Added Value of Shear-Wave Elastography for Evaluation of Breast Masses Detected with Screening US Imaging[1]

Su Hyun Lee, MD
Jung Min Chang, MD
Won Hwa Kim, MD
Min Sun Bae, MD
Mirinae Seo, MD
Hye Ryoung Koo, MD
A. Jung Chu, MD
Hye Mi Gweon, MD
Nariya Cho, MD
Woo Kyung Moon, MD

Purpose: To evaluate the additional value of shear-wave elastography (SWE) to B-mode ultrasonography (US) and to determine an appropriate guideline for the combined assessment of screening US-detected breast masses.

Materials and Methods: This study was conducted with institutional review board approval, and written informed consent was obtained. From March 2010 to February 2012, B-mode US and SWE were performed in 159 US-detected breast masses before biopsy. For each lesion, Breast Imaging Reporting and Data System (BI-RADS) category on B-mode US images and the maximum stiffness color and elasticity values on SWE images were assessed. A guideline for adding SWE data to B-mode US was developed with the retrospective cohort to improve diagnostic performance in sensitivity and specificity and was validated in a distinct prospective cohort of 207 women prior to biopsy.

Results: Twenty-one of 159 masses in the development cohort and 12 of 207 breast masses in the validation cohort were malignant. In the development cohort, when BI-RADS category 4a masses showing a dark blue color or a maximum elasticity value of 30 kPa or less on SWE images were downgraded to category 3, specificity increased from 9.4% (13 of 138) to 59.4% (82 of 138) and 57.2% (79 of 138) ($P < .001$), respectively, without loss in sensitivity (100% [21 of 21]). In the validation cohort, specificity increased from 17.4% (34 of 195) to 62.1% (121 of 195) and 53.3% (104 of 195) ($P < .001$) respectively, without loss in sensitivity (91.7% [11 of 12]).

Conclusion: The addition of SWE to B-mode US improved diagnostic performance with increased specificity for screening US-detected breast masses. BI-RADS category 4a masses detected at US screening that showed a dark blue color or a maximum elasticity value of 30 kPa or less on SWE images can be safely followed up instead of performing biopsy.

[1] From the Department of Radiology, Seoul National University Hospital, Seoul National University College of Medicine, 28 Yongon-dong, Chongno-gu, Seoul 100-744, Korea. Received October 18, 2013; revision requested November 20; revision received March 25, 2014; accepted April 14; final version accepted April 22. Supported by the Seoul National University Hospital Research Fund (grant 04-2011-0280). **Address correspondence to** J.M.C. (e-mail: *imchangjm@gmail.com*).

Ultrasonography (US) is used as a supplemental screening test in women with dense breast tissue and in high-risk women who cannot undergo magnetic resonance (MR) imaging for any reason (1). Although mammography remains the reference standard for breast cancer detection, in women with dense breasts, mammographic sensitivity may decrease to as low as 30%–48% (2,3) because noncalcified breast cancers are often obscured by the surrounding and overlying dense parenchyma. Supplemental screening US has the potential to depict these small, node-negative breast cancers not seen at mammography and has been shown to increase cancer detection by an additional 4.2 cancers per 1000 high-risk women versus mammography alone. However, a substantial increase in the number of false-positive findings has remained a major limitation of screening breast US (4,5). Therefore, reducing unnecessary biopsies and short-term follow-up is an important consideration for breast screening with US.

Advances in Knowledge

- Shear-wave elastography (SWE) improved the diagnostic performance of B-mode US to distinguish benign from malignant breast masses detected by screening: when Breast Imaging Reporting and Data System category 4a masses showing a dark blue color or a maximum elasticity value of 30 kPa or less on SWE images were downgraded to category 3, specificity of B-mode US increased from 17.4% (34 of 195) to 62.1% (121 of 195) and 53.3% (104 of 195), respectively ($P <$ .001), without loss in sensitivity.
- In our prospective cohort, screening US-detected breast cancers were small (mean, 0.8 cm; range, 0.4–1.7 cm) with relatively low elasticity values (mean, 84.3 kPa; range, 15.3–192.6 kPa).

Shear-wave elastography (SWE) is a quantitative elastographic technique based on the local estimation of shear-wave propagation speed. The elastic modulus of soft tissue, which is proportional to the square of shear-wave speed, is calculated and displayed as a color-coded image in real time (6). The SWE features that were studied thus far include both quantitative (ie, mean and maximum elasticity, lesion-to-fat elasticity ratio, size ratio relative to B-mode imaging) and qualitative assessments (ie, color assessment of maximum elasticity and homogeneity of elasticity in the mass or surrounding tissue), and several studies have shown that these SWE features can improve the specificity of B-mode US without loss of sensitivity (7–10). Recently, a multinational study of 939 masses showed that the best-performing SWE feature was a color assessment of maximum elasticity, which is correlated to maximum elasticity measured in kilopascals (7). The study suggested that a light blue color or stiffness less than 80 kPa were sufficient criteria to downgrade Breast Imaging Reporting and Data System (BI-RADS) category 4a masses to category 3, and a red color or stiffness greater than 160 kPa were sufficient criteria to upgrade BI-RADS category 3 masses to category 4a. However, those criteria were set for breast masses detected at diagnostic US, including palpable masses (376 of 939, 40%). We hypothesized that different guidelines may be needed for screening US-detected breast masses because most breast masses detected at screening US are small in size and relatively soft (4,10).

Therefore, the purpose of our study was to evaluate the additional value of SWE to B-mode US and to determine an appropriate guideline for the combined assessment of screening US-detected breast masses.

Implication for Patient Care

- SWE is useful to evaluate screening US-detected breast masses as an addition to conventional B-mode US because it decreases the need for biopsy of US category 4a lesions.

Materials and Methods

The institutional review board approved this prospective study and written informed consent forms were provided by all participating women. Our hospital offered SWE in addition to B-mode US for women who were scheduled for biopsy or surgery since March 2010.

Patients and Breast Lesions

Development cohort.—Between March 2010 and February 2012, 166 women who underwent SWE before percutaneous core needle biopsy for masses detected by using US imaging were retrospectively identified. Four women with breast implants and three women without quantitative elasticity measurements were excluded. In nine women with multiple breast lesions, only the most suspicious single mass was included from each woman. Finally, 159 women (mean age, 45.6 years ± 9.9; age range, 21–79 years) with 159 breast masses (mean size, 1.1 cm ± 0.6; range, 0.3–4.0 cm) constituted the initial development cohort. All women were asymptomatic, and 25 women (15.7%) had a family history or personal history of breast cancer. Mammography was performed in 143 women with negative findings. The interval between mammography and breast US was a mean of 14 days (range, 0–64 days).

Published online before print
10.1148/radiol.14132443 **Content codes:** BR US

Radiology 2014; 273:61–69

Abbreviations:
BI-RADS = Breast Imaging Reporting and Data System
PPV = positive predictive value
SWE = shear-wave elastography

Author contributions:
Guarantor of integrity of entire study, J.M.C.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, S.H.L., J.M.C.; clinical studies, S.H.L., J.M.C., W.H.K., M.S.B., M.S., H.R.K., A.J.C., H.M.G., N.C.; statistical analysis, S.H.L., J.M.C.; and manuscript editing, S.H.L., J.M.C.

Conflicts of interest are listed at the end of this article.

Of 159 breast masses, 21 masses (13.2%) were malignant and 138 masses (86.8%) were benign. All lesions were confirmed through US-guided core needle biopsy (14-gauge automated gun or 11-gauge vacuum assisted). Among the lesions, 45 lesions were surgically excised after US-guided needle localization. The mean duration of imaging follow-up with US for the lesions with benign biopsy findings was 25 months (range, 18–42 months), and lesion stability was confirmed in all cases. This retrospective data from the development cohort were used to build a guideline for combining B-mode US and SWE, which was applied prospectively in a validation cohort.

Validation cohort.—After an interval of 2 months, we prospectively enrolled patients with screening US-detected breast masses from April to October 2012 to validate the guidelines determined from the development cohort. Among 972 women screened, we included a total of 207 consecutive women (mean age, 45.5 years ± 9.9; age range, 21–74 years) who underwent percutaneous core needle biopsy for 207 breast masses detected by using US imaging (mean size, 1.0 cm ± 0.5; range, 0.2–3.3 cm) and fulfilled the inclusion and exclusion criteria mentioned above. Among them, 23 women (11.1%) had a family history or personal history of breast cancer. As in the development cohort, all lesions were confirmed through US-guided core needle biopsy (14-gauge automated gun or 11-gauge vacuum-assisted). Surgical excision was performed in 23 lesions. The duration of imaging follow-up with US for lesions with benign biopsy findings was a mean of 15 months (range, 11–18 months), and lesion stability was confirmed in all cases.

US Examinations

B-mode US and SWE images were obtained with a US system (Aixplorer; SuperSonic Imagine, Aix en Provence, France) equipped with a 15–4-MHz linear-array transducer (SL 15-4; SuperSonic Imagine) in both development and validation cohorts. US examinations were performed by one of five radiologists (S.H.L., J.M.C., W.H.K., M.S.B., and H.R.K.) with knowledge of the clinical and mammographic findings. All radiologists had 3–7 years of breast US experience and at least 3 months of experience (at least 50 cases) with SWE of breast lesions before this study. At least two orthogonal B-mode US images were obtained for each breast mass, and the BI-RADS final assessment categories were recorded before elastographic imaging. The expected malignancy rate of BI-RADS categories are as follows: category 3 (probably benign), 2% likelihood of malignancy or less; category 4a (low suspicion of malignancy), greater than 2% to 10% likelihood of malignancy; category 4b (intermediate suspicion of malignancy), greater than 10% to 50% likelihood of malignancy; category 4c (moderate suspicion of malignancy), greater than 50% to 95% likelihood of malignancy; and category 5 (highly suggestive of malignancy), 95% or greater likelihood of malignancy. SWE images were acquired at a plane that showed the largest diameter of the breast mass. Customized presets of SWE parameters were used, with the preference of penetration mode in cases of poor penetration. A color-coded map of tissue elasticity representing the elastic modulus in kilopascals at each pixel was obtained with a default color scale ranging from 0 (dark blue; soft) to 180 kPa (red; stiff). Quantitative elasticity values were measured by using two 2-mm circular quantification regions of interest (Q-box). One was placed on the stiffest portion of the mass including immediate adjacent stiff tissue and the other, on normal fatty tissue.

Development of Guideline for Combining B-Mode US and SWE

To develop a guideline for combining B-mode US and SWE, maximum stiffness color with a four-point scoring system (1, dark blue; 2, light blue; 3, green to orange; 4, red) and maximum elasticity values of breast masses, considered to be the two best-performing SWE parameters according to a breast elastography multinational study (7), were reviewed by two radiologists (S.H.L. and J.M.C.).

We established a guideline for recommending biopsy by adding SWE to B-mode BI-RADS assessment in which BI-RADS category 4a lesions were downgraded to BI-RADS category 3 by using specific cut-off points of maximum stiffness color and maximum elasticity in the development cohort. We selected cut-off points of maximum stiffness color and maximum elasticity in which no malignancy was included (ie, conservative strategy) or that had positive predictive values (PPVs) of less than 2% (ie, aggressive strategy). We also analyzed SWE features of masses that were BI-RADS category 3 for possible recommendation of biopsy. In the validation cohort, radiologists were instructed to use these guidelines derived from the development cohort when they combined B-mode US and SWE.

Image Evaluation in Validation Cohort

For the validation cohort, radiologists who performed breast US imaging assessed the likelihood of malignancy by using B-mode US findings alone and by using the combined dataset of B-mode US and SWE. The likelihood of malignancy score was recorded from 0% to 100% combined with the BI-RADS assessment categories. Guidelines for the combined assessment of B-mode US and SWE, derived from the development cohort, were provided at the time of the examination. The radiologists were allowed to choose either conservative or aggressive strategies for a breast mass according to the examiner's decision.

Data Analysis

Mean maximum elasticity values were compared between benign and malignant masses by using the independent samples *t* test. In the development cohort, the PPV for malignancy according to the BI-RADS category and qualitative or quantitative SWE features were evaluated, and the sensitivity and specificity of B-mode US were compared with those of combined B-mode US and SWE, with the guidelines hypothetically applied, by using the McNemar test. BI-RADS category 4a or higher was considered to indicate a positive test result

for malignancy. For the validation cohort, prospective reader performance was assessed by receiver operating characteristic curve analysis with the likelihood of malignancy score (0%–100%), and it was compared between B-mode US and combined B-mode US and SWE images by using the method described by DeLong et al (11). In addition, the sensitivity and specificity of B-mode US were compared with those of combined B-mode US and SWE under the prospectively or hypothetically applied guideline.

Two-tailed *P* values of less than .05 were considered to indicate a statistically significant difference. All statistical analyses were performed by using software (MedCalc version 12.5.0.0 for Microsoft; MedCalc Software, Mariakerke, Belgium).

Results

Diagnostic Performance in the Development Cohort

Of 159 screening US-detected breast lesions, 21 were malignant (Table 1). As for SWE features, the maximum stiffness color was dark blue for 78 masses (49.0%), light blue for 47 masses (29.6%), green to orange for 25 masses (15.7%), and red for nine masses (5.7%). Malignant masses showed higher quantitative elasticity values (maximum elasticity, 119.0 kPa ± 52.2 [standard deviation]; range, 55.2–220.3 kPa) than benign masses (41.4 kPa ± 32.1; range, 5.3–172.4 kPa]) ($P < .001$) (Fig 1). The PPV of maximum stiffness color was 0% (0 of 78) for dark blue, 8.5% (four of 47) for light blue, 40.0% (10 of 25) for green to orange, and 78% (seven of nine) for red colors (Table 2). The PPV of maximum elasticity value of 30 kPa or less was 0% (zero of 74) and the PPV of maximum elasticity value of 65 kPa or less was 1.8% (two of 114). Of 114 masses with maximum elasticity of 65 kPa or less, two were malignant (one category 4a and the other category 4b). For conservative strategy, cutoff points of maximum stiffness color of dark blue or maximum elasticity of 30 kPa or less were selected, whereas

Table 1

Demographic and US Features of Development and Validation Cohorts

Parameter	Development Cohort	Validation Cohort	*P* Value
Age (y)*	45.6 ± 9.9 (21–79)	45.5 ± 9.9 (21–74)	.851
Family history of breast cancer	14 (8.8)	10 (4.8)	.140
Personal history of breast cancer	11 (6.9)	13 (6.3)	.834
Mammographic breast density			.211
Grade 1	0 (0)	0 (0)	
Grade 2	15 (10.5)	14 (7.6)	
Grade 3	90 (62.9)	105 (57.1)	
Grade 4	38 (26.6)	65 (35.3)	
US size of breast masses (cm)*			
Benign	1.1 ± 0.5 (0.3–2.9)	1.0 ± 0.5 (0.2–3.3)	.061
Malignant	1.4 ± 0.9 (0.6–4.0)	0.8 ± 0.3 (0.4–1.7)	.057
BI-RADS category			<.001
3	13 (8.2)	35 (16.9)	
4a	127 (79.9)	168 (81.2)	
4b	10 (6.3)	3 (1.4)	
4c	7 (4.4)	1 (0.5)	
5	2 (1.2)	0 (0)	
Maximum stiffness color			.900
Dark blue	78 (49.0)	108 (52.2)	
Light blue	47 (29.6)	59 (28.5)	
Green to orange	25 (15.7)	31 (15.0)	
Red	9 (5.7)	9 (4.3)	
Maximum elasticity			.657
0 to ≤30 kPa	74 (46.5)	90 (43.5)	
>30 to ≤65 kPa	40 (25.2)	64 (30.9)	
>65 to ≤150 kPa	36 (22.6)	44 (21.3)	
>150 kPa	9 (5.7)	9 (4.3)	
Pathologic result			.017
Benign	138 (86.8)	195 (94.2)	
Fibrocystic change	55 (39.9)	73 (37.4)	
Fibroadenoma	43 (31.2)	79 (40.5)	
Papilloma	15 (10.9)	15 (7.7)	
Adenosis	11 (8.0)	5 (2.6)	
Ductal epithelial hyperplasia	6 (4.3)	13 (6.7)	
Columnar cell change	3 (2.2)	6 (3.1)	
Benign phyllodes tumor	2 (1.4)	1 (0.5)	
Lobular carcinoma in situ	1 (0.7)	0 (0)	
Radial scar	1 (0.7)	0 (0)	
GLM	1 (0.7)	0 (0)	
Chronic granulomatous inflammation	0 (0)	3 (1.5)	
Malignant	21 (13.2)	12 (5.8)	
Invasive ductal carcinoma	13 (61.9)	9 (75.0)	
Invasive lobular carcinoma	1 (4.8)	0 (0)	
Tubular carcinoma	0 (0)	1 (8.3)	
Ductal carcinoma in situ	7 (33.3)	2 (16.7)	

Note.—Data are number of women unless otherwise indicated. Data in parentheses are percentages except where indicated. There were 159 women in the development cohort and 207 women in the validation cohort. GLM = granlumoatous lobular mastitis.

* Data are mean ± standard deviation; data in parentheses are ranges.

Figure 1

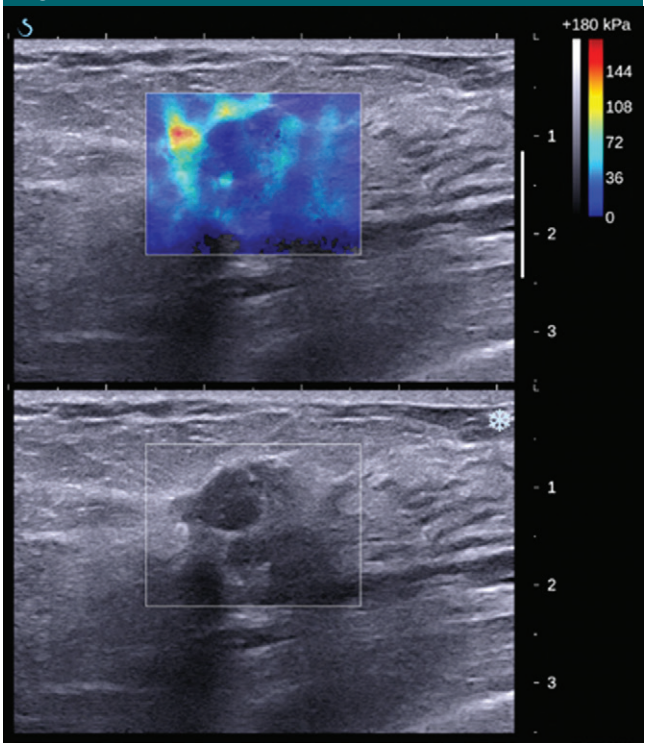

Figure 1: Images in a 66-year-old woman with grade III invasive ductal carcinoma in the development cohort. B-mode US image (bottom) shows a 1.2-cm irregular hypoechoic mass classified as BI-RADS category 4b. SWE image (top) shows high stiffness (maximum stiffness color, red) with a maximum elasticity of 153 kPa.

Table 2

Positive Predictive Values for Malignancy according to BI-RADS Category and SWE Parameters in the Development and Validation Cohorts

Parameter	Development Cohort (%)	Validation Cohort (%)
BI-RADS category		
3	0 (0/13)	2.9 (1/35)
4a	5.5 (7/127)	4.8 (8/168)
4b	60 (6/10)	66.7 (2/3)
4c	90 (6/7)	100 (1/1)
5	100 (2/2)	NA
SWE parameter		
Maximum stiffness color		
Dark blue	0 (0/78)	0.9 (1/108)
Light blue	8.5 (4/47)	10.2 (6/59)
Green to orange	40 (10/25)	9.7 (3/31)
Red	77.8 (7/9)	22.2 (2/9)
Maximum elasticity		
≤30 kPa	0 (0/74)	1.1 (1/90)
≤65 kPa	1.8 (2*/114)	3.2 (5†/154)
≥150 kPa	77.8 (7/9)	22.2 (2/9)

Note.—Data in parentheses are numerator and denominator. NA = not applicable (ie, no cases were found).

* Two malignancies included a 0.5-cm grade III invasive ductal carcinoma (BI-RADS category 4a) and a 3-cm high-grade ductal carcinoma in situ (BI-RADS category 4b).

† Five malignancies included a 1.7-cm ductal carcinoma in situ (BI-RADS category 3); a 1.0-cm low grade ductal carcinoma in situ, a 0.7-cm grade II invasive ductal carcinoma, a 0.5-cm grade I invasive ductal carcinoma (BI-RADS category 4a); and a 1.0-cm grade I invasive ductal carcinoma (BI-RADS category 4b).

maximum elasticity of 65 kPa or less was selected for aggressive strategy. We emphasized the potential effect of adding SWE features to B-mode US to downgrade BI-RADS category 4a lesions because there was no malignancy in BI-RADS category 3 masses.

The sensitivity and specificity of B-mode US were 100% (21 of 21) and 9.4% (13 of 138), respectively. Detailed information regarding sensitivities and specificities of breast US when these guidelines were hypothetically applied to the development cohort are summarized in Table 3. When BI-RADS category 4a masses were downgraded to category 3 if they had maximum stiffness color of dark blue or maximum elasticity of 30 kPa or less, the specificity increased to 59.4% (82 of 138) and 57.2% (79 of 138) ($P < .001$), respectively, without loss in sensitivity. When BI-RADS category 4a masses were downgraded to category 3 if they had maximum elasticity of 65 kPa or less, the specificity increased to 81.1% (112 of 138) ($P < .001$). There was a nonstatistically significant decrease in sensitivity to 95.2% (20 of 21) ($P = .999$). A 0.5-cm high-grade invasive ductal carcinoma with a maximum elasticity of 55.2 kPa was downgraded to category 3 by using this criterion (Fig 2).

Prospective Reader Performance in Validation Cohort

Of the 207 lesions in the validation cohort, 12 (5.8%) were malignant and 195 (94.2%) were benign. The rate of malignant masses in the validation cohort was lower than that in the development cohort ($P = .017$) (Table 1). There was one malignancy among the 35 BI-RADS category 3 masses (2.9%). The malignancy rates of BI-RADS category 4 were as follows: 4.8% (eight of 168) for category 4a, 67% (two of three) for category 4b, and 100% (one of one) for category 4c. For SWE, maximum stiffness color was dark blue for 108 masses (52.2%), light blue for 59 masses (28.5%), green to orange for 31 masses (15.0%), and red for nine masses (4.3%). Malignant masses showed higher maximum elasticity (mean, 84.3 kPa ± 54.2; range, 15.3–192.6 kPa) than did benign masses (mean, 47.1 kPa ± 38.1; range, 5.5–178.3 kPa) ($P = .002$). Of 12 malignancies, five (42%) showed maximum elasticity of 65 kPa or less (one category 3 mass with less than 30 kPa; three category 4a and one category 4b masses with 30–65 kPa) (Table 2).

The area under the receiver operating characteristic curve of B-mode US was 0.700 (95% confidence interval: 0.625, 0.761), the sensitivity was

Table 3

Diagnostic Performance of Conventional B-Mode US and B-Mode US Combined with SWE

Parameter	AUC	Sensitivity*	Specificity*
Development cohort			
B-mode US category			
3–4a	NA	100 (21/21)	9.4 (13/138)
B-mode US and SWE			
Hypothetical assessment, 4a downgrade			
E_{col} dark blue	NA	100 (21/21)	59.4 (82/138)†
$E_{max} \leq 30$ kPa	NA	100 (21/21)	57.2 (79/138)†
$E_{max} \leq 65$ kPa	NA	95.2 (20/21)	81.1 (112/138)†
Validation Cohort			
B-mode US Category			
3–4a	0.700	91.7 (11/12)	17.4 (34/195)
B-mode US and SWE			
Prospective assessment	0.879†	91.7 (11/12)	73.8 (144/195)†
Hypothetical assessment, 4a downgrade			
E_{col} dark blue	NA	91.7 (11/12)	62.1 (121/195)†‡
$E_{max} \leq 30$ kPa	NA	91.7 (11/12)	53.3 (104/195)†‡
$E_{max} \leq 65$ kPa	NA	66.7 (8/12)	79.0 (154/195)†

Note.—AUC = area under the receiver operating characteristic curve, NA = not applicable.

* Data are percentages with numbers of malignancies or numbers of masses in parentheses.

† Significant difference compared with B-mode US category ($P < .05$).

‡ Significant difference between maximum stiffness color (E_{col}) of dark blue or lighter and maximum elasticity (E_{max}) of 30 kPa or less ($P < .05$).

Figure 2

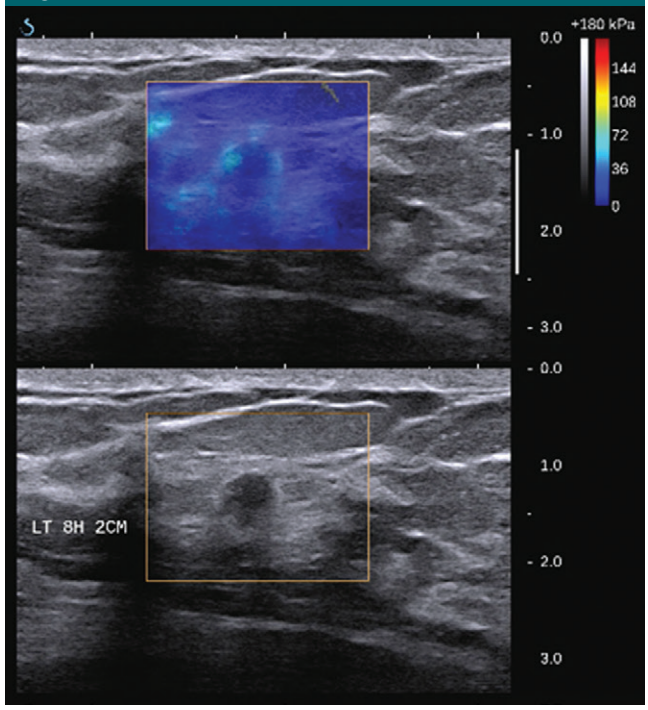

Figure 2: Images in a 57-year-old woman with grade III invasive ductal carcinoma in the development cohort. B-mode US image (bottom) shows a 0.7-cm hypoechoic mass with an indistinct margin classified as BI-RADS category 4a. SWE image (top) shows mild stiffness (maximum stiffness color, light blue) with maximum elasticity of 55.2 kPa. The mass would be downgraded to BI-RADS category 3 by using an aggressive strategy.

92% (11 of 12), and the specificity was 17.4% (34 of 195) (Table 3). When SWE elasticity data were combined with B-mode US by the radiologists by using one of the guidelines regarding downgrading BI-RADS category 4a mass that was derived from the development cohort (either conservative or aggressive strategies), the area under the receiver operating characteristic curve increased to 0.879 (95% confidence interval: 0.826, 0.920) compared with B-mode US alone ($P = .002$). Specificity increased to 73.8% (144 of 195) ($P < .001$). Regarding management decisions (ie, either follow-up or biopsy), biopsy was correctly changed to follow-up for 58.0% (110 of 195) of benign masses by SWE (Fig 3). Of 168 BI-RADS category 4a masses, 65.5% (110 of 168) were downgraded to category 3 (Fig 4). The radiologists downgraded all category 4a masses with maximum elasticity of 30 kPa or less ($n = 70$) to category 3. For 54 category 4a masses with maximum elasticity between 30 kPa and 65 kPa, radiologists chose the aggressive strategy (follow-up recommendation) in 40 masses (74%) and conservative strategy (biopsy recommendation) in 14 masses (26%). The maximum stiffness color of those 40 masses that underwent the aggressive strategy was dark blue in 17 (43%) and light blue in 23 (58%) masses, whereas the maximum stiffness color of all 14 masses that underwent the conservative strategy was light blue. As for all three category 4a cancers with maximum elasticity between 30 kPa and 65 kPa, the radiologists chose the conservative strategy. The maximum elasticity was 61.0 kPa for a 1.0-cm invasive ductal carcinoma with histologic grade II and 57.6 kPa for a 1.5-cm grade I invasive ductal carcinoma. A 1.0-cm low-grade DCIS showed a maximum elasticity of 44.9 kPa. There were no changes in the sensitivity (91.7%, 11 of 12), and biopsy decision for 12 malignant masses was the same. One malignancy with BI-RADS category 3 (low-grade ductal carcinoma in situ) showed dark blue color and maximum elasticity of 15.3 kPa (Fig 5).

Hypothetical reader performances that strictly followed one strategy are shown in Table 3. Higher specificity with lower sensitivity was achieved with the aggressive strategy than with the conservative strategy.

Discussion

Overall diagnostic performance and specificity of B-mode US for screening-detected breast masses were significantly improved without loss in sensitivity when BI-RADS category 4a

Figure 3

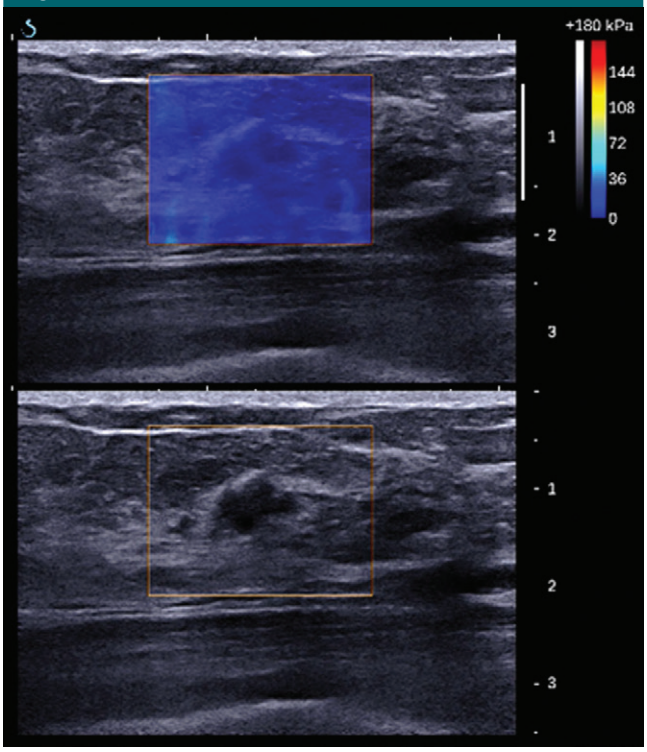

Figure 3: Images in a 44-year-old woman with fibroadenoma at core needle biopsy in the validation cohort. B-mode US image (bottom) shows an oval lobulated hypoechoic mass classified as BI-RADS category 4a. SWE image (top) shows a homogeneous maximum stiffness color of dark blue with a maximum elasticity of 21.7 kPa. When patient underwent combined assessment of B-mode US and SWE, the mass was downgraded to BI-RADS category 3 by the radiologist.

Figure 4

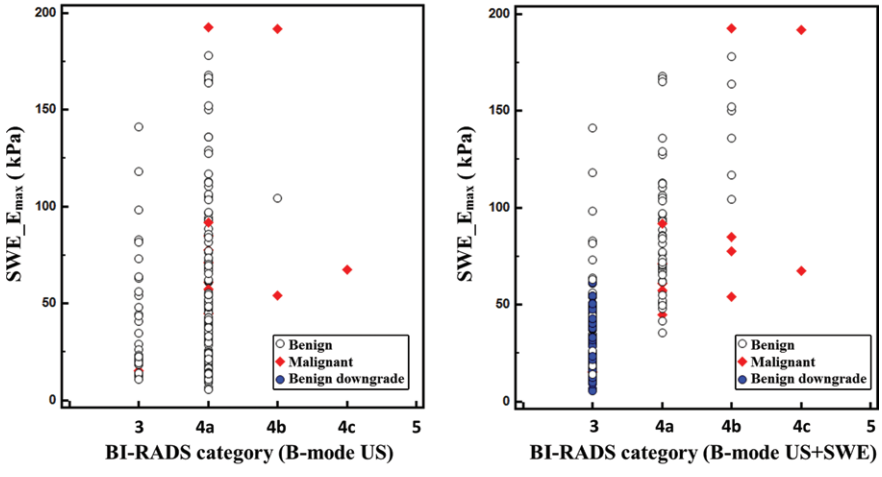

Figure 4: Scatterplots between maximum elasticity (E_{max}) values on SWE images and BI-RADS category assessed with **(a)** B-mode US and **(b)** combined dataset of B-mode US and SWE.

masses that showed a dark blue color or a maximum elasticity value of 30 kPa or less on SWE were downgraded to category 3. The radiologists were allowed to downgrade BI-RADS category 4a masses to category 3 in an aggressive strategy to achieve the highest specificity when the maximum elasticity was 30 kPa or less, or even when it was 65 kPa or less. Cut-off points of 30 kPa and 65 kPa were derived from our development cohort, in which the PPVs of SWE were 0% to 1.8%.

The strength of our study is that the results are not derived from a single cohort, they are prospectively validated with another cohort. In many studies on SWE, different cut-off values or criteria were used to differentiate malignancy and benign lesions (7–10), and even though the number of lesions included in several of the studies was large, it was not validated with a different study population. Because our development cohort was collected retrospectively and our validation group was collected prospectively, there were some differences in malignancy rate and BI-RADS final assessment category. However, the size of the screening US–detected lesions were not different among the groups, and both groups shared in common other imaging features, including the SWE pattern.

The implementation of screening US has yielded not only increased cancer detection, but also high numbers of unexpected small lesions with low PPV (4,5,12). Elastography has been suggested to help distinguish benign lesions from suspicious solid masses and should reduce false-positive findings (4). In many previous research studies, improvement in diagnostic performances occurred by adding elastography to increase specificity to discriminate benign lesions from malignancies (7,13–15). This was also noted in our study, and with the aggressive strategy, an increase in specificity was much larger than with the use of the conservative strategy.

Regarding the US-detected cancer characteristics, the results of our study revealed that screening US–detected malignancies were small in size and

Figure 5

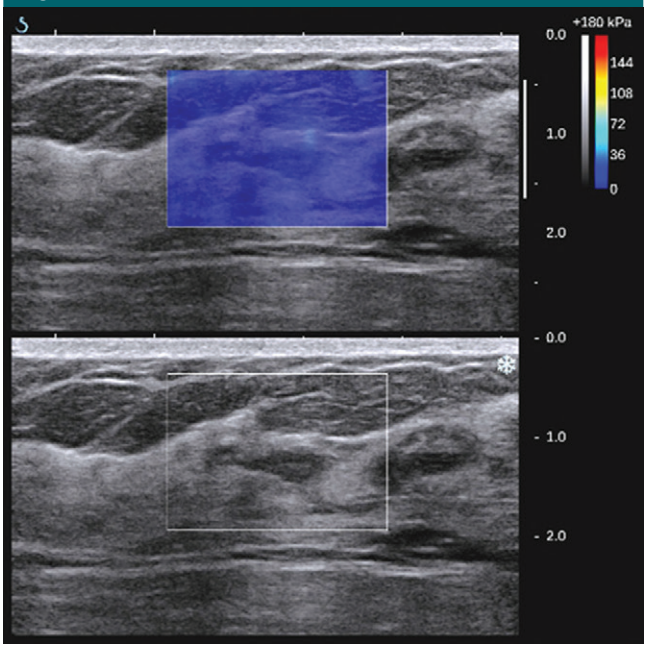

Figure 5: Images in a 43-year-old woman with low-grade ductal carcinoma in situ in the validation cohort. B-mode US image (bottom) shows an oval isoechoic mass classified as BI-RADS category 3. SWE image (top) shows a homogeneous maximum stiffness color of dark blue with maximum elasticity of 15.3 kPa. When patient underwent combined assessment of B-mode US and SWE, the mass was classified as BI-RADS category 3 by the radiologist.

showed relatively low quantitative elasticity values. This finding is consistent with previous reports (10,16–18) that showed that small lesions have lower elasticity values and higher false-negative rates. In addition, the mean elasticity values of malignant masses (84.3 and 119.0 kPa) were lower than those in previous studies (146.6 and 153.3 kPa) (8,10); the mean visual colors of malignant masses on SWE images were green to orange (72–144 kPa) in our study, whereas red colors (144–180 kPa) were seen in the previous studies when a 0–180-kPa color scale was used. As a result, our criteria for screening US–detected breast masses were lower compared with a previous study (7) that used 80 kPa to downgrade BI-RADS category 4a masses detected by using a diagnostic US setting.

In our study, the aggressive strategy was applied to more than 60% (110 of 168) of BI-RADS category 4a lesions in the validation cohort. However, all three category 4a cancers with elasticity values lower than 65 kPa were not downgraded, which was the examiner's decision. The experience of the examiner or patient factors could have affected the decision-making in terms of which guideline to use. By using the conservative strategy of downgrading BI-RADS category 4a masses with a maximum stiffness color of dark blue (or maximum elasticity $\leq$ 30 kPa) to category 3 in both the development and validation cohorts, a significant increase in specificity was achieved without sensitivity loss. As for the comparison of the two scales (maximum stiffness color and maximum elasticity) that were used for the conservative strategy, downgrading BI-RADS category 4a masses with maximum stiffness color of dark blue showed a significantly higher specificity than that of maximum elasticity of 30 kPa or less ($P < .001$). In the breast elastography multinational study, a maximum stiffness color of red or maximum elasticity of 160 kPa or greater was used for upgrading BI-RADS category 3 masses detected by using diagnostic US (7). In our study, there were no masses that fulfilled the criteria among the BI-RADS category 3 masses detected at screening US. There was one malignancy with BI-RADS category 3 that showed low elasticity value (maximum stiffness color, dark blue; maximum elasticity, 15.3 kPa). In the screening setting, the malignancy rate of BI-RADS category 3 lesions is by definition, very low, and there may be controversies in upgrading category 3 lesions to category 4a based on high elasticity values on SWE for screening US-detected lesions (19,20). Unlike our results, there were several reports in which SWE was used complementarily to B-mode US imaging to identify benign-looking invasive breast cancers (7,16). More investigation on these BI-RADS category 3 masses is warranted in future studies.

There were several limitations to our study. First, it was a single institutional study and included a relatively small number of screening US-detected malignancies. However, the small number of malignancies may befit that of a screening population, although it may impair the reliability of the statistical analysis. Second, the number of BI-RADS category 3 masses was limited. In our study, we only included screening US-detected lesions before biopsy. As a result, we only included BI-RADS category 3 masses with histologic confirmation conducted by the decision of the patient or referring physician. Third, this study was performed in the real clinical setting, thus interobserver and intraobserver agreement were not evaluated. However, it is known (21,22) that SWE is highly reproducible, and all radiologists who participated in the SWE data acquisition were instructed to reduce technical errors such as that from compression or the motion of the transducer before participating in this study. Fourth, the reasons for the decisions of the radiologists regarding choice of the appropriate guideline (ie, conservative or aggressive) and the correlation with experience were not

evaluated. Last, our decision to perform biopsy or follow up the validation cohort did not affect the clinical practice of patient management.

In conclusion, the addition of SWE to B-mode US improved diagnostic performance with increased specificity for screening US-detected breast masses. When applying SWE to screening US, we suggest the use of a conservative strategy downgrading BI-RADS category 4a masses with a maximum stiffness color of dark blue (or maximum elasticity ≤ 30 kPa) to category 3, which will potentially lead to an increase in the PPV of screening US, without missing cancers.

Acknowledgment: We thank Chris Woo, BA, for his kind assistance in editing this manuscript.

Disclosures of Conflicts of Interest: S.H.L. disclosed no relevant relationships. **J.M.C.** disclosed no relevant relationships. **W.H.K.** disclosed no relevant relationships. **M.S.B.** disclosed no relevant relationships. **M.S.** disclosed no relevant relationships. **H.R.K.** disclosed no relevant relationships. **A.J.C.** disclosed no relevant relationships. **H.M.G.** disclosed no relevant relationships. **N.C.** disclosed no relevant relationships.

References

1. Berg WA. Supplemental screening sonography in dense breasts. Radiol Clin North Am 2004;42(5):845–851, vi.
2. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92(13):1081–1087.
3. Kolb TM, Lichy J, Newhouse JH. 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 2002;225(1):165–175.
4. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008;299(18):2151–2163.
5. Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. BMC Cancer 2009;9:335.
6. Bercoff J, Tanter M, Fink M. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. IEEE Trans Ultrason Ferroelectr Freq Control 2004;51(4):396–409.
7. Berg WA, Cosgrove DO, Doré CJ, et al. Shear-wave elastography improves the specificity of breast US: the BE1 multinational study of 939 masses. Radiology 2012;262(2):435–449.
8. Athanasiou A, Tardivon A, Tanter M, et al. Breast lesions: quantitative elastography with supersonic shear imaging—preliminary results. Radiology 2010;256(1):297–303.
9. Evans A, Whelehan P, Thomson K, et al. Quantitative shear wave ultrasound elastography: initial experience in solid breast masses. Breast Cancer Res 2010;12(6):R104.
10. Chang JM, Moon WK, Cho N, et al. Clinical application of shear wave elastography (SWE) in the diagnosis of benign and malignant breast diseases. Breast Cancer Res Treat 2011;129(1):89–97.
11. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
12. Hooley RJ, Greenberg KL, Stackhouse RM, Geisel JL, Butler RS, Philpotts LE. Screening US in patients with mammographically dense breasts: initial experience with Connecticut Public Act 09-41. Radiology 2012;265(1):59–69.
13. Yi A, Cho N, Chang JM, Koo HR, La Yun B, Moon WK. Sonoelastography for 1,786 non-palpable breast masses: diagnostic value in the decision to biopsy. Eur Radiol 2012;22(5):1033–1040.
14. Gweon HM, Youk JH, Son EJ, Kim JA. Clinical application of qualitative assessment for breast masses in shear-wave elastography. Eur J Radiol 2013;82(11):e680–e685.
15. Alhabshi SM, Rahmat K, Abdul Halim N, et al. Semi-quantitative and qualitative assessment of breast ultrasound elastography in differentiating between malignant and benign lesions. Ultrasound Med Biol 2013;39(4):568–578.
16. Chang JM, Park IA, Lee SH, et al. Stiffness of tumours measured by shear-wave elastography correlated with subtypes of breast cancer. Eur Radiol 2013;23(9):2450–2458.
17. Evans A, Whelehan P, Thomson K, et al. Invasive breast cancer: relationship between shear-wave elastographic findings and histologic prognostic factors. Radiology 2012;263(3):673–677.
18. Yoon JH, Jung HK, Lee JT, Ko KH. Shear-wave elastography in the diagnosis of solid breast masses: what leads to false-negative or false-positive results? Eur Radiol 2013;23(9):2432–2440.
19. Moon HJ, Kim MJ, Kwak JY, Kim EK. Probably benign breast lesions on ultrasonography: a retrospective review of ultrasonographic features and clinical factors affecting the BI-RADS categorization. Acta Radiol 2010;51(4):375–382.
20. Berg WA, Zhang Z, Cormack JB, Mendelson EB. Multiple bilateral circumscribed masses at screening breast US: consider annual follow-up. Radiology 2013;268(3):673–683.
21. Cosgrove DO, Berg WA, Doré CJ, et al. Shear wave elastography for breast masses is highly reproducible. Eur Radiol 2012;22(5):1023–1032.
22. Lee SH, Cho N, Chang JM, et al. Two-view versus single-view shear-wave elastography: comparison of observer performance in differentiating benign from malignant breast masses. Radiology 2014;270(2):344–353.