






# Monitoring Breast Cancer Response to Neoadjuvant Chemotherapy Using Probability Maps Derived from Quantitative Ultrasound Parametric Images

Piotr Karwat , Hanna Piotrkowska-Wróblewska , Ziemowit Klimonda ,  
Katarzyna S. Dobruch-Sobczak , and Jerzy Litniewski 

## I. INTRODUCTION

**Abstract—Objective:** Neoadjuvant chemotherapy (NAC) is widely used in the treatment of breast cancer. However, to date, there are no fully reliable, non-invasive methods for monitoring NAC. In this article, we propose a new method for classifying NAC-responsive and unresponsive tumors using quantitative ultrasound. **Methods:** The study used ultrasound data collected from breast tumors treated with NAC. The proposed method is based on the hypothesis that areas that characterize the effect of therapy particularly well can be found. For this purpose, parametric images of texture features calculated from tumor images were converted into NAC response probability maps, and areas with a probability above 0.5 were used for classification. **Results:** The results obtained after the third cycle of NAC show that the classification of tumors using the traditional method (area under the ROC curve  $AUC = 0.81\text{--}0.88$ ) can be significantly improved thanks to the proposed new approach ( $AUC = 0.84\text{--}0.94$ ). This improvement is achieved over a wide range of cutoff values (0.2–0.7), and the probability maps obtained from different quantitative parameters correlate well. **Conclusion:** The results suggest that there are tumor areas that are particularly well suited to assessing response to NAC. **Significance:** The proposed approach to monitoring the effects of NAC not only leads to a better classification of responses, but also may contribute to a better understanding of the microstructure of neoplastic tumors observed in an ultrasound examination.

**Index Terms—**Breast cancer, neoadjuvant chemotherapy, quantitative ultrasound, treatment monitoring.

Manuscript received 20 February 2024; accepted 22 March 2024. Date of publication 1 April 2024; date of current version 22 August 2024. This work was supported in part by the National Science Centre of Poland under Grant 2019/35/B/ST7/03792 and in part by the National Centre for Research and Development under Grant LIDER/24/0088/L-11/19/NCBR/2020. (Corresponding author: Piotr Karwat.)

Piotr Karwat is with the Ultrasound Department, Institute of Fundamental Technological Research, Polish Academy of Sciences 02-106 Warsaw, Poland (e-mail: pkarwat@ippt.pan.pl).

Hanna Piotrkowska-Wróblewska, Ziemowit Klimonda, and Jerzy Litniewski are with the Ultrasound Department, Institute of Fundamental Technological Research, Polish Academy of Sciences, Poland.

Katarzyna S. Dobruch-Sobczak is with the Ultrasound Department, Institute of Fundamental Technological Research, Polish Academy of Sciences, Poland, and also with the Radiology Department II, Maria Skłodowska-Curie National Research Institute of Oncology, Poland.

Digital Object Identifier 10.1109/TBME.2024.3383920

PREOPERATIVE chemotherapy (neoadjuvant chemotherapy - NAC), introduced in 1970, was initially used for locally advanced breast cancer (LABC) and inflammatory breast cancer (BC). The aim was to reduce the size of the tumor and, as a result, limit the scope of surgical treatment of both the breast tumor and the axillary lymph nodes.

Currently, NAC is also recommended in the early stage of BC, in the following subtypes: triple negative breast cancer (TNBC), with the presence of HER-2+ receptors (Luminal B HER2-positive and HER-positive non-Luminal subtype) and in cases of Luminal B HER2-negative cancer with low expression of hormone receptors, high grade of malignancy (G3) and in patients at a young age (up to 35 years old) in stage II or III [1].

As with other cancers, pathological response to NAC treatment, especially pathological complete response (pCR), has been considered a surrogate for favorable overall survival, event-free survival and long-term survival for TNBC and HER2+ subtypes [2].

Unfortunately, the assessment of response to NAC based on methods commonly used in clinical practice (ultrasound, mammography, magnetic resonance) is not sufficiently accurate and generates false positive and false negative results [3], [4], [5]. Complete tumor regression, confirmed histopathologically as pCR, occurs in an average of 19% of patients and is highly dependent on the immunohistochemical subtype of the tumor [6], [7]. In patients whose BC remains insensitive to NAC, i.e., in approximately 20–30% of patients [6], [7], [8], chemotherapy delays necessary surgery, increases the risk of metastases, and may contribute to side effects.

Monitoring of BC during treatment with NAC in-vivo provides information regarding the sensitivity of the cancer to therapy. Currently, most of the methods used are based on the assessment of changes in tumor size estimated on the basis of imaging studies. Compared to MMG (mammography), contrast-enhanced MMG, or US (ultrasonography), the most accurate method of treatment monitoring is MRI (magnetic resonance imaging) [6].

However, the assessment of the size of tumors during treatment, both in MRI and US, in accordance with RECIST 1.1

(response evaluation criteria in solid tumors) [3], is not a sufficiently sensitive feature due to the appearance of necrotic lesions, which are a good response to treatment, but may mask a decrease in tumor dimension [6]. Similarly, MRI assessment of tumor vascularity provides false positive and false negative results. Another disadvantage of methods related to the assessment based on tumor size is the long period of time that elapses from the onset of NAC to the apparent change in tumor size. Thus, a rapid assessment of BC during NAC based on a non-invasive method would be particularly useful in this group of patients.

Quantitative ultrasound (QUS) allows the assessment of tissue structure and its scattering properties at the cellular level. This is important in monitoring the effectiveness of NAC because changes in tumor size caused by the use of therapy appear after many weeks [9], whereas changes at the cellular level after the first dose of the drug [10].

Quantitative ultrasound methods have proven to be very useful in the classification of breast cancers and in the monitoring of chemotherapy. A number of publications show the results showing the relationship between the parameters determined by quantitative ultrasound and the pathological response of the tumor to NAC. Papers [11], [12] describe parametric images generated with the use of spectral and scattering parameters of signals received from the tumor, from which texture features were then determined and included in the multi-parameter model used to predict the pathological response of the tumor.

Changes in the amplitude distribution of the ultrasonic signal scattered in the tumor after successive cycles of chemotherapy were also studied. The distributions were differentiated by the Kullback–Leibler divergence determined with respect to the distributions after the first NAC cycle [13] or with respect to the reference phantom [14]. The highest AUC (area under the receiver operating characteristic curve) was achieved after the 3rd NAC cycle, 0.92 and 0.91 respectively.

The usefulness of parameters determined from the power spectrum of backscattered ultrasound in a breast tumor, such as the integrated backscatter coefficient (IBC), average scatterer diameter (ASD), and average acoustic concentration (AAC), was also analyzed in the assessment of NAC. Parameters were determined from data from 30 LABC patients collected before and after subsequent doses of NAC and compared with tumor pathological response [15]. At week 8 NAC, the composite parameter (IBC, ASD, AAC) predicted no tumor response with sensitivity = 80%, specificity = 100% and accuracy = 85%.

The parameters of statistical distributions, which describe the amplitude distributions of scattered signals, were also evaluated. It has been shown that the addition of the shape parameter of the homodyne K-distribution to the IBC classifier is beneficial in predicting the NAC result [16], and their assessment is most accurate after the 3rd NAC course (AUC = 0.91).

In addition to examining QUS features in the tumor mass, as has been done in most studies, QUS parameters in the surrounding tumor tissue were also analyzed. In the paper [17], the authors demonstrated the great usefulness of the analysis

of tissue extending 3–10 mm from the focal lesion visible in ultrasound.

The effectiveness of multi-parametric QUS imaging in predicting breast tumor response to chemotherapy before treatment was also studied. Features were extracted from segmented areas within the tumor and tumor margin in various parametric images. The results showed that prior to treatment, patient response could be predicted with an accuracy of 85.4% and AUC = 0.89 [18].

Despite the high efficiency noted in studies using QUS, these techniques are still not sufficiently sensitive in predicting response to treatment. Limitations include, among others, small groups of tested tumors and their diversity in each of the published research results. Each new method that effectively predicts the response to NAC is valuable because it validates the usefulness of using quantitative ultrasonic methods to monitor chemotherapy for breast tumors.

Based on microscopic evaluation, it is known that the distribution of tumor cells and cell clusters is not homogeneous [19]. Cancer cells may be solitary or may form small groups or multicellular structures. Malignant breast tumors are characterized by various morphological structures, solid, tubular, follicular and trabecular structures. The morphological structures of the tumor consist of a different number of tumor cells, and they also differ in their distribution. It has been shown, for example, that alveolar structures had about 30 neoplastic cells, whereas tubular structures contained single rows of neoplastic cells, and trabecular structures had only one or two such rows. The largest clusters of tumor cells, up to hundreds of cells, could be observed in solid structures [20], [21].

Chemotherapy destroys cancer cells, so it can be assumed that its effects are particularly visible in areas of the tumor with a large number of such cells. Then, the observable signs of NAC therapy would not be evenly distributed throughout the tumor. The methods used to monitor the effects of NAC therapy based on quantitative ultrasound most often use the average values of tissue characteristics determined for the entire tumor or the tumor and its surroundings. If the assumption of heterogeneous distribution of the NAC effects is correct, then such averaging approach may not be optimal.

In this study, we hypothesize that tumor areas can be found that characterize the effects of NAC particularly well. Therefore, we propose a new approach that involves limiting the averaging to these specific areas. As a validation of this approach, we present results for selected QUS parameters that classify well using data from the entire tumor, and even better when only data from selected tumor regions are included.

We based our research on the images of texture features obtained from the tumor amplitude images. When searching for tumors that did not respond to NAC, we only used data collected from tumor areas pre-qualified as having a high probability of poor response. This was possible thanks to the proposed new type of parametric image processing, which results in the probability distribution of non-response to NAC in the tumor. Additionally, we present the classification results obtained directly using parametric texture images for comparison.

TABLE I  
PATIENTS AND TUMORS CHARACTERISTICS

Characteristics	Count/Value	
	before NAC	after 3rd NAC course
<b>Patients</b>		
Number of patients	40	38
Mean age (years)	56.6	56.0
Age range (years)	32-83	32-83
<b>Tumor</b>		
Number of tumors	51	48
Size (max diameter)	22,55 ( $\pm 9,89$ )	16,12 ( $\pm 8,23$ )
Multifocal	10	10
<b>Tumor histology</b>		
Invasive ductal carcinoma (IDC)	51	48
IDC with ductal carcinoma in situ	20	19
<b>Receptor Status</b>		
Luminal A	7	6
Luminal B	24	22
TNBC	9	9
HER 2+	11	11

Characteristics of the group of patients and the results of histopathological evaluation before the start of chemotherapy and after the third NAC course.

## II. METHODS

### A. Patients

The study included patients of the National Institute of Oncology, Scientific Centre in Warsaw, who were diagnosed with breast cancer and were referred to treatment in the form of neoadjuvant chemotherapy. Data from subsequent patients have been collected over the last few years. The latest version of our database was used in [22] and is also the basis for the current study. The study involved 40 patients aged 32 to 83 (average age 47). In accordance with the established protocol, patients were included in the study whose tumor size was larger than 5 mm and did not exceed 40 mm, and the number of multifocal lesions in one patient did not exceed three. These were the only criteria and patients meeting them were enrolled in the study sequentially. Ten patients were diagnosed with multifocal cancer. A total of 51 tumors were monitored. After the 3rd cycle of NAC, the number of tumors decreased to 48 (in 1 case the tumor regressed completely, in 2 cases data were not recorded for random reasons).

All study participants gave voluntary consent to participate in the study and signed the appropriate declaration. The research was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the Bioethics Committee (project identification code 49/2018).

Before qualifying for NAC, all patients underwent a core-needle biopsy (14GA biopsy needle) after administration of anesthesia in the form of 2% lidocaine. Three to five cores were taken during the biopsy. Based on the obtained tissue material, a pathologist with over 25 years of experience in the histopathological evaluation of focal breast lesions determined the type of cancer (molecular subtype and grade of malignancy, Table I).

Neoadjuvant chemotherapy was administered according to international guidelines. Treatment with doxorubicin and cyclophosphamide was given from the first to the fourth course. Thereafter, treatment with taxol was continued, and in HER2+ positive patients, trastuzumab plus taxol. During the interview, it was established that one patient was treated with doxorubicin and docetaxel, 5 years before starting the current therapy. The frequency of drug administration was selected individually for each patient. In most patients, the interval between particular NAC courses was 3 weeks. Two-week intervals were used in 2 patients.

After the treatment was completed, the patients underwent surgery. In 38 cases, mastectomy was performed, in two patients breast-conserving therapy was applied.

### B. Histopathology

The removed residual tumors were subjected to postoperative histopathological evaluation by the same pathologist who examined the biopsy samples. The residual malignant cells (RMC) parameter was used as an indicator of histopathological response to NAC treatment. This parameter is one of the analyzed elements of the residual cancer burden (RCB) scale, which determines the level of residual disease after completion of neoadjuvant chemotherapy. The RMC parameter ranges from 0% to 100%, where 0% represents a complete histopathological response (no tumor cells after treatment) and 100% represents a complete lack of response to treatment. In the group of patients participating in the study, 11 patients achieved a complete histopathological response (RMC = 0%) and 7 patients did not respond to treatment at all (RMC = 100%).

### C. Ultrasonic Data Registration

The patients' ultrasound data were recorded using the Ultrasonix SonixTOUCH ultrasound scanner (Ultrasonix Medical Corporation, Richmond, BC, Canada) and the L14-5/38 linear transducer with a center frequency of 7.2 MHz. The scanner, in addition to the standard functions of B-mode imaging, color Doppler and elastography, also had a research interface that enabled the recording of post-beamformed RF echoes.

Ultrasound evaluation of the tumors was performed according to the guidelines of the American College of Radiology (BI-RADS-lexicon) and the standards of the Polish Ultrasound Society. Recording of data from each patient took place before the start of treatment and one week after each course of chemotherapy. During the test, data from four scan-planes were recorded: radial, radial + 45°, anti-radial, and anti-radial + 45°. The doctor tried to reproduce a similar position of the transducer during each subsequent examination, based on visual assessment of tissue structures on images acquired after previous NAC course. All examinations were performed using the same scanner preset, with the transmit frequency set to 10 MHz. Measurement protocol allowed the doctor to adjust the focal depth to the location of the tumor. No other parameters affecting the collected data could be modified. In order to avoid scanner-embedded image enhancement that would affect our results, we acquired the raw B-mode images from the post-beamformed RF data.

Then, the radiologist indicated on the B-mode images the areas of the tumor from which data were collected for quantification.

#### D. Ultrasonic Data Preprocessing

B-mode images were obtained through a processing of the post-beamformed RF lines. First, the signal envelope was detected as an absolute value of the complex signal obtained through Hilbert transform. Then, the resulting envelope amplitude was log-compressed. As the scanning was performed at 4 lines per probe pitch (0.3048 mm), the horizontal pixel size was 0.076 mm. For the sampling rate of 40 Msps and the speed of sound of 1540 m/s the vertical pixel size would be 0.0192 mm. To equalize the vertical and horizontal pixel size, the images were decimated vertically by a factor of 4, so that the pixel size was 0.076 mm  $\times$  0.077 mm (width  $\times$  height). The dynamic range of the images was limited to 0–128 dB and was scaled to 256 gray levels (the final images were saved in an 8-bit format).

#### E. Quantitative Ultrasound Maps

The parametric images (maps of the parameters values) of tumors were generated using the sliding window technique. Each pixel in the parametric map was estimated based on the 4 mm  $\times$  4 mm (53  $\times$  52 pixels) block of B-mode data centered around the pixel. A set of textural features of the US images were extracted including autocovariance coefficient [23], features of the gray level co-occurrence matrix (GLCM) [24], [25], and Law's texture energy measures [26]. To extract features efficiently in the sliding window mode, we developed a dedicated library in Matlab 2021b (MathWorks, Inc., Natick, MA). The library was validated using the BUSAT Toolbox [27] as a reference.

#### F. Autocovariance

The autocovariance coefficient was already used in the context of breast tumor classification in [23]. The authors justify its use in place of the auto-correlation coefficient with the fact that the resulting predictors are unaffected by the overall brightness of an input image. The autocovariance coefficient  $acov$  is defined as follows:

$$acov = \frac{A(\Delta x, \Delta y)}{A(0, 0)}, \quad (1)$$

where

$$A(\Delta x, \Delta y) = \frac{1}{(X - \Delta x)(Y - \Delta y)} \times \sum_{x=1}^{X-\Delta x} \sum_{y=1}^{Y-\Delta y} \{ [f(x, y) - \bar{f}] \cdot [f(x + \Delta x, y + \Delta y) - \bar{f}] \}. \quad (2)$$

$X$  and  $Y$  in the above formulas denote the sliding window's horizontal and vertical size in pixels,  $x$  and  $y$  are the corresponding pixel indices within the window, and  $\Delta x$  and  $\Delta y$  stand for the horizontal and vertical offsets, respectively. The amplitudes (gray levels) are denoted as  $f$  whereas  $\bar{f}$  is the mean amplitude in the window.

TABLE II  
FILTERS FOR LAW'S TEXTURE ENERGY MEASURES

Low-pass (L5):	1	4	6	4	1
Edges (E5):	-1	-2	0	2	1
Spots (S5):	-1	0	2	0	-1
Waves (W5):	-1	2	0	-2	1
Ripples (R5):	1	-4	6	-4	1

In our study we calculated the autocovariance coefficients for  $\Delta x$  and  $\Delta y$  ranging from 0 to 5 pixels, for all  $\{\Delta x, \Delta y\}$  combinations except  $\{0, 0\}$ , for a total of 35 parameters. All combinations provided strongly correlated parameter maps, thus in this paper we present the results only for  $\{0, 1\}$  pair that leads to the best predictor.

#### G. Laws Texture Energy

Texture energy measures proposed by Laws [26] are the energies of an image after its proper filtration (convolution) with a set of center-weighted vectors. For 5-tap filters, the vectors are shown in Table II. The image is filtered vertically and horizontally by a certain pair of vectors. To make the texture energy measures rotationally invariant, images filtered with the same vectors but for different directions (e.g.,  $\{S5, L5\}$  and  $\{L5, S5\}$ ) are compounded. Next, to make the results independent of the overall brightness of the original image, each filtered image is pixel-wise divided by an image filtered with  $\{L5, L5\}$  vector pair. After this normalization the energy  $ener$  of the resulting image  $g(x, y)$  can be calculated as follows:

$$ener = \frac{\sum_{x,y} g(x, y)^2}{XY}. \quad (3)$$

In our study we calculated energies for all possible vector pairs except the redundant ones (e.g.,  $\{L5, S5\}$  is already included in  $\{S5, L5\}$ ) and also excluding the  $\{L5, L5\}$  used for normalization. This resulted in a total of 14 parameters. In this paper we present the results for three parameters that provided the best predictors. These were the energies of images filtered with vector pairs:  $\{S5, L5\}$ ,  $\{W5, L5\}$ , and  $\{R5, L5\}$ .

#### H. GLCM Features

GLCM parameters are often used in research on the classification of breast tumors [28], [29] and the prediction and monitoring of their response to treatment [11], [12], [17], [30], [31]. GLCM is a matrix of probabilities that a pair of certain gray tones occurs in a pair of pixels being in a particular relative spatial position [24], [25]. In our study the GLCM was calculated for 64 gray level intervals (from 0 to 255 with step of 4), resulting in 64  $\times$  64 GLCM size. The considered spatial relationship was a vertical or horizontal displacement by 4 pixels (0.3 mm). Based on the GLCM a number of texture parameters can be calculated. In our study, we analyzed contrast, correlation, energy, homogeneity, and variance. In this paper we present the results for the variance ( $var$ ) parameter as it provided good classification results in the classical approach and showed clear improvement when the new

approach was used:

$$var = \sum_{i,j} (i - \mu_i) GLCM(i, j), \quad (4)$$

where

$$\mu_i = \frac{\sum_{i,j} [GLCM(i, j) \cdot i]}{\sum_{i,j} GLCM(i, j)}. \quad (5)$$

In the above formulas,  $i$  and  $j$  are the indices of GLCM elements (they indicate the discrete gray levels) whereas  $\mu_i$  denotes the expected value of  $i$  for the probability distribution in the GLCM. The variance was calculated separately for vertical and horizontal spatial relationships, resulting in 2 GLCM parameters. Like other features described above, also the features determined from the GLCM matrix were independent of the brightness of the image.

### I. Processing of the QUS Maps

A typical approach when using QUS parameters to classify tumors or their response to NAC involves either calculating a texture parameter value from the whole tumor simultaneously or generating a texture parameter map and then averaging it over all pixels in the tumor. The latter was used as a reference approach to the method proposed in this paper. For each tumor, maps of each of the considered QUS parameters were determined from data collected in four scan planes of the tumor. The mean value of the parameter was then calculated and used as a predictor of tumor response to NAC.

The new method presented in this paper aims to create predictors using only those parts of the tumor which bring the most reliable information on the tumor response to NAC. For this purpose, each parametric map is translated to a map of probability of tumor non-response to NAC. It is done using functions that assign probabilities to parameter values. For convenience, we refer to these functions as “dictionaries” throughout the text, as they translate parametric images into probability maps. The dictionaries are calculated individually for each parameter and for each NAC therapy stage. They are determined based on “model” cases of responding and non-responding tumors, that is, tumors that we know responded very well or very badly to NAC. We assumed that tumors with  $RMC = 0\%$  (10 tumors) are the “model” of responders and tumors with  $RMC \geq 70\%$  (13 tumors) are the “model” of non-responders.

To compensate for a different number of model tumors in both groups and for different tumor sizes, all pixels within each model tumor are assigned a weight  $w$ :

$$w = (n_{pixel} \cdot n_{tumor})^{-1}, \quad (6)$$

where  $n_{pixel}$  is a total number of tumor pixels in all scan-planes of the tumor whereas  $n_{tumor}$  stands for the number of model responding or non-responding tumors ( $n_{tumor}$  equals 10 when the tumor belongs to the group of responding tumors, or 13 when it belongs to non-responding tumors).

The weights  $w$  are then used for calculating of the weighted probability density functions ( $w_{PDF}$ ) for individual parameters (amplitude of a  $w_{PDF}$  refers to the accumulated pixel weights

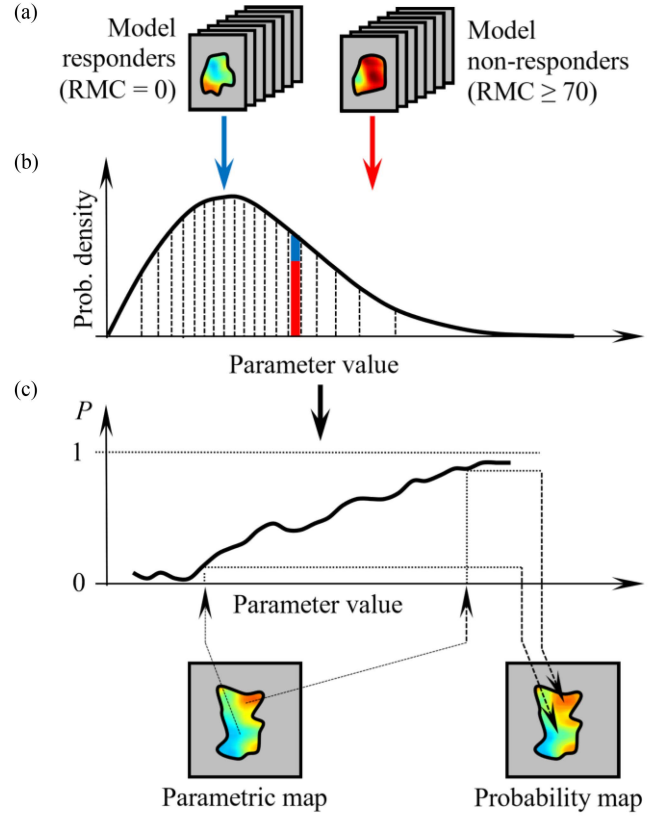


Fig. 1. Scheme of building and using a dictionary. Parametric maps of model tumors (a) are used to obtain the weighted probability density function  $w_{PDF}$  (b). The domain of the QUS parameter is divided into equally representative intervals. For each interval the probability  $P$  of non-response to NAC is determined according to (7). The resulting “dictionary” (c) is used to translate the parametric map of the tumor into a map of a probability that the tumor does not respond to NAC.

$w$ , not a pixel count). Then, the domain of the  $w_{PDF}$  is divided into a number of intervals so that the area under the  $w_{PDF}$  for each of them (sum of weights in each interval) is the same, as shown in Fig. 1(b). The first and last intervals extend to  $-\infty$  and  $+\infty$ , respectively. In this study the domain was divided into 100 intervals. Next, for each parameter value interval, two sums are computed, the sum of pixel weights from non-responders ( $w_{nr}$ ) and the sum of pixel weights from non-responders and responders together. The ratio of these sums determines the probability  $P$  that a pixel with a parameter value falling within the considered  $i$ -th range belongs to non-responding tumors:

$$P(i) = \frac{\sum w_{nr}(i)}{\sum w(i)}. \quad (7)$$

In this way, we obtain a dictionary that we use to translate parametric images into probability maps. It is worth noting that the above methodology ensures that all elements (parameter value intervals together with assigned probabilities) of the dictionary are equally representative. The data processing scheme is presented in Fig. 2.

As our goal is to detect non-responding tumors, the final predictors should be derived with special attention to areas of

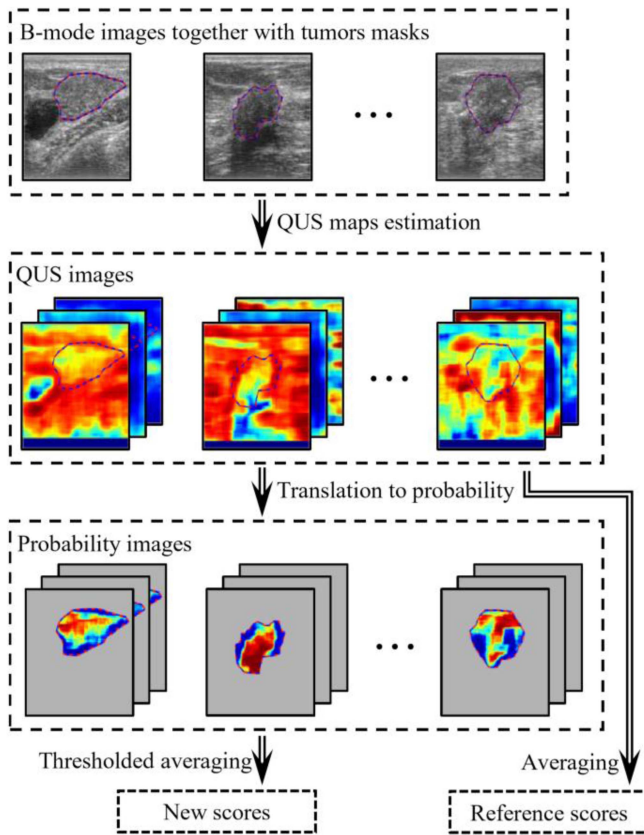


Fig. 2. Data processing scheme. First, the B-mode images are converted to QUS parametric maps, second, the parametric maps serve two purposes, classification - the results are the reference for the new approach, and building probability maps using a "dictionary". Third, the new approach classification results are determined by averaging the probability maps over a selected area with a probability above a certain threshold.

high-probability of no response to NAC. In this study, predictors were obtained as average probabilities calculated jointly from all four probability maps derived from the four tumor scanning planes. The averaging was limited to tumor regions with a probability above a certain threshold, which was set at 0.5 in this work. As the predictors are a direct measure of probability that the related tumors belong to the non-responding class, they were used as classification scores in the further statistical analysis.

### J. Statistical Analysis

Based on the Miller-Payne grading scale [32], tumors with  $RMC < 70\%$  and  $RMC \geq 70\%$  were considered responders and non-responders, respectively (poor response is represented by Miller-Payne grades 1 and 2, which corresponds to a tumor cell loss of up to 30%). Individual predictors were tested for their effectiveness in classifying tumors into the correct group. This was assessed using the receiver operating characteristic (ROC) curves, i.e., plots of true positive rate (TPR) versus false positive rate (FPR) for varying classification cutoff value. Each classifier's sensitivity and specificity were determined for the ROC point closest to the perfect classifier (i.e., the (0, 1) point in the ROC space) in the Euclidean sense [33]. For the purpose

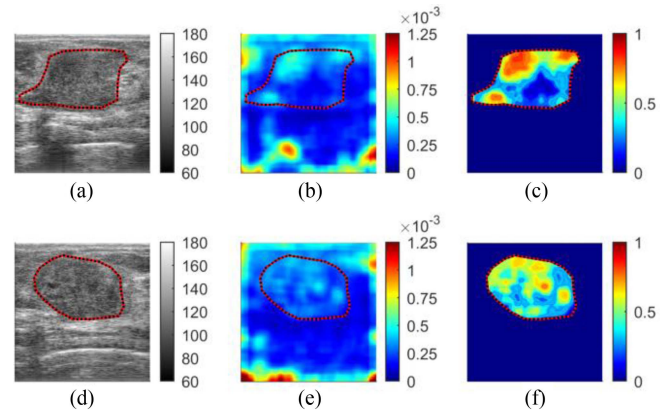


Fig. 3. Examples of image data for two sections of a "model" responding tumor after 3rd NAC course: B-mode images (a), (d), QUS parametric maps of the Law's R5L5 energy (b), (e), and related probability maps (c), (f).

of the overall evaluation of the entire ROC curve, area under the ROC curve (AUC) was used [34]. The estimation of the confidence intervals was done using the bootstrap method [35] with 1000 bootstrap samples and the confidence level of 0.95.

Both procedures, the reference method and the new approach, were cross-validated using the leave-one-out technique [36]. In case of the reference method the cross-validation was done at the statistical analysis stage, whereas in case of the new approach the cross-validation was performed at the stage of dictionary determination and translation to the probability-maps. As our database contains some multifocal lesions (multiple tumors from the same patient), the cross-validation was performed at the patient, not tumor, level.

All calculations were done using Matlab.

### III. RESULTS

The results presented in this section were obtained for data collected after the third course of NAC. The results corresponding to the first and second NAC courses are included in the Appendix.

Examples of image data at each processing step: B-mode images, QUS parametric images, and probability images are shown in Fig. 3 for two sections of a model responding tumor. Corresponding images for a model non-responding tumor are shown in Fig. 4.

The classification performance for the reference and the new approach is compared with use of the ROC curve in Fig. 5. The AUC values together with their confidence intervals are compared in Fig. 6. Numerical values of the classification performance measures are given in Table III.

The above results are obtained for the arbitrarily chosen probability threshold of 0.5. We have checked how the choice of the threshold value affects the classification performance. The dependence of AUC on the value of the selected probability threshold is shown in Fig. 7. We also studied the pixel value correlation between tumor probability maps obtained from different parametric maps. The obtained results are presented in Table IV.

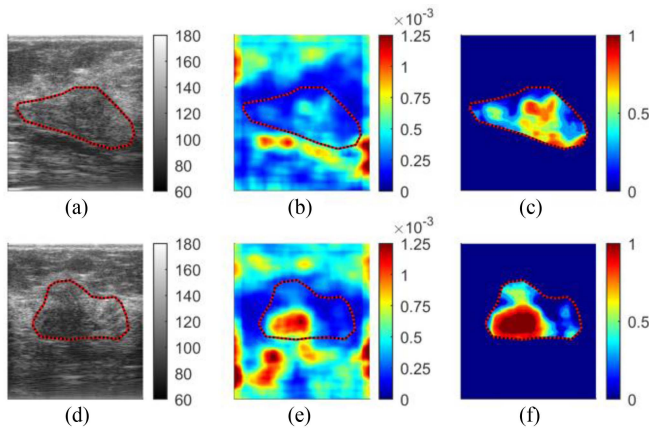


Fig. 4. Examples of image data for two sections of a “model” non-responding tumor after 3rd NAC course: B-mode images (a), (d), QUS parametric maps of the Law’s R5L5 energy (b), (e), and related probability maps (c), (f).

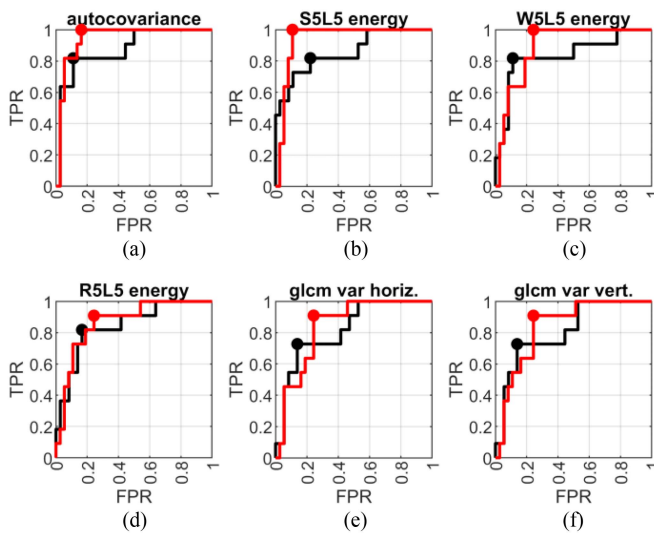


Fig. 5. Comparison of the ROC curves for the reference method (black line) and the new approach (red line) for each of the analyzed QUS parameters after the 3rd NAC course. The optimal operating points, for which the sensitivity and specificity were determined, are marked for each ROC curve.

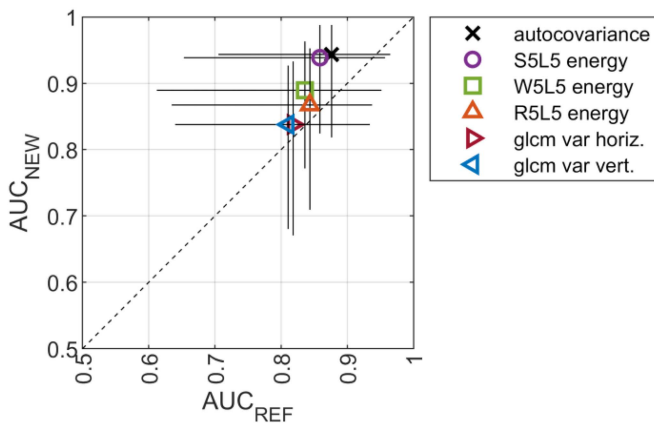


Fig. 6. AUC comparison between the reference method ( $AUC_{REF}$ ) and the new approach ( $AUC_{NEW}$ ) after 3rd NAC course.

TABLE III  
CLASSIFICATION PERFORMANCE AFTER THE 3RD NAC COURSE

	sensitivity		specificity		AUC		cutoff
	Ref	New	Ref	New	Ref	New	
autocovariance	0.82 <sup>1.00</sup> <sub>0.43</sub>	1.00 <sup>1.00</sup> <sub>1.00</sub>	0.89 <sup>0.97</sup> <sub>0.74</sub>	0.84 <sup>0.93</sup> <sub>0.69</sub>	0.88 <sup>0.96</sup> <sub>0.71</sub>	0.94 <sup>0.99</sup> <sub>0.82</sub>	0.75
S5L5 energy	0.82 <sup>1.00</sup> <sub>0.50</sub>	1.00 <sup>1.00</sup> <sub>1.00</sub>	0.78 <sup>0.90</sup> <sub>0.63</sub>	0.80 <sup>0.97</sup> <sub>0.76</sub>	0.80 <sup>0.96</sup> <sub>0.65</sub>	0.94 <sup>0.99</sup> <sub>0.82</sub>	0.75
W5L5 energy	0.82 <sup>1.00</sup> <sub>0.40</sub>	1.00 <sup>1.00</sup> <sub>1.00</sub>	0.89 <sup>0.97</sup> <sub>0.75</sub>	0.76 <sup>0.88</sup> <sub>0.59</sub>	0.84 <sup>0.95</sup> <sub>0.61</sub>	0.89 <sup>0.96</sup> <sub>0.77</sub>	0.68
R5L5 energy	0.82 <sup>1.00</sup> <sub>0.47</sub>	0.91 <sup>1.00</sup> <sub>0.50</sub>	0.83 <sup>0.94</sup> <sub>0.69</sub>	0.76 <sup>0.87</sup> <sub>0.58</sub>	0.84 <sup>0.94</sup> <sub>0.63</sub>	0.87 <sup>0.95</sup> <sub>0.71</sub>	0.66
glcm var horiz.	0.73 <sup>0.99</sup> <sub>0.37</sub>	0.91 <sup>1.00</sup> <sub>0.46</sub>	0.86 <sup>0.95</sup> <sub>0.71</sub>	0.76 <sup>0.87</sup> <sub>0.60</sub>	0.82 <sup>0.93</sup> <sub>0.64</sub>	0.84 <sup>0.93</sup> <sub>0.67</sub>	0.74
glcm var vert.	0.73 <sup>1.00</sup> <sub>0.37</sub>	0.91 <sup>1.00</sup> <sub>0.52</sub>	0.86 <sup>0.95</sup> <sub>0.71</sub>	0.76 <sup>0.89</sup> <sub>0.61</sub>	0.81 <sup>0.93</sup> <sub>0.64</sub>	0.84 <sup>0.93</sup> <sub>0.68</sub>	0.75

Comparison of the classification performance after the 3rd NAC course between the reference method (Ref) and the new method (New). The comparison includes the following classification parameters: sensitivity, specificity, and area under the ROC curve (AUC), each with confidence intervals for confidence level of 0.95. Additionally, classification cutoff values associated with the reported sensitivities and specificities are provided for the new method.

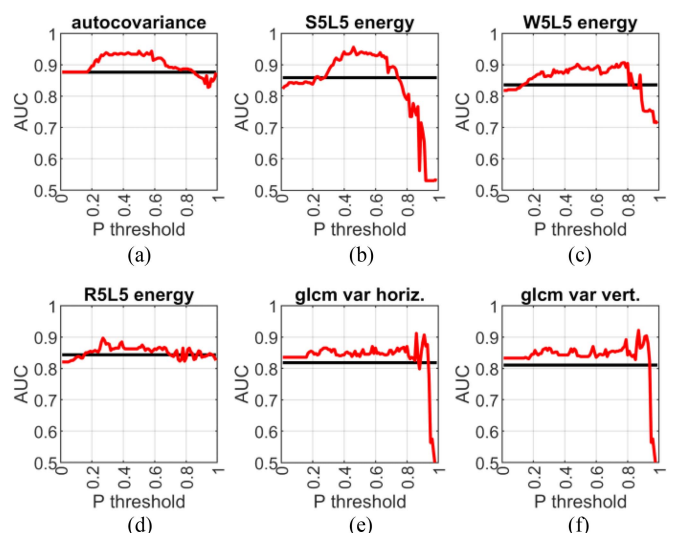


Fig. 7. Comparison of the AUC values for the reference method (black line) and the new approach (red line) as a function of the probability threshold for each of the analyzed QUS parameters after the 3rd NAC course.

#### IV. DISCUSSION

According to the literature [31], [37], the assessment of the tumor response to NAC may be inaccurate after the first and fourth week of treatment and improve after the eighth week of treatment. It is in line with our own previous observations [13], [14] that the results after the first and second NAC courses (first and fourth week of treatment, respectively) are poor and improve after the third NAC course (seventh week of treatment). This is also the case in this study, where results obtained before the third NAC course (Appendix) are significantly worse for both the reference method and the new one, and no improvement was observed compared to the reference method. Therefore, in this paper we focus on the results obtained after the third NAC course.

**TABLE IV**  
CORRELATIONS BETWEEN PROBABILITY MAPS BASED ON VARIOUS QUS PARAMETRIC IMAGES

	autoco variance	S5L5 energy	W5L5 energy	R5L5 energy	gcm var horiz.	gcm var vert.
autocovariance	1	0,56	0,62	0,61	0,78	0,79
S5L5 energy	0,56	1	0,88	0,78	0,17	0,17
W5L5 energy	0,62	0,88	1	0,93	0,16	0,17
R5L5 energy	0,61	0,78	0,93	1	0,18	0,19
gcm var horiz.	0,78	0,17	0,16	0,18	1	0,99
gcm var vert.	0,79	0,17	0,17	0,19	0,99	1

Correlation coefficients (Pearson) between probability maps based on various QUS parametric images after the 3rd NAC course.

The results obtained after the third course of NAC show that the 6 features selected to predict tumor response to NAC allow for good classification of tumors using the traditional method, which is used in this work as a reference for the proposed new approach. The determined AUC values ranged from 0.81 to 0.88, whereas the sensitivities and specificities were in the range of 0.73–0.82 and 0.78–0.89, respectively. The best predictors in terms of AUC were autocovariance (AUC = 0.88) and S5L5 energy (AUC = 0.86), with a sensitivity and specificity of 0.82 and 0.89 for autocovariance and 0.82 and 0.78 for S5L5 energy, respectively. However, these results can be further improved by the proposed new approach of selecting the area of tumor tissue used for classification.

First, the parametric maps of the tumors were used to determine the probability maps of the tumor response to NAC. Then, assuming that the information relevant for classification is contained mainly in tumor tissue, for which the probability map shows values above 0.5, a simple approach was adopted to transform the probability maps into scalar predictors of the classification outcome. This allowed us to easily focus on high-probability areas (assuming they contain useful information) and ignore low-probability areas. It should be noted, however, that there is room for improvement with other more sophisticated approaches.

Based on the AUC values, it can be concluded that the new method of assessing the effects of NAC improved the tumor response prediction (AUC = 0.84–0.94) compared to the prediction method using the average values of parametric maps (AUC = 0.81–0.88). In the new approach, the best predictors were also S5L5-energy and autocovariance but with higher AUC = 0.94 and sensitivity = 1.00 for both, and specificities of 0.89 and 0.84, respectively. The ROC curves for most of the predictors reached the TPR of 1.0 at lower FPR values than in the reference method (Fig. 5). This in turn leads to overall higher sensitivity of the new approach. The optimal operating points for the new approach show strong improvement in sensitivity (0.91–1.00) with slightly decreased specificity (0.76–0.89).

It is possible that this sensitivity-oriented improvement is a result of focusing on the high probability areas in the process of averaging. It may also be caused by a better representation of the non-responding tumors by the corresponding group of model tumors than it is in the case of responding tumors. Regardless of the reasons, this is a favorable situation because in the context of

NAC monitoring, false negative results have much more serious consequences than false positive results. In the cases of unidentified non-responders the ineffective therapy is continued. This unnecessarily delays surgery or the introduction of alternative therapy, giving the cancer more time to metastasize. Also, the toxic effect of the therapy on the patient remains unreasonably long. As a result, the patient’s chance of survival is reduced. On the other hand, in the cases of responders incorrectly classified as non-responders, verification procedures (e.g., MRI, biopsy) are performed unnecessarily.

In the layout presented in Fig. 6 the AUC markers that are above the diagonal indicate the parameters that perform better in the new approach than in the reference method. Even though the confidence intervals are wide, some parameters exhibit clear improvement, especially the autocovariance and the Laws S5L5 and W5L5 energies, where the AUC values improve by 0.05–0.08.

As the results of the new method were obtained for arbitrarily chosen probability threshold equal 0.5, it is natural to ask about the impact of the threshold on the classification performance. This was also investigated and is shown in Fig. 7. In most cases, AUC initially increases with increasing threshold. For a threshold value of 0, which corresponds to averaging of the probabilities over the whole tumor, the AUC value is lower than the AUC obtained at the selected higher threshold values. This seems to confirm that the approach of omitting some tumor areas from determining the final probability may lead to a better assessment of the tumor response to NAC. Further increasing of the threshold reveals a plateau in the AUC characteristic, so threshold tuning is of less importance here. This, in turn, justifies the arbitrary and approximate selection of the threshold level. Approaching each of the QUS parameters individually would allow for optimization of the threshold values to maximize the AUC. However, in our opinion, it would make sense if we had a larger database.

The results of correlation between the probability maps obtained from different tumor features (Table IV) showed a strong correlation for the maps that were calculated from the autocovariance and GLCM parameters (0.78–0.79). Moderate and strong correlations (0.56–0.62) were observed for maps obtained using autocovariance and lawsEnergy parameters. Maps obtained using the parameters of the lawsEnergy and the GLCM groups were very weakly correlated (0.16–0.19). The strong correlation of autocovariance images with other image groups may indicate the presence of tumor tissue areas that are characteristic of the therapy effect. It is worth noting that the classification of tumors into NAC-responders and non-responders was best when using the autocovariance parameter (AUC = 0.94).

In this study the classification performance of individual predictors, for which the interesting results (AUC above 0.9) were obtained, was investigated. Typically, to achieve further classification improvement, multi-parameter models are used. In case of the new approach, it is possible to build the multi-parameter classification models in two ways. The first one is to combine the probability predictors obtained as described in this paper. The other way is to determine the “dictionaries” for multiple



parameters. For example, for a pair of parameters the dictionary would be a function of two variables. The two parameter maps would still be translated into a single map of local probability of unresponsiveness to NAC, and the further processing would be performed without changes.

Apart from the multi-parameter approach, we also recognize a number of other concepts worth investigating. First, other QUS parameters could be examined, especially those based on RF data, such as spectral parameters. It is possible that they would have yielded better results with earlier doses of NAC. Another problem that needs analysis concerns alternative ways of selecting the most informative areas of the tumor. A comparison of the resulting selections with the saliency maps for some deep learning models would also be interesting. Furthermore, one could perform a mapping of the US-related images (QUS images, probability maps, masks of selected areas, or neural network saliency maps) to histopathological images. This could improve our understanding of the link between ultrasound signal properties and local tissue structures associated with tumor response to NAC. The challenge, however, is to collect the database in a way that would allow for such mapping. Nevertheless, such a comprehensive analysis would further improve the accuracy of the assessment of response to NAC.

We also believe that the application of the presented approach should not be limited to monitoring NAC treatment in breast cancer. It could be used to evaluate the results of chemotherapy in other cancer types that are accessible to ultrasound, e.g., liver cancer. It is also justified to consider other use-cases as long as the sought information is distributed heterogeneously over the area of interest. This may include, but is not limited to, classification of tumors as benign or malignant.

## V. CONCLUSION

We observed improved classification efficiency when using data from selected tumor regions. Increased AUC values were obtained for a wide range of probability cutoffs. It supports our hypothesis that areas characterizing the effect of therapy particularly well can be found. Also, the results of correlation studies between probability images may suggest the existence of such areas.

We believe that the proposed new approach to assessing the effects of NAC not only leads to a better classification of responses, but also may contribute to a better understanding of the microstructure of cancerous tumors seen with ultrasound.

The number of patients in this study seems sufficient to demonstrate the importance of tumor site selection for the classification of NAC non-responsive tumors. In the future, the statistical power of these studies will be improved by a larger cohort of participating patients. A fully developed method will be able to predict the tumor response for NAC in more objective way. In the event of treatment failure, this will enable more effective therapy to be implemented at an earlier stage of treatment than is currently the case. The consequence will be an increase in the survival rate of cancer patients.

## APPENDIX

TABLE V  
CLASSIFICATION PERFORMANCE AFTER THE 1ST NAC COURSE

	sensitivity		specificity		AUC		cutoff
	Ref	New	Ref	New	Ref	New	
autocovariance	0.92 <sup>1.00</sup> <sub>0.58</sub>	1.00 <sup>1.00</sup> <sub>1.00</sub>	0.68 <sup>0.81</sup> <sub>0.51</sub>	0.66 <sup>0.79</sup> <sub>0.50</sub>	0.81 <sup>0.91</sup> <sub>0.66</sub>	0.81 <sup>0.90</sup> <sub>0.67</sub>	0.64
S5L5 energy	0.69 <sup>0.91</sup> <sub>0.33</sub>	0.77 <sup>1.00</sup> <sub>0.42</sub>	0.89 <sup>0.97</sup> <sub>0.76</sub>	0.71 <sup>0.85</sup> <sub>0.56</sub>	0.80 <sup>0.91</sup> <sub>0.63</sub>	0.80 <sup>0.90</sup> <sub>0.65</sub>	0.67
W5L5 energy	0.77 <sup>0.94</sup> <sub>0.40</sub>	0.77 <sup>0.97</sup> <sub>0.46</sub>	0.78 <sup>0.91</sup> <sub>0.62</sub>	0.74 <sup>0.87</sup> <sub>0.58</sub>	0.82 <sup>0.93</sup> <sub>0.65</sub>	0.82 <sup>0.92</sup> <sub>0.68</sub>	0.73
R5L5 energy	0.85 <sup>1.00</sup> <sub>0.59</sub>	0.77 <sup>1.00</sup> <sub>0.45</sub>	0.78 <sup>0.90</sup> <sub>0.62</sub>	0.76 <sup>0.89</sup> <sub>0.62</sub>	0.81 <sup>0.92</sup> <sub>0.61</sub>	0.80 <sup>0.91</sup> <sub>0.63</sub>	0.70
glcm var horiz.	0.62 <sup>0.88</sup> <sub>0.31</sub>	0.69 <sup>0.92</sup> <sub>0.40</sub>	0.81 <sup>0.92</sup> <sub>0.66</sub>	0.76 <sup>0.88</sup> <sub>0.61</sub>	0.75 <sup>0.87</sup> <sub>0.58</sub>	0.71 <sup>0.85</sup> <sub>0.54</sub>	0.66
glcm var vert.	0.62 <sup>0.88</sup> <sub>0.30</sub>	0.69 <sup>0.92</sup> <sub>0.38</sub>	0.81 <sup>0.92</sup> <sub>0.65</sub>	0.74 <sup>0.86</sup> <sub>0.56</sub>	0.75 <sup>0.87</sup> <sub>0.59</sub>	0.70 <sup>0.85</sup> <sub>0.51</sub>	0.66

Comparison of the classification performance after the 1st NAC course between the reference method (Ref) and the new method (New). The comparison includes the following classification parameters: sensitivity, specificity, and area under the ROC curve (AUC), each with confidence intervals for confidence level of 0.95. Additionally, classification cutoff values associated with the reported sensitivities and specificities are provided for the new method.

TABLE VI  
CLASSIFICATION PERFORMANCE AFTER THE 2ND NAC COURSE

	sensitivity		specificity		AUC		cutoff
	Ref	New	Ref	New	Ref	New	
autocovariance	0.85 <sup>1.00</sup> <sub>0.47</sub>	0.62 <sup>0.88</sup> <sub>0.33</sub>	0.64 <sup>0.79</sup> <sub>0.47</sub>	0.76 <sup>0.86</sup> <sub>0.60</sub>	0.75 <sup>0.88</sup> <sub>0.59</sub>	0.75 <sup>0.85</sup> <sub>0.58</sub>	0.76
S5L5 energy	0.85 <sup>1.00</sup> <sub>0.53</sub>	0.85 <sup>1.00</sup> <sub>0.55</sub>	0.50 <sup>0.66</sup> <sub>0.33</sub>	0.65 <sup>0.79</sup> <sub>0.49</sub>	0.67 <sup>0.83</sup> <sub>0.47</sub>	0.70 <sup>0.85</sup> <sub>0.51</sub>	0.69
W5L5 energy	0.85 <sup>1.00</sup> <sub>0.50</sub>	0.77 <sup>0.94</sup> <sub>0.46</sub>	0.50 <sup>0.66</sup> <sub>0.33</sub>	0.59 <sup>0.75</sup> <sub>0.43</sub>	0.65 <sup>0.79</sup> <sub>0.45</sub>	0.69 <sup>0.82</sup> <sub>0.49</sub>	0.73
R5L5 energy	0.77 <sup>0.94</sup> <sub>0.42</sub>	0.85 <sup>1.00</sup> <sub>0.54</sub>	0.61 <sup>0.77</sup> <sub>0.45</sub>	0.54 <sup>0.71</sup> <sub>0.38</sub>	0.66 <sup>0.82</sup> <sub>0.49</sub>	0.68 <sup>0.81</sup> <sub>0.50</sub>	0.70
glcm var horiz.	0.85 <sup>1.00</sup> <sub>0.54</sub>	0.85 <sup>1.00</sup> <sub>0.51</sub>	0.58 <sup>0.75</sup> <sub>0.42</sub>	0.59 <sup>0.75</sup> <sub>0.44</sub>	0.72 <sup>0.85</sup> <sub>0.56</sub>	0.73 <sup>0.85</sup> <sub>0.56</sub>	0.70
glcm var vert.	0.85 <sup>1.00</sup> <sub>0.50</sub>	0.85 <sup>1.00</sup> <sub>0.56</sub>	0.58 <sup>0.76</sup> <sub>0.41</sub>	0.59 <sup>0.75</sup> <sub>0.42</sub>	0.72 <sup>0.86</sup> <sub>0.52</sub>	0.72 <sup>0.84</sup> <sub>0.56</sub>	0.70

Comparison of the classification performance after the 2nd NAC course between the reference method (Ref) and the new method (New). The comparison includes the following classification parameters: sensitivity, specificity, and area under the ROC curve (AUC), each with confidence intervals for confidence level of 0.95. Additionally, classification cutoff values associated with the reported sensitivities and specificities are provided for the new method.

## REFERENCES

- [1] B. Asselain et al., "Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: Meta-analysis of individual patient data from ten randomised trials," *Lancet Oncol.*, vol. 19, pp. 27–39, Jan. 2018, doi: [10.1016/s1470-2045\(17\)30777-5](https://doi.org/10.1016/s1470-2045(17)30777-5).
- [2] M. R. Chica-Parrado et al., "Resistance to neoadjuvant treatment in breast cancer: Clinicopathological and molecular predictors," *Cancers*, vol. 12, Jul. 2020, Art. no. 2012, doi: [10.3390/cancers12082012](https://doi.org/10.3390/cancers12082012).
- [3] E. A. Eisenhauer et al., "New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)," *Eur. J. Cancer*, vol. 45, pp. 228–247, Jan. 2009, doi: [10.1016/j.ejca.2008.10.026](https://doi.org/10.1016/j.ejca.2008.10.026).
- [4] I. Gounaris et al., "Accuracy of unidimensional and volumetric ultrasound measurements in predicting good pathological response to neoadjuvant chemotherapy in breast cancer patients," *Breast Cancer Res. Treat.*, vol. 127, pp. 459–469, Jun. 2011, doi: [10.1007/s10549-011-1454-x](https://doi.org/10.1007/s10549-011-1454-x).
- [5] I. V. Gruber et al., "Measurement of tumour size with mammography, sonography and magnetic resonance imaging as compared to histological tumour size in primary breast cancer," *BMC Cancer*, vol. 13, Jul. 2013, Art. no. 328, doi: [10.1186/1471-2407-13-328](https://doi.org/10.1186/1471-2407-13-328).
- [6] W. Haque et al., "Response rates and pathologic complete response by breast cancer molecular subtype following neoadjuvant chemotherapy," *Breast Cancer Res. Treat.*, vol. 170, pp. 559–567, Apr. 2018, doi: [10.1007/s10549-018-4801-3](https://doi.org/10.1007/s10549-018-4801-3).

- [7] B. Reig et al., "Role of MRI to assess response to neoadjuvant therapy for breast cancer," *J. Magn. Reson. Imag.*, vol. 52, pp. 1587–1606, Mar. 2020, doi: [10.1002/jmri.27145](https://doi.org/10.1002/jmri.27145).
- [8] L. M. Spring et al., "Pathologic complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: A comprehensive meta-analysis," *Clin. Cancer Res.*, vol. 26, pp. 2838–2848, Jun. 2020, doi: [10.1158/1078-0432.ccr-19-3492](https://doi.org/10.1158/1078-0432.ccr-19-3492).
- [9] V. Romeo et al., "Assessment and prediction of response to neoadjuvant chemotherapy in breast cancer: A comparison of imaging modalities and future perspectives," *Cancers*, vol. 13, Jul. 2021, Art. no. 3521, doi: [10.3390/cancers13143521](https://doi.org/10.3390/cancers13143521).
- [10] Y. H. Park et al., "Chemotherapy induces dynamic immune responses in breast cancers that impact treatment outcome," *Nature Commun.*, vol. 11, Dec. 2020, Art. no. 6175, doi: [10.1038/s41467-020-19933-0](https://doi.org/10.1038/s41467-020-19933-0).
- [11] A. Dasgupta et al., "Quantitative ultrasound radiomics using texture derivatives in prediction of treatment response to neo-adjuvant chemotherapy for locally advanced breast cancer," *Oncotarget*, vol. 11, pp. 3782–3792, Oct. 2020, doi: [10.18632/oncotarget.27742](https://doi.org/10.18632/oncotarget.27742).
- [12] K. Quiaoit et al., "Quantitative ultrasound radiomics for therapy response monitoring in patients with locally advanced breast cancer: Multi-institutional study results," *PLoS One*, vol. 15, Jul. 2020, Art. no. e0236182, doi: [10.1371/journal.pone.0236182](https://doi.org/10.1371/journal.pone.0236182).
- [13] K. S. Dobruch-Sobczak et al., "Quantitative assessment of the echogenicity of a breast tumor predicts the response to neoadjuvant chemotherapy," *Cancers*, vol. 13, Jul. 2021, Art. no. 3546, doi: [10.3390/cancers13143546](https://doi.org/10.3390/cancers13143546).
- [14] Z. Klimonda et al., "Assessment of breast cancer response to neoadjuvant chemotherapy based on ultrasound backscattering envelope statistics," *Med. Phys.*, vol. 49, pp. 1047–1054, 2022, doi: [10.1002/mp.15428](https://doi.org/10.1002/mp.15428).
- [15] L. Sannachi et al., "Non-invasive evaluation of breast cancer response to chemotherapy using quantitative ultrasonic backscatter parameters," *Med. Image Anal.*, vol. 20, pp. 224–236, Feb. 2015, doi: [10.1016/j.media.2014.11.009](https://doi.org/10.1016/j.media.2014.11.009).
- [16] H. Piotrkowska-Wróblewska et al., "Monitoring breast cancer response to neoadjuvant chemotherapy with ultrasound signal statistics and integrated backscatter," *PLoS One*, vol. 14, Mar. 2019, Art. no. e0213749, doi: [10.1371/journal.pone.0213749](https://doi.org/10.1371/journal.pone.0213749).
- [17] H. Tadayyon et al., "A priori prediction of neoadjuvant chemotherapy response and survival in breast cancer patients using quantitative ultrasound," *Sci. Rep.*, vol. 7, Apr. 2017, Art. no. 45733, doi: [10.1038/srep45733](https://doi.org/10.1038/srep45733).
- [18] H. Taleghamar et al., "Characterizing intra-tumor regions on quantitative ultrasound parametric images to predict breast cancer response to chemotherapy at pre-treatment," *Sci. Rep.*, vol. 11, Jul. 2021, Art. no. 14865, doi: [10.1038/s41598-021-94004-y](https://doi.org/10.1038/s41598-021-94004-y).
- [19] F. A. Spanhol et al., "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: [10.1109/TBME.2015.2496264](https://doi.org/10.1109/TBME.2015.2496264).
- [20] I. V. Balalaeva et al., "Spheroids of HER2-positive breast adenocarcinoma for studying anticancer immunotoxins in vitro," *Acta Naturae*, vol. 9, pp. 38–44, Mar. 2017, doi: [10.32607/20758251-2017-9-1-38-44](https://doi.org/10.32607/20758251-2017-9-1-38-44).
- [21] E. V. Denisov et al., "Clinically relevant morphological structures in breast cancer represent transcriptionally distinct tumor cell populations with varied degrees of epithelial-mesenchymal transition and CD44+CD24-stemness," *Oncotarget*, vol. 8, pp. 61163–61180, May 2017, doi: [10.18632/oncotarget.18022](https://doi.org/10.18632/oncotarget.18022).
- [22] Z. Klimonda et al., "On the assessment of local tumor response to neoadjuvant chemotherapy," in *Proc. IEEE Int. Ultrason. Symp.*, 2023, pp. 1–4, doi: [10.1109/IUS51837.2023.10307249](https://doi.org/10.1109/IUS51837.2023.10307249).
- [23] R. F. Chang et al., "Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis," *Ultrasound Med. Biol.*, vol. 29, pp. 679–686, May 2003, doi: [10.1016/s0301-5629\(02\)00788-3](https://doi.org/10.1016/s0301-5629(02)00788-3).
- [24] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: [10.1109/TSMC.1973.4309314](https://doi.org/10.1109/TSMC.1973.4309314).
- [25] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Boston, MA, USA: Addison-Wesley Longman Publishing, 1992, pp. 459.
- [26] K. I. Laws, "Texture energy measures," in *Proc. Image Understanding Workshop*, 1979, pp. 47–51.
- [27] A. Rodriguez-Cristerna et al., "BUSAT: A Matlab toolbox for breast ultrasound image analysis," in *Proc. 9th MCPR*, 2017, pp. 268–277.
- [28] W. Gomez, W. C. A. Pereira, and A. F. C. Infantosi, "Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound," *IEEE Trans. Med. Imag.*, vol. 31, no. 10, pp. 1889–1899, Oct. 2012, doi: [10.1109/TMI.2012.2206398](https://doi.org/10.1109/TMI.2012.2206398).
- [29] Z. Klimonda et al., "Breast-lesions characterization using quantitative ultrasound features of peritumoral tissue," *Sci. Rep.*, vol. 9, May 2019, Art. no. 7963, doi: [10.1038/s41598-019-44376-z](https://doi.org/10.1038/s41598-019-44376-z).
- [30] D. Bhardwaj et al., "Early changes in quantitative ultrasound imaging parameters during neoadjuvant chemotherapy to predict recurrence in patients with locally advanced breast cancer," *Cancers*, vol. 14, Feb. 2022, Art. no. 1247, doi: [10.3390/cancers14051247](https://doi.org/10.3390/cancers14051247).
- [31] A. Sadeghi-Naini et al., "Chemotherapy-response monitoring of breast cancer patients using quantitative ultrasound-based intra-tumour heterogeneities," *Sci. Rep.*, vol. 7, Sep. 2017, Art. no. 10352, doi: [10.1038/s41598-017-09678-0](https://doi.org/10.1038/s41598-017-09678-0).
- [32] K. N. Ogston et al., "A new histological grading system to assess response of breast cancers to primary chemotherapy: Prognostic significance and survival," *Breast*, vol. 12, pp. 320–327, Oct. 2003, doi: [10.1016/s0960-9776\(03\)00106-1](https://doi.org/10.1016/s0960-9776(03)00106-1).
- [33] I. Unal, "Defining an optimal cut-point value in ROC analysis: An alternative approach," *Comput. Math. Methods Med.*, vol. 2017, May 2017, Art. no. 3762651, doi: [10.1155/2017/3762651](https://doi.org/10.1155/2017/3762651).
- [34] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, pp. 29–36, Apr. 1982, doi: [10.1148/radiology.143.1.7063747](https://doi.org/10.1148/radiology.143.1.7063747).
- [35] B. Efron, "The bootstrap," in *The Jackknife, The Bootstrap, and Other Resampling Plans*. Philadelphia, PA, USA: Society Industrial Applied Mathematics, 1982, pp. 27–36, doi: [10.1137/1.9781611970319](https://doi.org/10.1137/1.9781611970319).
- [36] A. M. Molinaro et al., "Prediction error estimation: A comparison of resampling methods," *Bioinformatics*, vol. 21, pp. 3301–3307, Aug. 2005, doi: [10.1093/bioinformatics/bti499](https://doi.org/10.1093/bioinformatics/bti499).
- [37] H. Tadayyon et al., "Quantitative ultrasound assessment of breast tumor response to chemotherapy using a multi-parameter approach," *Oncotarget*, vol. 7, pp. 45094–45111, Apr. 2016, doi: [10.18632/oncotarget.8862](https://doi.org/10.18632/oncotarget.8862).