European Journal of Radiology 88 (2017) 141–147

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

3D quantitative breast ultrasound analysis for differentiating fibroadenomas and carcinomas smaller than 1 cm

A.S.S. Meel-van den Abeelen[a,b,*], G. Weijers[b], J.C.M. van Zelst[c], J.M. Thijssen[b], R.M. Mann[c], C.L. de Korte[b]

[a] Department of Biomechanical Engineering, MIRA-Institute, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
[b] Medical UltraSound Imaging Center (MUSIC), department of Radiology and Nuclear Medicine, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
[c] Radboud University Nijmegen Medical Centre, Department of Radiology and Nuclear Medicine, PO Box 9101, 6500 HB Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 13 May 2016
Received in revised form 2 September 2016
Accepted 5 January 2017

Keywords:
Breast cancer
Biopsies
Ultrasound
Quantitative analysis
Diagnosis

ABSTRACT

Purpose: In (3D) ultrasound, accurate discrimination of small solid masses is difficult, resulting in a high frequency of biopsies for benign lesions. In this study, we investigate whether 3D quantitative breast ultrasound (3DQBUS) analysis can be used for improving non-invasive discrimination between benign and malignant lesions.

Methods and materials: 3D US studies of 112 biopsied solid breast lesions (size <1 cm), were included (34 fibroadenomas and 78 invasive ductal carcinomas). The lesions were manually delineated and, based on sonographic criteria used by radiologists, 3 regions of interest were defined in 3D for analysis: ROI (ellipsoid covering the inside of the lesion), PER (peritumoural surrounding: 0.5 mm around the lesion), and POS (posterior-tumoural acoustic phenomena: region below the lesion with the same size as delineated for the lesion). After automatic gain correction (AGC), the mean and standard deviation of the echo level within the regions were calculated. For the ROI and POS also the residual attenuation coefficient was estimated in decibel per cm [dB/cm].

The resulting eight features were used for classification of the lesions by a logistic regression analysis. The classification accuracy was evaluated by leave-one-out cross-validation. Receiver operating characteristic (ROC) curves were constructed to assess the performance of the classification.

All lesions were delineated by two readers and results were compared to assess the effect of the manual delineation.

Results: The area under the ROC curve was 0.86 for both readers. At 100% sensitivity, a specificity of 26% and 50% was achieved for reader 1 and 2, respectively. Inter-reader variability in lesion delineation was marginal and did not affect the accuracy of the technique. The area under the ROC curve of 0.86 was reached for the second reader when the results of the first reader were used as training set yielding a sensitivity of 100% and a specificity of 40%. Consequently, 3DQBUS would have achieved a 40% reduction in biopsies for benign lesions for reader 2, without a decrease in sensitivity.

Conclusion: This study shows that 3DQBUS is a promising technique to classify suspicious breast lesions as benign, potentially preventing unnecessary biopsies.

1. Introduction

Breast cancer is one of the three most commonly diagnosed types of cancer among women [1]. During the past three decades, both X-ray mammographic screening for early stages of breast cancer and improvement in treatment options have significantly reduced breast cancer mortality and improved quality of life [2]. However, the sensitivity of mammography in women with dense breasts is seriously impaired [3], due to the fact that fibroglandular and stromal tissues (dense tissues) have the same X-ray attenuation properties as tumours and thus show equally bright on mammographic images. For extremely dense breasts, a sensitivity as low as 36% has been reported [3]. Therefore, more and more screening programs offer additional ultrasound breast screening to women with

* Corresponding author at: Department of Biomechanical Engineering, MIRA-Institute, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
E-mail addresses: aisha.vandenabeelen@radboudumc.nl, a.s.s.meel@gmail.com (A.S.S. Meel-van den Abeelen).

http://dx.doi.org/10.1016/j.ejrad.2017.01.006
0720-048X/

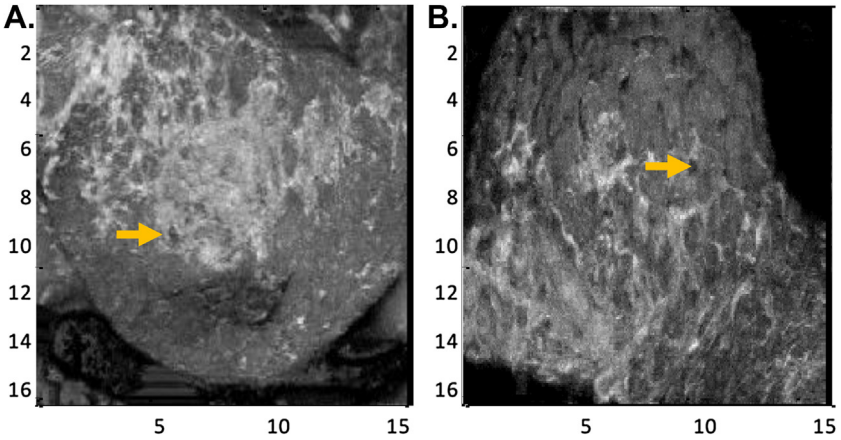

Fig. 1. Representative coronal 2D plane of two 3D ABVS scans of breasts containing suspected masses. Yellow arrow points towards the suspected lesion (A = fibroadenoma; B = invasive ductal carcinoma).

dense breasts. It has been shown that this additional ultrasound breast screening yields an additional 1.9–5.2 detected cancers per 1000 women [3–5].

Breast ultrasound has been performed for over 60 years [6]. Conventional handheld ultrasound (HHUS) is the gold standard for performing the examination. However, HHUS of the whole breast is time consuming and as it is operator-dependent, it should be performed by a trained physician. Thanks to recent advances in computer technology and 3D visualisation techniques, an automated breast volume scanner (ABUS), which is more operator independent, has been developed. The ABUS allows an automated assessment of a complete 3D breast volume and provides the possibility of an evaluation of the reconstructed breast in multiple orientations, including the standard transverse, sagittal and coronal planes. The ABUS has shown to have a high sensitivity for breast lesions [7]. However, a high number of false positives, due to considerable overlap in echographic appearance between benign and malignant masses, limits the use of ABUS for screening. The high number of false positives results in an excessive recall rate and a large number of biopsies of benign lesions [7]. This in turn yields adverse effects such as an increased anxiety among patients [8], and higher costs of the screening and clinical work-up procedures itself.

Quantitative analysis of echograms for the improvement (support) of medical diagnosis has received broad interest over the past 30 years. By analysing the texture of echographic images, additional information is gathered on the tissue type [9,10]. Quantitative analysis can be based on the spectral analysis of radio frequency echograms [11,12] or on the analysis of B-mode image statistics and texture characteristics [13,14]. In the present paper, we investigated whether 3D quantitative breast ultrasound (3DQBUS) analysis of B-mode images can be used for improving non-invasive discrimination of benign and malignant lesions.

2. Material and methods

The use of ABUS exams in this retrospective study has been approved by the institutions' local ethics committees and the need for informed consent was waived.

2.1. Equipment

The ABUS is an ultrasound system combining automation and 3D scanning of the breast. The ABUS used in this study was the ACUSON S2000 ABUS system (Siemens Medical Solutions, Mountain View, CA, USA) and consists of a scanning unit and diagnostic workstation. The scanning unit contains a high-frequency linear transducer (14L5BV linear array transducer), which is able to capture a maximum volume of $17 \times 15 \times 6\,cm^3$ in one scan. The ABUS provides optimized imaging pre-sets for volume acquisition based on the estimated size of the breast (A is the smallest size; D+ is the largest size). The central frequency of the transducer varies from 9 to 11 MHz. During imaging a frontal compression is performed on the breast by a membrane, this allows the transducer to move smoothly with a constant speed over the breast surface. In the course of a single scan the ABUS generates 318 two-dimensional slices with a layer thickness of 0.5 mm in the axial direction, of which a 3D image can be generated. Fig. 1 shows a representative coronal 2D plane of two 3D ABUS scans of breasts containing suspected masses. Technical details of the ABUS have already been described previously [7,15].

2.2. Datasets

The 3D ABUS images used in this study are a consecutive series of diagnostic cases obtained in routine clinical care at the Radiology department of the Radboud University Medical Centre (Nijmegen, The Netherlands) between October 2010 and February 2015. The whole series contains a total of 2241 studies of patients of which 423 studies contain a biopsy-confirmed mass. Biopsy confirmation was only performed by the radiologist when there was any doubt, and a reasonable suspicion of a potential malignancy, even after target ultrasound. All biopsied lesions were annotated in the 3D ABUS images using the primary radiology images and reports and biopsy results from the pathology reports. The annotation was performed by a clinical researcher (vJZ) with 3-years' experience with ABUS. Only benign fibroadenomas (n = 34) and invasive ductal carcinomas (IDC) (n = 78) appearing smaller than 1 cm on the ABUS image (as found by manual delineation of two readers; see section Quantitative analysis) were included for this initial study. We recorded Breast Imaging Reporting and Data System (BI-RADS), tumour grade, immunohistochemistry results for hormonal receptors, and the results from in-situ hybridisation for human epidermal growth factor receptor 2 (HER2) expression performed on surgical specimens. Tumour grade was assessed according to the Nottingham grading system (Bloom-Richardson-Elston). All data were anonymised. For each lesion, one single volumetric view was randomly selected. Only one dataset of a patient was used in this study if there were multiple volume data sets acquired per examination. The final outcome measure of the study was the histological diagnosis after biopsy.

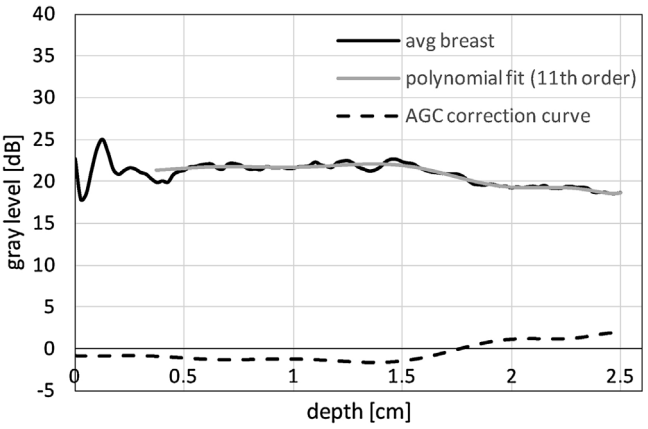

Fig. 2. Estimation of AGC. Black solid line: grey level depth profile, averaged over 20 breast images; grey line polynomial fit (11th order) to this profile; black dashed line: AGC correction curve derived from mean depth profile.

2.3. Image pre-processing

A problem with the analysis of echographic images is the depth dependence of the grey level, due to beam diffraction and focusing, as well as to the attenuation within the tissues [16]. Therefore, a calibration of the images to correct for this depth dependence is necessary. For each patient undergoing a scan with the ABUS, the US settings of the ABUS are, depending on the size of the breast of the patient, set to one of the predefined cup size settings: setting A, B, C, D, and D+.

As the main difference in cup size settings is the depth range over which the scan is taken (2.5, 3.5, 4.5, 5.0, and 5.5 cm respectively), the calibration of the images for depth dependence was done for each cup size separately. Per cup size, 20 reference ABUS images of breasts without abnormalities were used. The grey levels were plotted vs. pixel position for each reference measurement and an empirically chosen 11th-order polynomial was fitted to this curve and averaged over all reference measurements [16]. The resulting curve is called automatic gain compensation (AGC) curve.

Fig. 2 shows the AGC curve for the A cup. The scan lines of the ultrasound images used in this study are AGC corrected by subtracting the corresponding AGC curve. This results in images that no longer have an overall echo level depth dependence.

In addition to the depth dependence, the number of image grey levels per dB of the ultrasound equipment (Gamma, γ) was estimated using a tissue mimicking phantom with nominal contrast objects and was found to be 4.4 grey levels per dB. This Gamma is used to convert the grey levels of the images to echo levels (dB).

2.4. Quantitative analysis

In the past, many research groups have attempted to define reliable sonographic criteria for malignancy and benignancy. Most examiners analysed image parameters such as shape, intensity, and grey level co-occurrence matrix (GLCM) textural features of the lesion [17,18] and the coronal retraction patterns [19] around the lesion. Our 3DQBUS analysis is based on four commonly reported sonographic features for visual interpretation of lesions on breast US: internal echo pattern, echogenicity, posterior features, and the presence or absence of an echogenic halo [20–22]. To analyse these features, a custom-written Matlab (The MathWorks Inc., Natick, Massachusetts, USA) algorithm was used. The suspected mass was covered with an ellipsoid region of interest (ROI). This ellipsoid was manually chosen, by positioning two ellipses covering the suspected mass (one in the transverse plane and one in the coronal reconstruction plane; Fig. 3) and recombining these into an ellipsoid.

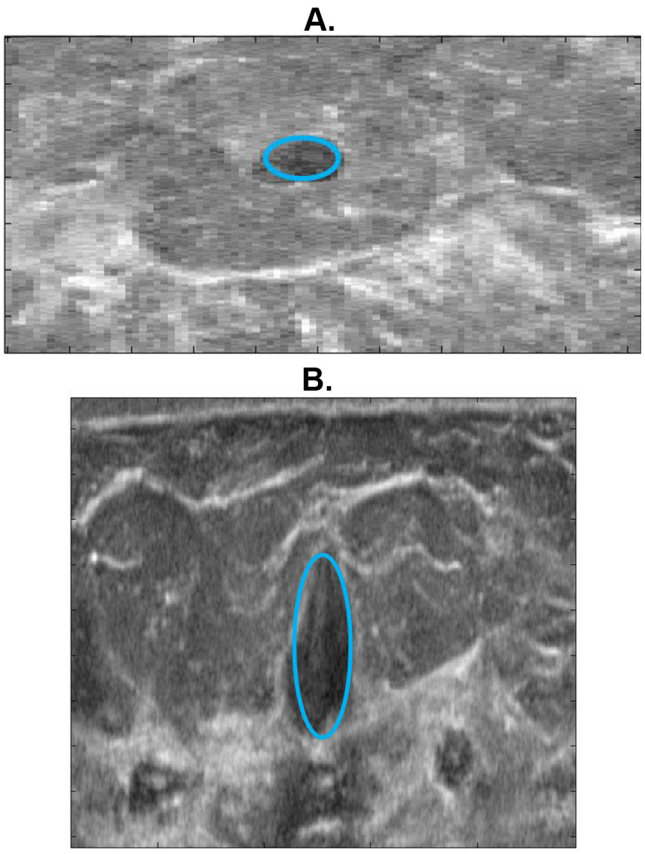

Fig. 3. Representative example of transverse (A) and coronal (B) reconstruction plane of the breast of a patient with an invasive ductal carcinoma. Scales are in cm. Ellipses in blue show the tumour delineated by Reader 1 in both the planes.

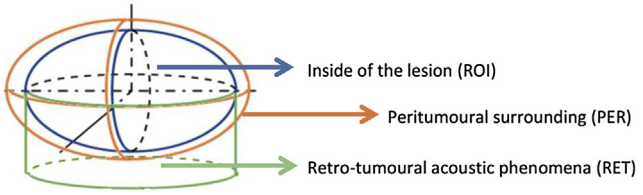

Fig. 4. Schematic view of regions of which quantitative ultrasound parameters are obtained. Blue shows the inside of the lesion (ROI) obtained by recombining the two ellipses delineated by the reader. Orange shows the peritumoural surrounding (PER) defined as a region covering 0.5 mm around the ROI. Green shows the posterior-tumoural acoustic phenomena (POS), a cylindrical region (having the same diameter as the corresponding ROI) centrally below the ROI.

Next, a region covering 0.5 mm around the ROI (peritumoural surrounding (PER)), was taken as the region in which a possible echogenic halo may be present (Fig. 4). Furthermore, a cylindrical region (having the same diameter as the corresponding ROI) centrally below the ROI was analysed as the posterior-tumoural acoustic phenomena (POS).

Eight 3DQBUS parameters were derived using the three above mentioned regions: the mean US grey level (MU [dB]) and amount of variation (VAR [dB]) in grey levels inside ellipsoid, PER, and POS, plus the residual attenuation coefficient (ATT [dB/cm]) [16] within the ellipsoid ROI and POS. After AGC correction, the attenuation coefficient was determined by the inverse slope of the linear fit through the US grey level vs. depth profile obtained by laterally averaging the grey level depth profiles within ROI and RET.

3. Comparative performance analysis

Two readers (authors AM and GW) delineated the lesions independently from each other to investigate the influence of the delineation. The performance measures reported for each reader are the area under the ROC curve (AUC), the best classification accuracy (set at no false negatives; no case labelled as benign while being malignant) (TP + TN)/(TP + TN + FP + FN) along with the associated specificity TN/(TN + FP), and positive predictive value TP/(TP + FP). TP is the number of true-positive findings (malignant lesions which are considered to be malignant), TN is the number of true-negative findings, FP is the number of false-positive findings, and FN the number of false-negative findings.

4. Statistics

Descriptive statistics (mean, SD, minimum, and maximum) of the 3DQBUS parameters were estimated for both groups (the benign and malignant lesions) and both readers using SPSS (SPSS 22, IBM Inc., Somers, NY, USA). Bar graphs with 95% confidence intervals, correlation coefficients, and *t*-tests were obtained using the same statistical package.

A logistic regression model was constructed to obtain a value for the probability of malignancy. This was done by leave-one-out cross validation, where the logistic regression model-fit procedure was iteratively applied to all samples minus the left-out sample. Using the logistic regression model outcome of the left-out sample of each iteration, sensitivity and specificity pairs for different values for the probability of malignancy were determined. These sensitivity and specificity values were used to construct a receiver operating characteristic (ROC) curve, in which the optimal sensitivity and specificity combination is visualized. An optimal threshold for the logistic regression model was set at 100% sensitivity (no false negatives allowed).

To investigate whether results are dependent on the delineation of the reader, results were validated using leave-one-out cross-validation in which the model again was trained using N-1 lesions delineated by Reader 1, but tested on the lesions as delineated by Reader 2. Statistical analysis on differences between Reader 1 and Reader 2 were carried out using the Paired-Samples *t*-test for parametric and Wilcoxon Matched-Pairs test for non-parametric data. A p value < 0.01, instead of 0.05, was considered statistically significant to correct for multiple testing.

5. Results

Table 1 summarizes the characteristics of the lesions. No differences were found in size of the lesion between the fibroadenomas and invasive ductal carcinomas. The average age of the patient having the suspected lesion did differ between the fibroadenomas and IDCs ($p < 0.01$).

Descriptive statistics of 3DQBUS parameters found by Reader 1 and Reader 2 are shown in Table 2. Paired-samples *t*-test did not reveal any significant differences in 3DQBUS parameters between Reader 1 and Reader 2. Differences were found in three of the 3DQBUS parameters: VAR of the ROI, ATT of the POS, and VAR of the PER, respectively ($p < 0.01$).

Logistic regression revealed an AUC for the discrimination between fibroadenomas and invasive ductal carcinomas of 0.86 for the first reader (Fig. 5, black line). When no false negative was allowed a specificity 26% was reached.

A weak correlation was found between the BI-RADS score and the 3DQBUS model outcome probability of malignancy (spearman's rho = 0.33, $p < 0.01$) (Fig. 6).

Table 1
Lesion characteristics.

Item	Fibroadenoma	Invasive ductal carcinoma
Number of lesions	34	78
Number of patients	20	55
Mean age of patient (y ± SD)	43 ± 9	57 ± 12*
Size lesion† (cm ± SD)		
Reader 1	0.59 ± 0.2	0.65 ± 0.2
Reader 2	0.61 ± 0.2	0.68 ± 0.2
BIRADS (n)		
2	1	1
3	18	2
4	14	28
5	1	43
6	0	1
Grade (n)		
I		14
II		39
III		23
X		3
ER positive		57
HER2 positive		15

† Size refers to the longest diameter of the mass in cm found by the delineation of the reader. ER = estrogen receptor. HER2 = human epidermal growth factor receptor 2. X is not present or not reported.

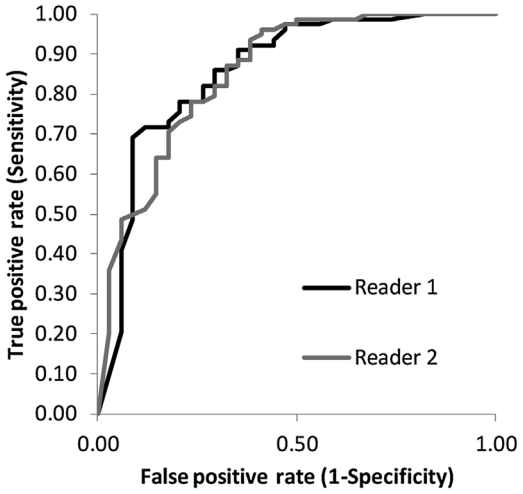

Fig. 5. Receiver operating characteristic curve (ROC) of the predictive value for the probability of malignancy. For the first reader (black line), the area under the curve equals 0.86; 95% confidence interval is 0.78–0.94; $p < 0.01$. The second reader (grey line) shows an area under the curve of 0.86; 95% confidence interval is 0.78–0.94; $p < 0.01$.

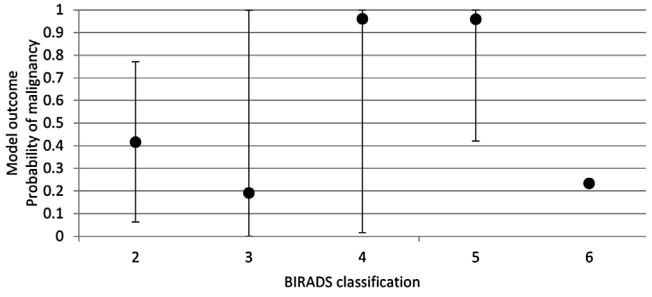

Fig. 6. Probability of malignancy for different BI-RADS classifications. Dots show group median and bars show 95% confidence intervals.

Both subgroups (the fibroadenomas and IDCs) showed no correlations between the 3DQBUS parameters and the size of the delineated lesion or age of the patient, nor was a correlation found between with the probability of having a malignant breast tumour

Table 2
Descriptive statistics of the quantitative ultrasound parameters.

			ROI			POS			PER	
Item		N	MU [dB]	VAR [dB]	ATT [dB/cm]	MU [dB]	VAR [dB]	ATT [dB/cm]	MU [dB]	VAR [dB]
Reader 1	Fibroadenoma	34								
	Mean		−5.1	17.1	−11.9	−0.4	26.9	19.4	−0.9	33.0
	SD		4.4	6.3	23.4	6.2	13.4	19.8	5.2	15.3
	Minimum		−12.5	7.4	−120.7	−8.3	5.8	−14.9	−9.4	9.3
	Maximum		3.6	30.7	32.0	13.4	60.1	99.7	8.5	61.7
	Invasive ductal carcinomas	78								
	Mean		−6.5	11.8	−9.3	−4.1	21.5	3.7	−3.9	19.9
	SD		5.5	4.4	10.0	6.5	15.8	17.0	5.6	9.8
	Minimum		−16.0	4.2	−41.3	−15.5	4.3	−29.6	−14.3	6.0
	Maximum		8.8	23.4	13.3	7.8	73.3	47.1	10.6	52.8
Independent samples *t*-test of mean (Reader 1: fibroadenomas vs. invasive ductal carcinomas)			0.19	0.00*	0.40	0.01	0.09	0.00*	0.01	0.00*
Reader 2	Fibroadenoma	34								
	Mean		−1.3	−3.6	−4.9	−0.1	−3.3	−3.1	−0.3	−3.2
	SD		1.1	0.2	0.8	1.6	0.3	1.4	1.3	0.3
	Minimum		−3.1	−4.0	−6.5	−2.7	−3.9	−4.4	−2.3	−3.8
	Maximum		1.0	−3.1	−3.3	4.4	−2.7	2.4	3.3	−2.5
	Invasive ductal carcinomas	78								
	Mean		−1.7	−3.7	−5.1	−1.0	−3.5	−4.3	−1.0	−3.4
	SD		1.4	0.2	0.5	1.6	0.4	0.8	1.4	0.3
	Minimum		−4.0	−4.0	−6.8	−4.5	−4.0	−6.4	−3.4	−3.9
	Maximum		1.8	−3.1	−4.1	2.2	−2.4	−1.9	2.3	−2.8
Independent samples *t*-test of mean (Reader 2: fibroadenomas vs. invasive ductal carcinomas)			0.40	0.01*	0.36	0.02	0.09	0.00*	0.05	0.00*
Paired samples *t*-test of mean (Reader1 vs. Reader 2)			0.97	0.34	0.81	0.78	0.76	0.43	0.87	0.61

ROI = Ellipsoid region of interest, covering the inside of the lesion. POS = posterior-tumoural acoustic phenomena. PER = peritumoural surrounding. * indicate $p < 0.01$.

Table 3
Correlations of quantitative ultrasound parameters with modelled probability of malignancy.

	ROI			POS			PER	
Item	mu [dB]	sd [dB]	att [dB/cm]	mu [dB]	sd [dB]	att [dB/cm]	mu [dB]	sd [dB]
Spearman's rho	−0.17	−0.48	0.09	−0.29	−0.28	−0.49	−0.30	−0.51
Significance (two-tailed)	0.07	0.00*	0.34	0.00*	0.00*	0.00*	0.00*	0.00*

ROI = Ellipsoid region of interest, covering the inside of the lesion. POS = posterior-tumoural acoustic phenomena. PER = peritumoural surrounding. * indicate $p < 0.01$.

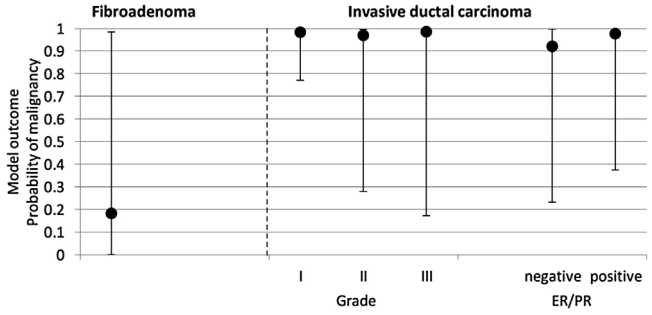

Fig. 7. Probability of malignancy for different grades of IDC and for estrogen receptor status. Dots show group median and bars show 95% confidence intervals.

and the size of the delineated lesion or age of the patient ($p > 0.01$). Relationships of the 3DQBUS parameters to the probability of having a malignant breast tumour according to the logistic regression model are shown in Table 3.

Fig. 7 shows the median and 95% confidence intervals of the model outcome probability of malignancy for IDC's of different grade and for different estrogen receptor status. No significant differences were found between the different grades nor the estrogen receptor status.

When leave-one-out cross-validation was performed in which the model was trained using N-1 lesions delineated by Reader 1 and tested on the remaining lesion delineated by Reader 2 a similar AUC of 0.86 was reached (Fig. 5, grey line). When the optimal cut-off for no false negatives by Reader 1 was used for Reader 2, a sensitivity of 100% and a specificity of 40% was obtained.

Table 4 shows the best classification accuracy when no false negatives are allowed (sensitivity of 100%). Differences induced by the subjectivity of the delineation by Reader 2 did not result in a change of the performance.

6. Discussion

This study shows that 3DQBUS is a promising technique to reclassify a proportion of suspicious breast lesions as benign so that biopsy can be avoided. In fact, the discriminative value of the logistic regression model formed out of the 3DQBUS parameters is so strong that it would have been possible to reduce the number of biopsies of benign lesion with at least 26%, without causing a single false negative case.

The large amount of biopsies for benign lesions after breast ultrasound is a serious problem. Berg et al. showed in a review of the literature that an average of 4.5% of women was recommended for biopsy after physician-performed handheld screening with breast US and that only 5.8% of the biopsies actually showed malignancy (ductal carcinoma in situ or invasive cancer) [23]. Different studies have shown that the ABUS is similar to handheld screening with US in terms of sensitivity, specificity, and accuracy [7,24]. Therefore, the high rate of false positives is also a big limitation of the ABUS. This is caused by the fact that in (3D) ultrasound, accurate discrimination of benign from malignant solid masses is difficult because of the considerable overlap of the sonographic findings [25]. To

Table 4
Classification results when no false negatives are allowed.

		Model prediction				
		Fibroadenoma	Invasive ductal carcinoma	Accuracy	Specificity	PPV
Reader 1						
Histology outcome	Fibroadenoma	9	25	70%	26%	76%
	Invasive ductal carcinoma	0	78			
Reader 2						
Histology outcome	Fibroadenoma	17	17	70%	50%	82%
	Invasive ductal carcinoma	0	78			

PPV = positive predictive value.

exclude malignancy it has even been recommended that biopsies should be performed on all solid masses seen on sonography [26].

These biopsies of benign lesions are a known harm of breast cancer screening (by evoking pain, fear and anxiety) [27,28], and lead to significant health care spendings. This study aimed to evaluate whether quantitative ultrasound parameters can be used to reduce the number of biopsies taken from fibroadenomas (the most common benign breast tumour). Considerable effort is being devoted to improve breast ultrasound for detecting solid malignant and benign masses [17–19,29,30]. Our study adds to these published systems by not focusing on the detection but on reducing the number of false-positives of suspected lesions annotated by the radiologist. Furthermore, quantitative evaluation of the peritumoural surrounding and posterior-tumoural acoustic phenomena were not included in analyses before.

For daily clinical practice, one of the great advantages of 3DQBUS is that it is non-invasive and does not require complicated analyses. It is a simple method which can be easily integrated into the workflow of a breast cancer centre with access to an ABUS. When a suspected lesion is seen on the 3D US images, only an ellipsoid has to be drawn covering the inside of lesion and the model will provide a probability for the lesion to be malignant or not. The value for this probability varies between 0 and 1. A value close to 0 means that the lesion is very unlikely to be malignant, and a value close to 1 means that the lesion is very likely to be malignant. To make it even easier, this probability can already be translated to a dichotomous classification tool for biopsy. Using the 3DQBUS analysis directly after the ABUS of the breast, it will provide extra information in just a few seconds which can be used by the radiologist for his/her decision on taking a biopsy.

There are several limitations to our study. First, only a limited number of fibroadenomas and invasive ductal carcinomas were included. As many other kinds of benign and malignant breast lesions exist, the logistic model, as presented in this paper, cannot yet be generalized for the whole range of breast cancer. However, as fibroadenomas account for a significant number of breast biopsies and invasive ductal carcinoma is the most common type of breast cancer, extending the model to other kinds of lesions will probably still allow a strong reduction of the number of biopsies. Further study with more and different kinds of malignant breast lesions is required in order to incorporate 3DQBUS in clinical practice as a complementary tool for diagnoses.

Second, breast density wasn't included in our study. As the age of the patients diagnosed with a fibroadenoma was significantly lower than those diagnosed with IDC, there is a change that the patients diagnosed with fibroadenoma also had denser breast tissue which may affect the entire echopattern surrounding the lesion. However, as no correlations were found between age and any of the 3DQBUS parameters, breast density would probably have not influenced the results of this study.

Third, it was not investigated whether the 3DQBUSUS parameters incorporated in the logistic regression model are the best choice or combination for the discrimination of fibroadenomas from IDCs. As the regions of interest were based on four often reported sonographic features for visual interpretation of breast US, we believe that the chosen 3DQBUS parameters are a good first step to extend the clinical visual interpretation. However, further study is needed to investigate whether other 3DQBUS parameters provide even better results.

Fourth, the minimal size of an ROI yielding a particular accuracy and precision for the 3DQBUS parameters was not determined. As the ROI becomes smaller also the number of independent samples that can be measured decreases, possibly reducing the accuracy and precision of the quantified 3DQBUS parameters. However, with this uncertainty it was still possible to reduce the number of biopsies using logistic regression modelling. Further study on optimal resolution for obtaining 3DQBUS may even improve the results.

Fifth, in our study the ellipses of the suspected lesions are drawn manually, the induced effect of subjectivity may reduce the accuracy and reproducibility of the model. However, both readers in our study showed good results, and the effect of manual segmentation was only marginal. On the basis of this result, the proposed model potentially may be capable of tolerating the variations in ellipsoid definition due to manual delineation by different persons.

Sixth, for this study we only used annotated suspected lesions. As literature shows that 10% to 30% of breast cancers is missed by the radiologist [31–33], detection of the lesion is also important. Further study should point out whether 3DQBUS can also be used for the detection of breast cancer or could be combined with a computer-aided detection system as described in literature [34].

Finally, the retrospective design of the study made that ABUS settings could not be modified to minimize bias due to influences of the breast size settings on the US images. While we tried to reduce this bias by including breast size as a categorical covariate (A, B, C, D, D+) in the logistic regression model, it is still possible that not all influences of the settings hereby are corrected. However, generalizing the US settings would probably only lead to a higher specificity.

7. Conclusion

This study shows that 3D QBUS is a promising technique to classify suspicious breast lesions as benign, paving the road to decreasing the number of unnecessary biopsies. It is a simple method which can be easily integrated into the workflow of a breast cancer centre with access to an ABUS. Further study on different kinds of breast lesions and extension of the model with other quantitative parameters (i.e. speckle size) may improve the biopsy reduction rate even further.

Contributors

AMA collected, analysed, and interpreted the data, and wrote and revised the manuscript for the article.

GW analysed, and interpreted the data and revised the manuscript.

CdK, RM, JvZ, and HT worked on the conception and design of the study and revised the manuscript.

All authors gave final approval of the version to be submitted.

Conflicts of interest

None.

Funding

This project has received funding from the European Union's Horizon 2020 research and *innovation programme* under grant agreement No 688188.

Role of the funding source

None.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, CA : Cancer J. Clin. 65 (1) (2015) 5–29.
[2] D.A. Berry, K.A. Cronin, S.K. Plevritis, D.G. Fryback, L. Clarke, M. Zelen, J.S. Mandelblatt, A.Y. Yakovlev, J.D. Habbema, E.J. Feuer, I. Cancer, C. Surveillance, Modeling Network, Effect of screening and adjuvant therapy on mortality from breast cancer, N. Engl. J. Med. 353 (17) (2005) 1784–1792.
[3] T.M. Kolb, J. Lichy, J.H. Newhouse, Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations, Radiology 225 (1) (2002) 165–175.
[4] W.A. Berg, J.D. Blume, J.B. Cormack, E.B. Mendelson, D. Lehrer, M. Bohm-Velez, E.D. Pisano, R.A. Jong, W.P. Evans, M.J. Morton, M.C. Mahoney, L.H. Larsen, R.G. Barr, D.M. Farria, H.S. Marques, K. Boparai, A. Investigators, Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer, JAMA 299 (18) (2008) 2151–2163.
[5] R.F. Brem, L. Tabar, S.W. Duffy, M.F. Inciardi, J.A. Guingrich, B.E. Hashimoto, M.R. Lander, R.L. Lapidus, M.K. Peterson, J.A. Rapelyea, S. Roux, K.J. Schilling, B.A. Shah, J. Torrente, R.T. Wynn, D.P. Miller, Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight Study, Radiology 274 (3) (2015) 663–673.
[6] J.J. Wild, D. Neal, Use of high-frequency ultrasonic waves for detecting changes of texture in living tissues, Lancet 260 (Mar 24) (1951) 655–657.
[7] S. Wojcinski, A. Farrokh, U. Hille, J. Wiskirchen, S. Gyapong, A.A. Soliman, F. Degenhardt, P. Hillemanns, The Automated Breast Volume Scanner (ABVS): initial experiences in lesion detection compared with conventional handheld B-mode ultrasound: a pilot study of 50 cases, Int. J. Women's Health 3 (2011) 337–346.
[8] S.W. Fletcher, Why question screening mammography for women in their forties? Radiol. Clin. North Am. 33 (6) (1995) 1259–1271.
[9] H. Liu, T. Tan, J. van Zelst, R. Mann, N. Karssemeijer, B. Platel, Incorporating texture features in a computer-aided breast lesion diagnosis system for automated three-dimensional breast ultrasound, J. Med. imaging 1 (2) (2014) 024501.
[10] T. Tan, J.J. Mordang, J. van Zelst, A. Grivegnee, A. Gubern-Merida, J. Melendez, R.M. Mann, W. Zhang, B. Platel, N. Karssemeijer, Computer-aided detection of breast cancers using Haar-like features in automated 3D breast ultrasound, Med. Phys. 42 (4) (2015) 1498–1504.
[11] J.M. Thijssen, Spectroscopy and image texture analysis, Ultrasound Med. Biol. 26 (2000) S41–S44.
[12] F.L. Lizzi, S.K. Alam, S. Mikaelian, P. Lee, E.J. Feleppa, On the statistics of ultrasonic spectral parameters, Ultrasound Med. Biol. 32 (11) (2006) 1671–1685.
[13] F.M.J. Valckx, J.M. Thijssen, A.J. van Geemen, J.J. Rotteveel, R. Mullaart, Calibrated parametric medical ultrasound imaging, Ultrason. Imaging 22 (1) (2000) 57–72.
[14] B.S. Knipp, J.A. Zagzebski, T.A. Wilson, F. Dong, E.L. Madsen, Attenuation and backscatter estimation using video signal analysis applied to B-mode images, Ultrason. Imaging 19 (3) (1997) 221–233.
[15] M. Tozaki, S. Isobe, M. Yamaguchi, Y. Ogawa, M. Kohara, C. Joo, E. Fukuma, Optimal scanning technique to cover the whole breast using an automated breast volume scanner, Japanese J. Radiol. 28 (4) (2010) 325–328.
[16] J.M. Thijssen, A. Starke, G. Weijers, A. Haudum, K. Herzog, P. Wohlsein, J. Rehage, C.L. De Korte, Computer-aided B-mode ultrasound diagnosis of hepatic steatosis: a feasibility study, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 55 (6) (2008) 1343–1354.
[17] W.K. Moon, Y.-W. Shen, C.-S. Huang, L.-R. Chiang, R.-F. Chang, Computer-aided diagnosis for the classification of breast masses in automated whole breast ultrasound images, Ultrasound Med. Biol. 37 (4) (2011) 539–548.
[18] J.H. Kim, J.H. Cha, N. Kim, Y. Chang, M.-S. Ko, Y.-W. Choi, H.H. Kim, Computer-aided detection system for masses in automated whole breast ultrasonography: development and evaluation of the effectiveness, Ultrasonography 33 (2) (2014) 105–115.
[19] T. Tan, B. Platel, R. Mus, L. Tabar, R.M. Mann, N. Karssemeijer, Computer-aided Detection of Cancer in Automated 3d Breast Ultrasound, 2013.
[20] P.B. Guyer, K.C. Dewbury, D. Warwick, J. Smallwood, I. Taylor, Direct contact B-scan ultrasound in the diagnosis of solid breast masses, Clin. Radiol. 37 (5) (1986) 451–458.
[21] J.C. Bamber, L. De Gonzalez, D.O. Cosgrove, P. Simmons, J. Davey, J.A. McKinna, Quantitative evaluation of real-time ultrasound features of the breast, Ultrasound Med. Biol. 14 (Suppl 1) (1988) 81–87.
[22] A.T. Stavros, D. Thickman, C.L. Rapp, M.A. Dennis, S.H. Parker, G.A. Sisney, Solid breast nodules: use of sonography to distinguish between benign and malignant lesions, Radiology 196 (1) (1995) 123–134.
[23] W.A. Berg, E.B. Mendelson, Technologist-performed handheld screening breast US imaging: how is it performed and what are the outcomes to date? Radiology 272 (1) (2014) 12–27.
[24] H.Y. Wang, Y.X. Jiang, Q.L. Zhu, J. Zhang, Q. Dai, H. Liu, X.J. Laj, Q. Sun, Differentiation of benign and malignant breast lesions: a comparison between automatically generated breast volume scans and handheld ultrasound examinations, Eur. J. Radiol. 1 (11) (2012) 3190–3200.
[25] Y.F. Zhou, Ultrasound diagnosis of Breast cancer, J. Med. Imaging Health Inf. 3 (2) (2013) 157–170.
[26] R.J. Hooley, L.M. Scoutt, L.E. Philpotts, Breast ultrasonography state of the art, Radiology 268 (3) (2013) 642–659.
[27] J. Brodersen, V.D. Siersma, Long-term psychosocial consequences of false-positive screening mammography, Ann. Fam. Med. 11 (2) (2013) 106–115.
[28] E.A. Sickles, Probably benign breast lesions: when should follow-up be recommended and what is the optimal follow-up protocol? Radiology 213 (1) (1999) 11–14.
[29] W.K. Moon, Y.W. Shen, M.S. Bae, C.S. Huang, J.H. Chen, R.F. Chang, Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images, IEEE Trans. Med. Imaging 32 (7) (2013) 1191–1200.
[30] R.P. Candelaria, L. Hwang, R.R. Bouchard, G.J. Whitman, Breast ultrasound: current concepts, Seminars in ultrasound, Semin. ultrasound CT MR 34 (3) (2013) 213–225.
[31] R.E. Bird, T.W. Wallace, B.C. Yankaskas, Analysis of cancers missed at screening mammography, Radiology 184 (3) (1992) 613–617.
[32] M.L. Giger, Computer-aided diagnosis in radiology, Acad. Radiol. 9 (1) (2002) 1–3.
[33] K. Kerlikowske, P.A. Carney, B. Geller, M.T. Mandelson, S.H. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg, R. Ballard-Barbash, Performance of screening mammography among women with and without a first-degree relative with breast cancer, Ann. Intern. Med. 133 (11) (2000) 855–863.
[34] T. Tan, B. Platel, R. Mus, L. Tabar, R.M. Mann, N. Karssemeijer, Computer-aided detection of cancer in automated 3-D breast ultrasound, IEEE Trans. Med. Imaging 32 (9) (2013) 1698–1706.