Clinical Radiology 74 (2019) 306–313

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Utilizing size-based thresholds of stiffness gradient to reclassify BI-RADS category 3–4b lesions increases diagnostic performance

J. Shang [a], L.-T. Ruan [a,*], Y.-Y. Wang [a], X.-J. Zhang [b], Y. Dang [a], B. Liu [a,c], W.-L. Wang [a,d], Y. Song [a], S.-J. Chang [a]

[a] Department of Ultrasound Medicine, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an 710061, China
[b] Department of Pathology, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an 710061, China
[c] Department of Ultrasound Diagnosis, Shaanxi Province People's Hospital, Xi'an 710068, China
[d] Department of Ultrasound Diagnosis, Xi'an NO.3 Hospital, Xi'an 710018, China

ARTICLE INFORMATION

Article history:
Received 4 September 2018
Accepted 8 January 2019

AIM: To investigate the role of utilizing size-based thresholds of stiffness gradient in diagnosing solid breast lesions and optimizing original Breast Imaging-Reporting And Data System (BI-RADS) classifications.

MATERIALS AND METHODS: Two-hundred and twenty-seven consecutive women underwent shear-wave elastography (SWE) before ultrasound-guided biopsy, and 234 solid breast lesions categorized as BI-RADS 3–5 were analysed. Receiver operating characteristic curve analysis was performed based on histopathology. Diagnostic performance among SWE, BI-RADS, and their combination were compared.

RESULTS: The stiffness gradient correlated with the standard deviation of elasticity (SD, r=0.90), and with Tozaki's pattern classification (r=0.64). The area under the receiver operating characteristic curves (AUC) for stiffness gradient (0.939) outperformed SD (0.897) or colour pattern (0.852). Due to significant association with lesion size (r=0.394, p<0.001), stiffness gradient's size-based thresholds (lesions >15 mm: 82.5 kPa; lesions ≤15 mm: 51.1 kPa) were established to reclassify BI-RADS 3–4b lesions. Upgrading category 3 lesions (over the corresponding cut-off value, 3 to 4a) and downgrading categories 4a-4b lesions (less than or equal to the corresponding cut-off value, 4b to 4a, 4a to 3), yielded significant improvement in specificity (90.28% versus 77.78%, p<0.001) and AUC (0.948 versus 0.926, p=0.035) than BI-RADS alone. No significant loss emerged in the sensitivity (88.89% versus 91.11%, p=0.500).

CONCLUSION: Stiffness gradient exhibited better discriminatory ability than SD or four-colour pattern classification in determining solid breast lesions and applying its size-specific thresholds to categorize BI-RADS 3–4b lesions could improve diagnostic performance.

Introduction

Breast cancer is currently the most frequently diagnosed malignancy among Chinese women, and a trend of

* Guarantor and correspondent: L.-T. Ruan, Department of Ultrasound Medicine, The First Affiliated Hospital, Xi'an Jiaotong University, No. 277, Yanta West Road, Xi'an, Shaanxi Province, 710061, China. Tel.: +(86) 029 85323640.

E-mail addresses: rlt555@mail.xjtu.edu.cn, ruanlitao@163.com (L.-T. Ruan).

https://doi.org/10.1016/j.crad.2019.01.004
0009-9260/

increasing incidence and mortality has been noted in China.[1] Due to the lower sensitivity of mammography for cancer detection in dense breasts, ultrasound (US) has been considered an important imaging technique to screen and characterize breast lesions, particularly in Asian women with dense breasts.[2] The American College of Radiology released the Breast Imaging-Reporting And Data System (BI-RADS US) so as to standardize findings on breast ultrasonography.[3] Although US possess advantages of portability, non-invasiveness, and cheapness, false-positive cases, increasing biopsies, and costs accompanies by increased breast cancer detection have been recognized.[4] As a recent meta-analysis reported, the pooled sensitivity for US is relatively high (0.87; 95% confidence interval [CI]: 0.85–0.90), and its specificity is low (0.72; 95% CI: 0.69–0.75).[5] Unlike conventional US, the newly developing shear-wave elastography (SWE) can convey tissue mechanical properties,[6] which increases confidence in breast lesion characterization.

Tumour genetic and histopathological heterogeneity is highly related to its treatment and prognosis.[7] Correspondingly, variability in Young's modulus between normal and malignant breast specimens can be observed *ex vivo*, and breast cancer displays obvious elastic heterogeneity.[8] Several *in vivo* studies have also confirmed that malignant lesions are more heterogeneous than benign lesions using qualitative and quantitative analysis of SWE.[9–11] On SWE, although elasticity heterogeneity can be reflected by qualitative classification, standard deviation of elasticity (SD) is its only quantitative measurement. Nevertheless, the diagnostic performance of SD has been mentioned to be modest.[12] Another parameter, the stiffness gradient[13], was first proposed by a recent meta-analysis, based on the consideration that the difference between mean elasticity (E_{mean}) and maximum elasticity (E_{max}) is greater in heterogeneous lesions than in homogeneous ones. The focus of the present prospective study was the effect of adding the stiffness gradient to conventional US in determining solid breast lesions.

Materials and methods

Patients and breast lesions

Two-hundred and twenty-seven consecutive female individuals scheduled for US-guided biopsy were recruited prospectively from January 2016 to July 2017. This work was approved by the Institutional Review Board of the hospital and all participants signed the informed consent.

Superficial lesions were more easily affected by transducer compression, therefore, the World Federation for Ultrasound in Medicine & Biology (WFUMB) guideline suggests that SWE should not be used when a lesion is <3 mm from the skin surface.[14] Lesions size >30 mm were also excluded due to their size being larger than the maximum sampling frame of SWE. In addition, patients with previous treatment for breast cancer and low-quality elastograms were excluded. Finally, a total of 234 solid breast lesions were assessed with conventional US and SWE. Most (n=197) were BI-RADS 4a or higher, as biopsy was recommended for these lesions according to BI-RADS criteria.[3] Thirty-seven other BI-RADS 3 lesions were also evaluated before biopsy, either because of patient request or clinician proposal.

Conventional US and SWE examinations

Both routine US and SWE examinations for breast lesions were obtained using an US machine (Aixplorer System®, Supersonic Imagine, Aix-en-Provence, France) equipped with a 4–15 MHz linear array transducer. Each targeted lesion was viewed carefully in two orthogonal planes by one junior and one senior breast sonographer, and its US characteristics, such as shape, margin, orientation, echo pattern, and posterior features, were recorded. BI-RADS classification was judged based on BI-RADS criteria,[3] and a consensus was reached in cases of discrepancy.

Two orthogonal elastographic images were obtained in each lesion. SWE mode was activated on identifying an optimal US image. The elastic scale was fixed 0–180 kPa, and the transducer was lightly and vertically placed on the skin.[14] The sampling area was set to include the lesion itself and surrounding normal fatty tissue. SWE acquisition was carried out after asking the patient to hold their breath for approximately 3–5 seconds. This procedure was repeated three times by the same two examiners with experience in SWE, and the pooled values were averaged for analysis. SWE analyses were performed by qualitative and quantitative methods. Qualitative analysis based on Tozaki's four-colour pattern[15]: no findings (Pattern 1), vertical stripe pattern artefacts (Pattern 2), a localized coloured area at the margin of the lesion (Pattern 3), and heterogeneously coloured areas in the interior of the lesion (Pattern 4). Quantitative analysis was realized by placing a 2 mm circular region of interest (ROI) over the stiffest part of the lesion, then E_{max}, E_{mean}, and SD were generated automatically.[16] The stiffness gradient was calculated according to the following equation[13]:

$$\text{Stiffness Gradient} = (E_{max} - E_{mean}) / E_{max}$$

Statistical analysis

Continuous variables were expressed as mean ± SD. Spearman's correlation coefficient was used to investigate the association among stiffness gradient, SD, and SWE pattern. In the same way, correlations between stiffness gradient and lesion depth or maximum size were also analysed. Differences between benign and malignant lesions in these elasticity values were compared using the Mann–Whitney test.

Receiver operating characteristics (ROC) curves were generated to evaluate diagnostic performances of SWE, BI-RADS, and BI-RADS combined with SWE. Histopathology results were used as the reference standard. The diagnostic power was determined by the area under the curve (AUC),

and its binomial exact confidence interval (CI) was also calculated. Their thresholds were determined by the maximum Youden's index.

There were three combined methods for stiffness gradient to assist traditional US in diagnosing breast masses. A unified standard was adopted as follows: If the lesion was over its corresponding size-specific criteria (>15 mm: 82.5 kPa; ≤15 mm: 51.1 kPa) it would be upgraded to the next grade, otherwise, it would be downgraded one level. In combination 1, categories 3–4a were upgraded (3 to 4a; 4a to 4b) and categories 4a–4b downgraded (4a to 3; 4b to 4a). Compared to combination 1, the only difference in combination 2 was that BI-RADS 4a was not upgraded. In combination 3, only BI-RADS 3 was upgraded and category 4a downgraded. AUC comparisons were performed using the method described by DeLong.[16] Differences in sensitivity or specificity were compared using the McNemar test.

Analyses were performed using commercial software (SPSS version 18, IBM, Armonk, NY, USA and MedCalc version 10.1.6 MedCalc, Mariakerke, Belgium). A p-value of <0.05 was considered statistically significant.

Results

Two-hundred and twenty-two consecutive women (mean age, 44.2±11.7 years; range, 23–84 years) underwent SWE examination before US-guided biopsy. Most patients had one lesion warranting biopsy whereas seven patients had two lesions. A total of 234 lesions (mean diameter, 16±6.2 mm; range, 4–30 mm) were analysed.

Of those 234 breast lesions, 90 (38.5%) were confirmed as malignant, and 144 (61.5%) were benign. Their histopathological types were shown in Table 1. The mean size of malignant lesions was 17.8±6.3 mm, compared with 14.8±5.9 mm for benign lesions ($p<0.001$). The depth for malignant lesions was comparable to benign lesions (0.75± 0.28 versus 0.78± 0.31 cm, $p=0.346$).

BI-RADS assessment

Thirty-nine lesions were classified as category 3 (probably benign), and 81 in category 4a (low suspicion), 37 in category 4b (intermediate suspicion), 41 in category 4c (moderate suspicion), and 36 in category 5 (highly suggestive). Malignancy rates were 2.6% (1/39) for category 3, 8.7% (7/81) for category 4a, 29.7% (11/37) for category 4b, 87.8% (36/41) for category 4c, and 97.2% (35/36) for category 5. Assuming the cut-off value of BI-RADS 4a, the AUC was 0.926 (95% CI: 0.884–0.956). Its sensitivity and specificity were 91.11% and 77.78%, respectively.

Table 1
Histopathology results for 234 solid breast lesions.

Benignancy (N=144)	Malignancy (N=90)
Fibroadenoma (n=93)	Invasive ductal carcinoma (n=76)
Ductal or epithelial hyperplasia (n=24)	Invasive lobular carcinoma (n=6)
Granulomatous mastitis (n=14)	Ductal carcinoma in situ (n=4)
Abscess (n=4)	Phylloides tumour (n=3)
Adenosis (n=3)	Mucinous carcinoma (n=1)
Phylloides tumour (n=1)	

Correlations between stiffness gradient, SD and colour pattern

The typical SWE maps for benign and malignant breast lesions are seen in Fig 1a,b, respectively. As expected, the malignant exhibited greater stiffness gradient and SD (Fig 1c), as well as higher grade of the four-colour pattern classifications (Fig 1d). The stiffness gradient was 46.8±36.9 kPa for benign lesions, significantly lower than that for malignant lesions (183.6±84.9 kPa, $p<0.001$). The same trend was also observed in SD and SWE pattern. The value of SD was 3.9±3.9 kPa for the benign group, and 16.8±10.9 kPa for the malignant group. Twenty-nine (12.3%) lesions were categorized as colour pattern 1, 93 (39.7%) in pattern 2, 94 (40.1%) in pattern 3, and 18 (7.6%) in pattern 4. The malignancy rate of each pattern was 6.9% (2/29) for pattern 1, 7.5% (7/93) for pattern 2, 70.2% (66/94) for pattern 3, and 83.3% (15 of 18) for pattern 4 ($p<0.001$).

The stiffness gradient displayed significantly high correlation with SD ($r=0.90$, $p<0.001$; Fig 2a), but only moderate correlation with colour pattern ($r=0.64$, $p<0.001$; Fig 2b). Meanwhile, SD also correlated significantly with SWE pattern classifications ($r=0.55$, $p<0.001$; Fig 2c).

Diagnostic performance of elasticity

For stiffness gradient, the AUC was 0.939 (95% CI: 0.900–0.966), with a cut-off value, sensitivity, and specificity of 82.5 kPa, 85.56%, and 88.89%, respectively. SD was 0.897 (95% CI: 0.850–0.933) in the AUC, with a cut-off value, sensitivity, and specificity of 6.8 kPa, 81.11%, and 84.72%, respectively. Regarding pattern 1 or 2 as negative and pattern 3 or 4 as positive, the AUC was 0.852 (95% CI: 0.799–0.895), with sensitivity of 90% and specificity of 78.47%.

The colour pattern showed similar diagnostic performance to SD ($Z=1.880$, $p=0.060$), but both of them displayed less diagnostic power compared with the stiffness gradient (SD: $Z=3.079$, $p=0.002$; colour pattern: $Z=4.629$, $p<0.001$).

Diagnostic performance of stiffness gradient added to BI-RADS

Considering that stiffness gradient was significantly correlated with maximal lesion size ($r=0.394$, $p<0.001$) rather than lesion depth (($r=-0.026$, $p=0.697$), diagnostic performance was further analysed according to lesion size (Table 2). For smaller lesions (≤15 mm), the values of cut-off, sensitivity, and specificity were 51.1 kPa, 87.18%, and 80.23%, respectively. The corresponding values for the remaining larger lesions (>15 mm) were 82.5 kPa, 98.04%, and 84.48%. Accordingly, the cut-off value of SD was 10.5 kPa for larger lesions and 5.6 kPa for the smaller lesions; however, the cut-off value for SWE pattern remained unchanged. Those elasticity features unexceptionally displayed a higher value of AUC and Youden's index in the

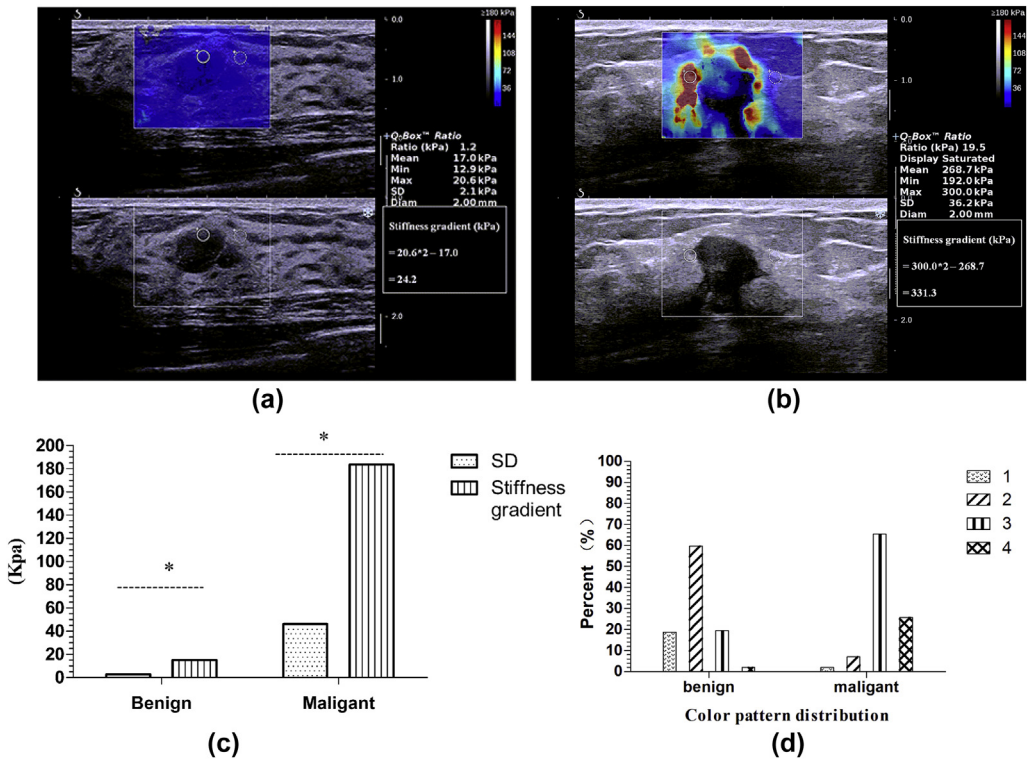

Figure 1 Typical SWE colour maps for benign and malignant breast lesions. (a) A 35-year-old woman with fibroadenoma. SWE assessment: BI-RADS 4a lesion located at 11 o'clock in the left breast was classified as colour pattern 1, and its stiffness gradient and SD values were 24.2 and 2.1 kPa, respectively. (b) A 57-year-old woman with invasive ductal carcinoma. SWE assessment: BI-RADS 4c lesion located at 3 o'clock in the right breast was classified as colour pattern 3, and its stiffness gradient and SD values were 331.3 and 36.2 kPa, respectively. (c) Malignant group displayed higher values of SD and stiffness gradient than benign group. (d) Most malignancies were classified into pattern 3 or 4, which was quite different from the benign group.

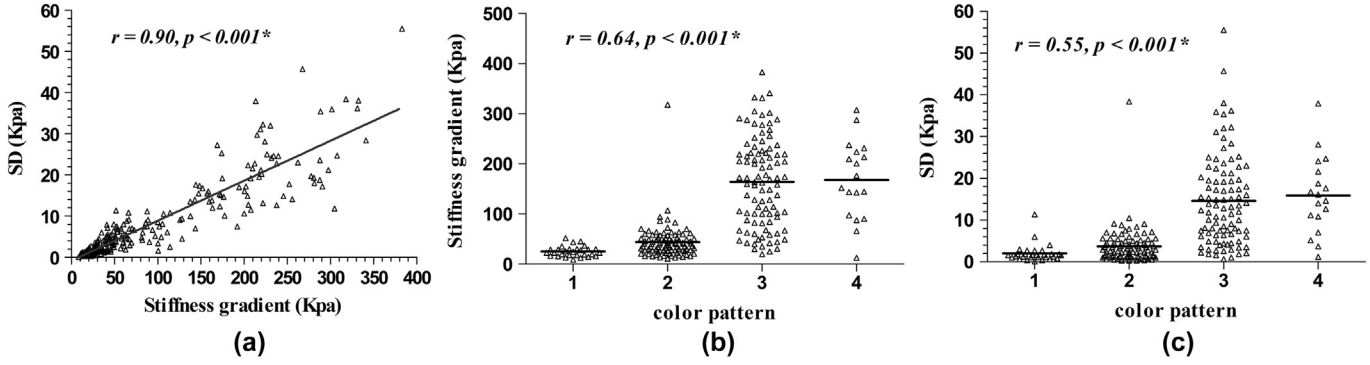

Figure 2 Spearman's correlation analyses between SWE features and stiffness gradient. (a) Correlation between stiffness gradient and SD. (b) Correlation between stiffness gradient and four-colour pattern. (c) Correlation between SD and the four-colour pattern.

bigger lesion group than in the entire or smaller lesion group.

Owing to its higher specificity and AUC, the stiffness gradient was, therefore, selected as the best SWE parameter to optimize BI-RADS classification. A unified size-based standard was adopted as follows: a lesion over its corresponding size-specific criteria (>15 mm: 82.5 kPa; ≤15 mm: 51.1 kPa) would be upgraded to the next grade. In contrast, it would be downgraded for one level. Categories 4c to 5 remained unchanged. In combination 1, categories 3–4a were upgraded (3 to 4a; 4a to 4b) and categories 4a–4b were downgraded (4a to 3; 4b to 4a). Combination 2 was similar to combination 1 except that BI-RADS 4a lesions were not upgraded to 4b. As shown in Table 3, combination 2 acquired a greater improvement in specificity (90.28% versus 77.78%, Z=2.105, p<0.001) and AUC (0.948 versus 0.926, p=0.035), without any significant loss in sensitivity (88.89% versus 91.11%, p=0.500) than conventional US. This combination exhibited increased AUC than BI-RADS itself in Fig 3.

Table 2
Diagnostic performances of shear-wave elastography (SWE) with or without considering lesion size.

Elasticity	Lesion	Threshold	AUC (95%CI)	Sensitivity	Specificity	Youden index
Colour pattern	≤15 mm	2	0.810 (0.730–0.875)	82.05%	80.03%	0.62
	>15 mm	2	0.880 (0.804–0.935)	96.08%	75.86%	0.72
	entire	2	0.852 (0.799–0.895)	90.00%	78.47%	0.69
Stiffness gradient	≤15 mm	51.1 kPa	0.913 (0.849–0.956)	87.18%	80.23%	0.67
	>15 mm	82.5 kPa	0.965 (0.911–0.991)	98.04%	84.48%	0.83
	entire	82.5 kPa	0.939 (0.900–0.966)	85.56%	88.89%	0.75
SD	≤15 mm	5.6 kPa	0.857 (0.783–0.913)	71.79%	86.05%	0.58
	>15 mm	10.5kPa	0.934 (0.870–0.972)	86.27%	93.10%	0.79
	entire	6.9 kPa	0.897 (0.850–0.933)	81.11%	84.72%	0.66

Table 3
Adding stiffness gradient to adjust original BI-RADS classification.

Diagnostic methods	Cut-off		AUC (95%CI)	Sensitivity	Specificity
BI-RADS	4a		0.926 (0.884–0.956)	91.11%	77.78%
Combination 1	Downgrade 4a, 4b Upgrade 3, 4a	4a	0.947 (0.910–0.972)	94.44%	80.56%
Combination 2	Downgrade 4a, 4b Upgrade 3	4b	0.948 (0.912–0.973)[a]	88.89%	90.28%[b]
Combination 3	Downgrade 4a Upgrade 3	4b	0.940 (0.901–0.967)	78.89%[c]	95.83%[b]

There were three types of stiffness gradient combined with traditional US.
A unified standard was adopted: if lesion is over its corresponding size-specific criteria (lesion >15 mm: 82.5 kPa; lesion ≤15 mm: 51.1 kPa) would be upgraded to next grade, otherwise, it would be downgraded for one level.
Compared to original BI-RADS.
[a] Significant increase in AUC (p=0.035).
[b] Significant improvement in specificity (p<0.001).
[c] Significant loss in sensitivity (p=0.001).

In combination 2, two BI-RADS 3 lesions (benign and malignant for each) were upgraded to category 4a. Most benign BI-RADS 4a lesions (83.78%, 60/74) were correctly downgraded to category 3; however, two malignant cases were also downgraded. In BI-RADS 4b, 18 benign and two malignant lesions were downgraded to classification 4a. The false-positive cases were a fibroadenoma (n=1) in BI-RADS 3, and the false-negative cases were mucinous carcinomas (n=2) in BI-RADS 4a and invasive ductal carcinomas (n=2) in BI-RADS 4b. Although two malignant BI-RADS 4b lesions were mistakenly downgraded, they were also biopsied. In combination 1, both the sensitivity and specificity were slightly increased without significant improvement compared to conventional US. When upgrading BI-RADS 4a to 4b, two malignant lesions were upgraded successfully, but 11 in 74 benign BI-RADS 4a lesions were adjusted incorrectly. As for combination 3, the specificity was significantly increased, but the sensitivity was obviously decreased.

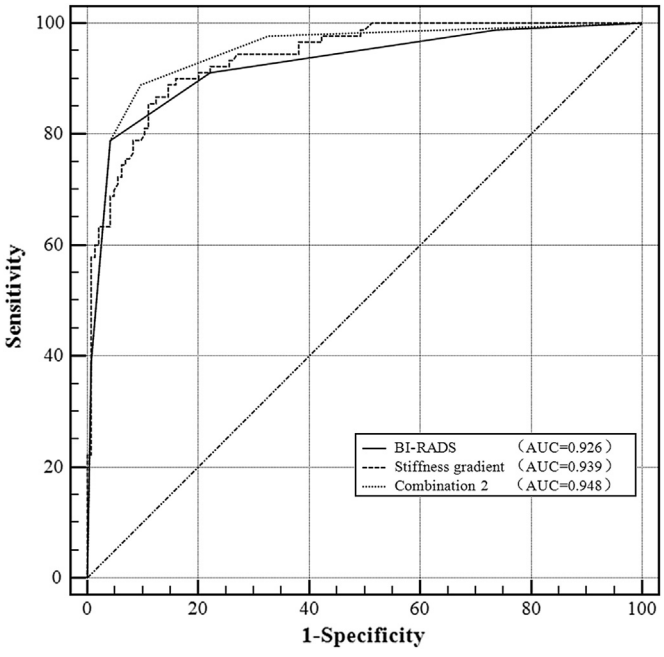

Figure 3 ROC curves for BI-RADS, stiffness gradient, and their combination. Here, the combination refers to the combination 2.

Discussion

A heterogeneous appearance on SWE has been acknowledged as an important malignant characteristic,[9–11] which has been demonstrated to be helpful in identifying breast cancer.[11,17] Although many researches have studied the supplementary value of SWE for breast cancer diagnosis, only few have specifically discussed the heterogeneous manifestation.[9–11] After careful review of the literature, lesion heterogeneity could be assessed qualitatively (two or three or four colour-pattern) and quantitatively (SD) on SWE. The three-colour pattern (very homogeneous, reasonably homogeneous, and

heterogeneous) has worse diagnostic performance than SD or the four-colour pattern,[10] with relative lower specificity (60%) resulting from dichotomous views (homogeneous or heterogeneous).[9] The stiffness gradient was proposed by a meta-analysis, with the hypothesis of bigger gap between E_{mean} and E_{max} in a heterogeneous lesion than a homogeneous lesion.[13] The present study applied this parameter in a prospective study and found that malignant lesions exhibited greater stiffness gradient than benign ones. The stiffness gradient correlated significantly with SD. This is consistent with the study of Lee *et al*., in which malignant lesions with high E_{mean} or E_{max} values tended to be histologically heterogeneous and have a high level of SD.[16] Only a moderate correlation between SD and SWE pattern (r=0.55, p<0.001) was found, which is similar to 0.545 in the study of Yoon *et al*.[18] A moderate correlation was also observed between the stiffness gradient and colour pattern. The results of the present study indicate that the stiffness gradient possibly reflects lesion heterogeneity similar to four-colour pattern categories or SD.[10]

A meta-analysis determined that the values of AUC, sensitivity, and specificity for SWE were 0.94 (95% CI: 0.91–0.96), 0.897 (95% CI: 0.863–0.923), and 0.863 (95% CI: 0.831–0.889), respectively (24). This is in good agreement with the present result that the AUC of stiffness gradient was 0.939, with a sensitivity of 85.56% and specificity of 88.89%. The strength of the stiffness gradient is that it involves the two most effective quantitative parameters and emphasizes the most efficient parameter of Emax (16). In the present study, the AUC value of the stiffness gradient was significantly higher than that of the SD or SWE pattern (0.939 versus 0.897, 0.939 versus 0.852, all p<0.050). This led us to apply the stiffness gradient to improve diagnostic performance of traditional US in the differential diagnosis of solid breast masses.

Most studies hold the view that SWE combined with BI-RADS can improve the AUC,[19–21] whereas some have argued that no significantly better performance was obtained.[16] The evidence for this has been less certain, and it was speculated that this might be caused by multiple factors such as the accuracy of BI-RADS classifications.[16,18] Rui *et al*. reported that most breast lesions were correctly categorized, and concluded that it had a high value of AUC (0.94) for conventional ultrasound, with little possibility for its improvement by SWE.[22] In contrast, diagnostic improvement tended to appear in the modest performance on US alone, such as the specificity of 13% in the study of Feldmann *et al*.,[9] and an obvious increase in malignancy rate in BI-RADS 4a (19%, 16/86) in the research of Zhou *et al*.[23]

Unlike in other studies,[11,24] no significant association between stiffness gradient and lesion depth was found in the present study. This could be explained by the similar and narrow range of lesion depth between the benign and malignant groups (benign: 7.8±3.1 mm; malignant: 7.5±2.9 mm). The BE1 multinational study of 939 masses has reported that there was a 4.5 kPa decrease in lesion stiffness for each 5 mm increase in lesion depth.[24] As was reported by several previous studies,[22,25–27] the SWE value also correlated significantly with lesion size, and its threshold increased with lesion size. Yoon *et al*.[24] further demonstrated that breast lesions of larger size usually resulted in false-positive results on SWE, and vice visa.

In the light of these results, a threshold based on lesion size might be necessary for using SWE to optimize BI-RADS classification. This decision was also noted by Skerl *et al*.[26] and Cong *et al*.[22] In the present study, the identical size-specific criterion (82.5 kPa for lesion >15 mm; 51.1 kPa for lesion ≤15 mm) was utilized to optimize in three different ways. Upgrading BI-RADS 3 and downgrading 4a and 4b (combination 2), both the specificity and AUC significantly improved, without any loss of sensitivity. In combination 1, the diagnostic parameters increased slightly, but with no significant improvement, which was consistent with the conclusion of Cong *et al*. that the performance of the combinations using the subgroups' thresholds did not differ significantly from those based on the entire study group's thresholds.[22] In combination 3, although the specificity significantly increased, the sensitivity decreased obviously. The present results could be explained by the malignant rate among the three categories growing disproportionately (BI-RADS 3:≤2%; 4a: 3–9%; 4b: 10–49%). For example, upgrading category 3 to 4a must be less likely to incur mistakes than upgrading category 4a to 4b. As the multinational study and WFUMB guidelines clearly state that using SWE features to upgrade BI-RADS category 3 and downgrade 4a masses is reasonable, but to change treatment of masses with moderate-to-high suspicion (BI-RADS category 4c or 5) was not advocated.[11,14] To upgrade category 4a and downgrade category 4b was not clearly illustrated. According to the BI-RADS standard, it is much more difficult to detect malignant BI-RADS 3 lesions than to detect benign BI-RADS 4b lesions. As in the present study, Feldmann *et al*. showed that the specificity improved from 13% to 51% without loss of sensitivity (100%) on using benign SWE signs to selectively downgrade category 4a and 4b lesions[9]; however, Wang *et al*. have proven that BI-RADS classification in conjunction with SWE (upgrade 3 to 4a; upgrade 4a to 4b; downgrade 4b to 4a) effectively increased sensitivity and specificity.[28] Therefore, the present strategy needs to be further validated.

In combination 2, there were still one upgraded BI-RADS 3 lesion and two wrongly downgraded lesions for each in BI-RADS 4a and 4b. The benign case was ultimately proven to be fibroadenoma of 16 mm in size, which could be explained as fibroadenoma showing false-positive SWE results tended to be larger in size.[29] Four malignant cases, including three case of invasive carcinoma and one case of mucinous carcinoma, were mistakenly downgraded. Although two lesions originally classified as BI-RADS 4b were adjusted to BI-RADS 4a, they still warranted biopsy and avoided being missed. In combination 2, up to 81% (60/74) of benign BI-RADS 4a lesions could avoid biopsy. This is in line with some authors' findings that a large proportion (46–85.7%) of benign biopsies can be reduced by the addition of SWE to ultrasound.[30,31]

There were several limitations to the present research. First, inter- and intra-observer variability were not assessed as E_{mean} and SD have been proven with perfect

interobserver agreement.[32] Second, SD is obtained through a fixed ROI located at the stiffest part of the lesion rather than a freehand ROI encompassing the entire lesion. Although this ROI method might not be ideal, the typical method of 2 mm ROI over the stiffest area was adopted by many studies[11,16,29] and was also suggested as a good compromise in evaluating several parameters so as to avoid different sets of measurements.[32] No statistically significant differences in the AUC of SD between these two ROI methods were found by the prospective study of Liu *et al.*[19] Third, three colour-patterns (very homogeneous, reasonably homogeneous, and heterogeneous)[11] and the dichotomy method (homogeneous or heterogeneous)[9] were not included in the present study due to their lower effectiveness than SD or the four-colour pattern categories. Finally, the malignancy rate of the present BI-RADS 3 lesions was 2.6%, which was higher than the guideline.[3] This could be explained by the fact that the authors' institution is a referral centre. SD on routine two-dimensional SWE cannot truly reflect an entire lesion's heterogeneity because of ill-defined boundaries as well as inner anechoic areas on elastograms in some malignant lesions.

In conclusion, the stiffness gradient showed better diagnostic power than SD or qualitative patterns, and both the AUC and specificity were improved after using the lesion size-based threshold to optimize BI-RADS 3–4b lesions. These findings suggest that it might be a better quantitative parameter of elasticity heterogeneity in breast lesions.

Conflict of interest

None declared.

Acknowledgements

The authors thank Libby Cone, MD and MA, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac) for editing a draft of this manuscript. This work was supported by a grant (no. 2017SF-132) from the Programme in Science and Technology for Social Development of Shaanxi province, China.

References

1. Chen W, Zheng R, Baade PD, *et al.* Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;**66**(2):115–32.
2. Brem RF, Lenihan MJ, Lieberman J, *et al.* Screening breast ultrasound: past, present, and future. *AJR Am J Roentgenol* 2015;**204**(2):234–40.
3. American College of Radiology. Breast imaging reporting and data system. In: *Breast imaging atlas*. 5th ed. Reston, VA: American College of Radiology; 2013.
4. Meuwly JY. [Ultrasound for breast cancer screening: an effective tool in a personalized screening]. *Praxis* 2015;**104**(25):1399–404.
5. Li Q, Hu M, Chen Z, *et al.* Meta-analysis: contrast-enhanced ultrasound versus conventional ultrasound for differentiation of benign and malignant breast lesions. *Ultrasound Med Biol* 2018;**44**(5):919–29.
6. Bercoff J, Tanter M, Fink M. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Trans Ultrason Ferr* 2004;**51**(4):396–409 [e].
7. Martelotto LG, Ng CK, Piscuoglio S, *et al.* Breast cancer intra-tumour heterogeneity. *Breast Cancer Res* 2014;**16**(3):210.
8. Samani A, Zubovits J, Plewes D. Elastic moduli of normal and pathological human breast tissues: an inversion-technique-based investigation of 169 samples. *Phys Med Biol* 2007;**52**(6):1565–76.
9. Feldmann A, Langlois C, Dewailly M, *et al.* Shear wave elastography (SWE): an analysis of breast lesion characterization in 83 breast lesions. *Ultrasound Med Biol* 2015;**41**(10):2594–604.
10. Gweon HM, Ji HY, Son EJ, *et al.* Visually assessed colour overlay features in shear-wave elastography for breast masses: quantification and diagnostic performance. *Eur Radiol* 2013;**23**(3):658–63.
11. 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):435–49.
12. Zhang Q, Xiao Y, Chen S, *et al.* Quantification of elastic heterogeneity using contourlet-based texture analysis in shear-wave elastography for breast tumour classification. *Ultrasound Med Biol* 2015;**41**(2):588–600.
13. Blank MAB, Antaki JF. Breast lesion elastography region of interest selection and quantitative heterogeneity: a systematic review and meta-analysis. *Ultrasound Med Biol* 2017;**43**(2):387–97.
14. Barr RG, Nakashima K, Amy D, *et al.* WFUMB guidelines and recommendations for clinical use of ultrasound elastography. Part 2: breast. *Ultrasound Med Biol* 2015;**41**(5):1148–60.
15. Tozaki M, Fukuma E. Pattern classification of shear-wave elastography images for differential diagnosis between benign and malignant solid breast masses. *Acta Radiol* 2011;**52**(10):1069–75.
16. Lee EJ, Jung HK, Ko KH, *et al.* Diagnostic performances of shear wave elastography: which parameter to use in differential diagnosis of solid breast masses? *Eur Radiol* 2013;**23**(7):1803–11.
17. 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.
18. Yoon JH, Ko KH, Jung HK, *et al.* Qualitative pattern classification of shear wave elastography for breast masses: how it correlates to quantitative measurements. *Eur J Radiol* 2013;**82**(12):2199–204.
19. Liu B, Zheng Y, Huang G, *et al.* Breast lesions: quantitative diagnosis using ultrasound shear wave elastography—a systematic review and meta-analysis. *Ultrasound Med Biol* 2016;**42**(4):835–47.
20. Li DD, Xu HX, Guo LH, *et al.* Combination of two-dimensional shear wave elastography with ultrasound breast imaging reporting and data system in the diagnosis of breast lesions: a new method to increase the diagnostic performance. *Eur Radiol* 2016;**26**(9):3290–300.
21. Xiao Y, Yu Y, Niu L, *et al.* Quantitative evaluation of peripheral tissue elasticity for ultrasound-detected breast lesions. *Clin Radiol* 2016;**71**(9):896–904.
22. Cong R, Li J, Wang X. Comparing performance of combinations of shear wave elastography and B-mode ultrasound in diagnosing breast masses: is it influenced by mass size? *Ultrasound Med Biol* 2017;**43**(10):2133–43.
23. Zhou J, Zhan W, Chang C, *et al.* Breast lesions: evaluation with shear wave elastography, with special emphasis on the "stiff rim" sign. *Radiology* 2014;**272**(1):63–72.
24. Yoon JH, Jung HK, Lee JT, *et al.* Shear-wave elastography in the diagnosis of solid breast masses: what leads to false-negative or false-positive results? *Eur Radiol* 2013;**23**(9):2432–40.
25. 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.
26. Skerl K, Vinnicombe S, Giannotti E, *et al.* Influence of region of interest size and ultrasound lesion size on the performance of 2D shear wave elastography (SWE) in solid breast masses. *Clin Radiol* 2015;**70**(12):1421–7.
27. Song EJ, Sohn YM, Seo M. Tumour stiffness measured by quantitative and qualitative shear wave elastography of breast cancer. *Br J Radiol* 2018;**91**(1086):20170830.
28. Wang M, Yang Z, Liu C, *et al.* Differential diagnosis of breast category 3 and 4 nodules through BI-RADS classification in conjunction with shear wave elastography. *Ultrasound Med Biol* 2017;**43**(3):601–6.
29. Elseedawy M, Whelehan P, Vinnicombe S, *et al.* Factors influencing the stiffness of fibroadenomas at shear wave elastography. *Clin Radiol* 2016;**71**(1):92–5.

30. Youk JH, Gweon HM, Son EJ, *et al*. Diagnostic value of commercially available shear-wave elastography for breast cancers: integration into BI-RADS classification with subcategories of category 4. *Eur Radiol* 2013;**23**(10):2695–704.
31. Evans A, Whelehan P, Thomson K, *et al*. Differentiating benign from malignant solid breast masses: value of shear wave elastography according to lesion stiffness combined with greyscale ultrasound according to BI-RADS classification. *Br J Cancer* 2012;**107**:224.
32. Hong S, Woo OH, Shin HS, *et al*. Reproducibility and diagnostic performance of shear wave elastography in evaluating breast solid mass. *Clin Imag* 2017;**44**:42–5.