Evaluation of Screening US–detected Breast Masses by Combined Use of Elastography and Color Doppler US with B-Mode US in Women with Dense Breasts: A Multicenter Prospective Study[1]

Su Hyun Lee, MD
Jin Chung, MD
Hye Young Choi, MD
Seon Hyeong Choi, MD
Eun Bi Ryu, MD
Kyung Hee Ko, MD
Hye Ryoung Koo, MD
Jeong Seon Park, MD
Ann Yi, MD
Ji Hyun Youk, MD
Eun Ju Son, MD
A Jung Chu, MD
Jung Min Chang, MD
Nariya Cho, MD
Myoung-jin Jang, PhD
Shin Ho Kook, MD
Eun Suk Cha, MD
Woo Kyung Moon, MD

Purpose: To investigate the value of the combined use of elastography and color Doppler ultrasonography (US) with B-mode US for evaluation of screening US–detected breast masses in women with dense breasts.

Materials and Methods: This prospective, multicenter study included asymptomatic women with dense breasts who were referred for screening US between November 2013 and December 2014. Eligible women had a newly detected breast mass at conventional B-mode US screening, for which elastography and color Doppler US were performed. The following outcome measures were compared between B-mode US and the combination of B-mode US, elastography, and color Doppler US: area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and the number of false-positive findings at screening US.

Results: Among 1021 breast masses (mean size, 1.0 cm; range, 0.3–3.0 cm) in 1021 women (median age, 45 years), 68 were malignant (56 invasive). Addition of elastography and color Doppler US to B-mode US increased the AUC from 0.87 (95% confidence interval [CI]: 0.82, 0.91) to 0.96 (95% CI: 0.95, 0.98; $P < .001$); specificity from 27.0% (95% CI: 24.2%, 29.9%) to 76.4% (95% CI: 73.6%, 79.1%; $P < .001$) without loss in sensitivity (95% CI: −1.5%, 1.5%; $P > .999$); and PPV from 8.9% (95% CI: 7.0%, 11.2%) to 23.2% (95% CI: 18.5%, 28.5%; $P < .001$), while avoiding 67.7% (471 of 696) of unnecessary biopsies for nonmalignant lesions.

Conclusion: Addition of elastography and color Doppler US to B-mode US can increase the PPV of screening US in women with dense breasts while reducing the number of false-positive findings without missing cancers.

Online supplemental material is available for this article.

[1] From the Department of Radiology (S.H.L., J.M.C., N.C., W.K.M.) and Medical Research Collaborating Center (M.J.J.), Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea; Department of Radiology, Ewha Womans University Mokdong Hospital, Seoul, Korea (J.C., E.S.C.); Department of Radiology, Gyeongsang National University Hospital, Jinju, Gyeongsangnam-do, Korea (H.Y.C.); Department of Radiology, Kangbuk Samsung Hospital, Seoul, Korea (S.H.C., S.H.K.); Department of Radiology, Dongnam Institute of Radiological and Medical Science, Busan, Korea (E.B.R.); Department of Radiology, Bungdang CHA Hospital, Seongnam, Gyeonggi-do, Korea (K.H.K.); Department of Radiology, Hanyang University Hospital, Seoul, Korea (H.R.K., J.S.P.); Department of Radiology, Seoul National University Hospital Healthcare System Gangnam Center, Seoul, Korea (A.Y.); Department of Radiology, Gangnam Severance Hospital, Seoul, Korea (J.H.Y., E.J.S.); and Department of Radiology, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Korea (A.J.C.). Received October 21, 2016; revision requested January 3, 2017; revision received March 15; accepted April 6; final version accepted April 19. **Address correspondence to** W.K.M. (e-mail: *moonwk@snu.ac.kr*).

Study supported by Korean Society of Breast Imaging & Korean Society For Breast Screening (KSBI&KSFBS-2013-02).

Mammography is a standard screening test that has been proven to reduce breast cancer–related mortality (1,2). However, dense breast parenchyma reduces the sensitivity of mammography by masking noncalcified breast cancers, which causes delayed diagnosis and worse outcomes (3). Ultrasonography (US) is the most common supplemental screening modality in women with dense breasts because it is widely available, relatively inexpensive, and well tolerated by patients because it does not involve the use of ionizing radiation or intravenous contrast material injection (4,5). Screening US can depict small, node-negative invasive cancers that are occult at mammography; therefore, screening US as an adjunct to mammography can increase the sensitivity and detection rate of early cancers while reducing interval cancers in women with dense breasts (6–10). However, a low positive predictive value (PPV) with a substantial number of false-positive findings that cause unnecessary biopsies or short-interval follow-ups is a major limitation of screening US (11–13). According to the American College of Radiology Imaging Network protocol 6666 (known as ACRIN 6666), the PPV for biopsies performed after screening US in high-risk women was 7.4% (18 of 242) for incident screening, which is the current benchmark for breast US screening (14,15). However, the PPV for recall leading to biopsy is 24%–37% for tomosynthesis and 24%–50% for magnetic resonance (MR) imaging (16–19).

US elastography and color Doppler US are additional techniques used to further characterize breast masses at US and have been incorporated into the new edition of the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) lexicon (20). In addition to morphologic assessment at B-mode US, acquisition of elasticity and vascularity information of a breast mass can increase the diagnostic accuracy of US for differentiation of benign masses from malignancies. Many previous single-center studies reported that the specificity of breast US increases by either adding elastography alone or combining color Doppler US with B-mode US (21–26). A multinational Breast Elastography 1 study performed in the United States and Europe showed that the number of unnecessary biopsies for nonmalignant lesions can be reduced by the addition of elastography to B-mode US imaging, and elastography was highly reproducible to help assess elastographic features of breast masses within and across observers (27,28). However, they included a heterogeneous group of women with breast masses revealed by palpation, mammography, US, or MR imaging who were referred for breast US.

Our purpose was to investigate the value of the combined use of elastography and color Doppler US with B-mode US for evaluation of screening US–detected breast masses in women with dense breasts. We hypothesized that the diagnostic accuracy and PPV of screening US would increase with the addition of elastography and color Doppler US in women with dense breasts.

Advances in Knowledge

- The combined use of elastography and color Doppler US with B-mode US can increase the positive predictive value (PPV) for biopsy recommendation after screening US in women with dense breasts; PPV increased from 8.9% (95% confidence interval [CI]: 7.0%, 11.2%) to 23.2% (95% CI: 18.5%, 28.5%; $P < .001$) while avoiding 67.7% (471 of 696) of unnecessary biopsies for nonmalignant lesions without losing sensitivity (95% CI: −1.5%, 1.5%; $P > .999$).
- Breast Imaging Reporting and Data System (BI-RADS) category 3 and 4a masses, which are the most frequent findings and the major source of false-positive findings at screening US, can be downgraded to BI-RADS category 2 when they show negative results at both elastography and color Doppler US; up to 79% (752 of 953) of false-positive findings, unnecessary biopsies, and short-term follow-ups for nonmalignant lesions can then be avoided without missing cancers.

Implication for Patient Care

- Elastography and color Doppler US are useful for the evaluation of breast masses detected at screening US in women with dense breasts.

Materials and Methods

This prospective multicenter study was approved by the institutional review board of each recruiting site, and written informed consent was obtained from all participants between November 2013 and December 2014. Ten academic breast centers across four provinces in South Korea, where elastography and color Doppler US are routinely used for both screening and diagnostic breast US, participated this study. Investigators from each recruiting site were fully instructed regarding the study protocol, including eligibility criteria, standardized data acquisition,

https://doi.org/10.1148/radiol.2017162424

Content codes: BR US

Radiology 2017; 285:660–669

Abbreviations:
AUC = area under the receiver operating characteristic curve
BI-RADS = Breast Imaging Reporting and Data System
CI = confidence interval
DCIS = ductal carcinoma in situ
PPV = positive predictive value

Author contributions:
Guarantors of integrity of entire study, S.H.L., S.H.C., E.B.R., M.J.J., W.K.M.; 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; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, S.H.L., J.C., S.H.C., W.K.M.; clinical studies, S.H.L., J.C., H.Y.C., S.H.C., E.B.R., K.H.K., H.R.K., J.S.P., A.Y., J.H.Y., E.J.S., A.J.C., J.M.C., N.C., S.H.K., E.S.C., W.K.M.; experimental studies, J.C., S.H.C.; statistical analysis, S.H.L., M.J.J., W.K.M.; and manuscript editing, S.H.L., E.J.S., W.K.M.

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

and interpretation procedures before the start of enrollment (29). This study is registered at *ClinicalTrials.gov* (Identifier: NCT01963624).

Study Participants

Eligible participants were asymptomatic women at least 30 years of age who presented for bilateral whole-breast screening US that revealed one or more newly-detected breast masses at B-mode US classified as BI-RADS category 3 (probably benign), 4a (low suspicion for malignancy), 4b (moderate suspicion for malignancy), 4c (high suspicion for malignancy), or 5 (highly suggestive of malignancy) with a recommendation of short-interval follow-up or biopsy. Our institutions recommend screening mammography to women over 35 years of age. All patients were required to have a mammogram negative for cancer within 3 months of the US examination and heterogeneously dense or extremely dense parenchyma in at least one quadrant (30). An index lesion, in which elastography and color Doppler US were additionally performed, was defined as the single breast mass most suspicious for cancer with a maximum diameter of 3 cm in each participant. Women who had any known current malignancy, had signs or symptoms of breast cancer including palpable mass, focal pain, bloody nipple discharge, skin redness, skin retraction or nipple inversion, had undergone breast surgery within the previous 12 months, were pregnant or lactating, or had breast implants were excluded.

Image Acquisition and Interpretation

All US examinations were performed by one of 20 breast radiologists, with 2–25 years of experience with breast US and color Doppler US and 2–5 years of experience with elastography. Bilateral whole-breast screening US was performed by using one of two types of hand-held US systems equipped with a 14–6-MHz linear transducer in five sites (EUB or Hi-Vision system; Hitachi Medical, Chiba, Japan) and a system equipped with a 15–4-MHz linear transducer in six sites (Aixplorer; SuperSonic Imagine, Aix-en-Provence, France). In one of 10 study sites, both US systems were used for study enrollment. Lesions were evaluated in at least two orthogonal planes (the radial and antiradial planes or transverse and longitudinal planes) and sonographic features were described. Assessments for each lesion and for each breast overall were recorded on the basis of the expanded seven BI-RADS categories: category 1, negative; category 2, benign; category 3, probably benign; category 4a, low suspicion; category 4b, moderate suspicion; category 4c, high suspicion; and category 5, highly suggestive of malignancy (20). Eligibility was determined at the time of the examination by the radiologist who performed the screening US, and women with a final assessment category of 3, 4a, 4b, 4c, or 5 were invited to participate the study. The BI-RADS category and the probability of malignancy on a percentage scale from 0% to 100% were prospectively recorded for the index breast mass at B-mode US: category 3, 2% or less; category 4a, greater than 2% to 10%; category 4b, greater than 10% to 50%; category 4c, greater than 50% to less than 95%; and category 5, 95% or greater. Solid breast masses without features suspicious for cancer were classified as BI-RADS category 3 (31). The US criteria suggested by Yoon et al (32) was used for the subcategorization of BI-RADS category 4. Typically, an oval mass with slightly indistinct margins was considered to be BI-RADS category 4a.

After acquiring informed consent from the eligible participants, elastography and color Doppler US were performed by the same radiologist for the index breast mass, defined as the single breast lesion on B-mode US in each participant most suspicious for cancer. The order of elastography and color Doppler US was determined by the radiologist. Either strain elastographic (in five sites) or shear-wave elastographic (in six sites) images were acquired according to a US system of recruiting sites. The parameters were equal among institutions by using the same type of US system. Color Doppler US was performed by using standardized parameter settings (pulse repetition frequency between 700–1000 Hz, wall filter as low as possible [50–100 Hz], appropriate algorithm to remove motion artifacts, maximum gain [85%–90%], medium persistence, and box without angulation). Elastographic images were acquired at least twice and the most representative data with higher image quality were determined by the radiologist (33,34). The data acquisition procedure for elastography and color Doppler US took 3–5 minutes per case. Elasticity and vascularity for a lesion were interpreted according to the BI-RADS atlas (20) on the basis of qualitative assessment: elasticity was classified as soft, intermediate, and hard (Table E1 [online]); vascularity was classified as absent, vessels in rim, and internal vascularity. For intraductal and complex cystic and solid lesions, elasticity and vascularity of solid portion of the index lesion were evaluated. Failure of elastographic data acquisition was defined as inadequate or poor image quality by qualitative assessment previously described (33,34). Reasons for failure were reported subjectively by the radiologists.

Combination Criteria of B-Mode US and Elastography and/or Color Doppler US

The BI-RADS category and the probability of malignancy for the index breast mass were reassessed in consideration of elasticity and vascularity according to the predefined criteria. For combined assessment of B-mode US and elastography or B-mode US and color Doppler US, the BI-RADS category and the probability of malignancy were upgraded by one category (ie, category 3 to 4a, 4a to 4b, 4b to 4c, and 4c to 5) if a breast mass showed hard elasticity or internal vascularity, remained unchanged for a mass with intermediate elasticity or vessels in rim, and were downgraded by one category if a breast mass showed soft elasticity or absent vascularity. When both elastography and color Doppler US were combined with B-mode US, upgrading was performed for cases that showed results positive for cancer on both tests, defined as hard elasticity and internal vascularity, respectively. The BI-RADS category was not changed for cases showing discrepant results (ie, positive for cancer on one test and

negative for cancer on the other test). For cases that showed negative results on both tests, the BI-RADS category and the probability of malignancy were downgraded by one category. In cases of failed elastographic acquisition, combined assessments of B-mode US and elastography and B-mode US, elastography, and color Doppler US were determined by the results of B-mode US alone and B-mode US and color Doppler US, respectively. Management was based on recommendations of the higher of the two BI-RADS categories of B-mode US, and combined assessment of B-mode US, elastography, and color Doppler US.

Outcome Measures and Reference Standards

The primary outcome was the area under the receiver operating characteristic curve (AUC) of breast US for differentiating benign masses from malignancies. The secondary outcomes were lesion-level sensitivity, specificity, PPV, and negative predictive value of breast US when BI-RADS category 4a or higher were considered as test-positive results. We also quantified false-positive findings of screening US by calculating the number of short-term follow-ups and biopsies recommended for nonmalignant breast lesions. The reference standard was a combination of biopsy results and US findings at 2-year follow-up. Biopsy results that showed breast cancer (ie, invasive carcinoma or ductal carcinoma in situ [DCIS]) were considered to be malignant. Excision was prompted for core biopsy results of atypical or high-risk lesions including atypical ductal or lobular hyperplasia, atypical papilloma, lobular carcinoma in situ, and radial sclerosing lesion. Lesions stable or decreased at 2-year follow-up US without biopsy were considered to be nonmalignant. For increasing lesions during follow-up, biopsies were performed. A data and safety monitoring board was established to monitor the study progress every 6 months.

Statistical Analysis

The sample size was calculated according to the hypothesis that the combined use of elastography and color Doppler US would improve the AUC of B-mode US. A previous study (23) showed that the AUC of breast US increased from 0.771 to 0.844 by adding elastography and color Doppler US. We estimated that 1256 samples would be needed to show this difference with 5% significance (two sided) and 80% power while allowing for 10% missing data on the basis of a disease prevalence of 10% in the study population.

Comparisons were performed for the difference between B-mode US alone and three combined tests; B-mode US and elastography, B-mode US and color Doppler US, and B-mode US and both elastography and color Doppler US. AUC was compared by using the method of DeLong et al (35). Sensitivity and specificity were compared with the McNemar test. A generalized estimating equation was performed to compare PPV and negative predictive value. All tests were two sided, and *P* values of less than .017 were considered to indicate statistical significance by using Bonferroni correction for three comparisons in each outcome. The increase in AUC by elastography and color Doppler US was compared between subgroups according to participant and imaging characteristics by using an independent *t* test on the basis of estimates and standard errors for AUC increments. Statistical analyses were performed with statistical software (SAS version 9.3; SAS Institute, Cary, NC).

Figure 1

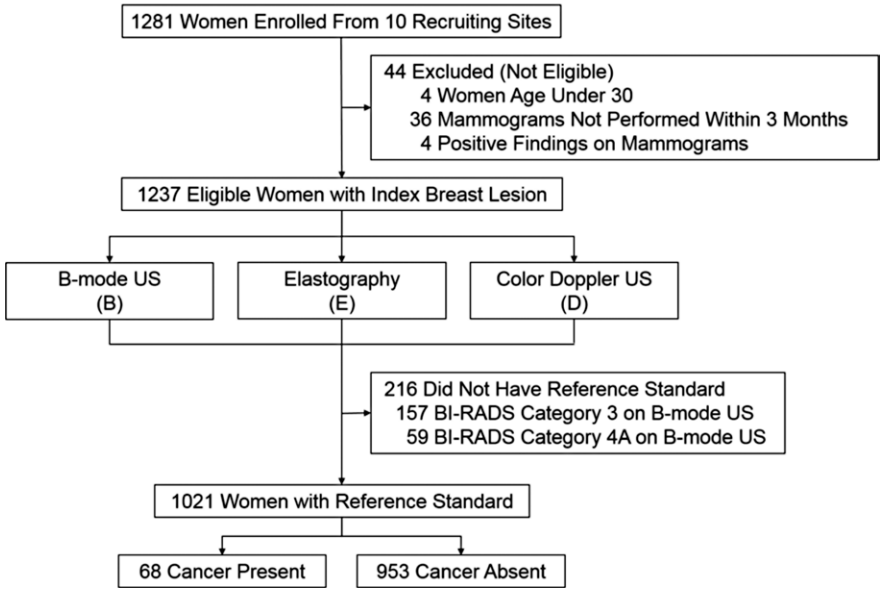

Figure 1: Study flow diagram.

Results

Of 1281 women enrolled, 1237 were eligible (Fig 1). Among them, a total of 1021 women who underwent B-mode US, elastography, and color Doppler US for the index breast lesion and had a reference standard were included in the analysis. Of 1021 women (median age, 45 years; age range, 30–84 years), 574 (56.2%) underwent prevalent US screening and the other 447 (43.8%) underwent incident US screening. Demographic and clinical characteristics of the study population are summarized in Table 1.

US Findings of Index Breast Lesions

The mean size of 1021 index breast lesions at B-mode US was 1.0 cm ± 0.4 (standard deviation; median, 0.9 cm; range, 0.3–3.0 cm). US features according to BI-RADS lexicon are summarized in Table E2 (online).

Table 1

Participant Characteristics

Parameter	Eligible (n = 1237)	Analysis Set (n = 1021)	Excluded (n = 216)
Age at enrollment (y)			
Mean	46.2 ± 8.8	46.4 ± 8.9	45.0 ± 8.6
Median*	45 (30–84)	45 (30–84)	44 (30–75)
Age group at enrollment (y)			
30–39	277 (22.4)	224 (21.9)	53 (24.5)
40–49	582 (47.0)	471 (46.1)	111 (51.4)
≥50	378 (30.6)	326 (31.9)	52 (24.1)
Menopausal status			
Premenopause	820 (66.3)	665 (65.1)	155 (71.8)
Postmenopause	355 (28.7)	313 (30.7)	47 (21.8)
Unknown	62 (5.0)	43 (4.2)	14 (6.5)
Family history of breast cancer			
Absent	1174 (94.9)	965 (94.5)	209 (96.8)
Present	63 (5.1)	56 (5.5)	7 (3.2)
First-degree relative	43 (3.5)	39 (3.8)	4 (1.8)
Other relative	20 (1.6)	17 (1.7)	3 (1.4)
Personal history of breast cancer			
Absent	1183 (95.6)	971 (95.1)	212 (98.1)
Present	54 (4.4)	50 (4.9)	4 (1.9)
Mammographic breast density†			
Heterogeneously dense	776 (62.7)	630 (61.7)	146 (67.6)
Extremely dense	461 (37.3)	391 (38.3)	70 (32.4)
Interval between mammography and US (d)			
0	727 (58.8)	541 (53.0)	186 (86.1)
1–30	406 (32.8)	376 (36.8)	30 (13.9)
>30	104 (8.4)	104 (10.2)	0 (0)
US screening			
Prevalent	713 (57.6)	574 (56.2)	139 (64.4)
Incident	524 (42.4)	447 (43.8)	77 (35.6)
Background echotexture at US‡			
Homogeneous	569 (46.0)	447 (43.8)	122 (56.5)
Heterogeneous	668 (54.0)	574 (56.2)	94 (43.5)
Size of index breast lesion on US image (cm)			
Mean	1.0 ± 0.4	1.0 ± 0.4	0.8 ± 0.3
Median*	0.9 (0.3–3.0)	0.9 (0.3–3.0)	0.8 (0.3–2.3)
Elastographic technique			
Strain elastography	527 (42.6)	428 (41.9)	99 (45.8)
Shear-wave elastography	710 (57.4)	593 (58.1)	117 (54.2)

Note.—Data are number of women, with percentages in parentheses unless otherwise indicated.

* Data in parentheses are range.

† Mammographic density was visually estimated according to the fifth edition of Breast Imaging Reporting and Data System (30).

‡ Background echotexture was classified according to the composition of fat and fibroglandular tissues on US image (20).

BI-RADS assessments according to the morphologic feature at B-mode US imaging were category 3 for 257 lesions (25.2%), category 4a for 675 lesions (66.1%), category 4b for 71 lesions (7.0%), category 4c for 11 lesions (1.1%), and category 5 for 7 lesions (0.7%) (Table 2). The elasticity of the index breast lesions was soft in 695 (68.1%), intermediate in 226 (22.1%), and hard in 88 (8.6%). For the other 12 lesions (1.2%), elastographic data acquisition failed because of deep lesion location (n = 8) or a thick breast (n = 4) (median lesion depth, 1.8 cm [range, 1.0–2.4 cm]; median breast thickness, 2.5 cm [range, 1.3–3.4 cm]). The vascularity of the index breast lesions evaluated at color Doppler US was absent in 552 lesions (54.1%), vessels in rim in 305 lesions (29.9%), and internal vascularity in 164 lesions (16.1%). When hard elasticity and internal vascularity were considered to be positive results for each test, a total of 793 among 1021 index lesions (77.7%) showed negative results at both tests; 128 lesions (12.5%) showed negative results at elastography but positive results at color Doppler US; 53 lesions (5.2%) showed positive results at elastography but negative results at color Doppler US; and 35 lesions (3.4%) showed positive results at both tests. Among the 12 cases with failed elastography acquisition, 11 showed negative results and the other one showed positive result at color Doppler US.

Histopathologic and Follow-up Results

Core-needle biopsy was performed for 844 index lesions including 17 lesions that showed increased size at follow-up. Subsequent surgical excision was performed in 178 biopsied lesions. A total of 68 lesions were diagnosed as malignant: 12 DCIS, 47 invasive ductal carcinoma, five invasive lobular carcinoma, two mixed invasive ductal and lobular carcinoma, one mucinous carcinoma, and one adenoid cystic carcinoma. More detailed histopathologic features of the detected cancers are shown in Table E3 (online). None of 17 lesions biopsied during the follow-up period were malignant. The most common histologic type of nonmalignant lesions was fibroadenoma followed by fibrocystic change and intraductal papilloma (Table E4 [online]). Follow-up US imaging was performed for 177 index lesions; 127 lesions were stable and the remaining 50 lesions decreased or disappeared. The mean follow-up duration was 27 months (range, 23–38 months). During follow-up in a 49-year-old woman, a new lesion associated with calcifications developed in a quadrant other than the index lesion of the ipsilateral breast. The lesion was diagnosed as DCIS at core-needle biopsy and subsequent surgical excision.

Table 2

Findings with B-Mode US and Elastography, Color Doppler US, and Combined Elastography and Color Doppler US

Parameter	BI-RADS Category 3		BI-RADS Category 4a		BI-RADS Category 4b		BI-RADS Category 4c		BI-RADS Category 5		Total No. of Nonmalignant Findings	Total No. of Malignant Findings
	Nonmalignant	Malignant	Nonmalignant	Malignant	Nonmalignant	Malignant	Nonmalignant	Malignant	Nonmalignant	Malignant		
Elastography												
Soft	195	0	456	5	26	11	1	0	0	1	678	17
Intermediate	51	0	153	3	11	4	0	2	0	2	215	11
Hard	9	0	32	18	6	11	3	5	0	4	50	38
Failed	2	0	8	0	0	2	0	0	0	0	10	2
Color Doppler US												
Absent	165	0	349	3	25	7	1	1	0	1	540	12
Vessels in rim	60	0	214	8	11	8	2	1	0	1	287	18
Internal vascularity	32	0	86	15	7	13	1	5	0	5	126	38
Combined E*/D†												
Negative/negative	218	0	534	0	31	8	1	0	0	1	784	9
Negative/positive	28	0	75	8	6	7	0	2	0	2	109	19
Positive/negative	5	0	22	11	5	5	2	2	0	1	34	19
Positive/positive	4	0	10	7	1	6	1	3	0	3	16	19
Fail/negative	2	0	7	0	0	2	0	0	0	0	9	2
Fail/positive	0	0	1	0	0	0	0	0	0	0	1	0
Total	257	0	649	26	43	28	4	7	0	7	953	68

Note.—Data are number of findings. The findings were from 953 nonmalignant and 68 malignant index breast masses. D = color Doppler US, E = elastography.

* Positive result at elastography was defined as hard elasticity; negative result at elastography was defined as other than hard elasticity (ie, soft or intermediate elasticity).

† Positive result at color Doppler US was defined as internal vascularity; negative result at color Doppler US was defined as other than internal vascularity or vessels in rim).

Diagnostic Performance of Breast US

The AUC of B-mode US with the seven-point BI-RADS system and the 0%–100% scale of probability of malignancy was 0.84 (95% confidence interval [CI], 0.80, 0.88) and 0.87 (95% CI: 0.82, 0.91), respectively (Table 3). When a BI-RADS category of 4a or higher was considered to be test-positive, lesion-level sensitivity, specificity, PPV (for biopsy recommendation), and negative predictive value were 100% (68 of 68), 27.0% (257 of 953), 8.9% (68 of 764), and 100% (257 of 257), respectively.

By the addition of both elastography and color Doppler US to B-mode US, the AUC for the percentage of probability of malignancy increased from 0.87 (95% CI: 0.82, 0.91) to 0.96 (95% CI: 0.95, 0.98; $P < .001$; Fig 2); specificity increased from 27.0% (95% CI: 24.2%, 29.9%) to 76.4% (95% CI: 73.6%, 79.1%; $P < .001$) without a loss in sensitivity; and PPV increased from 8.9% (95% CI: 7.0%, 11.2%) to 23.2% (95% CI: 18.5%, 28.5%; $P < .001$). When either elastography or color Doppler US was added to B-mode US, the AUC, specificity, and PPV of B-mode US also significantly increased; however, sensitivity decreased because five soft cancers and three avascular cancers assessed as BI-RADS category 4a on B-mode US were downgraded to category 3 (Table 2). Five soft cancers that showed morphologic features with low suspicion for cancer at B-mode US included three DCIS, a microinvasive carcinoma, and a 0.5-cm invasive ductal carcinoma associated with DCIS, and all of them showed internal vascularity at color Doppler US. However, the three cancers among the BI-RADS category 4a masses that showed absent vascularity at color Doppler US included one DCIS with sclerosing adenosis and two invasive ductal carcinomas (0.6 cm and 1.5 cm, tumor grade II), and all three of these cancers showed hard elasticity at elastography. Therefore, none of these cancers showed negative findings at both elastography and color Doppler US, which prevented a loss in sensitivity.

Table 3

Diagnostic Performance of B-Mode US and Combined Tests with Elastography and/or Color Doppler US

Parameter	B-Mode US Value	B + E			B + D			B + E + D		
		Combined Value	Difference (B + E − B)		Combined Value	Difference (B + D − B)		Combined Value	Difference (B + E + D − B)	
			Estimate	*P* Value		Estimate	*P* Value		Estimate	*P* Value
AUC										
BI-RADS category	0.84 (0.80, 0.88)	0.91 (0.88, 0.95)	0.07 (0.03, 0.12)	.001	0.89 (0.86, 0.93)	0.05 (0.01, 0.10)	.014	0.95 (0.93, 0.96)	0.11 (0.07, 0.15)	<.001
Probability of malignancy (%)	0.87 (0.82, 0.91)	0.92 (0.89, 0.96)	0.05 (0.01, 0.10)	.009	0.91 (0.88, 0.95)	0.04 (0.00, 0.09)	.049	0.96 (0.95, 0.98)	0.09 (0.05, 0.14)	<.001
Sensitivity (%)	100 (94.7, 100) [68/68]	92.6 (83.7, 97.6) [63/68]	−7.4 (−15.0, 0.3)	.062	95.6 (87.6, 99.1) [65/68]	−4.4 (−10.8, 1.9)	.250	100 (94.7, 100) [68/68]	0 (−1.5, 1.5)	>.999
Specificity (%)	27.0 (24.2, 29.9) [257/953]	74.0 (71.1, 76.7) [705/953]	47.0 (43.6, 50.4)	<.001	60.2 (57.0, 63.4) [574/953]	33.2 (29.7, 36.8)	<.001	76.4 (73.6, 79.1) [728/953]	49.4 (45.9, 52.9)	<.001
PPV (%)	8.9 (7.0, 11.2) [68/764]	20.3 (15.9, 25.2) [63/311]	11.4 (8.6, 14.2)	<.001	14.6 (11.5, 18.3) [65/444]	5.7 (4.2, 7.3)	<.001	23.2 (18.5, 28.5) [68/293]	14.3 (11.2, 17.4)	<.001
NPV (%)	100 (98.6, 100) [257/257]	99.3 (98.4, 99.8) [705/710]	−0.7 (−1.3, −0.1)	.025	99.5 (98.5, 99.9) [574/577]	−0.5 (−1.1, 0.1)	.082	100 (99.5, 100) [728/728]	0 (−0.6, 0.6)	>.999

Note.—Data in parentheses are 95% CIs; data in brackets are numerators and denominators. B = B-mode US, E = elastography, D = color Doppler US, NPV = negative predictive value.

False-Positive Findings at Screening US

Among 953 nonmalignant breast lesions, short-term follow-up was recommended (BI-RADS category 3) for 257 lesions, and biopsy was recommended (BI-RADS category 4a, 4b, 4c, and 5) for 696 lesions by using B-mode US alone (Table 4). By following our study protocol, upgrading or downgrading one category according to the results of elastography and color Doppler US, 67.7% (471 of 696) of unnecessary biopsies for nonmalignant lesions could be avoided by the addition of both elastography and color Doppler US to B-mode US. However, the number of short-interval follow-ups for nonmalignant lesions moderately increased, mostly by the downgrading of BI-RADS category 4a masses to category 3. Therefore, the overall number of false-positive findings from screening US modestly decreased from 953 to 749 (21.4% decrease).

The majority of masses (91.3% [932 of 1021]) were assessed as BI-RADS category 3 or 4a at B-mode US and included 26 malignancies. None of the 26 cancers showed negative findings at both elastography and color Doppler US. Therefore, a hypothetical criterion was derived: downgrade BI-RADS category 3 and category 4a masses at B-mode US that showed results negative for cancer at both elastography and color Doppler US to BI-RADS category 2 and recommend routine follow-up. When BI-RADS category 3 and category 4a masses that showed results negative for cancer at both elastography and color Doppler US were downgraded to BI-RADS category 2, the number of biopsies and short-term follow-ups recommended for nonmalignant lesions decreased, and a considerable number of false-positive findings (78.9% [752 of 953]) could be eliminated without missing cancers. The AUC and PPV of breast US also significantly increased without a loss in sensitivity according to the hypothetical criteria (Table E5 [online]).

Subgroup Analyses by Participant and Imaging Characteristics

The AUC increment by elastography and color Doppler US was not significantly related to the participant's age group (<50

Figure 2

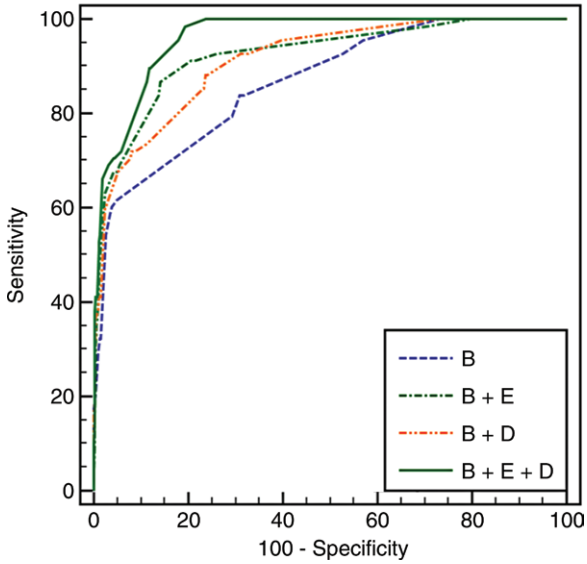

Figure 2: Receiver operating characteristic curves for B-mode US and three sets of combined tests. The AUC for B-mode US *(B)* was 0.87 (95% CI: 0.82, 0.91); AUC increased to 0.92 (95% CI: 0.89, 0.96) (P = .009) by addition of elastography to B-mode US *(B + E)*, 0.91 (95% CI: 0.88, 0.95) (P = .049) by addition of color Doppler US to B-mode US *(B + D)*, and 0.96 (95% CI: 0.95, 0.98) ($P < .001$) by addition of both elastography and color Doppler US to B-mode US *(B + E + D)*.

Table 4

Management Recommendation for 68 Malignant and 953 Nonmalignant Screening US–detected Lesions

Management Recommendations	B-Mode US	Combined (B+E+D)	
		Criteria 1*	Criteria 2†
All			
Short-interval follow-up	257	524 (+267)	39 (−218)
Biopsy	764	293 (−471)	230 (−534)
Malignant			
Short-interval follow-up	0	0 (0)	0 (0)
Biopsy	68	68 (0)	68 (0)
Nonmalignant			
Short-interval follow-up	257	524 (+267)	39 (−218)
Biopsy	696	225 (−471)	162 (−534)
FP at screening US	953	749 (−204)	201 (−752)

Note.—Data are number of lesions; changes in the number of lesions compared with B-mode US are shown in parentheses. B = B-mode US, E = elastography, D = color Doppler US, FP = false-positive finding.

* Criteria 1 represents the study protocol, upgrading or downgrading one category according to the results of elastography and color Doppler US.

† Criteria 2 represents the hypothetical guideline, selective downgrading of BI-RADS category 3 and 4a masses to category 2 when both elastography and color Doppler US show negative results.

years vs ≥50 years), mammographic density (heterogeneously dense vs extremely dense), background echotexture (homogeneous vs heterogeneous), lesion size at US (≤1 cm vs >1 cm), and elastographic technique used (strain elastography vs shear-wave elastography; all, $P > .50$). However, there is insufficient statistical power because the number of lesions included in the subgroups was small (Table E6 [online]). Elasticity of an index breast lesion was evaluated by using strain elastography technique in 41.9% (428 of 1021) and shear-wave elastography technique in 58.1% (593 of 1021) of cases. The failure rate of data acquisition at elastography was 2.3% (10 of 428) for strain elastography and 0.3% (two of 593) for shear-wave elastography. The specificity and PPV of B-mode US significantly increased by either elastographic technique in subgroup analyses ($P < .001$, all; Fig E1 [online]).

Discussion

Elastography and color Doppler US are additional diagnostic imaging techniques for the evaluation of breast lesions at B-mode US. Elastography and color Doppler US cannot reduce the number of recalls at screening US for potentially abnormal findings for which additional imaging is needed; however, it may change the management for lesions detected at screening US. In our study, elastography and color Doppler US were performed for breast masses smaller than 3 cm detected at screening US in women older than 30 years who had dense breasts. According to our prospective multicenter results, the PPV for biopsy recommendation significantly increased from 8.9% to 23.2% by the addition of both elastography and color Doppler US to B-mode US, and 67.7% (471 of 696) of unnecessary biopsies for nonmalignant lesions could be avoided. This would allow screening performance to be within the acceptable range of recommendation for screening modalities (15). The PPV for biopsy recommendation of screening US without elastography and color Doppler US (8.9%) was slightly higher than the current benchmark for incident screening US (7.4%), which was derived from a large prospective study (7) performed in women at high risk for breast cancer, although our study included women

at average risk for breast cancer and of low prevalence of breast cancer. However, a recent study (16) showed a high PPV of 48% for biopsy performed at screening US in women at average risk for breast cancer with dense breasts. Most of the screening US examinations were incident rounds of screening, and the authors attributed the high PPV of screening US to the experience of the radiologists.

Elasticity and vascularity are two different characteristics of breast masses; therefore, both can have a complementary role for each other when breast masses are evaluated. Indeed, the highest AUC and specificity of screening US were achieved without loss in sensitivity when both elastography and color Doppler US were added to B-mode US, which is concordant with the results of a previous study (23) for nonpalpable breast lesions. As shown in our study, DCIS or small invasive cancer can show soft elasticity but increased internal vascularity. However, some cancers with sclerotic tumor stroma can show hard elasticity but absent internal vascularity at color Doppler US. By considering both the elasticity and vascularity of breast masses, sensitivity loss can be prevented while achieving higher specificity and PPV.

The major sources of false-positive findings at screening US are BI-RADS category 3 and 4a masses, which were also the most frequent findings at screening US (91.3% [932 of 1021]) in our study. According to our study results, up to 79% of false-positive findings at screening US can be reduced by selective downgrading of BI-RADS category 3 and 4a masses that show results negative for cancer at both elastography and color Doppler US to category 2 in asymptomatic women with dense breasts.

It is clear that women with dense breasts might benefit from supplemental screening. Among the currently used supplemental screening modalities, MR imaging is usually recommended in women at high risk for cancer who have any breast density and with a lifetime risk greater than 20% (36,37). In women with dense breasts without other risk factors, US and tomosynthesis are the most commonly used supplemental screening options (4). According to a recent report, tomosynthesis is less effective than US for finding cancers that are entirely masked by dense tissue, and still it missed a substantial number of invasive cancers in women with dense breasts (16). In this regard, reducing the false-positive findings at screening US by the combined use of elastography and color Doppler US may facilitate implementation of screening US for women with dense breasts.

There are several limitations to this study. First, elastography and color Doppler US were performed by the same radiologist. Therefore, results of B-mode US combined with elastography or color Doppler US were not blinded to each other. However, we had a strict guideline for combined assessments according to the results of elastography and color Doppler US. Second, we used a qualitative imaging feature for elasticity evaluation at shear-wave elastography. The qualitative color assessment of maximum elasticity is known to best perform in differentiation of benign from malignancy and correlate to maximum elasticity measured in kilopascals (27). Third, the PPV of screening US was calculated on the basis of lesion level because the single breast mass most suspicious for cancer was included from each participant with multiple breast lesions to reduce the clustering effect. There was one cancer diagnosed at follow-up examination that developed in the quadrant other than the index lesion in the ipsilateral breast. Fourth, the management of participants with either follow-up or biopsy was determined by the clinicians on the basis of recommendations from the higher of the two BI-RADS categories of B-mode US and combined assessments of B-mode US, elastography, and color Doppler US. Therefore, the actual clinical effects of the combined tests could not be evaluated. A case-control study with and without the combined use of elastography and color Doppler US will be required in the future. Finally, audit parameters for all supplemental screening US examinations cannot be determined because the study design only included screening US examinations that were BI-RADS category 3 or higher at B-mode US for which elastography and color Doppler US were indicated. Assessment of the cancer detection rate, abnormal interpretation rate, or interval cancer rate of supplemental screening US were not the primary aims of the study.

To conclude, the combined use of elastography and color Doppler US with B-mode US resulted in a reduction of false-positive findings at screening US without missing cancers in women with dense breasts and the PPV for biopsy recommendation increased from 8.9% to 23.2%. BI-RADS category 3 and 4a masses that showed results negative for cancer at both elastography and color Doppler US can be downgraded to BI-RADS category 2 and recommended for routine screening; therefore, a considerable number of false-positive findings (up to 79%), including benign biopsies and unnecessary short-interval follow-ups, can be avoided.

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

References

1. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380(9855):1778–1786.

2. Tabár L, Vitak B, Chen TH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011;260(3):658–663.

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. Current status of supplemental screening in dense breasts. J Clin Oncol 2016 Mar 9. pii: JCO658674. [Epub ahead of print]
5. Mendelson EB, Berg WA. Training and standards for performance, interpretation, and structured reporting for supplemental breast cancer screening. AJR Am J Roentgenol 2015;204(2):265–268.
6. Berg WA. Supplemental screening sonography in dense breasts. Radiol Clin North Am 2004;42(5):845–851, vi.
7. 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.
8. Corsetti V, Houssami N, Ghirardi M, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up. Eur J Cancer 2011;47(7):1021–1026.
9. 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.
10. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anticancer Randomized Trial (J-START): a randomised controlled trial. Lancet 2016;387(10016):341–348.
11. Corsetti V, Houssami N, Ferrari A, et al. Breast screening with ultrasound in women with mammography-negative dense breasts: evidence on incremental cancer detection and false positives, and associated cost. Eur J Cancer 2008;44(4):539–544.
12. 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.
13. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. Ann Intern Med 2015;162(3):157–166.
14. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA 2012;307(13):1394–1404.
15. Sickles EA, D'Orsi CJ. ACR BI-RADS Follow-up and Outcome Monitoring. In: ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, Va: American College of Radiology, 2013.
16. Tagliafico AS, Calabrese M, Mariscotti G, et al. Adjunct screening with tomosynthesis or ultrasound in women with mammography-negative dense breasts: interim report of a prospective comparative trial. J Clin Oncol 2016 Mar 9. pii: JCO634147. [Epub ahead of print]
17. Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. Eur Radiol 2016;26(1):184–190.
18. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection-a novel approach to breast cancer screening with MRI. J Clin Oncol 2014;32(22):2304–2310.
19. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. J Clin Oncol 2005;23(33):8469–8476.
20. Mendelson EB, Bohm-Velez M, Berg WA, et al. ACR BI-RADS Ultrasound. In: ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, Va: American College of Radiology, 2013.
21. Itoh A, Ueno E, Tohno E, et al. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006;239(2):341–350.
22. 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.
23. Cho N, Jang M, Lyou CY, Park JS, Choi HY, Moon WK. Distinguishing benign from malignant masses at breast US: combined US elastography and color doppler US--influence on radiologist accuracy. Radiology 2012;262(1):80–90.
24. Yi A, Cho N, Chang JM, Koo HR, La Yun B, Moon WK. Sonoelastography for 1,786 nonpalpable breast masses: diagnostic value in the decision to biopsy. Eur Radiol 2012;22(5):1033–1040.
25. Lee SH, Chang JM, Kim WH, et al. Added value of shear-wave elastography for evaluation of breast masses detected with screening US imaging. Radiology 2014;273(1):61–69.
26. 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(10):3542–3549.
27. 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.
28. Cosgrove DO, Berg WA, Doré CJ, et al. Shear wave elastography for breast masses is highly reproducible. Eur Radiol 2012;22(5):1023–1032.
29. Lee SH, Chang JM, Cho N, et al. Practice guideline for the performance of breast ultrasound elastography. Ultrasonography 2014;33(1):3–10.
30. Sickles EA, D'Orsi CJ, Bassett LW, et al. ACR BI-RADS Mammography. In: ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, Va: American College of Radiology, 2013.
31. Stavros AT, Rapp CL, Parker SH. Breast ultrasound. Philadelphia, Pa: Lippincott Williams & Wilkins, 2004; 448–527.
32. Yoon JH, Kim MJ, Moon HJ, Kwak JY, Kim EK. Subcategorization of ultrasonographic BI-RADS category 4: positive predictive value and clinical factors affecting it. Ultrasound Med Biol 2011;37(5):693–699.
33. Chang JM, Moon WK, Cho N, Kim SJ. Breast mass evaluation: factors influencing the quality of US elastography. Radiology 2011;259(1):59–64.
34. 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.
35. 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.
36. Gradishar WJ, Anderson BO, Balassanian R, et al. Invasive Breast Cancer Version 1.2016, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2016;14(3):324–354.
37. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 2007;57(2):75–89.