The role of Shear-Wave elastography in the differentiation of benign and malign non-mass lesions of the breast

Ann Ital Chir, 2018 89, 5: 385-391
pii: S0003469X18027495
Epub Ahead of Print 23 July 2018
free reading: www.annitalchir.com

Hulya Aslan*, Aysin Pourbagher*, Merve Ozen**

**Baskent University Faculty of Medicine, Department of Radiology, Adana, Turkey*
***Merve Ozen, MD, Rush University Medical Center, Department of Interventional Radiology, Chicago, Illinois, USA*

The role of shear wave elastography in the differentiation of benign and malign non-mass lesions of the breast

AIM: *The aim of this study is to retrospectively assess the additional diagnostic value of shear wave elastography (SWE) added to ultrasound (US) versus US alone in differentiating malignant and benign non-mass lesions (NMLs) of the breast by readers with different experience levels and to assess interreader agreement.*
MATERIAL AND METHODS: *This retrospective study enrolled 53 NMLs (31 benign, 22 malignant). Two radiologists (readers 1 and 2 had 15 years and 1 year of experience in breast imaging, respectively) independently reviewed each study and determined the BI-RADS category using US alone and again after adding SWE to US. Diagnostic performances of US alone and US combined with SWE were compared for both readers. Areas under receiver operating characteristic curves (AUCs) were estimated. The levels of interobserver agreement were determined by the calculated kappa coefficient.*
RESULTS: *With the addition of SWE to US, AUCs for differentiation of benign and malignant lesions increased significantly for the less experienced reader (0.56 vs. 0.79; p=0.028), but not for the more experienced reader (0.60 vs. 0.75; p=0.170). While evaluating US alone, the interobserver agreement was moderate, and the interobserver variability among the readers was statistically significant (k=0.493, p<0.001). After adding SWE, the agreement increased to 0.773, and the interobserver variability among the readers became non-significant (k=0.773, p=0.688)*
CONCLUSIONS: *SWE increased the diagnostic performance of relatively less experienced reader significantly. SWE improved interobserver agreement of two readers with different levels of experience and reduced the interobserver variability in differentiating benign and malignant NMLs of the breast.*

KEY WORDS: Breast, Elastography, Ultrasound

Introduction

The American College of Radiology (ACR) developed the Breast Imaging-Reporting and Data System (BI-RADS) lexicon to standardize the characterization of breast lesions by mammography, ultrasound (US) and magnetic resonance Imaging (MRI) [1]. The BI-RADS US classification includes mass lesions, calcifications, and associated features like edema and skin retraction. However, not all breast lesions meet the criteria of a mass lesion when seen in two different views.

Previously reported non-mass lesions (NMLs) include not only benign lesions such as mastitis, epithelial hyperplasia, sclerosing adenosis, fibrocystic changes, and radial scarring, but also malignant lesions such as ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), and mucinous adenocarcinoma [2-4]. Therefore, deciding how to approach NMLs and when to biopsy NMLs is very important. Shear-Wave elastography (SWE) has been shown to have additional value in differentiating both mass lesions and

Pervenuto in Redazione Giugno 2017. Accettato per la pubblicazione Settembre 2017
Correspondence to: Hulya Aslan, MD, Department of Radiology, Baskent University Dr. Turgut Noyan Adana Teaching and Medical Research Center, Dadaloglu Mh, 01250, Adana, Turkey (e-mail: hul_yaaslan@hotmail.com)

NMLs of the breast [5-8]. In 2013, the ACR included elasticity criteria as an associated finding in the second edition of the BI-RADS US lexicon [1]. Presently, two main US elastography techniques are used in clinical practice: strain elastography and SWE. SWE uses shear waves generated by acoustic radiation forces, while strain elastography uses manual compression produced by the operator [9].

Reader experience level has been shown to affect the diagnostic value of mammograms and breast MRI [10-13]. Few reports describe the additional diagnostic value of SWE to US in differentiating benign and malignant NMLs of the breast [6-8]. However, there has been no representative study examining the influence of reader experience on diagnostic accuracy. To the best of our knowledge, this is the first study to evaluate the effect of reader experience level in differentiating NMLs with US alone and US plus SWE. BI-RADS aims to standardize reading and reporting of breast imaging studies, and reader experience has the potential to be an important factor to reduce interobserver variability. We therefore also aimed to compare the interobserver variability of radiologists with different levels of experience in assessing NMLs.

The aim of the present study was to assess the additional diagnostic value of SWE added to B-mode US versus US alone in differentiating malignant and benign NMLs of the breast by readers with different experience levels and to assess interreader agreement.

Materials and Methods

Study Desing and Patients

The Institutional Review Board of Baskent University Dr. Turgut Noyan Adana Teaching and Medical Research Center approved this retrospective study, and informed consent was not required. Baskent University is a Foundation University and our breast center in Adana, Turkey is a tertiary referral center. Initially, a total of 653 breast lesions from 631 consecutive patients examined by US-guided core needle biopsy and/or excisional biopsy in our breast cancer center between December 2014 and January 2017 were included in the study. All of these patients had available B-mode US and SWE images. B-mode US images of all patients were retrospectively evaluated by two radiologists (not readers 1 and 2) by consensus to reveal NMLs that did not meet the criteria of a mass lesion as seen in two different views. None of these lesions had clear margins, and none had any occupying effect in two different scanning planes. Our search revealed 61 patients with NMLs.

Patients who had been treated with neoadjuvant chemotherapy (n = 4), those who had an insufficient fine-needle core biopsy report suggesting excisional biopsy but did not undergo excisional biopsy (n = 3), and those with low-quality SWE and/or US images were excluded from the study (n = 1). Finally, we included 53 women in the study.

Image Acquisition and Image Analysis

The US equipment used at our tertiary breast cancer center consisted of an Acuson S2000 system (Siemens Medical Solutions, Erlangen, Germany) and a 9-4 linear array transducer (Siemens Medical Solutions, Erlangen, Germany). Both US machine and all the probes are checked periodically. B-mode breast US examinations and SWE were performed for all patients by one of two radiologists experienced in breast imaging. For each lesion, B-mode US and SWE were performed before biopsy or surgical excision. SWE images included virtual touch tissue imaging (VTI), virtual touch tissue quantification (VTQ), and virtual touch tissue imaging quantification (VTIQ). All core biopsies were performed under US guidance by one of two radiologists who specialized in breast imaging. All diagnoses were made by a pathologist with 15 years of experience in breast pathology.

B-mode US images of all patients were retrospectively evaluated by two radiologists (not readers 1 and 2) by consensus to reveal NMLs that did not meet the criteria of a mass lesion. These radiologists also evaluated the quality of images. The quality of the images was assessed by color-coded quality maps provided by the US system, in which the green areas were considered reliable, but the yellow and red color–coded areas were considered to be low-quality scans. Low quality scans were excluded from the study.

Two readers retrospectively and independently assessed all US and SWE images of the NMLs. Reader 1 was a breast radiologist with 15 years of experience in breast imaging. Reader 2 was a staff radiologist and had 1 year of experience interpreting breast studies (nonauthors). The readers independently estimated the likelihood of cancer and classified lesions according to the BI-RADS lexicon for US as NMLs (1: negative for malignancy; 2: benign findings; 3: probably benign; 4: suspicious abnormalities; 5: highly suggestive of malignancy). The readers were blinded to the clinical information, patient history, mammographic and other imaging findings, and histopathologic diagnoses of the lesions.

In the first session, readers evaluated only B-mode US images, and 3 weeks after the first session, they evaluated combined B-mode US and SWE images while blinded to the first reading. All of the images were evaluated on a workstation (Synapse version 4.0, Fujifilm Medical Systems Inc., Stamford, CT, USA). While evaluating SWE images, the authors made a final decision by both qualitative and quantitative images including VTI, VTQ and VTIQ. The readers downgraded BI-RADS category 3 or 4 masses with very soft elasticity to category 2 or 3. BI-RADS category 3 masses with

high elasticity were upgraded to BI-RADS category 4. Lesions with intermediate elasticity were categorized by the general impressions of the readers from B-mode and SWE images.
Diagnostic performances of B-mode US and US combined with SWE were calculated to differentiate between benign and malignant NMLs. Areas under receiver operating characteristic (ROC) curves (AUCs) were estimated.

Statistical Analysis

The normality of the distributions of continuous variables was determined by the Shapiro-Wilk test. Descriptive statistics are expressed as the mean ± standard deviation (SD) or as the number of cases and percentages where appropriate. When analyzing diagnostic performance, BI-RADS scores of 1-3 were considered benign, and scores of 4-5 were considered malignant. Interobserver agreement levels were determined by the calculated kappa coefficient. A kappa value of 0 corresponds to no agreement, a kappa value of 1.0 corresponds to complete agreement, and a kappa value of less than 0 corresponds to disagreement. Kappa values less than or equal to 0.20 indicate slight agreement, values of 0.21-0.40 indicate fair agreement, values of 0.41-0.60 indicate moderate agreement, values of 0.61-0.80 indicate substantial agreement, and values of 0.81-1.00 indicate almost perfect agreement. The differences in malignancy assessments between the first and second reader were evaluated using McNemar's test. The diagnostic performances of each reader were evaluated by ROC analyses calculating AUCs, giving the maximum sum of sensitivity and specificity for the significance test. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for each reader were also calculated. Data analysis was performed using SPSS Statistics version 17.0 software (SPSS Inc, Chicago, IL, USA). A p value less than 0.05 was considered statistically significant.

Results

The mean age of the study population was 43.9 years (SD, 10.2; range, 24–66 years). Among 53 breast lesions, 31 (58.5%) were benign and 22 (41.5%) were malignant. In our study population, IDC and DCIS were the most frequent malignancies. Among the benign lesions, mastitis was the most frequent diagnosis. Histologic diagnoses of the NMLs are shown in Table I.
Table II demonstrates the sensitivity, specificity, PPV, NPV and accuracy rates of both readers with US alone and US plus SWE for the differentiation of benign and malignant NMLs of the breast. ROC analyses (Figs. 1A, 1B) showed that adding SWE to US significantly improved diagnostic accuracy for reader 2 (the less experienced reader), increasing the AUC from 0.56 (95%

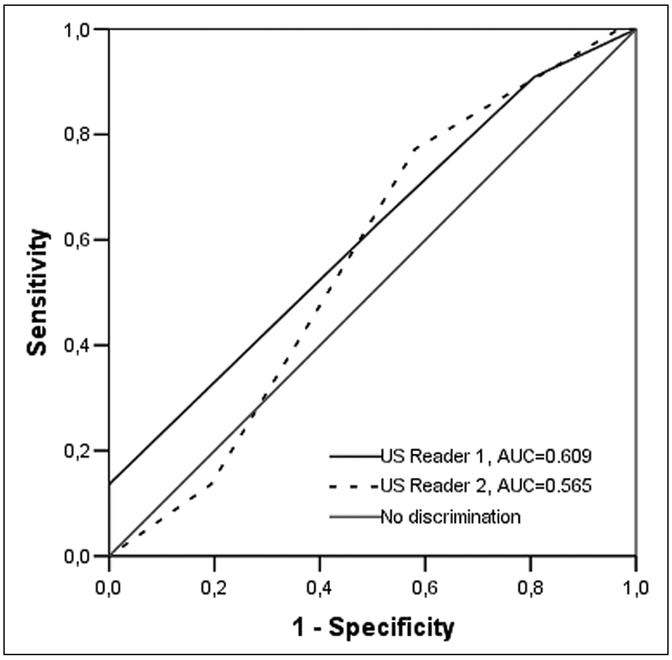

Fig. 1

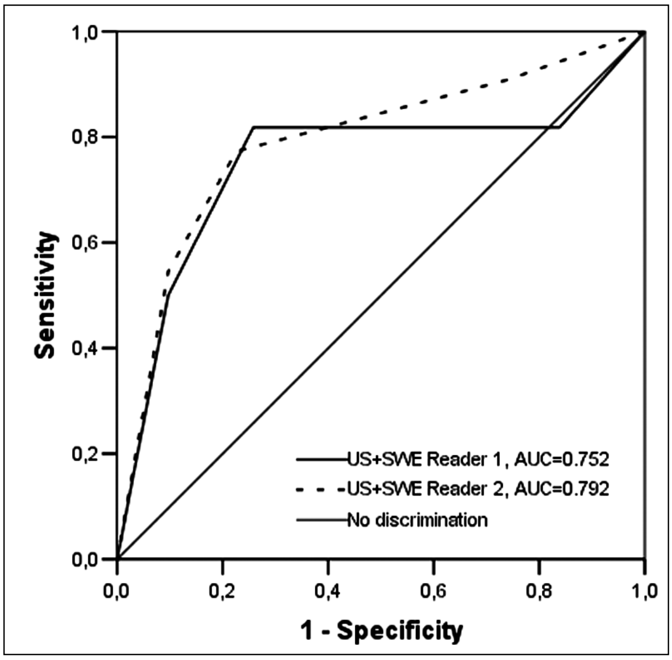

Fig. 2

Receiver operating characteristic (ROC) curves: Figures 1 and 2 show the accuracies of reader 1 and reader 2 for US alone (Figure 1) and US plus SWE (US+SWE) (Figure 2). With the addition of SWE, the area under the ROC curve (AUC) increased significantly for reader 2 (p=0.028) but not for reader 1 (p=0.170).

TABLE I - *Histologic Diagnosis of Benign and Malignant Non-mass Breast Lesions in 53 Patients*

Histopathologic Diagnosis	N. Lesions
BENIGN LESIONS (N=31)	
Fibroadenoma (complex or not)	3
Fibrocystic disease and/or ductal epithelial hyperplasia	3
Mastitis, chronic or granulomatous	11
Adenomyoepithelioma	1
Sclerosing adenosis	5
Complex sclerosing lesion	1
Normal breast tissue and ductal ectasia	2
Intraductal papilloma or papillomatosis	4
Atypical ductal epithelial hyperplasia	1
MALIGNANT LESIONS (N=22)	
Invasive ductal carcinoma (IDC)	8
Ductal carcinoma in situ (DCIS)	5
DCIS+IDC	1
Invasive labular carcinoma	2
DCIS+neuroendocrine tumor	1
Labular carcinoma in sito and adenosis	1
Tubular carcinoma	1
Angiosarcoma	1
T-cell lymphoblastic lymphoma	1
Adenoid cystic carcinoma	1
TOTAL	53

confidence interval [CI], 0.41–0.71) to 0.79 (95% CI, 0.66–0.92) (p=0.028), but not for reader 1 (p=0.170). For reader 1, the AUC increased from 0.60 (95% CI, 0.45–0.716) to 0.75 (95% CI, 0.60–0.90).
Table III summarizes the AUCs for each reader with their corresponding standard errors and 95% CIs.
The frequencies of the BI-RADS scores for US alone and US plus SWE for each reader are summarized in Table IV. By US alone, 39 of the 53 patients (73.6%) were scored in the same BI-RADS category. After adding SWE to US, 47 of the 53 patients (88.7%) were scored in the same BI-RADS category.
While differentiating benign and malignant NMLs, the interobserver agreement was moderate and the variability among the readers was statistically significant (k=0.493, p<0.001). After adding SWE the agreement increased to 0.773 and the interobserver variability among the readers became non-significant (k=0.773, p=0.688)
The interobserver agreement of the readers is summarized in Table V.
Representative images of patients were shown in figs. 3, 4.

Discussion

Our results showed that adding SWE to US increased the diagnostic performance of both readers; however the increase was only significant for the relatively less experienced reader. SWE improved the interobserver agreement of two readers with different levels of experience and reduced the interobserver variability in differentiating benign and malignant NMLs of the breast.
In routine clinical practice, radiologists are faced with an increased number of NMLs due to advances in US technologies. Both detecting and describing NMLs varies among radiologists. Frequent evaluation of NMLs of the breast by US may be difficult because of the risk of malignancy. NMLs have indistinct margins and shape. Further, it is more difficult to detect and differentially diagnose NMLs compared to mass lesions because there are no guidelines to categorize these lesions or to explain management strategies. Therefore, there is still confusion

TABLE II - *Diagnostic Performances of US Alone and US Plus SWE for Differentiation of NMLs of the Breast.*

Imaging Technique	Definitions	1st reader	2nd reader
US			
Sensitivity	TP/(TP+FN)	14/22 (63.6%)	5/22 (22.7%)
Specificity	TN/(TN+FP)	15/31 (48.4%)	18/31 (58.1%)
PPV	TP/(TP+FP)	14/30 (46.7%)	5/18 (27.8%)
NPV	TN/(TN+FN)	15/23 (65.2%)	18/35 (51.4%)
Accuracy	(TP+TN)/N	29/53 (54.7%)	23/53 (43.4%)
p-value †		0.556	0.246
US+SWE			
Sensitivity	TP/(TP+FN)	18/22 (81.8%)	17/22 (77.3%)
Specificity	TN/(TN+FP)	23/31 (74.2%)	24/31 (77.4%)
PPV	TP/(TP+FP)	18/26 (69.2%)	17/24 (70.8%)
NPV	TN/(TN+FN)	23/27 (85.2%)	24/29 (82.8%)
Accuracy	(TP+TN)/N	41/53 (77.4%)	41/53 (77.4%)
p-value †		<0.001	<0.001

PPV: positive predictive value, NPV: negative predictive value, TP: true positive, FN: false negative, TN: true negative, FP: false positive, N: number of total cases, † continuity-corrected chi-squared test.

TABLE III - *Area Under the Receiver Operating Characteristics Curve (AUC) for Each Reader and 95% Confidence Intervals.*

	1st reader	2nd reader
US		
AUC	0.609	0.565
95% CI	0.45–0.76	0.41–0.71
p-value	0.182	0.427
US+SWE		
AUC	0.752	0.792
95% CI	0.600–0.904	0.660–0.924
p-value	**0.002**	**<0.001**

AUC: Area under the curve, CI: confidence interval.

TABLE IV - *Frequency of each BI-RADS Score for US Alone and US Plus SWE* by Reader.

	1st reader	2nd reader
BI-RADS US	(n)	(n)
BI-RADS 2	8 (15.1%)	9 (17.0%)
BI-RADS 3	15 (28.3%)	26 (49.0%)
BI-RADS 4	27 (50.9%)	17 (32.1%)
BI-RADS 5	3 (5.7%)	1 (1.9%)
BI-RADS US+SWE		
BI-RADS 2	9 (17.0%)	10 (18.9%)
BI-RADS 3	18 (34.0%)	19 (35.8%)
BI-RADS 4	12 (22.6%)	9 (17.0%)
BI-RADS 5	14 (26.4%)	15 (28.3%)

TABLE V - *Interobserver Agreement of the Readers for US Alone and US Plus SWE.*

	1st reader				
	Benign	Malignant	Total	p-value †	K
2nd reader					
US				<0.001	0.493
Benign	22 (41.5%)	13 (24.5%)	35 (66.0%)		
Malignant	1 (1.9%)	17 (32.1%)	18 (34.0%)		
Total	23 (43.4%)	30 (56.6%)	53 (100.0%)		
2nd reader					
US+SWE				0.688	0.773
Benign	25 (47.2%)	4 (7.5%)	29 (54.7%)		
Malignant	2 (3.8%)	22 (41.5%)	24 (45.3%)		
Total	27 (51.0%)	26 (49.0%)	53 (100.0%)		

† McNemar's test, : kappa coefficient.

in describing and managing NMLs [2, 14]. B-mode US features of benign and malignant breast NMLs have been analyzed previously [2-4]. Previous studies included NMLs appearing as lesions that were mostly identified based on abnormalities of the ducts, hypoechoic areas in the mammary glands, architectural distortion, multiple small cysts, or echogenic foci without a hypoechoic area. Thus, it is suggested that a unified terminology should be developed to standardize image interpretation and reporting

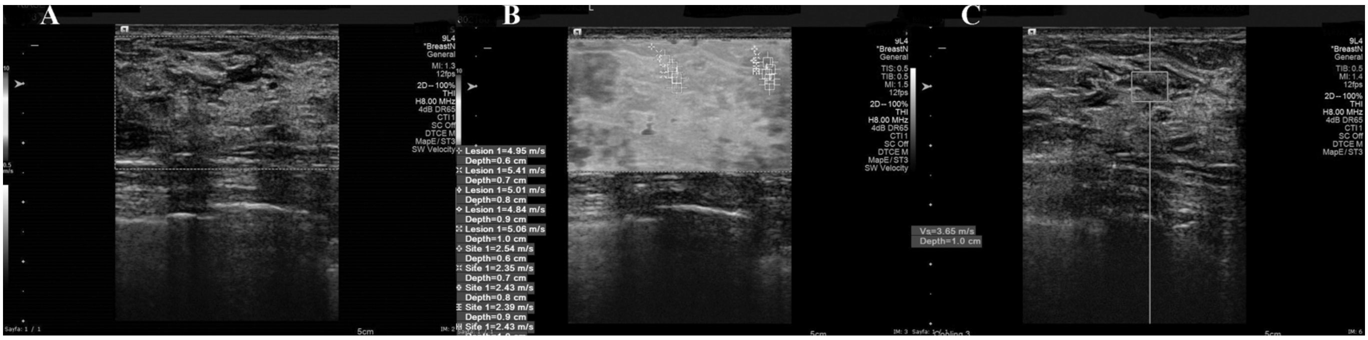

Fig. 3: Ductal epithelial hyperplasia in a 47-year-old woman (A) B-mode US images shows non-mass lesion assessed as BI-RADS category 3. After adding VTIQ (B) and VTQ (C) images, the lesion was categorized as BI-RADS 3 by both readers.

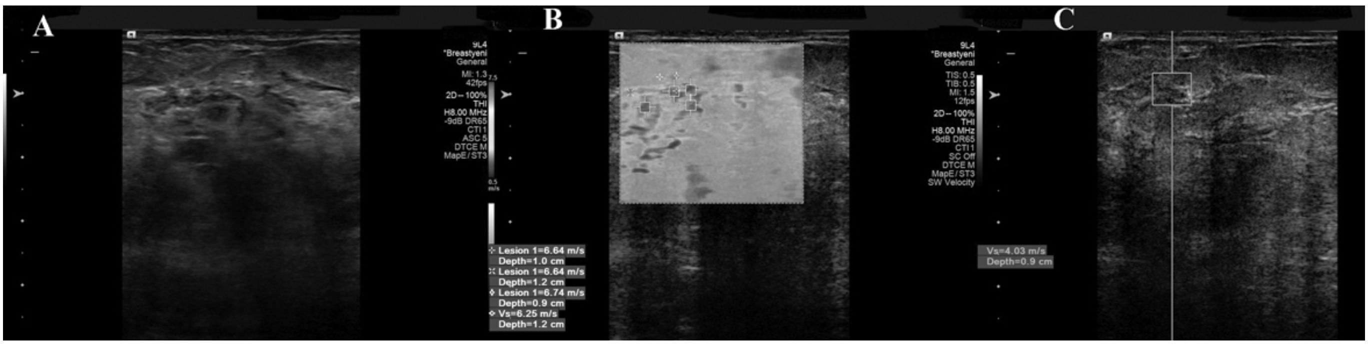

Fig. 4: Invasive ductal carcinoma in a 57-year-old woman (A) B-mode US images shows non-mass lesion assessed as BI-RADS category 3. After adding VTIQ (B) and VTQ (C) images, the lesion was categorized as BI-RADS 4 by both readers.

of NMLs on US and to facilitate communication of final assessment categories that clearly indicate management recommendations. Our findings show that SWE improved interobserver agreement in the BI-RADS categorization of these lesions and that SWE provided an appropriate categorization of NMLs and was useful in clarifying the indications for biopsy of these lesions.

SWE has been previously reported to provide additional functional information, improve diagnostic performances in differentiating benign and malignant NMLs, and avoid unnecessary biopsies [6]. Wang et al showed that the combination of conventional US and SWE can reduce unnecessary benign biopsies of NML s [8]. Our results are in agreement with these studies.

Correct image interpretation, also known as "observer performance", depends on various personal characteristics such as each reader's experiences and caseloads [12]. In studies assessing observer performance, experience factors such as the number of years reading mammograms, number of mammograms read per year, and hours reading mammograms per week were positively correlated with performance [12, 15-18]. A prior study in MRI reported that less reader experience negatively affected the diagnostic performance of breast, especially in differentiating NMLs, and especially for the least experienced radiologists [10]. The additional diagnostic value of SWE is more prominent in the less experienced radiologist. This finding is important because a less experienced reader has an increased risk of misevaluating lesions. Our results showed that sensitivity, specificity, PPV, NPV and accuracy rates of the less experienced reader was lower than the more experienced reader with US alone (Table II). However, after adding SWE all these parameters reached the more experienced reader's assessments. Based on these results, adding SWE to US should be recommended, especially for less experienced readers who are evaluating NMLs.

The BI-RADS lexicon is the most widely used and accepted reporting guideline. It was developed to ensure accurate communication and interpretation of clinical images [1,19]. In our study, the interreader agreement increased from moderate to substantial after adding SWE. Additionally, adding SWE resulted in interobserver variability becoming non-significant between readers with different levels of experience. This increase in agreement and decrease in interobserver variability for NMLs identification furthers the BI-RADS aims of standardizing the reporting of breast images. In the future, if the BI-RADS includes not only mass lesions of the breast but also NMLs, adding SWE findings as criteria may be helpful.

In this study, we evaluated only B-mode US and SWE images of the patients. However, in routine practice, clinicians do not evaluate US alone, but rather they are aware of each patient's history, physical examination findings, and other imaging findings such as mammography or MRI. In particular, most of our cases have mammograms available. Microcalcifications were shown to increase the risk of malignancy among NMLs [4]. Evaluating patient mammogram results, patient history, and physical examination findings may improve the diagnostic performance of US. For example, if an NML has microcalcifications that are suspicious for malignancy, we perform a biopsy regardless of the benign findings on B-mode US or SWE. Therefore, the additional value of combined imaging modalities may result in greater diagnostic improvement than shown in our results. In the future, combined BI-RADS classifications may be suggested.

Our study has several limitations, including its retrospective design and small sample size. In addition, we included only patients who had pathology results, but in routine clinical practice we sometimes detect NMLs in patients who do not have pathology results. We do not perform biopsies on all lesions, and we chose histopathology results as our standard of reference. These factors limited the numbers of BI-RADS 1, 2 and 3 lesions in the study.

Conclusions

In conclusion, our results are in agreement with those of the abovementioned studies and support the conclusion that adding SWE to US improves both accuracy and interreader agreement in differentiating benign and malignant NMLs. However, the increase was significant only for the relatively inexperienced reader. SWE improved agreement between readers with differing levels of experience and reduced interobserver variability

Riassunto

Scopo del presente studio è stato quello di determinare il valore diagnostico aggiunto della shear wave elastography (SWE) in aggiunta agli ultrasuoni (US) a confronto con i soli US per differenziare la natura maligna e benigna delle lesioni non solide (NML) della mammella da parte di esaminatori con diversi livelli di esperienza, e di definire la convergenza di interpretazione.

Si tratta di uno studio retrospettivo eseguito su 53 NML arruolate (31 benigne e 22 maligne). Due radiologi (esaminatore 1 e 2 con esperienza nell'imaging mammario rispettivamente di 15 e di 1 anno, hanno esaminato indipendentemente la singola documentazione determinando la categoria BI-RAD usando soltanto gli US e nuovamente dopo aver sommato la SWE agli US.

La efficienza degli US da soli ed associati alla SWE è stata paragonata per entrambi gli esaminatori. Sono state valutate le aree sotto le curve operative caratteristiche del ricevente. Il grado di concordanza tra i due osservatori è stato determinato mediante il calcolo del coefficiente K.

Con la sommazione delle SWE agli US , la differenziazione delle AUC delle lesioni benigne e maligne è risultata significativamente accresciuta per gli esaminatori con meno esperienza (0,56 vs 0,79; p=0.028) ma non per gli esaminatori con maggiore esperienza (0,60 vs 0,75; p=0.170).
Considerando i soli US l'accordo tra gli osservatori è risultato moderato, e la variabilità tra gli esaminatori è risultata statisticamente significativa (k=0,493, p<0,001). Aggiungendo le SWE la concordanza si è incrementata a 0,773 a la variabilità tra gli esaminatori è diventata non significativa (k=0,773, p=0.688).
Si conclude per un incremento significativo dell'efficienza diagnostica degli esaminatori relativamente meno esperti con le SWE, che hanno determinato un miglioramento della concordanza tra i due gruppi di osservatori di diversa esperienza e una riduzione della variabilità tra loro nel differenziare NML benigne e maligne della mammella.

References

1. D'Orsi CJ, Sickles EA, Mendelson EB, et al: Radiology; 2013.

2. Uematsu T: *Non-mass-like lesions on breast ultrasonography: a systematic review.* Breast Cancer, 2012; 19:295-301.

3. Wang ZL, Li N, Li M, Wan WB: *Non-mass-like lesions on breast ultrasound: classification and correlation with histology.* Radiol Med, 2015; 120:905-10.

4. Ko KH, Hsu HH, Yu JC, Peng YJ, Tung HJ, Chu CM, et al: *Non-mass-like breast lesions at ultrasonography: Feature analysis and BI-RADS assessment.* Eur J Radiol, 2015; 84:77-85.

5. Chang JM, Moon WK, Cho N, Yi A, Koo HR, Han W, et al:. *Clinical application of shear wave elastography (SWE) in the diagnosis of benign and malignant breast diseases.* Breast Cancer Res Treat, 2011; 129:89-97.

6. Choi JS, Han BK, Ko EY, Ko ES, Shin JH, Kim GR: *Additional diagnostic value of shear-wave elastography and color Doppler US for evaluation of breast non-mass lesions detected at B-mode US.* Eur Radiol, 2016; 26:3542-549.

7. Ko KH, Jung HK, Kim SJ, Kim H, Yoon JH: *Potential role of shear-wave ultrasound elastography for the differential diagnosis of breast non-mass lesions: Preliminary report*. Eur Radiol, 2014; 24: 305-11.

8. Wang ZL, Li Y, Wan WB, Li N, Tang J: *Shear-wave Elastography: Could it be helpful for the diagnosis of non-mass-like breast lesions?* Ultrasound Med Biol, 2017; 43:83-90.

9. Barr RG. *Sonographic breast elastography: A primer*. J Ultrasound Med, 2012. 31:773-83.

10. Baltzer PA, Kaiser WA, Dietzel M: *Lesion type and reader experience affect the diagnostic accuracy of breast MRI: A multiple reader ROC study*. Eur J Radiol, 2015; 84:86-91.

11. Marino MA, Clauser P, Woitek R, Wengert GJ, Kapetas P, Bernathova M, et al: *A simple scoring system for breast MRI interpretation: Does it compensate for reader experience?* Eur Radiol, 2016; 26: 2529-537.

12. Tavakoli Taba S, Hossain L, Heard R, Brennan P, Lee W, Lewis S: *Personal and network dynamics in performance of knowledge workers: a study of australian breast radiologists.* PLoS One. 2016;. 11:e0150186.

13. Lee AY, Wisner DJ, Aminololama-Shakeri S, Arasu VA, Feig SA, Hargreaves J, et al: *Inter-reader Variability in the Use of BI-RADS Descriptors for suspicious findings on diagnostic mammography: a multi-institution study of 10 academic radiologists*. Acad Radiol, 2017; 24:60-6.

14. Sotome K, Yamamoto Y, Hirano A, Takahara T, Hasegawa S, Nakamaru M, et al: *The role of contrast enhanced MRI in the diagnosis of non-mass image-forming lesions on breast ultrasonography*. Breast Cancer, 2007; 14:371-80.

15. Miglioretti DL, Smith-Bindman R, Abraham L, Brenner RJ, Carney PA, Bowles EJ, et al: *Radiologist characteristics associated with interpretive performance of diagnostic mammography.* J Natl Cancer Inst, 2007; 99:1854-863.

16. Reed WM, Lee WB, Cawson JN, Brennan PC: *Malignancy detection in digital mammograms: Important reader characteristics and required case numbers.* Acad Radiol, 2010; 17:1409-13.

17. Suleiman WI, Rawashdeh MA, Lewis SJ, McEntee MF, Lee W, Tapia K, et al.: *Impact of Breast Reader Assessment Strategy on mammographic radiologists' test reading performance.* J Med Imaging Radiat Oncol, 2016; 60:352-58.

18. Rawashdeh MA, Lee WB, Bourne RM, Ryan EA, Pietrzyk MW, Reed WM, et al.: *Markers of good performance in mammography depend on number of annual readings*. Radiology, 2013; 269:61-7.

19. Reston, VA: American College of Radiology (2010) *The ACR practice guideline for communication of diagnostic imaging findings*. (Revised 2010, Resolution 11) pp. 1-6.