Breast Cancer Research and Treatment (2019) 173:365–373
https://doi.org/10.1007/s10549-018-4984-7

CLINICAL TRIAL

Machine learning for diagnostic ultrasound of triple-negative breast cancer

Tong Wu[1,2] · Laith R. Sultan[2] · Jiawei Tian[1] · Theodore W. Cary[2] · Chandra M. Sehgal[2]

Received: 24 September 2018 / Accepted: 28 September 2018 / Published online: 20 October 2018

Abstract

Purpose Early diagnosis of triple-negative (TN) breast cancer is important due to its aggressive biological characteristics, poor clinical outcomes, and limited options for therapy. The goal of this study is to evaluate the potential of machine learning with quantitative ultrasound image features for the diagnosis of TN breast cancer.

Methods Ultrasonic and clinical data of 140 surgically confirmed breast cancer cases were analyzed retrospectively for the diagnosis of TN and non-TN (NTN) subtypes. The subtypes were classified based on the expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). Ultrasound image features were measured from the grayscale and color Doppler images and used with logistic regression for classification by machine learning. Leave-one-out cross validation was used to train and test the differentiation. Diagnostic performance was measured by the area under receiver operating characteristic (ROC) curve, and sensitivity and specificity determined at the Youdons index.

Results Of the twelve grayscale and Doppler features measured, eight were found to be statistically different for the TN and NTN subtypes ($p < 0.05$). The area under the ROC curve (AUC) of the statistically significant grayscale (GS) and color Doppler (CD) features was 0.85 and 0.65, respectively. The AUC increased to 0.88 when the GS and CD features were used together, with sensitivity of 86.96% and specificity of 82.91%. Consideration of patient age in the analysis did not improve discrimination of TN and NTN.

Conclusions The analysis of breast ultrasound images by machine learning achieves high level of differentiation between the TN and NTN subtypes, exceeding the diagnostic performance by standard visual assessments of the images.

Keywords Breast cancer · Triple negative breast cancer · Quantitative breast ultrasound · Machine learning · Computer-aided diagnosis

Introduction

Breast cancer is the most common malignancy and the leading cause of cancer-related death in women worldwide. Every year, there are an estimated 1.67 million newly diagnosed patients and 0.52 million deaths due to breast cancer around the world [1]. In the United States, 268,670 cases and 41,400 deaths are projected to occur in 2018 [2]. Breast cancer is not a single disease. Triple-negative (TN) breast cancer is a subtype which lacks the expression of hormone estrogen receptor (ER), progesterone receptor receptors (PR), and human epidermal growth factor receptor 2 (HER2). Although the TN subtype only accounts for 15–20% of all breast cancers, it is extremely aggressive compared to other subtypes and is generally associated with poor prognosis, early recurrence, and distant metastasis. At present, there is no reliable and effective method of treatment

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-018-4984-7) contains supplementary material, which is available to authorized users.

✉ Jiawei Tian
jwtian2004@163.com

✉ Chandra M. Sehgal
chandra.sehgal@uphs.upenn.edu

[1] Department of Ultrasound, Second Affiliated Hospital of Harbin Medical University, 246 Xuefu Road, Harbin City 150086, Heilongjiang Province, People's Republic of China

[2] Ultrasound Research Lab, Department of Radiology, University of Pennsylvania, 168B John Morgan Building, 3620 Hamilton Walk, Philadelphia, PA 19104, USA

except with adjuvant chemotherapy. It is widely recognized that early detection and diagnosis of this aggressive subtype are important for improving postoperative survival and quality of life of the patients [3, 4].

Ultrasound is a versatile diagnostic modality for diagnosing breast lesions with a relatively high sensitivity and specificity. It can effectively distinguish between benign and malignant breast masses with a high negative predictive value [5]. The quality of ultrasound images, however, is influenced by several factors including low contrast and low signal-to-noise ratio, inadequate compensation of signal attenuation, speckle noise, and shadowing and enhancement artifacts. At the same time, image acquisition is user dependent and subjective due to the varying experience and skill levels of the sonographers [6]. These effects could be even more pronounced in developing countries or rural areas that lack resources, leading to higher mortality and poorer prognosis. In addition, TN breast cancer presents atypically in different imaging, making it difficult to get an effective early diagnosis [7–9]. There is an emerging trend of using quantitative methods with machine learning to decrease diagnostic subjectivity, and to reduce misdiagnosis for ultrasound imaging so that costly and resource-intensive invasive biopsies might be avoided [6, 10].

In the current state of the art, many of the technological advances have been applied to differentiating malignant and benign tumors with all subtypes grouped together [10–13]. In this study, we go beyond this discrimination of solid breast masses by applying computer methods to separate TN and NTN subtypes. Toward this goal, we assess the role of grayscale (GS) and color Doppler (CD) ultrasound features extracted from images in distinguishing TN from NTN by machine learning.

Materials and methods

Patients and image acquisition

The study was approved by the institutional review board (IRB) for investigation. All patient information was de-identified and the requirement for informed consent was waived in accordance with the IRB protocol. Data were collected from 140 consecutive primary invasive breast cancer cases (140 masses), surgery-confirmed at the Second Affiliated Hospital of Harbin Medical University between January 2014 and March 2015. The ultrasound images were acquired with a broadband 12–15 MHz linear-array transducer (HITACHI Vision 900, Hitachi Medical System, Tokyo, Japan) by a sonographer with five years of experience in breast imaging. GS and CD images of breast masses were acquired in multiple standard tangent planes and used for offline analysis.

Breast imaging reporting and data system (BI-RADS) category and subtypes classification

Ultrasound image data were reviewed by two radiologists with at least five years of experience in interpreting breast ultrasound images. The radiologists were blinded to the pathology results. In cases of disagreement, a consensus was sought. Ultrasonic features of the breast masses, including tumor shape, orientation, margin, echo patterns, posterior acoustic features, calcifications, and vascularity were retrospectively analyzed according to the fifth edition of the BI-RADS criteria [14]. Using clinicopathologic and immunohistochemical analysis, breast cancer cases were further classified into two molecular subtypes, namely, TN subtype (ER−/PR−/HER2−) and NTN subtype, according to the recommendations of the 2013 St. Gallen consensus [15]. Among the NTN cases were luminal A [ER+/PR+,/HER2−/Ki-67 < 14%], luminal B (HER2−) subtype [ER+/HER2−/Ki-67 ≥ 14% or PR < 20%], luminal B (HER2+) subtype [ER+/HER2+/any Ki-67/any PR], and HER2+ subtype [HER2+/ER^{-}/PR^{-}].

Computerized feature extraction

Five GS images of each lesion from each subject were analyzed. Nine quantitative grayscale features describing margin and shape characteristics of the lesion were extracted from a manually drawn region of interest (ROI) using a custom application written in the IDL programming language (Harris Corporation, Herndon, VA, USA). These features have been described in our earlier studies [11, 16] and also included in the supplementary material. In short, they were computed by partitioning the lesion into *N* sectors and analyzing grayscale characteristics of the pixels within the lesion, at the margin and in surrounding tissue. These quantitative features are margin sharpness (MS), margin echogenicity difference (MED), angular variance in margin (AVM), depth-to-width ratio (DWR), axis ratio (AR), tortuosity, circularity, radius variation (RV), and elliptically normalized skeleton (ENS).

Three CD images of each lesion were used to measure Doppler vascularity within the manually drawn ROI [17]. The Doppler measurement involved two steps. First, the color bar on the Doppler image was used to set the color scale by dividing each directional flow into 100 equal levels. Each level was assigned a velocity value between 0 and maximum velocity in the color bar based on the position of the color level in the bar. The colored pixels were detected in the region of interest and assigned a velocity value based on color level. The number of pixels with flow (colored pixels) and their velocities were used to measure

vascular indices of fractional area of flow in the lesion (AI), mean flow velocity in the lesion (VI), and flow volume in the lesion (FVI).

Feature reduction

Standard descriptive statistics based on the arithmetic mean and SD values of the ultrasound features and age were derived for the TN and NTN subtypes. The two-tailed Student's *t*-test of unequal variance was used to determine the statistical significance of the difference between two groups. Features with significant difference of $p < 0.05$ were selected for characterizing lesions.

Model construction and evaluation

The grayscale features, Doppler features, and patient age were used to train and test the classification model by logistic regression using leave-one-out cross validation, in which $N-1$ samples (of the N samples in the data) were trained to predict the probability of the remaining Nth sample for being TN subtype. The TN probability of a lesion was compared with the biopsy results to perform ROC analysis using MedCalc [Version 17.9, Ostend, Belgium]. The area under the ROC curve (AUC), sensitivity (Se), and specificity (Sp) at Youden index for each ROC curve were used for evaluating diagnostic performance of the classification model.

Results

Of the 140 cases, 115 were invasive ductal carcinomas, 18 were invasive lobular carcinomas, and 7 were others. 23 (16.43%) of these cases were TN and the remaining 117 (83.57%) were NTN. In NTN subtypes, there were 30 cases of LA subtype, 27 of LB (HER2+), 35 of LB (HER2−), and 25 HER2+ subtype. The mean age at diagnosis of the all patient was 51.45 ± 9.39. Patients with TN tumors were slightly younger than patients with NTN (50.26 ± 9.64 vs. 51.68 ± 9.37 years, Table 2). The difference in age between the two groups was not significant ($p = 0.52$). The BI-RADS categories evaluated by visual inspection of the images showed no significant difference between TN and NTN subtypes. p was determined to be 0.17 by Chi square (Table 1).

Table 1 BI-RADS Category of TN and NTN subtypes

BI-RADS category	TN (%)	NTN (%)	*p*-value
Category 3	2 (8.70%)	4 (3.42%)	0.17
Category 4	7 (30.43%)	30 (25.64%)	
Category 5	14 (60.87%)	83 (70.94%)	
Total	23 (16.43%)	117 (83.57%)	

BI-RADS breast imaging reporting and data system, *TN* triple negative, *NTN* non-triple negative

Qualitative assessment of images

Figure 1 shows examples of triple-negative breast cancers. GS ultrasound images of a 36-year-old woman show a hypoechoic mass with irregular shape, microlobulated margin, posterior enhancement, and greater lesion depth compared to width (left panel, Fig. 1a). The Doppler image of the same mass shows no vascularity, however there is a small vascular region in the tissue surrounding the mass (right panel, Fig. 1a). The lesion was assessed to be BI-RADS category 5 by visual inspection of the images. The quantitative computerized features for the lesion were 94.72 for margin sharpness, 17.86 for margin echogenicity difference, 1.40 for tortuosity, 0.50 for circularity, 0.21 for elliptically normalized skeleton, 3.02 for AI, 1.27 for VI, and 0.04 for FVI. The GS ultrasound image of a 49-year-old woman with a TN mass shows a hypoechoic mass with an irregular shape, circumscribed and microlobulated margin (left panel Fig. 1b). Doppler imaging showed the mass to be moderately vascular with vascular regions distributed within and at the rim. The BI-RADS category for the mass was clinically assessed to be 4 by visual inspection of the images. The computerized quantitative features were 97.22 for margin sharpness, 18.87 for margin echogenicity difference, 1.43 for tortuosity, 0.47 for circularity, 0.21 for elliptically normalized skeleton, 6.85 for AI, 2.55 for VI, and 0.17 for FVI.

Figure 2 shows examples of non-triple-negative cancers. The GS ultrasound image of a 75-year-old woman with NTN (HER2+ subtype) shows a hypoechoic mass with an irregular shape, spiculated and microlobulated margin, and an indifferent posterior echo (left panel, Fig. 2a). The mass is highly vascularized within the interior and at the rim (right panel, Fig. 2a). The mass was clinically assessed to be BI-RADS category 5 by visual inspection of the images. The quantitative computerized features were 97.22 for margin sharpness, 15.78 for margin echogenicity difference, 1.59 for tortuosity, 0.37 for circularity, 0.26 for elliptically normalized skeleton; 7.21 for AI, 1.71 for VI, and 0.12 for FVI. GS ultrasound imaging of a 55-year-old woman with NTN (luminal B subtype) shows a hypoechoic mass with irregular shape, spiculated and microlobulated margin, and posterior echo. Similar to the case shown in Fig. 2a, the mass is highly vascular throughout. BI-RADS category of the mass was assessed to be 5 by visual inspection of the images. The quantitative computerized features for the mass were 93.61 for margin sharpness, 15.66 for margin echogenicity difference, 1.61 for tortuosity, 0.35 for circularity, 0.25 for elliptically normalized skeleton; 9.75 for AI, 2.85 for VI, and 0.28 for FVI.

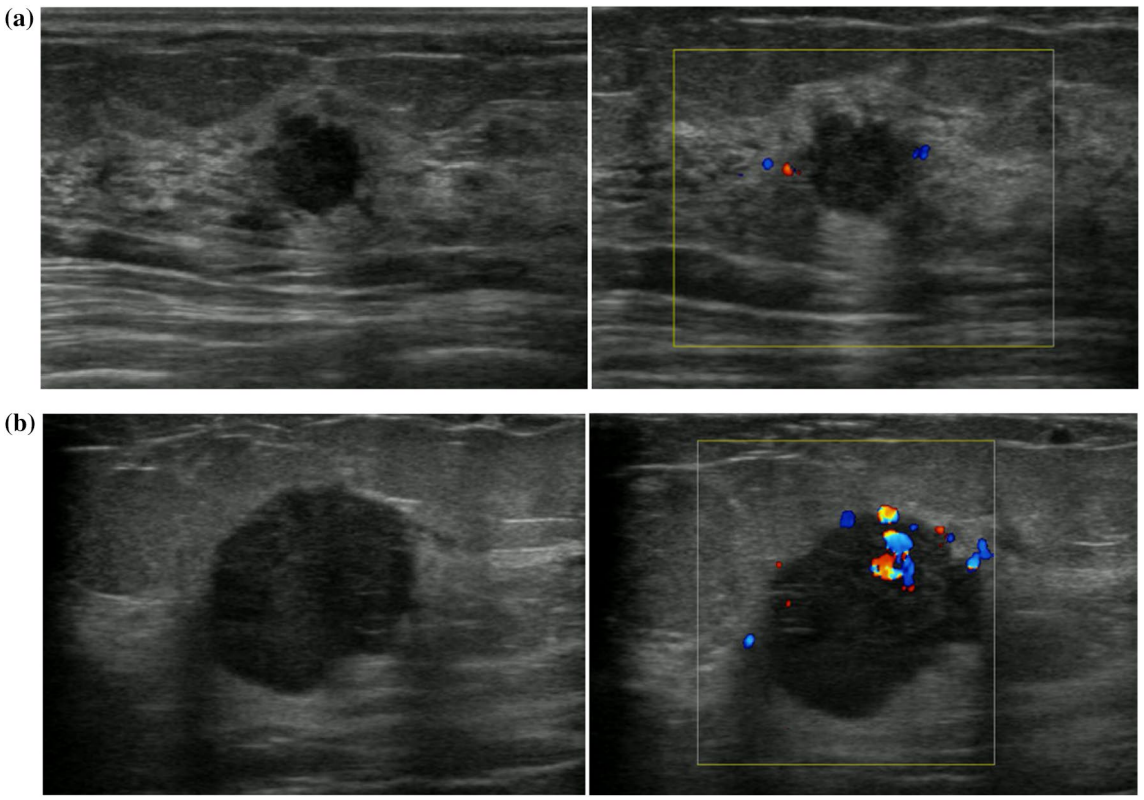

Fig. 1 Examples of grayscale and color Doppler images of triple-negative breast cancers. **a** TNBC in a 36-year-old woman; **b** TNBC in a 49-year-old woman

Quantitative features of images

Table 2 compares the magnitude of the quantitative computerized ultrasound features of all the TN lesions with those of NTN lesions. Of the various features studied, two margin, three shape, and three vascularity features showed statistical difference between the two subtypes ($p < 0.05$). Qualitatively, the margin of TN was found to be more sharply defined than in NTN cases. Quantitatively, both margin sharpness and margin echogenicity difference confirm that the margins of TN are better defined than those of the NTN subtype: MS was 94.37 ± 2.39 versus 91.67 ± 3.33, and MED was 16.58 ± 3.35 versus 13.80 ± 3.35, for TN versus NTN. The shape features showed TN lesions to be more regular with smoother margins than NTN lesions. Mean circularity of TN lesions (0.49 ± 0.09) was higher than that of NTN (0.42 ± 0.10) with $p = 0.0024$. Tortuosity, on the other hand, of TN (1.37 ± 0.12) was less than for NTN (1.49 ± 0.16) with $p = 0.0002$. The smaller value of elliptically normalized skeleton (ENS) in TN (0.20 ± 0.03) compared to NTN (0.23 ± 0.04) indicates that NTN shapes are more complex and irregular ($p = 0.0001$). The depth-to-width ratio (DWR) and axis ratio (AR), which to a certain extent reflect the growth pattern of breast tumors, were not found to be statistically different ($p > 0.05$), indicating that there was no preferential growth for these two subtypes. Marked difference in Doppler vascular features was observed between TN and NTL groups. Color Doppler vascular fraction area (AI), mean flow velocity index (VI), and flow volume index (FVI) in the lesion were all substantially lower for TN compared to NTN.

Performance of the diagnostic model

Table 3 summarizes the diagnostic performances of significant margin features (margin sharpness, margin echogenicity difference), shape features (tortuosity, circularity, elliptically normalized skeleton), and vascularity features (velocity index, area index, flow index) used as groups to characterize TN and NTN subtypes. Of the three feature groups, margin features performed the best with AUC of 0.73, followed by shape and vascularity groups with AUC of 0.68 and 0.65, respectively. Each feature group provided improvement over baseline random performance of AUC of 0.5 ($p < 0.002$). At the operation point defined by the Youden index, the vascularity feature group was most sensitive (Se = 95.65%) followed by the shape (Se = 86.96%) and margin (Se = 60.87%) feature groups. The specificity was highest

Table 2 Quantitative features of ultrasound images used in differentiating TN and NTN breast cancers

Features	TN (mean ± SD)	NTN (mean ± SD)	*p*-value
Margin			
MS	94.37 ± 2.39	91.67 ± 3.33	0.0000
MED	16.58 ± 3.35	13.80 ± 3.35	0.0010
AVM	4.18 ± 1.25	3.79 ± 1.17	0.1847
Shape			
DWR	0.72 ± 0.18	0.73 ± 0.17	0.6363
AR	1.57 ± 0.43	1.53 ± 0.38	0.6661
Tortuosity	1.37 ± 0.12	1.49 ± 0.16	0.0002
Circularity	0.49 ± 0.09	0.42 ± 0.10	0.0024
RV	0.18 ± 0.07	0.19 ± 0.06	0.8990
ENS	0.20 ± 0.03	0.23 ± 0.04	0.0001
Vascularity			
AI (%)	5.70 ± 3.46	7.90 ± 4.94	0.0136
VI (m/s)	1.86 ± 0.55	2.31 ± 0.79	0.0020
FVI (mL/mL)	0.11 ± 0.07	0.20 ± 0.15	0.0001
Age (year)	50.26 ± 9.64	51.68 ± 9.37	0.5205

SD standard deviation, *TN* triple negative, *NTN* non-triple negative, *MS* margin sharpness, *MED* margin echogenicity difference, *AVM* angular variation in margin, *DWR* depth to width ratio, *AR* axis ratio, *RV* radius variation, *ENS* elliptically normalized skeleton, *AI* vascular fraction area, *VI* mean flow velocity index, *FVI* blood flow volume index

for the margin feature group (Sp = 80.34%) followed by the shape (Sp = 63.25%) and vascularity (Sp = 38.46%) feature groups. A marked improvement in AUC was observed when two feature groups were used together. Combined use of margin and shape increased the performance from 0.73 and 0.68 for the individual groups to 0.85 ($p = 0.0421$ and 0.0001 respectively, Fig. 3). Similarly, combined use of margin and vascularity, and shape and vascularity, improved the performance to 0.78 and 0.78, respectively. When all the feature groups (margin, shape, and vascularity) were used together in the diagnostic model, AUC increased to 0.88 with overall sensitivity of 86.96% and specificity of 82.91%, Fig. 3.

The effect of age on differentiating NTN and TN

Age has often been found to have diagnostic value in differentiating malignant and benign masses. When age was added to the image feature groups, no improvement in performance was observed (Table 4). The predictive accuracy of the model, as measured by AUC, remained in the range of 0.67–0.88 when machine learning on age in addition to shape, margin, and Doppler features (Table 4).

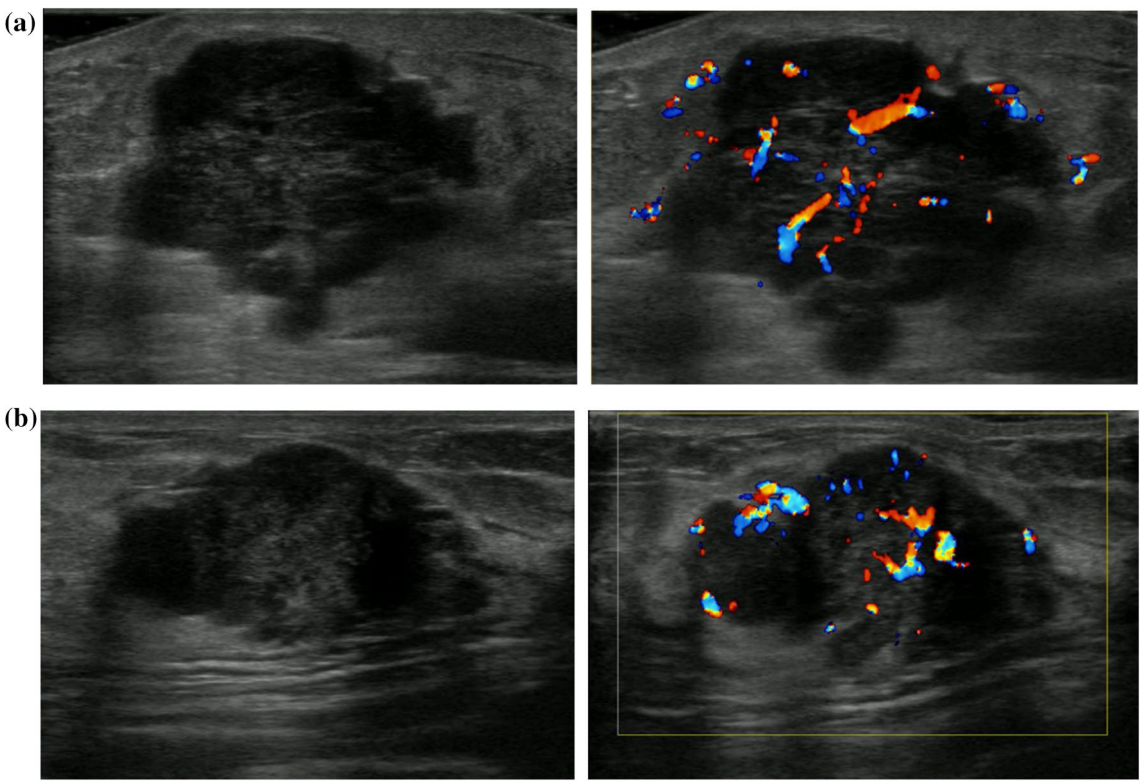

Fig. 2 Examples of grayscale and color Doppler images of non-triple-negative breast cancers. **a** NTN in a 75-year-old woman; **b** NTN in a 55-year-old woman

Table 3 Diagnostic performance of machine learning methods for differentiating TN and NTN masses using logistic regression with grayscale and color Doppler features

Groups	Margin	Shape	CD	GS (M and S)	Margin and CD	Shape and CD	GS and CD
AUC	0.73	0.68	0.65	0.85	0.78	0.78	0.88
SD	0.06	0.05	0.05	0.04	0.05	0.05	0.04
95% CI	0.65–0.80	0.60–0.76	0.57–0.73	0.78–0.90	0.71–0.85	0.70–0.84	0.82–0.93
Sensitivity (%)	60.87	86.96	95.65	78.26	82.61	82.61	86.96
Specificity (%)	80.34	63.25	38.46	82.05	70.94	62.39	82.91
YI	0.412	0.502	0.341	0.603	0.536	0.450	0.699
p value	0.0001	0.0007	0.0047	< 0.0001	< 0.0001	< 0.0001	< 0.0001

CD color Doppler, *GS* grayscale (margin + shape), *AUC* area under curve, *SD* standard error, *CI* confidence interval, *YI* Youden index

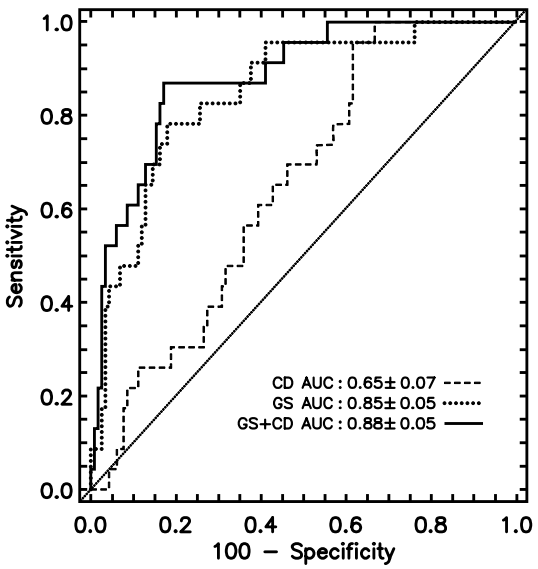

Fig. 3 ROC curves of different feature groups

Discussion

Most TN breast cancers are basal-like subtype tumors, characteristic of basal epithelial cells, which tend to be high-grade and more aggressive [7, 18]. These tumors are commonly believed to be associated with shorter patient survival, greater local recurrence, and a higher incidence of visceral and brain metastasis [19]. Furthermore, the absence of both hormone receptors and HER2 makes these cancers less susceptible to hormonal therapies like Tamoxifen, Arimidex, Aromasin, as well as to targeted therapy such as Herceptin or Tykerb [4, 15, 20]. In addition, our results showed that BI-RADS commonly used in clinical practice could not effectively differentiate TN and NTN subtypes ($p = 0.17$), which is consistent with the previous findings of Wojcinskiet et al. using ultrasound ($p = 0.62$, [21]), Krizmanich-Conniffet et al. using mammography ($p = 0.25$ [8]) and Martine Boisserie-Lacroix et al. using MRI ($p = 0.06$, [9]). Clearly, there is a need to improve the diagnosis of TN breast lesions. With the rapid advances in quantitative methods and computer technology, machine learning could play an important role in the diagnosis of TN. Toward this goal, several studies have focused on differentiating TN and NTN subtypes using subjective visual inspection of images for diagnosis. Ko et.al observed TN cancers were more likely to be circumscribed (43/75, 57.3%), and markedly hypoechoic (36/75, 48.0%), and less likely to exhibit posterior shadowing (4/75, 5.0%) on ultrasound images [7]. Krizmanich-Conniffet et al. found TN cancer to be mostly irregular, non-calcified masses with ill-defined or spiculated margins on mammography, which appeared as hypoechoic or complex with an irregular shape and non-circumscribed margins on ultrasound [8]. The studies of Boisserie-Lacroixeet et al. confirmed TN breast cancers to be more likely circumscribed or microlobulated with no posterior acoustic shadowing or enhancement on ultrasound. They also observed TN to often exhibit rim enhancement compared to other subtypes on MRI [9]. Since TN breast cancers do not have typical appearance on images, their early diagnosis is often not straightforward using subjective approaches.

Table 4 The effect of age using machine learning for the diagnostic performance of grayscale and color Doppler features

Groups	Margin	Shape	CD	GS	Margin and CD	Shape and CD	GS and CD
AUC	0.73	0.68	0.65	0.85	0.78	0.78	0.88
AUC and age	0.73	0.67	0.63	0.85	0.78	0.77	0.88

AUC area under curve, *GS* grayscale (margin and shape), *CD* color Doppler

To overcome subjectivity in the diagnosis of TN breast cancers, advances in digital image processing, pattern recognition, and artificial intelligence have been proposed to increase diagnostic accuracy [6, 11, 22]. The goal of these methods has been to reduce false positives and false negatives and to avoid unnecessary biopsies and transitional treatments [11, 13, 16, 17, 23, 24]. Much of the emphasis of the advanced methods has been on differentiating malignant and benign masses. In this study, we take the next step of differentiating malignant subgroups TN and NTN using quantitative ultrasonic features and machine learning. The results show that the ultrasound features characterizing margin sharpness (MS), margin echogenicity difference (MED), tortuosity, circularity (C), elliptically normalized skeleton (ENS), and vascularity features (AI, VI, and FVI), were statistically different for the two subtypes. While some feature like MS, MED, and circularity were greater for TN, other features like tortuosity, ENS, AI, VI, and FVI were lower compared to those for NTN. The use of quantitative features has not been previously reported and the discussion that follows compares these quantitative measurements with the qualitative assessment of the breast images reported in the earlier studies [7, 9, 25].

It is commonly known that most malignant breast masses have an irregular shape often characterized by ill-defined, spiculated, angular margins with posterior acoustic shadowing [5]. However, unlike other malignant masses, quantitative measurement of margin sharpness and margin echogenicity show TN masses to be more sharply defined with significant echogenicity differences between the interior and immediate exterior of the breast mass. This is consistent with the previous qualitative observations that TNs are more likely to have circumscribed or microlobulated margins on ultrasound [7, 9, 25]. The well-defined appearance of the TN subtype was considered to be associated with rapid and aggressive cell proliferation without infiltrating stromal reaction [26, 27]. Contrary to these studies, few studies have also noted that the TN subtype was accompanied with desmoplastic response and inflammation [28], along with irregular shape, and ill-defined or microlobulated margins on ultrasound images [8], possibly due to a low proliferative rate of the breast tumor, giving sufficient time for stromal interactions and fibroblasts, inflammatory cells and angiogenesis surrounding the invasive edge [29]. The lower tortuosity and elliptically normalized skeleton, and greater circularity of the TN subtype observed in this study, indicate that these masses were uniform, circular or oval in shape, without significant surface distortions. This is consistent with the qualitative studies that have noted that TN cancers are more likely to be regular and round/oval than NTN [9, 26], lacking angular/spiculated margins [29]. Although there are some studies reporting the TN subtype to be irregular, especially in small masses [30], the consensus is that TNs often have well-circumscribed margin and regular shape, common traits of benign tumors with rapid cellular proliferation [24, 31, 32].

The grayscale image features described above characterize margin and morphology of the lesions. There is increasing recognition that angiogenesis, abnormal blood vessel formation, and changes in blood flow patterns are often associated with malignant neoplastic changes in breast lesions, and could therefore provide independent information for improving diagnosis [17, 33, 34]. In this study, ultrasound Doppler features were measured to assess their role in improving differentiation of TN from NTN. Three different features characterizing vascular fractional area were measured: fractional area of flow in the lesion, AI, mean blood flow velocity, VI, and blood flow index, FVI. The results showed that all the three vascular features had lower magnitude for TN compared to NTN. Currently, only a few studies are available for comparison that report tumor vascularity in the different subtypes. In one study involving 1000 breast cases, Zhang et al. showed lack of vascularity in TN [35], which is in agreement with lower vascularity observed in our study. Also the HER2 gene contributes to angiogenesis by increasing the expression of vascular endothelial growth factor [36], which means that the absence of HER2 expression in TN should result in lower vascularity in this subtype, as observed in this study. Furthermore, TN breast cancers are often (24–41%) associated with posterior acoustic enhancement [37, 38], indicating accumulation of fluid within the lesion due to tumor necrosis [39] caused by the rapid proliferation of tumor cells. The necrosis in turn damages the vascular tissue and lowers vascularity. Thus there is overwhelming evidence supporting lower vascularity in TN, consistent with the findings of this study.

Each of the grayscale and vascular features described above are different for TN and NTN, aiding in the diagnosis. Machine learning optimizes the relative weights of these features to maximize lesion discrimination. Not surprisingly, the features that demonstrated significant difference yielded the highest correct classification. The diagnostic performance of these features as determined by ROC analysis was an AUC of 0.88 with 86.96% sensitivity and 82.91% specificity at Youden index. When only the grayscale features were used, the AUC was 0.85 with 78.26% sensitivity and 82.05% specificity at Youden index. The increase in diagnostic performance by 3.4% when adding Doppler features to the model is important, since it becomes more difficult to improve on performance as AUC approaches the maximum achievable area of 1 indicating perfect discrimination. Qualitative studies using ultrasound for differentiating TN breast cancers from fibroadenoma by visual assessment showed AUC with grayscale images alone and in combination with Doppler Ultrasound to be 0.65 and 0.58, respectively. The diagnostic performance increased significantly to

AUC of 0.87 when ultrasound was used with elastography [40]. Although there are differences in the study design and populations between the present and previous studies, the results consistently demonstrate that machine learning with quantitative features can achieve a high level of diagnostic performance, comparable to or exceeding performance of subjective visual assessment of the images. A closer inspection of the quantitative features shows the difference in gray-scale features between TN and NTN to be small, although statistically significant. While these small differences are easy to identify with quantitative methods, they may be difficult to detect reliably by visual inspection.

Several previous studies have found patient age to be a significant factor in classifying breast masses as malignant or benign [11]. In this study, we also evaluated whether age played a role in differentiating TN and NTN, adding patient age as one of the quantitative features for machine learning. The inclusion of patient age did not improve diagnosis, suggesting that triple-negative breast cancers have an equal likelihood of occurrence across all age groups in which they are present. This result was consistent with Qi Yang's study (46.5 years for TN vs. 49.2 years of or ER+ vs. 50.1 years for HER2+, $p = 0.223$) [41], and Hai-Yan Du's study (49 years for TN vs. 54 years for NTN, $p > 0.05$) [42]. In contrast to these findings, Krizmanich-Conniff et al. found that TN breast cancer was more commonly found in younger women (51.1 years of mean age) than NTN (55.9 years of mean age) with significant difference of $p \leq 0.0001$ [8]. Bauer et al. performed a large-scale study of the California Cancer Registry, concluding that the median age for TN cases ($n = 6370$) was 54 years, significantly less than 60 years for NTN ($n = 6370$) with $p < 0.001$; and TNs were especially likely to occur in younger women less than 40 (12.2% vs. 5.7% for NTN, $p < 0.001$) [43]. The reasons for the differences between studies are not known but could be related to population selection and size of the study.

The present study is not without limitations. First, tracing the ROI to define breast masses is user dependent and could be a source of variability between observers. This problem could be addressed in the future by computerized segmentation of masses [24]. The small sample size and retrospective nature of the study introduces sampling bias. Future prospective studies involving multiple observers and a larger number of patients are warranted to demonstrate the viability of the proposed approach in clinical settings.

Conclusion

In summary, using machine learning methods to analyze ultrasound images of breast masses provided a high level of differentiation between TN and NTN subtypes comparable to or improving upon standard differentiation by visual assessments. Future confirmation of this diagnostic technology could effectively assist the clinician's diagnosis and increase diagnostic confidence.

Acknowledgements This research was supported by NIH grant RO1 CA130946 and by the National Natural Science Foundation of China grant 81630048, 81271647.

Compliance with ethical standards

Conflict of interest All authors declare they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Torre LA, Bray F, Siegel RL et al (2015) Global cancer statistics, 2012. CA Cancer J Clin 65:87–108. https://doi.org/10.3322/caac.21262
2. Siegel RL, Miller KD, Jemal A (2018) Cancer statistics, 2018. CA Cancer J Clin 68:7–30. https://doi.org/10.3322/caac.21442
3. Lin NU, Vanderplas A, Hughes ME et al (2012) Clinicopathologic features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network. Cancer 118:5463–5472. https://doi.org/10.1002/cncr.27581
4. Sasaki Y, Tsuda H (2009) Clinicopathological characteristics of triple-negative breast cancers. Breast Cancer 16:254–259. https://doi.org/10.1007/s12282-009-0153-5
5. Stavros AT, Thickman D, Rapp CL et al (1995) Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 196:123–134. https://doi.org/10.1148/radiology.196.1.7784555
6. Sehgal CM, Weinstein SP, Arger PH, Conant EF (2006) A Review of Breast Ultrasound. J Mammary Gland Biol Neoplasia 11:113–123. https://doi.org/10.1007/s10911-006-9018-0
7. Ko ES, Lee BH, Kim H-A et al (2010) Triple-negative breast cancer: correlation between imaging and pathological findings. Eur Radiol 20:1111–1117. https://doi.org/10.1007/s00330-009-1656-3
8. Krizmanich-Conniff KM, Paramagul C, Patterson SK et al (2012) Triple receptor–negative breast cancer: imaging and clinical characteristics. Am J Roentgenol 199:458–464
9. Boisserie-Lacroix M, Macgrogan G, Debled M et al (2013) Triple-negative breast cancers: associations between imaging and pathological findings for triple-negative tumors compared with hormone receptor-positive/human epidermal growth factor receptor-2-negative breast cancers. Oncologist 18:802–811. https://doi.org/10.1634/theoncologist.2013-0380
10. Yassin NIR, Omran S, El Houby EMF, Allam H (2018) Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: a systematic review. Comput Methods Programs Biomed 156:25–45. https://doi.org/10.1016/j.cmpb.2017.12.012

11. Sehgal CM, Cary TW, Kangas SA et al (2004) Computer-based margin analysis of breast sonography for differentiating malignant and benign masses. J Ultrasound Med 23:1201–1209
12. Huang Y-S, Takada E, Konno S et al (2018) Computer-Aided tumor diagnosis in 3-D breast elastography. Comput Methods Progr Biomed 153:201–209. https://doi.org/10.1016/j.cmpb.2017.10.021
13. Bouzghar G, Levenback BJ, Sultan LR et al (2014) Bayesian probability of malignancy with BI-RADS sonographic features. J Ultrasound Med 33:641–648. https://doi.org/10.7863/ultra.33.4.641
14. Breast Imaging Reporting Data System | American College of Radiology. https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads#Ultrasound. Accessed 23 Jul 2018
15. Goldhirsch A, Winer EP, Coates AS et al (2013) Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol 24:2206–2223. https://doi.org/10.1093/annonc/mdt303
16. Sehgal CM, Cary TW, Cwanger A et al (2012) Combined Naïve Bayes and logistic regression for quantitative breast sonography. In: 2012 IEEE International Ultrasonics Symposium. IEEE, pp 1686–1689
17. Sultan LR, Xiong H, Zafar HM et al (2015) Vascularity assessment of thyroid nodules by quantitative color Doppler ultrasound. Ultrasound Med Biol 41:1287–1293. https://doi.org/10.1016/j.ultrasmedbio.2015.01.001
18. Prat A, Perou CM (2011) Deconstructing the molecular portraits of breast cancer. Mol Oncol 5:5–23. https://doi.org/10.1016/j.molonc.2010.11.003
19. Dairkee SH, Ljung BM, Smith H, Hackett A (1987) Immunolocalization of a human basal epithelium specific keratin in benign and malignant breast disease. Breast Cancer Res Treat 10:11–20
20. Carey LA, Dees EC, Sawyer L et al (2007) The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res 13:2329–2334. https://doi.org/10.1158/1078-0432.CCR-06-1109
21. Wojcinski S, Soliman AA, Schmidt J et al (2012) Sonographic features of triple-negative and non-triple-negative breast cancer. J Ultrasound Med 31(10):1531–1541
22. Tang J, Rangayyan RM, Xu J et al (2009) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Trans Inf Technol Biomed 13:236–251. https://doi.org/10.1109/TITB.2008.2009441
23. Dromain C, Boyer B, Ferré R et al (2013) Computed-aided diagnosis (CAD) in the detection of breast cancer. Eur J Radiol 82:417–423. https://doi.org/10.1016/j.ejrad.2012.03.005
24. Xiong H, Sultan LR, Cary TW et al (2017) The diagnostic performance of leak-plugging automated segmentation versus manual tracing of breast lesions on ultrasound images. Ultrasound 25:98–106. https://doi.org/10.1177/1742271X17690425
25. Dogan BE, Turnbull LW (2012) Imaging of triple-negative breast cancer. Ann Oncol 23(Suppl 6):vi23–29. https://doi.org/10.1093/annonc/mds191
26. Shin HJ, Kim HH, Huh MO et al (2011) Correlation between mammographic and sonographic findings and prognostic factors in patients with node-negative invasive breast cancer. Br J Radiol 84:19–30. https://doi.org/10.1259/bjr/92960562
27. Yang W-T, Dryden M, Broglio K et al (2008) Mammographic features of triple receptor-negative primary breast cancers in young premenopausal women. Breast Cancer Res Treat 111:405–410. https://doi.org/10.1007/s10549-007-9810-6
28. Elkabets M, Gifford AM, Scheel C et al (2011) Human tumors instigate granulin-expressing hematopoietic cells that promote malignancy by activating stromal fibroblasts in mice. J Clin Invest 121:784–799. https://doi.org/10.1172/JCI43757
29. Li J-W, Zhang K, Shi Z-T et al (2018) Triple-negative invasive breast carcinoma: the association between the sonographic appearances with clinicopathological feature. Sci Rep 8:9040. https://doi.org/10.1038/s41598-018-27222-6
30. Yoon GY, Cha JH, Kim HH et al (2018) Sonographic features that can be used to differentiate between small triple-negative breast cancer and fibroadenoma. Ultrasonography 37:149–156. https://doi.org/10.14366/usg.17036
31. Costantini M, Belli P, Bufi E et al (2016) Association between sonographic appearances of breast cancers and their histopathologic features and biomarkers. J Clin Ultrasound 44:26–33. https://doi.org/10.1002/jcu.22312
32. Schrading S, Kuhl CK (2008) Mammographic, US, and MR imaging phenotypes of familial breast cancer. Radiology 246:58–70. https://doi.org/10.1148/radiol.2461062173
33. Verheul HMW, Voest EE, Schlingemann RO (2004) Are tumours angiogenesis-dependent? J Pathol 202:5–13. https://doi.org/10.1002/path.1473
34. Sehgal CM, Arger PH, Rowling SE et al (2000) Quantitative vascularity of breast masses by Doppler imaging: regional variations and diagnostic implications. J Ultrasound Med 19:427 – 40; quiz 441–2
35. Zhang L, Li J, Xiao Y et al (2015) Identifying ultrasound and clinical features of breast cancer molecular subtypes by ensemble decision. Sci Rep 5:11085. https://doi.org/10.1038/srep11085
36. Kumar R, Yarmand-Bagheri R (2001) The role of HER2 in angiogenesis. Semin Oncol 28:27–32
37. Dogan BE, Gonzalez-Angulo AM, Gilcrease M et al (2010) Multimodality imaging of triple receptor-negative tumors with mammography, ultrasound, and MRI. AJR Am J Roentgenol 194:1160–1166. https://doi.org/10.2214/AJR.09.2355
38. Kojima Y, Tsunoda H (2011) Mammography and ultrasound features of triple-negative breast cancer. Breast Cancer 18:146–151. https://doi.org/10.1007/s12282-010-0223-8
39. Lerma E, Peiro G, Ramón T et al (2007) Immunohistochemical heterogeneity of breast carcinomas negative for estrogen receptors, progesterone receptors and Her2/neu (basal-like breast carcinomas). Mod Pathol 20:1200–1207. https://doi.org/10.1038/modpathol.3800961
40. Yeo SH, Kim GR, Lee SH, Moon WK (2018) Comparison of ultrasound elastography and color Doppler ultrasonography for distinguishing small triple-negative breast cancer from fibroadenoma. J Ultrasound Med. https://doi.org/10.1002/jum.14564
41. Yang Q, Liu H-Y, Liu D, Song Y-Q (2015) Ultrasonographic features of triple-negative breast cancer: a comparison with other breast cancer subtypes. Asian Pac J Cancer Prev 16:3229–3232
42. Du H-Y, Lin B-R, Huang D-P (2015) Ultrasonographic findings of triple-negative breast cancer. Int J Clin Exp Med 8:10040–10043
43. Bauer KR, Brown M, Cress RD et al (2007) Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. Cancer 109:1721–1728. https://doi.org/10.1002/cncr.22618