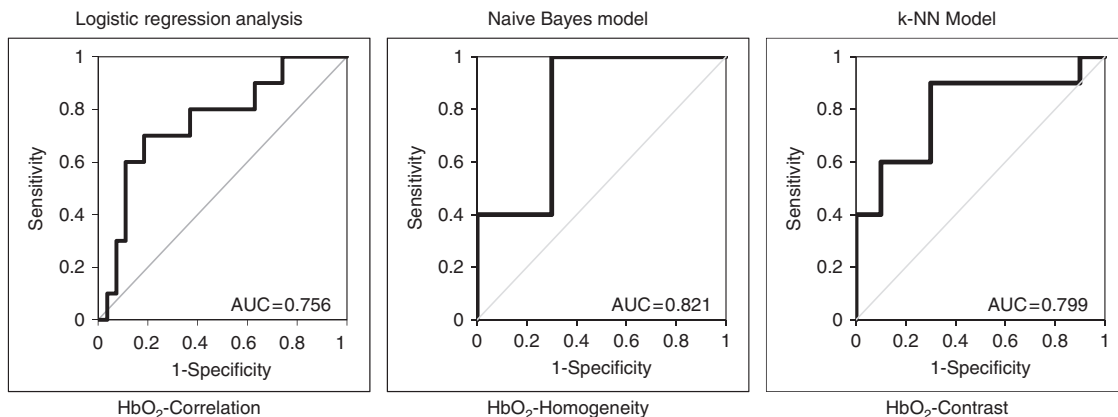


Figure 4. Receiver-operating characteristic (ROC) curves for univariate DOS texture features. ROC curves for the best performing single DOS texture parameter are presented.

non-responders. This relationship was also observed across patient subgroups separated by ER +, triple-negative, and chemotherapy treatments. Receiver-operating characteristic analysis of those features also indicated a good classification accuracy that represents heterogeneities in blood perfusion and StO₂ in breast tumours. A previous study by Ueda *et al* (2012) demonstrated that pretreatment StO₂ using the mean DOS values could predict treatment response with a %Sn and %Sp of 75.0% and 73.3%,

respectively. The study here complements those results; demonstrating that texture-based analysis of the StO₂-contrast may yield similar statistical differences between response groups ($P = 0.044$). Grey-level co-occurrence matrices analyses here, provided discriminant features by using volumetric tumour analysis, in addition to second-order statistical analyses that examined the pixel-by-pixel relationships of tumour heterogeneities within the parametric maps. Measures of spatial heterogeneity in tumour

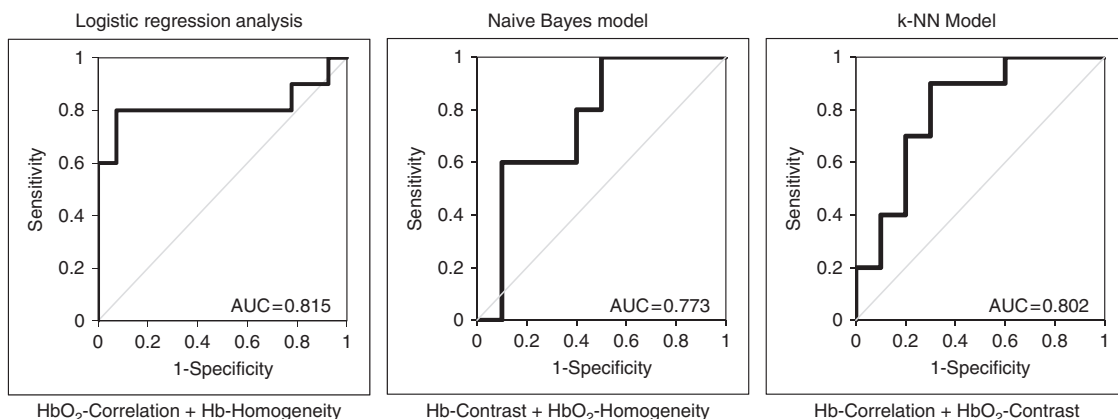


Figure 5. Receiver-operating characteristic (ROC) curves for multivariate DOS texture features. ROC curves for the best performing pairwise DOS texture parameters are presented.

Table 3. Regression coefficients (*r*) of the multiple regression analysis for DOS–GLCM features and corresponding regression

DOS–GLCM feature	Comparison	<i>r</i>	F-value	P-value
Hb-homogeneity	Age	− 0.130	0.599	0.444
	ER/PR status	− 0.087	0.267	0.608
	Her2 status	− 0.104	0.382	0.540
	Tumour size	+ 0.231	1.967	0.170
	Miller–Payne grade	− 0.358	5.137	0.030
HbO ₂ -correlation	Age	− 0.116	0.475	0.495
	ER/PR status	− 0.003	0.000	0.988
	Her2 status	− 0.109	0.418	0.522
	Tumour size	− 0.295	3.335	0.076
	Miller–Payne grade	+ 0.375	5.172	0.022
HbT-homogeneity	Age	− 0.142	0.715	0.403
	ER/PR status	+ 0.007	0.002	0.969
	Her2 status	+ 0.206	1.544	0.222
	Tumour size	+ 0.085	0.257	0.616
	Miller–Payne grade	− 0.233	2.015	0.165
St-contrast	Age	− 0.231	1.972	0.169
	ER/PR status	+ 0.056	0.111	0.741
	Her2 status	+ 0.095	0.322	0.574
	Tumour size	− 0.164	0.971	0.331
	Miller–Payne grade	+ 0.177	1.138	0.293
StO ₂ -contrast	Age	− 0.083	0.241	0.626
	ER/PR status	− 0.074	0.190	0.665
	Her2 status	− 0.213	1.661	0.206
	Tumour size	+ 0.279	2.966	0.094
	Miller–Payne grade	− 0.325	4.140	0.050

Abbreviations: DOS = diffuse optical spectroscopy; ER = oestrogen receptor; GLCM = grey-level co-occurrence matrices; Hb = deoxy-haemoglobin; HbO₂ = oxy-haemoglobin; HbT = total haemoglobin; PR = progesterone receptor; StO₂ = tumour oxygen saturation; St = oxygen desaturation. F-values are presented. Clinical features such as age, ER/PR status, Her2 status, and tumour size were not significantly correlated to DOS–GLCM features in this patient cohort. However, DOS–GLCM features such as the Hb-hom, HbO₂-cor, StO₂-con were correlated to Miller–Payne pathologic response grade. Statistically significant values are in bold.

Table 4. A subgroup analysis was completed based on ER/PR + and triple-negative tumours

Subgroup	Best feature	Model	%Sn	%Sp	AUC
ER/PR +	Hb-con	Logistic regression	76.2	66.7	0.746
	HbO ₂ -hom	Naive Bayes	93.3	90.1	0.883
	HbO ₂ -con	k-NN	85.8	82.5	0.851
Triple negative	Hb-hom	Logistic regression	100.0	33.3	0.917
	Hb-ene	Naive Bayes	100.0	66.7	0.667
	Hb-hom	k-NN	75.0	66.7	0.917
FEC-D	TOI-hom	Logistic regression	100.0	92.3	0.949
	Hb-con	Naive Bayes	60.0	81.7	0.722
	Hb-hom	k-NN	80.0	80.0	0.806
AC-T	HbO ₂ -cor	Logistic regression	100.0	71.4	0.837
	HbO ₂ -hom	Naive Bayes	96.4	90.7	0.882
	HbO ₂ -hom	k-NN	83.6	85.0	0.896

Abbreviations: AC-T = adriamycin, cyclophosphamide, taxol; AUC = area under curve; ER = oestrogen receptor; FEC-D = fluorouracil, epirubicin, cyclophosphamide, docetaxel; Hb = deoxy-haemoglobin; HbO₂ = oxy-haemoglobin; k-NN = k-nearest neighbour; PR = progesterone receptor; Sn = sensitivity; Sp = specificity; TOI = tissue optical index. Patients were also grouped according to chemotherapy type for analysis. Three classification models were used (logistic regression, naive Bayes, and k-NN) and the best predictive features are presented.

physiology as conducted here, could potentially provide good characterisation of biological traits that influence tumour response to treatment. Such features include tumour hypoxia (Hockel and Vaupel, 2001), and haematological characteristics such as blood flow and vascular density (Folkman, 2002). These features have been shown to influence tumour cell proliferation and metabolism, and therefore may also affect chemosensitivity (Folkman, 2002). The use of such measures better reflects tumour physiology, which is not homogeneous but rather spatially heterogeneous.

Additionally, multiparametric analysis resulted in sensitive and specific combined markers for response classification. Logistic regression analysis demonstrated ~10% improvement in all performance measures by using pairwise features compared to the case of using only one single feature. However, the naive Bayes and k-NN did not show a significant improvement. This may be related to the small sample size used and peaking phenomena (Jain *et al*, 2000). Features into the pairwise models included: HbO₂-cor, HbO₂-hom, Hb-con, HbO₂-con, Hb-hom, and Hb-con. Individually, those non-texture DOS parameters were previously correlated to tumour vasculature (Intes, 2005). Additionally, the heterogenic tumour vasculature has been linked to mediating drug resistance; caused by structural scaffolds that inhibit effective drug delivery (Teicher *et al*, 1990; Galmarini *et al*, 2000; Tredan *et al*, 2007). These include poor vascular flow, increased interstitial fluid, and a tightly bound cellular matrix that may constrain drugs from reaching into the tumour stroma thereby affecting the efficacy of chemotherapies.

In comparison to other studies, texture analysis of MRI (Ahmed *et al*, 2013; Golden *et al*, 2013; Teruel *et al*, 2014), ultrasound

spectroscopy (Sadeghi-Naini *et al*, 2014), and DOS (Sadeghi-Naini *et al*, 2015) images have been used to assess and monitor chemotherapy response in breast tumours during the course of treatment. Textural analysis of pretreatment MRI-based kinetic maps have indicated positive results for predicting chemotherapy response in 'triple-negative' breast tumours (Golden *et al*, 2013). Those results also strongly suggest that pretreatment tumour heterogeneity can influence drug resistance (Golden *et al*, 2013). Other similar studies have examined texture features of dynamic contrast-enhanced MRI images to predict NAC response (Ahmed *et al*, 2013; Teruel *et al*, 2014). Results have indicated significant differences in GLCM texture features between responders and non-responders at pretreatment (Ahmed *et al*, 2013) and have reported an increase in textural heterogeneity caused by necrotic tumour areas (Ahmed *et al*, 2013). Those studies demonstrated comparable frameworks to the present study. Specifically, that heterogeneous tumour features caused by pathophysiology, and initial biochemical composition might play an important role in chemoresistance.

In terms of novelty, the results indicate that selecting volumetric tumour-based ROIs may improve the method for DOS texture analysis to predict NAC response. Additionally, we compared the performance of several classification methods and found that using naive Bayes classifier demonstrated high accuracy in predicting chemotherapy treatment response. The preliminary work in this study highlights an important phase in the 'imaging biomarker roadmap' outlined by Cancer Research UK (CRUK) and the European Organisation for Research and Treatment of Cancer (EORTC) (O'Connor *et al*, 2017). Diffuse optical spectroscopy-based biomarkers have surpassed the initial translational gap outlined within this roadmap; specifically, as a useful tool in medical research (O'Connor *et al*, 2017). Several works by Cerussi *et al* (2007), Tromberg *et al* (2005), and Roblyer *et al* (2011) have made significant contributions towards crossing the second translational gap; particularly, for DOS to be used as a clinical decision-making tool (Tromberg *et al*, 2005; Cerussi *et al*, 2007; Roblyer *et al*, 2011). In order for DOS-based imaging biomarkers to be considered clinically useful in personalised medicine, further research is required to build standard operating procedures for DOS imaging biomarkers for treatment response with validated cutoff points, removing bias, setting-up standard practices for image processing techniques and ensuring comparability to pathology (O'Connor *et al*, 2017). The work here builds on determining optimal techniques for where DOS could be used as a predictive or prognostic marker. Using DOS imaging biomarkers to guide cancer therapies can potentially reduce costs to the health-care system by mitigating ineffective treatments. Therefore, DOS imaging used for routine cancer care would necessitate a cost-effectiveness analysis to quantify the added value to the health-care system (O'Connor *et al*, 2017).

Limitations in this study include a limited sample size; although the patient cohort size has been improved compared to our previous study (Sadeghi-Naini *et al*, 2015). The statistical test of significance was performed with a relatively small sample size in each responding group ($n_1 = 27$ and $n_2 = 10$) (Siegel and Castellan, 1988; de Winter, 2013). In order to evaluate the reliability of the performed tests, the statistical power (SP) was calculated using PASS¹⁴ (NCSS, LLC. Kaysville, Utah, USA). The results obtained are reported in the rightmost column in Table 2A. As can be seen from these results, the SP for the statistically significant parameters varied between 71.8 and 85.6%. This is one of the limitations of this study, as a threshold of 70% has to be considered for the SP instead of a commonly used threshold of 80%. By fixing the number of responders ($n_1 = 27$), we have also estimated the number of non-responders (n_2) required to achieve a minimum SP of 80%, which are reported inside parentheses in the rightmost column in Table 2A. As can be seen from this analysis, at most four additional

non-responders are needed to achieve an 80% SP. This change would permit stronger conclusions to be drawn from this study. We note that this study used binary response classifications (R vs NR) and that the ratio of response groups in this cohort is biased towards the responders (complete pathologic and partial responders were grouped into one category). This was due to the chosen MP cutoff point.

Within the framework of the CRUK and EORTC imaging biomarker discovery roadmap, future work would require an increase in the number of patients to strengthen the statistical conclusions. This would subsequently allow classifying patients into multiple response categories (MP 1–5). Increasing the number of patients would also permit testing for repeatability, reproducibility, and would increase precision via multicentre trials (O'Connor *et al*, 2016). Baseline DOS imaging biomarkers could also be used to predict long-term survival data for patients. Taken together, these early results suggest that there is an opportunity for DOS-GLCM analysis to improve analysis and classification of tumour response prior to the start of NAC in breast patients.

In conclusion, this study investigated baseline DOS functional maps using GLCM texture analysis to classify patient response to NAC. In comparison to our previous study where we analysed DOS textures of the whole breast; here, we analysed the DOS texture features within the tumour region only (Sadeghi-Naini *et al*, 2015). The results indicated that such indices can differentiate between response groups prior to the start of treatment with high %Sn and %Sp. Although further studies are required, this first report demonstrates promising potential for DOS-based textural parameters to evaluate baseline tumour vascular heterogeneity, and subsequently as markers for response to chemotherapy. The use of these markers may help guide treatments to personalise patient-care plans by potentially predicting chemoresponse. Ultimately if used to guide therapy, DOS-based texture analysis may help improve breast cancer therapeutics, and may potentially improve overall disease-free survival.

ACKNOWLEDGEMENTS

This project was funded by the Terry Fox Research Institute, Canada. We thank Professor Heidi Probst from Sheffield Hallam University (Sheffield, United Kingdom).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Ahmed A, Gibbs P, Pickles M, Turnbull L (2013) Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging* **38**: 89–101.
- Cance WG, Carey LA, Calvo BF, Sartor C, Sawyer L, Moore DT, Rosenman J, Ollila DW, Graham 2nd M (2002) Long-term outcome of neoadjuvant therapy for locally advanced breast carcinoma: effective clinical downstaging allows breast preservation and predicts outstanding local control and survival. *Ann Surg* **236**: 295–302discussion 302-3.
- Castellano G, Bonilha L, Li LM, Cendes F (2004) Texture analysis of medical images. *Clin Radiol* **59**: 1061–1069.
- Cerussi A, Hsiang D, Shah N, Mehta R, Durkin A, Butler J, Tromberg BJ (2007) Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy. *Proc Natl Acad Sci USA* **104**: 4014–4019.
- Cerussi A, Shah N, Hsiang D, Durkin A, Butler J, Tromberg BJ (2006) *In vivo* absorption, scattering, and physiologic properties of 58 malignant breast

- tumors determined by broadband diffuse optical spectroscopy. *J Biomed Opt* **11**: 044005.
- Cerussi AE, Tanamai VW, Hsiang D, Butler J, Mehta RS, Tromberg BJ (2011) Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy. *Philos Trans A Math Phys Eng Sci* **369**: 4512–4530.
- Chen W, Giger ML, Li H, Bick U, Newstead GM (2007) Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med* **58**: 562–571.
- Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, COOK GJ (2013) Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging* **40**: 133–140.
- de Winter JCF (2013) Using the Student's *t*-test with extremely small sample sizes. *Pract Assess Res Eval* **18**: 1–10.
- Desantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS, Jemal A (2014) Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* **64**: 252–271.
- Duda RO, Hart PE, Stork DG (2001) *Pattern Classification*. Wiley: New York.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancy J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, SHANKAR L, DODD L, KAPLAN R, Lacombe D, Verweij J (2009) New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**: 228–247.
- Evans A, Armstrong S, Whelehan P, Thomson K, Rauchhaus P, Purdie C, Jordan L, Jones L, Thompson A, Vinnicombe S (2013) Can shear-wave elastography predict response to neoadjuvant chemotherapy in women with invasive breast cancer? *Br J Cancer* **109**: 2798–2802.
- Folkman J (2002) Role of angiogenesis in tumor growth and metastasis. *Semin Oncol* **29**: 15–18.
- Galmarini FC, Galmarini CM, Sarchi MI, Abulafia J, Galmarini D (2000) Heterogeneous distribution of tumor blood supply affects the response to chemotherapy in patients with head and neck cancer. *Microcirculation* **7**: 405–410.
- Gibbs P, Turnbull LW (2003) Textural analysis of contrast-enhanced MR images of the breast. *Magn Reson Med* **50**: 92–98.
- Giordano SH (2003) Update on locally advanced breast cancer. *Oncologist* **8**: 521–530.
- Golden DI, Lipson JA, Telli ML, Ford JM, Rubin DL (2013) Dynamic contrast-enhanced MRI-based biomarkers of therapeutic response in triple-negative breast cancer. *J Am Med Assoc* **309**: 1059–1066.
- Gupta R, Undrill PE (1995) The use of texture analysis to delineate suspicious masses in mammography. *Phys Med Biol* **40**: 835–855.
- Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans Syst Man Cybern* **3**: 610–621.
- Hockel M, Vaupel P (2001) Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J Natl Cancer Inst* **93**: 266–276.
- Intes X (2005) Time-domain optical mammography SoftScan: initial results. *Acad Radiol* **12**: 934–947.
- Jain AK, Duin RPW, Mao J (2000) Statistical pattern recognition: a review. *IEEE Trans Pattern Anal Mach Intell* **22**: 4–37.
- Jakubowski DB, Cerussi AE, Bevilacqua F, Shah N, Hsiang D, Butler J, Tromberg BJ (2004) Monitoring neoadjuvant chemotherapy in breast cancer using quantitative diffuse optical spectroscopy: a case study. *J Biomed Opt* **9**: 230–238.
- Jiang S, Pogue BW, Kaufman PA, Gui J, Jermyn M, Frazee TE, Poplack SP, Diflorio-Alexander R, Wells WA, Paulsen KD (2014) Predicting breast tumor response to neoadjuvant chemotherapy with diffuse optical spectroscopic tomography prior to treatment. *Clin Cancer Res* **20**: 6006–6015.
- Lee MC, Newman LA (2007) Management of patients with locally advanced breast cancer. *Surg Clin North Am* **87**: 379–398ix.
- Lerski RA, Straughan K, Schad LR, Boyce D, Blum S, Zuna I (1993) MR image texture analysis—an approach to tissue characterization. *Magn Reson Imaging* **11**: 873–887.
- Li H, Giger ML, Olopade OI, Margolis A, Lan L, Chinander MR (2005) Computerized texture analysis of mammographic parenchymal patterns of digitized mammograms. *Acad Radiol* **12**: 863–873.
- Mathew J, Asgeirsson KS, Cheung KL, Chan S, Dahda A, Robertson JF (2009) Neoadjuvant chemotherapy for locally advanced breast cancer: a review of the literature and future directions. *Eur J Surg Oncol* **35**: 113–122.
- Mourant JR, Canpolat M, Brocker C, Esponda-Ramos O, Johnson TM, Matanock A, Stetter K, Freyer JP (2000) Light scattering from cells: the contribution of the nucleus and the effects of proliferative status. *J Biomed Opt* **5**: 131–137.
- National Institute For Health And Clinical Excellence (NICE) (2009) *Early and Locally Advanced Breast Cancer: Diagnosis and Treatment*. NICE: Cardiff, UK.
- O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, Boellaard R, Bohndiek SE, Brady M, Brown G, Buckley DL, Chenevert TL, Clarke LP, Collette S, Cook GJ, Desouza NM, Dickson JC, Dive C, Evelhoch JL, Faivre-Finn C, Gallagher FA, Gilbert FJ, Gillies RJ, Goh V, Griffiths JR, Groves AM, Halligan S, Harris AL, Hawkes DJ, Hoekstra OS, Huang EP, Hutton BF, Jackson EF, Jayson GC, Jones A, Koh DM, Lacombe D, Lambin P, Lassau N, Leach MO, Lee TY, Leen EL, Lewis JS, Liu Y, Lythgoe MF, Manoharan P, Maxwell RJ, Miles KA, Morgan B, Morris S, Ng T, Padhani AR, Parker GJ, Partridge M, Pathak AP, Peet AC, Punwani S, Reynolds AR, Robinson SP, Shankar LK, Sharma RA, Soloviev D, Stroobants S, Sullivan DC, Taylor SA, Tofts PS, Tozer GM, Van Herk M, Walker-Samuel S, Wason J, Williams KJ, Workman P, Yankeelov TE, Brindle KM, Mcshane LM, Jackson A, Waterton JC (2017) Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* **14**: 169–186.
- Orgston KN, Miller ID, Payne S, Hutcheon AW, Sarkar TK, Smith I, Schofield A, Heys SD (2003) A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast* **12**: 320–327.
- Provenzano E, Bossuyt V, Viale G, Cameron D, Badve S, Denkert C, Macgrogan G, Penault-Llorca F, Boughey J, Curigliano G, Dixon JM, Esserman L, Fastner G, Kuehn T, Peintinger F, von Minckwitz G, White J, Yang W, Symmans WF, Residual Disease Characterization Working Group Of The Breast International Group-North American Breast Cancer Group, C (2015) Standardization of pathologic evaluation and reporting of postneoadjuvant specimens in clinical trials of breast cancer: recommendations from an International Working Group. *Mod Pathol* **28**: 1185–1201.
- Roblyer D, Ueda S, Cerussi A, Tanamai W, Durkin A, Mehta R, Hsiang D, Butler JA, McLaren C, Chen WP, Tromberg B (2011) Optical imaging of breast cancer oxyhemoglobin flare correlates with neoadjuvant chemotherapy response one day after starting treatment. *Proc Natl Acad Sci USA* **108**: 14626–14631.
- Sadeghi-Naini A, Sannachi L, Pritchard K, Trudeau M, Gandhi S, Wright FC, Zubovits J, Yaffe MJ, Kolios MC, Czarnota GJ (2014) Early prediction of therapy responses and outcomes in breast cancer patients using quantitative ultrasound spectral texture. *Oncotarget* **5**: 3497–3511.
- Sadeghi-Naini A, Vorauer E, Chin L, Falou O, Tran WT, Wright FC, Gandhi S, Yaffe MJ, Czarnota GJ (2015) Early detection of chemotherapy-refractory patients by monitoring textural alterations in diffuse optical spectroscopic images. *Med Phys* **42**: 6130–6146.
- Schaafsma BE, Van De Giessen M, Charehbil A, Smit VT, Kroep JR, Lelieveldt BP, Liefers GJ, Chan A, Lowik CW, Dijkstra J, Van De Velde CJ, Wasser MN, Vahrmeijer AL (2015) Optical mammography using diffuse optical spectroscopy for monitoring tumor response to neoadjuvant chemotherapy in women with locally advanced breast cancer. *Clin Cancer Res* **21**: 577–584.
- Siegel S, Castellan NJ (1988) *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill: New York.
- Teicher BA, Herman TS, Holden SA, Wang YY, Pfeffer MR, Crawford JW, Frei 3rd E (1990) Tumor resistance to alkylating agents conferred by mechanisms operative only *in vivo*. *Science* **247**: 1457–1461.
- Teruel JR, Heldahl MG, Goa PE, Pickles M, Lundgren S, Bathen TF, Gibbs P (2014) Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *NMR Biomed* **27**: 887–896.
- Tredan O, Galmarini CM, Patel K, Tannock IF (2007) Drug resistance and the solid tumor microenvironment. *J Natl Cancer Inst* **99**: 1441–1454.
- Tromberg BJ, Cerussi A, Shah N, Compton M, Durkin A, Hsiang D, Butler J, Mehta R (2005) Imaging in breast cancer: diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy. *Breast Cancer Res* **7**: 279–285.
- Ueda S, Roblyer D, Cerussi A, Durkin A, Leproux A, Santoro Y, Xu S, O'Sullivan TD, Hsiang D, Mehta R, Butler J, Tromberg BJ (2012) Baseline tumor oxygen saturation correlates with a pathologic complete response in breast cancer patients undergoing neoadjuvant chemotherapy. *Cancer Res* **72**: 4318–4328.

- von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, Gerber B, Eiermann W, Hilfrich J, Huober J, Jackisch C, Kaufmann M, Konecny GE, Denkert C, Nekljudova V, Mehta K, Loibl S (2012) Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* **30**: 1796–1804.
- Whitman GJ, Strom EA (2009) Workup and staging of locally advanced breast cancer. *Semin Radiat Oncol* **19**: 211–221.
- Xu RX, Povoski SP (2007) Diffuse optical imaging and spectroscopy for cancer. *Expert Rev Med Devices* **4**: 83–95.
- Yang X, Tridandapani S, Beitler JJ, Yu DS, Yoshida EJ, Curran WJ, Liu T (2012) Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an *in vivo* study of late toxicity. *Med Phys* **39**: 5732–5739.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)