BMJ Open

A combination of qualitative and quantitative measurements of shear wave elastography to predict malignancy in patients with pathological nipple discharges: a diagnostic accuracy study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008848
Article Type:	Research
Date Submitted by the Author:	21-May-2015
Complete List of Authors:	Guo, xiao-bo Liu, Ying Li, Wanhu
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Radiology and imaging
Keywords:	Breast tumours < ONCOLOGY, ONCOLOGY, Adult oncology < ONCOLOGY

A combination of qualitative and quantitative measurements of shear wave elastography to detect malignancy in patients with pathological nipple discharges: a diagnostic accuracy study

Xiaobo Guo[1] Ying Liu[1] Wanhu Li[2*]

1. Department of Ultrasonography, 476 Clinical Department of Fuzhou General Hospital of Nanjing military region, China
2. Department of Radiology, Shandong Cancer Hospital and Institute, No. 440, Jinan Road, Jinan 250117, China

Email address: Xiaobo Guo: 656726103@qq.com; Ying Liu : 1637824435@qq.com; Wanhu Li :lwhvzg@126.com.

Corresponding author: Xiaobo Guo, phone number：0531-86833526, E-mail: 656726103@qq.com.

ABSTRACT

Objectives: To evaluate the diagnostic accuracy of a combination of qualitative and quantitative measurements of shear wave elastography (SWE) to detect malignancy in patients with pathological nipple discharge (PND).

Design: Prospective diagnostic accuracy study comparing a combination of qualitative and quantitative measurements of SWE (index test) to a microdochectomy for histological diagnosis (reference test).

Setting: Fuzhou general hospital of Nanjing military command.

Participants: A total of 227 patients with PND were finally included from January, 2011 to March 2014, after we screened 1084 possible candidates. All participants were evaluated through a SWE with qualitative parameters generated by Virtual TouchTM tissue imaging (VTI) and quantitative parameters generated by Virtual TouchTM tissue quantification (VTQ). All the patients were consented to receive microdochectomy for histological diagnosis, and the results were set as a reference test.

Outcome measures: Sensitivity and specificity of the combined VTI and VTQ of the SWE for detection of malignancy in patients with PND.

Results: The 227 participants presented with 237 lesions. The results of pathological examination showed that 159(67.1%) of the 237 lesions were benign and the other 78 (32.9%) were malignant. An area under the curve (AUC) of elasticity score, VTQm and VTQc were 0.854, 0.826, 0.856 respectively, with the corresponding cut-off point as 2.50, 3.285m/s and 3.445m/s respectively. After a combination of these measurements, the sensitivity, specificity, positive and negative

predictive value (PPV and NPV) were 82.1%, 96.9%, 92.8% and 91.7%, respectively.

Conclusion: A combination of qualitative and quantitative measurements of SWE is predictive for detecting malignancy in patients with PND. However, the generalization of this method is limited because of the highly selected study population.

Key words: Shear wave elastography; Pathological nipple discharge; Microdochectomy; Breast Imaging- Reporting and Data System.

Strengths and limitations of this study

- Accuracy of the combined measures of shear wave elastography (SWE) for detecting malignancy of PND patients has never previously been studied.
- For the first time, this study tested diagnostic accuracy of the combined measures of SWE for detection of malignancy in PND patients.
- Limitations include that we excluded patients without a positive ductoscopy findings, which might overestimate the diagnostic accuracy of elastography in PND patients

Introduction

According to the 2014 cancer statistics of the American cancer society, the new cases of breast cancer are 235030, accounting for 40430 deaths[1]. The prevalence of breast cancer ranks the highest in Chinese women with malignant carcinoma[2]. The number of Chinese women with breast cancer account for 12.2% of all newly diagnosed breast cancer and 9.6% of deaths from it around the world[2].

Pathological nipple discharge (PND) is the third most common complaint of patients prompting

referral to doctors for breast diseases[3]. PND is believed to be an indicator for breast carcinoma, especially when the nipple discharge is bloody[3]. So mastectomies were carried out unnecessarily without obtaining a histopathological diagnosis before the year 1950. Although mastectomies are not necessarily needed nowadays, patients with PND would be normally suggested to take further examinations to rule out malignancies. These examinations involve mammography, ultrasonography, ductoscopy with or without galactography, as well as cytological examination. Mammography and B-mode ultrasonography are used for screening patients with high risk of breast carcinoma, but they are not sensitive enough for detecting malignancy in PND patients, since small lesions may be occult on these 2 modalities[4]. Ductoscopy may miss the malignant lesions outside a depth of 2cm[5], and cytological could only be used as supportive tests [6].

Ultrasonography elastography work through the assumption that tissue compression produces strain within the tissue, and that the strain is smaller in firm tissue than in soft tissue. So through measuring tissue strain ratio, we could evaluate the firmness of the tissue to predict malignancy of masses. Recently, several studies indicated that ultrasonography elastography may be a useful tool for differentiating benign and malignant masses. The sensitivity of the predicting a malignancy in patients complaining breast lesions ranges from 80% to 98%, while the specificity ranges from 66% to 84%[7-12]. Recently, a shear wave elastography is reported to be less operator-dependent and predictive of ductal or intraductal carcinoma[13-15]. A five-level elasticity score is normally used for qualitative assessment of the images of elastography, while a quantitative measurement of elastography is also becoming popular; a combination of these ultrasonic measurements seems to be beneficial for improving diagnostic accuracy of breast malignancies[14 15]. We noticed that the

diagnostic value of a combination of qualitative and quantitative measurements for elastography to detect malignancies in patients with PND has rarely been studied. So we conducted a study aiming to find out the sensitivity and specificity of the combination measurements for predicting malignancy in patients with PND.

Methods and materials

This is a study designed for evaluating diagnostic sensitivity and specificity of shear wave elastography (SWE) for detection of malignancy in patients with PND. This study protocol was approved by the ethical review board in Fuzhou general hospital of Nanjing military command (No. FZJQ2011018).

Patients

Patients who complained a PND within 1 year were recruited from outpatient settings in the Fuzhou general hospital of Nanjing military command. The patients were screened for eligibility according to the following inclusion criteria: (1) with an age over 18 years; (2) with or without palpable breast masses; (3) with breast masses detected by B-mode ultrasonography; (4) agree to participate in this study; (4) providing a written consent for further examination, biopsy and use of the study data. We excluded the participants who (1) refused to receive a microdochectomy or a biopsy; (2) had serious damage in internal organs that may bias the results; (3) could not receive local anesthesia because of allergy; (4) were taking or have taken chemotherapy or radiotherapy for suspicious malignant breast tumors.

Ultrasonography imaging

A B-mode ultrasonography (BUS) and an elastography using acoustic radiation force Impulse (ARFI)

were performed with a Siemens ACUSON S2000 ultrasonography system (Siemens Ltd., China, product standard: YZB/USA 3876-2010). One operator (with 10 years of experience in ultrasonography imaging) ran the BUS and the ARFI to get the ultrasonography and elastography images, using a superficial probe (9L4, Siemens Ltd., China) with a frequency of 4 to 9 Hz.

The BUS images of the breast lesions were obtained and categorized according to Breast Imaging Reporting and Data System (BI-RADS). Moreover, we applied a subclassification scheme for BI-RADS 4: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy [16]. Two radiologists (XBG and YL who are with at least 5 years of experience in ultrasonography) independently assessed the BUS images in the following aspects: the shape of the mass, depth, orientation, margin, boundary of the lesions, echo pattern, posterior acoustic features, as well as surrounding tissue changes. All the detected masses were divided into two types (solid and cystic) according to the presence of a cystic portion. In cases involving a solid mass in a dilated duct (intraductal) or cyst (intracystic) with a predominant cystic portion, the radiologists considered the cases as cystic type lesions.

After the BUS examination, the operator switched the ultrasonic mode to the virtual touch^{TM} quantification (VTQ). The patients were told to hold the breath, when the operator detected the inside and margin of the lesion, with a previous set region of interest (ROI) in a fixed region of 5×5cm. The operator recorded the value of VTQ (m/s) and the depth of the ROI (cm) from skin for each measurement, in a total of 5 measurements. We calculated the mean value of the 5 measurements as the final VTQ score. Additionally, the value of VTQ was measured at both the margin (VTQm) and center (VTQc) of the breast lesions. Measurable shear wave velocity (SWE)

ranges from 0 to 9 m/s, a value of ROI exceed the range would be shown as "X.XX m/s" on the screen. The participants with this "X.XX m/s" value were excluded.

After finishing examination with the VTQ mode, the operators switched to virtual touch^{TM} imaging (VTI). ARFI sequences were generated to evaluate the elasticity scores of the breast lesions in VTI mode. The VTI images were evaluated and categorized through a five-level scoring system[17]. An elasticity score of 1 indicates strain in the entire hypoechoic lesion (the entire lesion was equally shaded in green); score 2 indicates that strain is not seen in part of the hypoechoic (the lesion is shown as a mosaic of green and blue); score 3 indicates that strain is in the peripheral areas only (peripheral areas in green); score 4 indicates no strain in the entire hypoechoic area (lesion is shown in blue); score 5 indicates no strain in the entire hypoechoic lesion or the surrounding area(both the entire hypoechoic lesion and its surrounding area were blue). A higher score indicates a harder tissue, vice versa. For cystic-type lesions, we analyzed elasticity scores in solid areas. After B-mode ultrasonography and ARFI elastography examination, 2 radiologists independently reviewed the images and the measurements without knowing the medical history of the patients; any discrepancy happened was solved by discussion.

Microdochectomy and pathological examination

Microdochectomy was performed to get a biopsy sample, which was set as a reference test (the golden standard). The patients were asked to be prepared in supine position with the arm preferably extended on a side board. After skin disinfection, the fluid-producing ducts were identified by massaging the breast towards the nipple. After dilating the target ducts with lacrimal

dilators, a surgeon (with 5 years of experience in microdochectomy) introduced endoscope (a fiberoptic ductoscopy, 0.8mm of sheath, 80mm of length; manufactured in Denzlingen, Germany) through the ducts. After insertion of the ductoscopy, saline was injected into the ducts to keep them expanded. The ductoscopy was advanced through the ducts under direct vision from the endoscope camera, all tiers of branches were examined until the endoscope could not be advanced further or an obstructing lesion was detected. When an obstructing lesion was found, the image of this lesion would be taken using an endoscope camera, moreover, an endoscopy guided biopsy would be performed as scheduled. After biopsy and image taken, the patient would be referred to a surgeon for further treatment advice.

After detection of ductal lesions, methylene blue was injected into the discharged mammary duct. Local anesthesia was performed and then an incision around the areola was implemented. The pathological duct was separated according to tissue with methylene-blue stain, and the duct was then dissected and excised. The operator then pulled the clip out of duct together with the scope and after the excision was made. The biopsy samples over 1 mm^3 were fixed using formalin after sample collection, and were sent for histologic examination.

Statistical analysis

Statistical analysis was performed to specify the diagnostic sensitivity and specificity of a combination of VTI and VTQ measurements for elastography to predict malignancy in patients with PND. First, we would assess the interobserver variability between the two radiologists (XBG and YL) using the Cohen's kappa test. The value of kappa over 0.75 was considered as excellent, 0.40 to

0.75 as fair to good, and below 0.40 as poor. Second, we evaluated if elasticity score, VTQm and VTQc were independent predictors using a logistic model. In the model, we used differentiation of benign and malignant breast lesions as dependent variable, and the elasticity score, VTQm and VTQc were set as covariates. Third, we calculated the sensitivity, specificity and the area under the curve (AUC) of elasticity score, VTQm and VTQc using receiver operating characteristics (ROC) curve. A valued of AUC of 0.5 was considered as no diagnostic value; 0.5 to 0.7, low diagnostic value; 0.7 to 0.9, mediate diagnostic value; 0.9 to 1.0, high diagnostic value[18]. We calculated the cut-off point of VTQ and gave the respective sensitivity and specificity, based on the highest sum of sensitivity and specificity. According to the cut-off point of each quantitative measurements, we synthesized the measurements by the following formula: the synthesized score=elasticity score + VTQm + VTQc. A synthesized score of 1 was given, if a value over the cut-off point was observed in a breast lesion; otherwise, a score of -1 would be given. A breast lesion with a final score below 0 was classified as benign, otherwise it would be categorized as malignant. Fourth, after differentiating benign and malignant lesions using the synthesized score, we evaluated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The diagnostic value of this synthesized quantitative measurement was compared with the biopsy results using McNemar analysis. All statistically analysis was performed using Statistical Package for the Social Sciences (SPSS, Version 20.0, Armonk, NY: IBM Corp.) and R software (www.r-project.org, version 3.1.1).

Results

A total of 227 patients with PND were included in this study, after 1084 possible candidates were screened and 857 of them were excluded. The 227 participants presented with 237 lesions, and 10

participants showed breast lesions on both sides of the breasts. The whole procedure was shown in figure 1. The age of the included participants ranged from 25 to 67 years. The results of pathological examination showed that 159(67.1%) of the 237 lesions were benign and the other 78 (32.9%) were malignant. The types of the benign and malignant lesions were shown in table 1. Two imaging examples of an invasive ductal carcinoma diagnosed on both B-mode ultrasonography and elastography were shown in figure2. Of all the 237 breast lesions, 191 (80.6%) were palpable masses; 127 of them were benign and 64 were malignant. Sixty (25.3%) lesions were complained of a history of bloody discharge; 7 of them were benign and 53 were malignant. The size of the lesions ranged from 3.2mm to 38.7mm. The benign lesions showed a mean size of 8.16mm, while the malignant lesions showed a mean size of 15.58mm. The agreement between the two radiologists for ultrasound elastography was good (kappa=0.76, $p<0.01$).

Predictors for SWE to detect malignant breast lesions

To evaluate whether elasticity scores, VTQ in the marginal area of the breast lesions (VTQm) and VTQ in the central area of the breast lesions (VTQc) are potential predictors for detecting malignancies in PND patients, a logistic regression was performed. The results showed that all of the parameters were good predictors, malignant breast lesions illustrated a higher value of elasticity score (malignant versus benign, 2.17 versus 3.51), VTQm (malignant versus benign, 2.54 m/s versus 3.39 m/s), and VTQc (malignant versus benign, 2.74 m/s versus 3.72 m/s). Table 2 showed the results of the logistic regression analysis, which showed a significant association between higher values of elasticity score, VTQm and VTQc and higher possibility of malignant

breast lesions.

The diagnostic cut-off points of elasticity score, VTQm and VTQc

We calculated the cut-off points for elasticity scores, VTQm and VTQc. The results showed that a cut-off point of 2.50 for elasticity score yielded a sensitivity of 0.872(95%CI, 0.794 to 0.949), a specificity of 0.736 (95%CI, 0.667 to 0.805) and an area under the curve of 0.854. We also calculated a cut-off point of 3.285m/s for VTQm with a sensitivity of 0.615 (95%CI, 0.500 to 0.718), a specificity of 0.994(95%CI, 0.981 to 1.000) and an AUC of 0.826. Additionally, a cut-off point of 3.445 m/s was calculated for VTQc with a sensitivity of 0.628 (95%CI, 0.513 to 0.731), a specificity of 0.981 (95%CI, 0.956 to 1.000) and an AUC of 0.856. Figure 3 showed the details of each measurement.

Predictive value of a combination of measurements for elastography

We used a synthesized quantitative measurement to predict malignant breast lesions. The measurement categorized 168 (70.9%) breast lesions as benign and 69 (29.1%) as malignant. The biopsy results showed that 159 (66.7%) breast lesions were classified as benign while 78 (33.3%) were malignant. We used the McNemar test to compare the synthesized measurement with the biopsy results, the result showed no significant difference (McNemar's chi-squared = 3.3684, p=0.066). The sensitivity, specificity, positive and negative predictive value (PPV and NPV) were 0.820, 0.968, 0.927 and 0.917, respectively. However, we compared elasticity sore, VTQm and VTQc to the biopsy results respectively, using also the McNemar test. We found significant differences

between these quantitative results and the biopsy results (p<0.01). More details were shown in table 3 and table 4.

Discussion

The results of our study showed that a combination of qualitative and quantitative measurements for SWE elastography is highly sensitive and specific in predicting malignancy in PND patients, which is consistent with several recent studies testing the diagnostic value of elastography in breast masses[14 15 19]. However, several studies reported that ultrasonography was not predictive to malignancy in PND patients[20-22]. We discussed our major findings and the remaining questions below.

Predicting malignancy in patients with PND is still challenging, Sabel and colleagues found that 95% of the PND lesions were benign, which need not be excised. However, most patients will choose duct excision for palliative purposes[23]. In the screen phase of our study, 53 (5%) patients refuse to take ductoscopy or duct excision and therefore were excluded. The patients refuse to take the examinations for that no evidence of malignancy was found and they were afraid of the contraindications of ductoscopy or duct excision. Based on these facts, a non-invasive method with accurate diagnose are needed by patients with PND. Alcock found no predictive value of ductoscopy and ultrasonography in differentiating benign and malignant breast lesions in PND patients. The authors of this study suggested that ductal surgery was the only reliable way of providing a diagnosis[24]. In our study, we found the PPV of the synthesized score for SWE elastography was 0.927. Moreover, our study showed a sensitivity of 0.820 and a specificity of 0.968 for SWE

elastography. This result indicated that SWE elastography could promote both the sensitivity and specificity in ultrasonography examination, and therefore SWE elastography is of better diagnostic performance than B-mode ultrasonography. So our study, for the first time, revealed that ultrasonography elastography is valuable in predicting malignant lesions in patients with PND. However, we should explain this result with caution. First, the incidence of malignancy among patients with PND is high in our study. The incidence is 33.3%, while a reported incidence ranged from 4% to 29%[25-28]. A higher incidence of malignancy may contribute to a higher positive prediction, which may partly explain a high PPV found in our study.

The size of the breast mass could also be an important factor contributing to the high PPV. We found that PDN patients with malignant lesions presented a bigger size of the breast mass (malignant versus benign, 15.58 mm versus 8.15 mm), which might be easier to be detected by B-mode ultrasonography and elastography. This findings are consistent with other studies using elastography to detect breast lesions[14 19]. On the other hand, this result indicated that PND patients presented with no breast masses or the size of the breast masses below 8mm might not be detected by elastography. So our study findings could only apply to PND patients with a breast masses over 8mm large.

It is interesting to find out that 513 patients with PND refused to participate in our study. Absence of these patients might contribute to a high incidence of malignant breast lesions, and thus made the diagnostic value of elastography overestimated. The high rate of rejection could be caused by the bad doctor-patient relationship in recent years in China[29]. Patients might go to several doctors until they could accept the suggestions for further examination and treatment.

The synthesized score we used in this study is similar to a published method for classification of benign and malignant breast lesions, except that we did not include the BI-RADS[14]. The BI-RADS is a classification system, which is in part coincident with the elasticity score[17]. We chose elasticity score instead of the BI-RADS into the synthesized score. The elasticity score is a qualitative measurement, which was treated as continuous data in several studies[14 19 30]. We also treated the elasticity score as a continuous data, so that we could calculate a cut-off point and include it into the synthesized score. The way that we handled the elasticity score may bring bias into our results. However, neither did we find a better way to solve the problem, nor did the previous studies. Therefore, in future studies, a qualitative measurement should be developed, which differentiate benign and malignant breast lesions more clearly than the elasticity score.

In this study, we incorporated both VTQ at the margin(VTQm) and in the center (VTQc) into the synthesized score. To the best of our knowledge, we are the first to use these two quantitative measurements, instead, previous studies normally used VTQc[7 14 19]. Inclusion of both VTQm and VTQc could increase the understanding of the breast lesions, since they were reported to be different[17 19]. The results of our study indicated that both VTQm and VTQc were significantly higher in malignant lesions than benign ones. So we might have assigned more weight in the VTQ measurement, and therefore made breast lesions with higher VTQ more easier to be malignant in the synthesized score.

We have several limitations in this study. First, in the screening stage of our study, we excluded patients without a positive ductoscopy findings, because we could not perform microdochectomy without an exact location of the lesion reported by ductoscopy. However, ruling out these patients

may overestimated the diagnostic accuracy of elastography in PND patients. We followed-up these patients to try to solve the problem in the future, but now we could not be definitive to what extent we have exaggerated the diagnostic value of elastography. Second, the weight of elasticity score, VTQm and VTQc were assigned evenly, and therefore made the value of VTQ a more important factor than elasticity score in predicting malignancy in PND patients. Third, we used a ductoscopy-guided microdochectomy to get sample for biopsy. We might have missed the malignant lesions, since ductoscopy could only reach a depth within 2cm[5]. The ideal way would be that each patient was followed-up for at least 5 years to achieve a more definitive result. However, now we still take the risk that we overestimate the PPV of elastography. Fourth, the ultrasonography and elastography were obtained by one operator, which may induce bias to the results, despite the operator was with 10 years' experience. Therefore, our study results should be explained with caution.

In conclusion, ultrasonography elastography was predictive for malignancy in PND patients, with breast mass over 8mm. However, the weight of the qualitative and quantitative measurements for elastography warrants more studies.

Authors' contributions

XBG and YL contributed to the design and manuscript writing. XBG and WHL conceptualized the study and data analysis. YL assisted with data collection. All authors read and approved the final manuscript.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments: This study was supported financially by the Key Project of science and technology of Fujian Province of China(No. 2009Y0043). We would like to thank all the participants who participated in this study.

Reference

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: A Cancer Journal for Clinicians 2014;**64**(1):9-29 doi: 10.3322/caac.21208[published Online First: Epub Date]|.
2. Fan L, Strasser-Weippl K, Li J-J, et al. Breast cancer in China. The Lancet Oncology 2014;**15**(7):e279-e89
3. Chen L, Zhou WB, Zhao Y, et al. Bloody nipple discharge is a predictor of breast cancer risk: a meta-analysis. Breast cancer research and treatment 2012;**132**(1):9-14 doi: 10.1007/s10549-011-1787-5[published Online First: Epub Date]|.
4. Jagmohan P, Pool FJ, Putti TC, Wong J. Papillary lesions of the breast: imaging findings and diagnostic challenges. Diagnostic and interventional radiology (Ankara, Turkey) 2013;**19**(6):471-8 doi: 10.5152/dir.2013.13041[published Online First: Epub Date]|.
5. Denewer A, El-Etribi K, Nada N, El-Metwally M. The role and limitations of mammary ductoscope in management of pathologic nipple discharge. Breast J 2008;**14**(5):442-9 doi: 10.1111/j.1524-4741.2008.00620.x[published Online First: Epub Date]|.
6. Kooistra BW, Wauters C, van de Ven S, Strobbe L. The diagnostic value of nipple discharge cytology in 618 consecutive patients. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2009;**35**(6):573-7 doi: 10.1016/j.ejso.2008.09.009[published Online First: Epub Date]|.
7. Tozaki M, Isobe S, Sakamoto M. Combination of elastography and tissue quantification using the acoustic radiation force impulse (ARFI) technology for differential diagnosis of breast masses. Japanese journal of radiology 2012;**30**(8):659-70 doi: 10.1007/s11604-012-0106-3[published Online First: Epub Date]|.
8. Tozaki M, Isobe S, Fukuma E. Preliminary study of ultrasonographic tissue quantification of the breast using the acoustic radiation force impulse (ARFI) technology. European journal of radiology 2011;**80**(2):e182-7 doi: 10.1016/j.ejrad.2011.05.020[published Online First: Epub Date]|.
9. Meng W, Zhang G, Wu C, Wu G, Song Y, Lu Z. Preliminary results of acoustic radiation force impulse (ARFI) ultrasound imaging of breast lesions. Ultrasound in medicine & biology 2011;**37**(9):1436-43 doi: 10.1016/j.ultrasmedbio.2011.05.022[published Online First: Epub Date]|.
10. Taylor K, O'Keeffe S, Britton PD, et al. Ultrasound elastography as an adjuvant to conventional ultrasound in the preoperative assessment of axillary lymph nodes in suspected breast cancer: a

pilot study. Clinical radiology 2011;**66**(11):1064-71 doi: 10.1016/j.crad.2011.05.015[published Online First: Epub Date]|.

11. Sadigh G, Carlos RC, Neal CH, Dwamena BA. Accuracy of quantitative ultrasound elastography for differentiation of malignant and benign breast abnormalities: a meta-analysis. Breast cancer research and treatment 2012;**134**(3):923-31 doi: 10.1007/s10549-012-2020-x[published Online First: Epub Date]|.
12. Gong X, Xu Q, Xu Z, Xiong P, Yan W, Chen Y. Real-time elastography for the differentiation of benign and malignant breast lesions: a meta-analysis. Breast cancer research and treatment 2011;**130**(1):11-8 doi: 10.1007/s10549-011-1745-2[published Online First: Epub Date]|.
13. Kılıç F, Ustabaşıoğlu FE, Samancı C, et al. Ductal Carcinoma In Situ Detected by Shear Wave Elastography within a Fibroadenoma. Journal of breast cancer 2014;**17**(2):180-3 doi: 25013441[published Online First: Epub Date]|.
14. Kim YS, Park JG, Kim BS, Lee CH, Ryu DW. Diagnostic value of elastography using acoustic radiation force impulse imaging and strain ratio for breast tumors. J Breast Cancer 2014;**17**(1):76-82 doi: 10.4048/jbc.2014.17.1.76[published Online First: Epub Date]|.
15. Youk JH, Son EJ, Gweon HM, Kim H, Park YJ, Kim JA. Comparison of Strain and Shear Wave Elastography for the Differentiation of Benign From Malignant Breast Lesions, Combined With B-mode Ultrasonography: Qualitative and Quantitative Assessments. Ultrasound in medicine & biology 2014;**40**(10):2336-44 doi: 10.1016/j.ultrasmedbio.2014.05.020[published Online First: Epub Date]|.
16. Sanders MA, Roland L, Sahoo S. Clinical implications of subcategorizing BI-RADS 4 breast lesions associated with microcalcification: a radiology-pathology correlation study. The breast journal 2010;**16**(1):28-31 doi: 10.1111/j.1524-4741.2009.00863.x[published Online First: Epub Date]|.
17. Itoh A, Ueno E, Tohno E, et al. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006;**239**(2):341-50 doi: 10.1148/radiol.2391041676[published Online First: Epub Date]|.
18. Wu S. Medlical Statistics. BeiJing: People's Medical Publishing House, 2002:164-78.
19. Li G, Li DW, Fang YX, et al. Performance of shear wave elastography for differentiation of benign and malignant solid breast masses. PloS one 2013;**8**(10):e76322 doi: 10.1371/journal.pone.0076322[published Online First: Epub Date]|.
20. Munoz Carrasco R, Alvarez Benito M, Rivin del Campo E. Value of mammography and breast ultrasound in male patients with nipple discharge. European journal of radiology 2013;**82**(3):478-84 doi: 10.1016/j.ejrad.2012.10.019[published Online First: Epub Date]|.
21. Hamed ST, Abdo MH, Ahmed HH. Breast discharge: ultrasound and Doppler evaluation. Journal of the Egyptian National Cancer Institute 2008;**20**(3):262-70
22. Parajuly SS, Lan PY, Yun MB, Gang YZ, Hua Z. Diagnostic potential of strain ratio measurement and a 5 point scoring method for detection of breast cancer: Chinese experience. Asian Pacific journal of cancer prevention : APJCP 2012;**13**(4):1447-52
23. Sabel MS, Helvie MA, Breslin T, et al. Is duct excision still necessary for all cases of suspicious nipple discharge? Breast J 2012;**18**(2):157-62 doi: 10.1111/j.1524-4741.2011.01207.x[published Online First: Epub Date]|.
24. Alcock C, Layer GT. Predicting occult malignancy in nipple discharge. ANZ journal of surgery 2010;**80**(9):646-9 doi: 10.1111/j.1445-2197.2010.05270.x[published Online First: Epub Date]|.

25. Seltzer MH. Breast complaints, biopsies, and cancer correlated with age in 10,000 consecutive new surgical referrals. Breast J 2004;**10**(2):111-7

26. Lau S, Kuchenmeister I, Stachs A, Gerber B, Krause A, Reimer T. Pathologic nipple discharge: surgery is imperative in postmenopausal women. Annals of surgical oncology 2005;**12**(7):546-51 doi: 10.1245/aso.2005.04.013[published Online First: Epub Date]|.

27. Vargas HI, Vargas MP, Eldrageely K, Gonzalez KD, Khalkhali I. Outcomes of clinical and surgical assessment of women with pathological nipple discharge. The American surgeon 2006;**72**(2):124-8

28. Morrogh M, Morris EA, Liberman L, Borgen PI, King TA. The predictive value of ductography and magnetic resonance imaging in the management of nipple discharge. Annals of surgical oncology 2007;**14**(12):3369-77 doi: 10.1245/s10434-007-9530-5[published Online First: Epub Date]|.

29. The L. Ending violence against doctors in China. The Lancet 2012;**379**(9828):1764

30. Choi JJ, Kang BJ, Kim SH, et al. Role of sonographic elastography in the differential diagnosis of papillary lesions in breast. Japanese journal of radiology 2012;**30**(5):422-9 doi: 10.1007/s11604-012-0070-y[published Online First: Epub Date]|.

Figure Lengend

Figure 1 Flowchart of this study

Figure 2 The image of a 45-years old patient with invasive ductal carcinoma

The left picture showed a B-mode ultrasonography locating the region of interest (ROI). The right picture showed the elastography of the breast lesion, with the blue color indicated a harder tissue and the red color indicated a softer tissue.

Figure 3 The receiver-operating-characteristics (ROC) curve of Elasticity score, VTQ.m and VTQ.c.

The figure 3 showed the area under the curve (AUC), specificity, sensitivity and the best cut-off point of Elasticity score, VTQ.m and VTQ.c. for example the curves showed that Elasticity score yielded an AUC of 0.854, with a cut-off point of 2.50. At this cut-off point, the specificity was 0.736, while the sensitivity was 0.872. So we presented as 2.50(0.736, 0.872) in the figure.

Table 1 The distribution of benign and malignant lesions

Diagnosis after biopsy	Number of lesions (%)
Benign	159
Fibrocystic breast disease	82(51.6)
Fibroadenoma	60 (37.7)
Intraductal pailloma	6(3.8)
Atypical ductal hyperlasia	11(6.9)
Maglinant	78
Invasive ductal carcinoma	61(78.2)
Ductal carcinoma in situ	6(7.7)
Paillary carcinoma	7(9.0)
Lobular carcinoma	4(5.1)

Table 2 Predictors for elastography to detect malignant lesions

Predictors	Estimate	Odds ratio(95%CI)	*P* value
Elasticity score	1.974	7.20(3.77 to 15.92)	<0.001
VTQm	2.542	12.71(4.43 to 46.04)	<0.001
VTQc	2.837	17.06(6.24 to 61.36)	<0.001

VTQ, virtual touch™ quantification. VTQm and VTQc referred to virtual touch™ quantification at the margin and the center of breast masses, respectively. The elasticity score was a qualitative measure of virtual touch™ imaging (VTI).

Table 3 Predictive value of the synthesized quantitative measurement

	Estimates	95%CI	P-value*
Elasticity score			<0.001
Sensitivity	0.872	0.798 to 0.946	
Specificity	0.736	0.765 to 0.883	
PPV	0.708	0.617 to 0.799	
NPV	0.929	0.887 to 0.971	
VTQm			<0.001
Sensitivity	0.615	0.521 to 0.735	
Specificity	0.994	0.951 to 0.999	
PPV	0.925	0.853 to 0.996	
NPV	0.842	0.790 to 0.895	
VTQc			<0.001
Sensitivity	0.628	0.521 to 0.735	
Specificity	0.981	0.941 to 0.996	
PPV	0.907	0.830 to 0.985	
NPV	0.842	0.789 to 0.894	
Synthesized score			-
Sensitivity	0.820	0.735 to 0.906	
Specificity	0.968	0.941 to 0.996	
PPV	0.927	0.866 to 0.989	

NPV	0.917	0.875 to 0.958

Abbreviations: SE, systematic error. CI, confidence interval. PPV, positive predictive value. NPV, negative predictive value. VTQm, virtual touch^{TM} quantification at the margin of a lesion. VTQc, virtual touch^{TM} quantification at the center of a lesion.

*Compared to Synthesized score. The comparison was made using area under the curve(AUC).

Table 4 BI-RADS classification and the synthesized measures

BI-RADS*	Elasticity score	VTQm	VTQc	Synthesized score	
	Mean(sd)	Mean(sd)	Mean(sd)	Benign(N, %)	Malignant(N, %)
3	1.98(0.64)	2.64(0.51)	2.86(0.55)	126(75.0)	13(18.8)
4A	2.84(0.68)	2.85(0.76)	3.02(0.76)	25(14.9)	12(17.4)
4B	3.91(0.29)	3.24(0.86)	3.48(0.78)	5(3.0)	17(24.6)
4C	3.71(0.61)	2.90(0.70)	3.34(0.84)	9(5.4)	5(7.2)
5	4.00(0.81)	3.31(0.62)	3.71(0.75)	3(1.8)	22(31.9)

Abbreviations: BI-RADS: the Breast Imaging Reporting and Data System. VTQm, virtual touch^{TM} quantification at margin of a lesion. VTQc, virtual touch^{TM} quantification at center of a lesion. SD, standardized deviation. N, number of counts.

* In the BI-RADS, 3 referred to probably benign, 4 referred to suspicious malignancy, 5 referred to highly suggestive of malignancy. In the 4 category, a subclassification was also used: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	1
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	5
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	5
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	5
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Yes
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	5
Test methods	7	The reference standard and its rationale.	7-8
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	5-8
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	5-8
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	6, 8
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	6, 8
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	8-9
	13	Methods for calculating test reproducibility, if done.	9
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	9-10
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	10
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	10
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	10
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	10
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	10
	20	Any adverse events from performing the index tests or the reference standard.	10
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	11-12
	22	How indeterminate results, missing data and outliers of the index tests were handled.	7, 11-12
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	11-12
	24	Estimates of test reproducibility, if done.	11-12
DISCUSSION	25	Discuss the clinical applicability of the study findings.	12-13

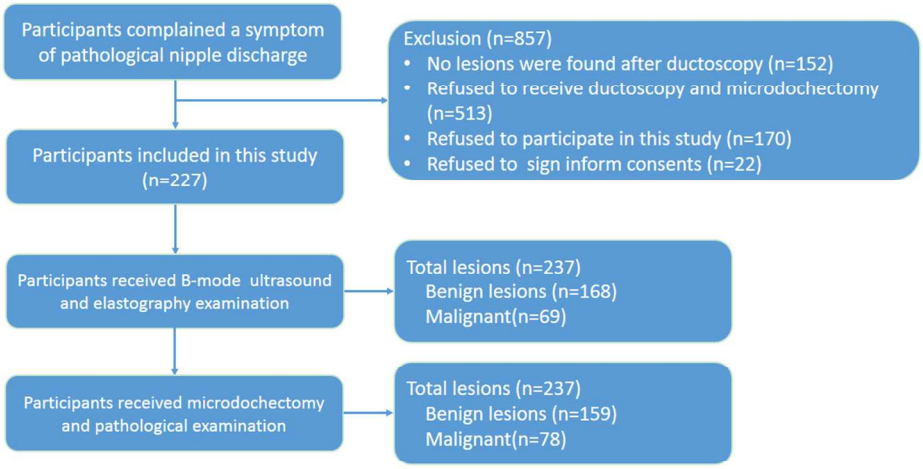

Figure 1 Flowchart of this study
338x190mm (96 x 96 DPI)

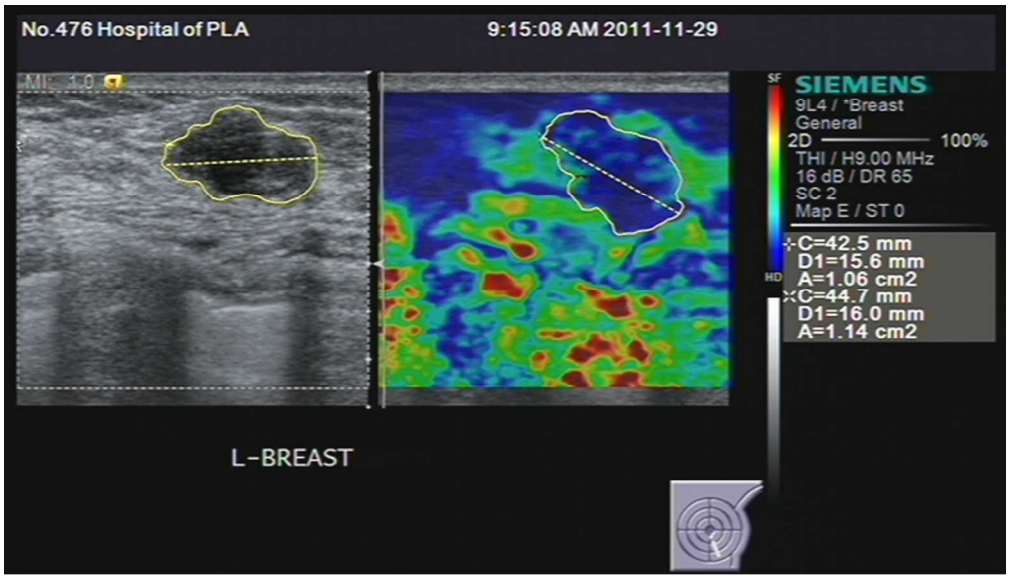

Figure 2 The image of a 45-years old patient with invasive ductal carcinoma
The left picture showed a B-mode ultrasonography locating the region of interest (ROI). The right picture showed the elastography of the breast lesion, with the blue color indicated a harder tissue and the red color indicated a softer tissue.

56x32mm (600 x 600 DPI)

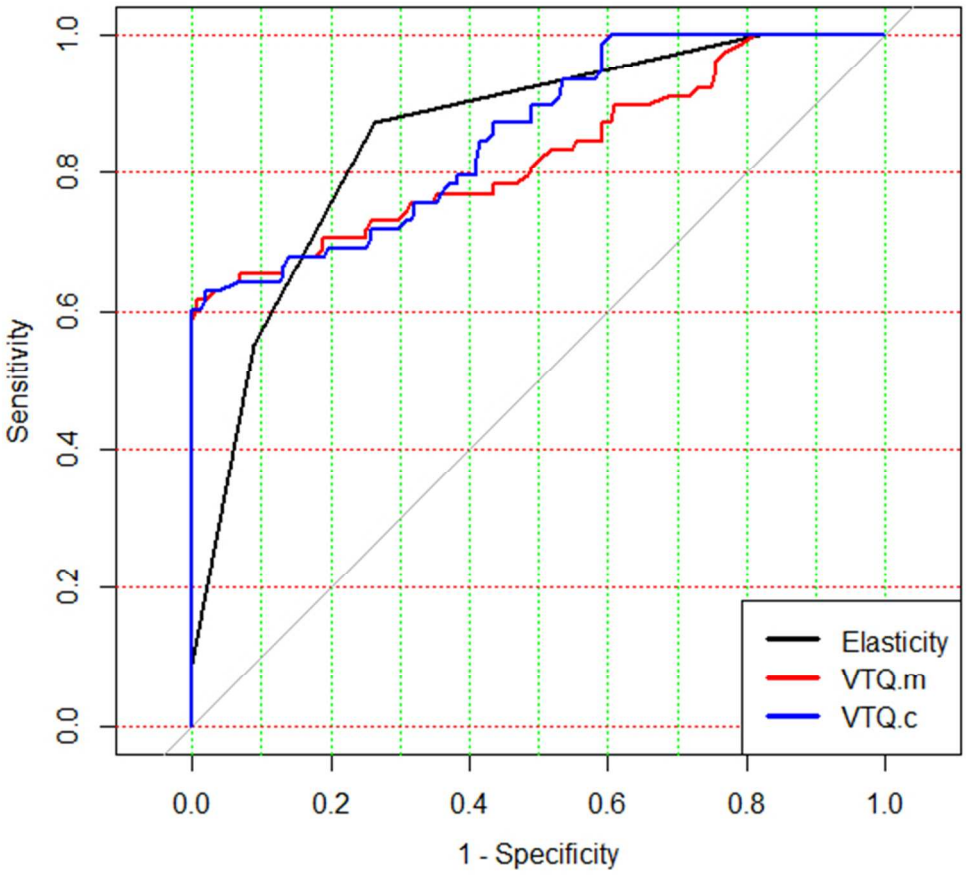

Figure 3 The receiver-operating-characteristics (ROC) curve of Elasticity score, VTQ.m and VTQ.c. The figure 3 showed the area under the curve (AUC), specificity, sensitivity and the best cut-off point of Elasticity score, VTQ.m and VTQ.c. for example the curves showed that Elasticity score yielded an AUC of 0.854, with a cut-off point of 2.50. At this cut-off point, the specificity was 0.736, while the sensitivity was 0.872. So we presented as 2.50(0.736, 0.872) in the figure.

BMJ Open

Diagnostic accuracy of shear wave elastography for prediction of breast malignancy in patients with pathological nipple discharge

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-008848.R1
Article Type:	Research
Date Submitted by the Author:	23-Aug-2015
Complete List of Authors:	Guo, xiao-bo Liu, Ying Li, Wanhu
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Radiology and imaging
Keywords:	Breast tumours < ONCOLOGY, ONCOLOGY, Adult oncology < ONCOLOGY

Diagnostic accuracy of shear wave elastography for prediction of breast malignancy in patients with pathological nipple discharge

Xiaobo Guo[1] Ying Liu[1] Wanhu Li[2*]

1. Department of Ultrasonography, 476 Clinical Department of Fuzhou General Hospital of Nanjing military region, China
2. Department of Radiology, Shandong Cancer Hospital and Institute, No. 440, Jinan Road, Jinan 250117, China

Email address: Xiaobo Guo: 782605810@qq.com; Ying Liu: 1637824435@qq.com; Wanhu Li :lwhvzg@126.com.

Corresponding author: Xiaobo Guo, phone number：0531-86833526, E-mail: 782605810@qq.com.

ABSTRACT

Objectives: To test the diagnostic accuracy of shear wave elastography (SWE) in detecting malignancy in patients with pathological nipple discharge (PND).

Design: Prospective diagnostic accuracy study comparing a combination of qualitative and quantitative measurements of SWE (index test) to a ductoscopy and microdochectomy for histological diagnosis (reference test).

Setting: Fuzhou general hospital of Nanjing military command.

Participants: A total of 379 patients with PND were finally included from January, 2011 to March 2014, after we screened 1084 possible candidates. All participants were evaluated through a SWE with qualitative parameters generated by Virtual Touch^{TM} tissue imaging (VTI) and quantitative parameters generated by Virtual Touch^{TM} tissue quantification (VTQ). All the patients were consented to receive a ductoscopy and microdochectomy for histological diagnosis, and the results were set as a reference test.

Outcome measures: Sensitivity and specificity of the combined VTI and VTQ of the SWE for detection of malignancy in patients with PND.

Results: Results: The 379 participants presented with 404 lesions. The results of pathological examination showed that 326(80.7%) of the 404 lesions were benign and the other 78 (19.3%) were malignant. An area under the curve (AUC) of elasticity score, VTQm and VTQc were 0.872, 0.825, 0.857 respectively, with the corresponding cut-off point as 2.50, 2.860m/s and 3.015m/s respectively. After a combination of these measurements, the sensitivity, specificity, positive and negative predictive value (PPV and NPV) were 90.0%, 72.0%, 43.5% and 96.7%, respectively. The

sensitivity analysis showed a consistent result.

Conclusion: Ultrasonography elastography is sensitive for PND patients and could be used as a triage test before ductoscopy examination. A study for further improvement of diagnostic sensitivity is warranted by assigning proper weights to the qualitative and quantitative measurements for elastography.

Key words: Shear wave elastography; Pathological nipple discharge; Microdochectomy; Breast Imaging- Reporting and Data System.

Strengths and limitations of this study

- Diagnostic accuracy of shear wave elastography (SWE) for detecting malignancy of PND patients is rarely studied.
- For the first time, this study tested diagnostic accuracy of a synthesized measurement of qualitative and quantitative measures of SWE for detection of malignancy in PND patients.
- Limitations include that the weight of each measurement in the synthesized score was assigned evenly and the surgeon was not blinded.

Introduction

In 2014, the American cancer society reported 235,030 new cases of breast cancer and 40,430 deaths because of it[1]. The prevalence rate of breast cancer is the highest in all types of malignant carcinoma in Chinese women [2]. The cases of Chinese women with breast cancer account for 12.2% of all newly diagnosed breast cancer around the world and 9.6% of deaths from breast cancer [2]. Pathological nipple discharge (PND) is the third most common complaint of patients prompting referral to doctors for breast diseases[3]. PND is believed to be an indicator for breast carcinoma, especially when the nipple discharge is bloody[3]. So mastectomies were carried out unnecessarily without obtaining a histopathological diagnosis before the year 1950. Although mastectomies are not necessarily needed nowadays, patients with PND would be normally suggested to take further examinations to rule out breast malignancy. These examinations involve mammography, ultrasonography, ductoscopy with or without galactography, as well as cytological tests. Mammography and B-mode ultrasonography are used for screening patients with high risk of breast carcinoma, but they are not sensitive enough for detecting malignancy in PND patients, since small lesions may be occult with these 2 modalities[4]. Ductoscopy and microdochectomy are the gold standard for diagnosis of malignancy in PND patients[5 6], but they are invasive examinations with high cost. Cytological tests could only be used as supportive evidence[7]. Therefore, a triage test is needed before the PND patients are suggested to take the ductoscopy and microdochectomy examination.

Ultrasonographic elastography works through the assumption that tissue compression produces strain within the tissue, and that the strain is smaller in firm tissue than in soft tissue. So we could

evaluate the firmness of the tissue to predict malignancy through measuring tissue strain ratio, which is usually measured as elasticity[8]. Recently, several studies indicated that ultrasonographic elastography may be a useful tool for differentiating benign and malignant masses. The sensitivity of predicting malignancy in patients complaining breast lesions ranges from 80% to 98%, while the specificity ranges from 66% to 84%[9-14]. Recently, a shear wave elastography (SWE) is reported to be less operator-dependent and predictive of ductal or intraductal carcinoma[8 15 16]. A five-level elasticity score is normally used for qualitative assessment of the images of elastography, while a quantitative measurement of elastography is also becoming popular; a combination of these ultrasonic measurements seems to be beneficial for improving diagnostic accuracy of breast malignancies[8 16]. We noticed that the diagnostic value of SWE to detect malignancy in patients with PND has rarely been studied, especially with a combination of qualitative and quantitative measurements of SWE. So we conducted a study aiming to find out the sensitivity and specificity accuracy of shear wave elastography (SWE) for predicting malignancy in patients with PND.

Methods and materials

This study is designed for evaluating diagnostic accuracy of shear wave elastography (SWE) for detection of malignancy in patients with PND. The study protocol is approved by the ethical review board in Fuzhou general hospital of Nanjing military command (No. FZJQ2011018), in which the study was conducted and the data were collected. Data collection was planned before the index and reference test.

Patients

Patients who complained of PND within 1 year were recruited from outpatient settings in the

Fuzhou general hospital of Nanjing military command. The PND was defined as nipple discharge in women who stopped breastfeeding and were not being pregnant[17]. The patients were screened for eligibility according to the following inclusion criteria: (1) with an age over 18 years; (2) with or without palpable breast masses; (3) agree to participate in this study; (4) providing a written consent for further examination, biopsy and use of the study data. We excluded the participants who (1) refused to receive a ductoscopy and microdochectomy or a biopsy; (2) had serious damage in internal organs that may bias the results; (3) could not receive local anesthesia because of allergy; (4) were taking or have taken chemotherapy or radiotherapy for malignant breast tumors.

Ultrasonographic imaging

The eligible participants were asked to take ultrasonographic examinations before receiving ductoscopy and microdochectomy. The ultrasonographic examinations included a B-mode ultrasound (BUS) and an elastography with acoustic radiation force Impulse (ARFI), which were performed with a Siemens ACUSON S2000 ultrasonographic system (Siemens Ltd., China, product standard: YZB/USA 3876-2010). One operator (with 10 years of experience in ultrasonography imaging) ran the BUS and the ARFI to get the BUS and elastography images, using a superficial probe (9L4, Siemens Ltd., China) with a frequency of 4 to 9 Hz.

The BUS images of the breast lesions were obtained and categorized according to Breast Imaging Reporting and Data System (BI-RADS). In the BI-RADS, number 1 refers to negative findings; number 2 refers to benign findings; number 3 refers to probably benign findings; number 4 refers to suspicious malignancy; number 5 is highly suggestive of malignancy. Moreover, we applied a subclassification scheme for BI-RADS 4: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate

suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy [18]. Two radiologists (XBG and YL who are with at least 5 years of experience in ultrasonography) independently assessed the BUS images in the following aspects: the shape of the mass, depth, orientation, margin, boundary of the lesions, echo pattern, posterior acoustic features, as well as surrounding tissue changes. All the detected masses were divided into two types (solid and cystic) according to the presence of a cystic portion. In cases involving a solid mass in a dilated duct (intraductal) or cyst (intracystic) with a predominant cystic portion, the radiologists considered the cases as cystic type lesions.

After the BUS examination, the operator switched the ultrasonic mode to the virtual touch™ quantification (VTQ). The patients were told to hold the breath, when the operator detected the inside and margin of the lesion, with a previous set region of interest (ROI) in a fixed region of 5×5cm. The operator recorded the value of VTQ (m/s) and the depth of the ROI (cm) from skin for each measurement, in a total of 5 measurements. We calculated the mean value of the 5 measurements as the final VTQ score. Additionally, the value of VTQ was measured at both the margin (VTQm) and center (VTQc) of the breast lesions. Measurable shear wave velocity (SWE) ranges from 0 to 9 m/s, a value of ROI exceed the range would be shown as “X.XX m/s” on the screen. The participants with this “X.XX m/s” value were excluded.

After finishing examination with the VTQ mode, the operators switched to virtual touch™ imaging (VTI). ARFI sequences were generated to evaluate the elasticity scores of the breast lesions in VTI mode. The VTI images were evaluated and categorized through a five-level scoring system[19]. An elasticity score of 1 indicates strain in the entire hypoechoic lesion (the entire lesion is equally

shaded in green); score 2 indicates that strain is not seen in part of the hypoechoic lesion (the lesion is shown as a mosaic of green and blue); score 3 indicates that strain is in the peripheral areas only (peripheral areas in green); score 4 indicates no strain in the entire hypoechoic area (lesion is shown in blue); score 5 indicates no strain in the entire hypoechoic lesion or the surrounding area(both the entire hypoechoic lesion and its surrounding area are shown in blue). A higher elasticity score indicates a stiffer tissue. For cystic-type lesions, we analyzed the elasticity scores in solid areas. After the B-mode ultrasonography and ARFI elastography examination, 2 radiologists independently reviewed the images and the measurements without awareness of the medical history of the participants. Any discrepancy happened was solved by discussion.

Microdochectomy and pathological examination

After finishing the examination of BUS imaging and elastography, the participants received a ductoscopy examination. Microdochectomy was performed to get a biopsy sample if the ductoscopy indicated a breast lesion. The result of microdochectomy and pathological examination was set as a reference test (the golden standard). The patients were asked to stay in supine position with the arm preferably extending on a side board. After skin disinfection, the fluid-producing ducts were identified by massaging the breast towards the nipple. After dilating the target ducts with lacrimal dilators, a surgeon (with 5 years of experience in microdochectomy) introduced endoscope (a fiberoptic ductoscopy, 0.8mm of sheath, 80mm of length; manufactured in Denzlingen, Germany) through the ducts. The surgeon was not blinded from the results of the ultrasonographic examinations. After inserting the ductoscopy, the surgeon injected saline into the ducts to keep them expanded. The ductoscopy was advanced through the ducts under direct vision from the

endoscope camera, all tiers of branches were examined until the endoscope could not be advanced further or an obstructing lesion was detected. When an obstructing lesion was found, the image of this lesion would be taken using an endoscope camera, moreover, an endoscopy guided biopsy would be performed as scheduled. After biopsy and image taken, the patient would be referred to a surgeon for further treatment advice.

After detection of ductal lesions, methylene blue was injected into the discharged mammary duct. Local anesthesia was performed and then an incision around the areola was implemented. The pathological duct was separated according to tissue with methylene-blue stain, and the duct was then dissected and excised. The operator then pulled the clip out of duct together with the scope after the excision was finished. The biopsy samples over 1 mm^3 were fixed using formalin after sample collection and were sent for histologic examination.

Statistical analysis

Statistical analysis was performed to specify the diagnostic sensitivity and specificity of a synthesized measurement of VTI (elasticity score) and VTQ measurements. First, we would assess the interobserver variability between the two radiologists (XBG and YL) using the Cohen's kappa test. The value of kappa over 0.75 was considered as excellent, 0.40 to 0.75 as fair to good, and below 0.40 as poor. Second, we calculated the sensitivity, specificity and the area under the curve (AUC) of elasticity score, VTQm and VTQc using receiver operating characteristics (ROC) curve. A valued of AUC of 0.5 was considered as no diagnostic value; 0.5 to 0.7, low diagnostic value; 0.7 to 0.9, mediate diagnostic value; 0.9 to 1.0, high diagnostic value[20]. We calculated the cut-off point of the elasticity score, VTQm and VTQc and gave the sensitivity and specificity respectively, based

on both the highest sum of sensitivity and specificity and a relatively high sensitivity of the measurements. According to the cut-off points, we synthesized the measurements by the following formula: the synthesized score=elasticity score + VTQm + VTQc. A synthesized score of 1 was given, if a value over the cut-off point was observed in a breast lesion; otherwise, a score of -1 was be given. A breast lesion with a final score below 0 was classified as benign, otherwise it was categorized as malignant. Third, after differentiating benign and malignant lesions using the synthesized score, we evaluated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The diagnostic value of this synthesized measurement was compared with the biopsy results using the McNemar's test. We also compared the BI-RADS with the synthesized measurement using the chi-square test. All statistical analysis was performed using Statistical Package for the Social Sciences (SPSS, Version 20.0, Armonk, NY: IBM Corp.) and R software (www.r-project.org, version 3.1.1).

Results

A total of 379 patients with PND were included in this study, after 1084 possible candidates were screened and 705 of them were excluded. The 379 participants presented with 404 lesions, while 25 participants showed breast lesions on both sides. The process of eligibility screening and a brief description of the results of the BUS, elastography, ductoscopy and microdochectomy examinations were shown in figure 1. The age of the included participants ranged from 25 to 67 years. The result of pathological examination showed that 326(80.7%) of the 404 lesions were benign and 78 (19.3%) were malignant. Table 1 shows the types of the benign and malignant lesions. Figure 2 and 3 show examples of invasive ductal carcinoma and fibroadenoma diagnosed with B-mode ultrasonography

and elastography. Of all the breast lesions, 191 (47.2%) were palpable masses; 127 of them were benign and 64 were malignant. A total of 59 participants complained of bloody discharge in 62 breast lesions(15.3%); 9 of them were benign and 53 were malignant. The size of the lesions ranged from 3.2mm to 38.7mm. The benign lesions showed a mean size of 8.16mm, while the malignant lesions showed a mean size of 15.58mm. Moderate or severe pain (with a score of visual analog scale over 50 mm) in the breast was reported by 133 of the 227 participants who received ductoscopy. Other adverse events were not reported. The agreement between the two radiologists for ultrasound elastography was good (kappa=0.76, p<0.01).

The diagnostic cut-off points of the elasticity score, VTQm and VTQc

We calculated the cut-off points for elasticity scores, VTQm and VTQc. The results showed that a cut-off point of 2.50 for elasticity score yielded a sensitivity of 0.872(95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. We also calculated a cut-off point of 2.860 m/s for VTQm with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of 3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857. Table 2 showed raw numbers and diagnostic estimates of the elasticity score, VTQm and VTQc. Figure 4 shows the ROC curve of each measurement.

Diagnostic accuracy of a synthesized measurement of the elasticity score, VTQm and VTQc

We used a synthesized measurement of both quantitative and qualitative parameters to predict malignant breast lesions. The measurement categorized 243 (60.1%) breast lesions as benign and

161 (39.9%) as malignant. Table 3 shows a comparison of BI-RADS with the synthesized measurement for elastography, and it showed that using BI-RADS diagnosed significantly less participants with malignancy than the synthesized measurement (X-squared = 72.035, p<0.001).

The biopsy results showed that 326 (80.6%) breast lesions were classified as benign while 78 (19.4%) were malignant. We used the McNemar's test to compare the synthesized measurement with the biopsy results, the result showed that significantly more malignant lesions were found by elastography with the synthesized measurement (McNemar's chi-squared = 67.919, p<0.001). The sensitivity, specificity, positive and negative predictive value (PPV and NPV) were 0.900, 0.720, 0.435 and 0.967, respectively. Table 2 shows the raw numbers and diagnostic estimates of the synthesized measurement. In the sensitivity analysis, we classified participants who were found with no pathological changes by ductoscopy into a malignant category. The result showed that the measurement of synthesized score categorized 212 (52.5%) breast lesions as benign and 192 (47.5%) as malignant. The sensitivity and specificity were 0.949 and 0.638, respectively.

Discussion

The results of our study showed that a combination of qualitative and quantitative measurements for SWE elastography is sensitive in predicting malignancy in PND patients, which is consistent with several recent studies testing the diagnostic value of elastography in breast masses[8 16 21]. However, several studies reported that ultrasonography was not predictive to malignancy in PND patients[22-24]. We discussed our major findings and the remaining questions below.

Predicting malignancy in patients with PND is still challenging, Sabel and colleagues found that 95% of the PND lesions were benign, which need not be excised. However, most patients will choose

duct excision for palliative purposes[25]. In the screen phase of our study, 53 patients refuse to take ductoscopy or duct excision and therefore were excluded. The patients refuse to take the examinations for that no evidence of malignancy was found and they were afraid of the contraindications of ductoscopy or duct excision. Therefore, a non-invasive method with accurate diagnosis is needed by patients with PND. Alcock and Layer found no predictive value of ductoscopy and ultrasonography in differentiating benign and malignant breast lesions in PND patients. The authors of this study suggested that ductal surgery was the only reliable way of providing a diagnosis[17]. In our study, we found the SWE elastography with a sensitivity of 90% for screening malignancy in PND patients. This result indicated that SWE elastography could be used as a helpful triage test before suggesting PND patients to ductoscopy examinations. And compared with B-mode ultrasound, SWE elastography is with lower false negative risk, according to the comparison of the BI-RADS and synthesized measurements in our study. And therefore SWE elastography is of better diagnostic performance than B-mode ultrasonography. So our study, for the first time, revealed that ultrasonographic elastography is valuable in predicting malignant lesions in patients with PND.

The size of the breast mass might be an important factor in differentiating benign and malignant lesions. We found that PND patients with malignant lesions presented a bigger size of the breast mass (malignant versus benign, 15.58 mm versus 8.15 mm), which might be easier to be detected by B-mode ultrasonography and elastography. This finding is consistent with other studies using elastography to detect breast lesions[8 21]. However, whether the size of the breast mass could be used as an independent indicator for predicting breast malignancy needs more studies.

The synthesized score we used in this study is similar to a published method for classification of benign and malignant breast lesions, except that we did not include the BI-RADS[8]. The BI-RADS is a classification system, which is in part coincident with the elasticity score[19]. We chose elasticity score instead of the BI-RADS into the synthesized score. The elasticity score is a qualitative measurement, which was treated as continuous data in several studies[8 21 26]. We also treated the elasticity score as a continuous data, so that we could calculate a cut-off point and include it into the synthesized score. The way that we handled the elasticity score may bring bias into our results. However, neither did we find a better way to solve the problem, nor did the previous studies. Therefore, in future studies, a qualitative measurement should be developed, which differentiate benign and malignant breast lesions more clearly than the elasticity score.

In this study, we incorporated both VTQ at the margin (VTQm) and in the center (VTQc) into the synthesized score. To the best of our knowledge, we are the first to use these two quantitative measurements, instead, previous studies used VTQc alone[8 9 21]. Inclusion of both VTQm and VTQc could increase the understanding of the breast lesions, since they were reported to be different[19 21]. The results of our study indicated that both VTQm and VTQc were significantly higher in malignant lesions than benign ones. So we might have assigned more weight in the VTQ measurement, and therefore made breast lesions with higher VTQ more easier to be malignant in the synthesized score.

We have several limitations in this study. First, a total of 152 participants were not found of pathological change after ductoscopy examination, and thus were not performed microdochectomy to get a biopsy result. These participants were classified as with benign lesions without solid

evidence. We might have missed the malignant lesions, since ductoscopy could only reach a depth within 2cm[5]. To ensure the reliability of our study, we ran a sensitivity analysis, in which all these participants were classified as with malignant lesions. And we found the result was consistent. Moreover, we found that the incidence of malignancy in PND patients is 19.3% in our study, while a reported incidence ranged from 4% to 29%[26-29]. This supported that the 152 participants were classified as with benign lesions. Second, the weight of elasticity score, VTQm and VTQc were assigned evenly, and therefore made the value of VTQ a more important factor than elasticity score in predicting malignancy in PND patients. Third, the surgeon who performed the ductoscopy was not blinded from the ultrasonographic results, which may introduce performance bias.

In conclusion, ultrasonography elastography is sensitive for PND patients and could be considered as a triage test before ductoscopy examination. However, a study for further improvement of diagnostic sensitivity is warranted by assigning proper weights to the qualitative and quantitative measurements for elastography.

Authors' contributions

XBG and YL contributed to the design and manuscript writing. XBG and WHL conceptualized the study and data analysis. YL assisted with data collection. All authors read and approved the final manuscript.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments: This study was financially supported by the Key Project of science and technology of Fujian Province of China (No. 2009Y0043). We would like to thank all the participants who participated in this study.

Reference

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: A Cancer Journal for Clinicians 2014;**64**(1):9-29 doi: 10.3322/caac.21208[published Online First: Epub Date]|.
2. Fan L, Strasser-Weippl K, Li J-J, et al. Breast cancer in China. The Lancet Oncology 2014;**15**(7):e279-e89
3. Chen L, Zhou WB, Zhao Y, et al. Bloody nipple discharge is a predictor of breast cancer risk: a meta-analysis. Breast cancer research and treatment 2012;**132**(1):9-14 doi: 10.1007/s10549-011-1787-5[published Online First: Epub Date]|.
4. Jagmohan P, Pool FJ, Putti TC, Wong J. Papillary lesions of the breast: imaging findings and diagnostic challenges. Diagnostic and interventional radiology (Ankara, Turkey) 2013;**19**(6):471-8 doi: 10.5152/dir.2013.13041[published Online First: Epub Date]|.
5. Denewer A, El-Etribi K, Nada N, El-Metwally M. The role and limitations of mammary ductoscope in management of pathologic nipple discharge. Breast J 2008;**14**(5):442-9 doi: 10.1111/j.1524-4741.2008.00620.x[published Online First: Epub Date]|.
6. Makita M, Akiyama F, Gomi N, Iwase T. Mammary ductoscopy and watchful follow-up substitute microdochectomy in patients with bloody nipple discharge. Breast cancer (Tokyo, Japan) 2014 doi: 10.1007/s12282-014-0561-z[published Online First: Epub Date]|.
7. Kooistra BW, Wauters C, van de Ven S, Strobbe L. The diagnostic value of nipple discharge cytology in 618 consecutive patients. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2009;**35**(6):573-7 doi: 10.1016/j.ejso.2008.09.009[published Online First: Epub Date]|.
8. Kim YS, Park JG, Kim BS, Lee CH, Ryu DW. Diagnostic value of elastography using acoustic radiation force impulse imaging and strain ratio for breast tumors. J Breast Cancer 2014;**17**(1):76-82 doi: 10.4048/jbc.2014.17.1.76[published Online First: Epub Date]|.
9. Tozaki M, Isobe S, Sakamoto M. Combination of elastography and tissue quantification using the acoustic radiation force impulse (ARFI) technology for differential diagnosis of breast masses. Japanese journal of radiology 2012;**30**(8):659-70 doi: 10.1007/s11604-012-0106-3[published Online First: Epub Date]|.
10. Tozaki M, Isobe S, Fukuma E. Preliminary study of ultrasonographic tissue quantification of the breast using the acoustic radiation force impulse (ARFI) technology. European journal of radiology 2011;**80**(2):e182-7 doi: 10.1016/j.ejrad.2011.05.020[published Online First: Epub Date]|.
11. Meng W, Zhang G, Wu C, Wu G, Song Y, Lu Z. Preliminary results of acoustic radiation force impulse (ARFI) ultrasound imaging of breast lesions. Ultrasound in medicine & biology 2011;**37**(9):1436-43 doi: 10.1016/j.ultrasmedbio.2011.05.022[published Online First: Epub Date]|.
12. Taylor K, O'Keeffe S, Britton PD, et al. Ultrasound elastography as an adjuvant to conventional ultrasound in the preoperative assessment of axillary lymph nodes in suspected breast cancer: a pilot study. Clinical radiology 2011;**66**(11):1064-71 doi: 10.1016/j.crad.2011.05.015[published Online First: Epub Date]|.
13. Sadigh G, Carlos RC, Neal CH, Dwamena BA. Accuracy of quantitative ultrasound elastography for differentiation of malignant and benign breast abnormalities: a meta-analysis. Breast cancer research and treatment 2012;**134**(3):923-31 doi: 10.1007/s10549-012-2020-x[published Online First: Epub Date]|.

14. Gong X, Xu Q, Xu Z, Xiong P, Yan W, Chen Y. Real-time elastography for the differentiation of benign and malignant breast lesions: a meta-analysis. Breast cancer research and treatment 2011;**130**(1):11-8 doi: 10.1007/s10549-011-1745-2[published Online First: Epub Date]|.
15. Kılıç F, Ustabaşıoğlu FE, Samancı C, et al. Ductal Carcinoma In Situ Detected by Shear Wave Elastography within a Fibroadenoma. Journal of breast cancer 2014;**17**(2):180-3 doi: 25013441[published Online First: Epub Date]|.
16. Youk JH, Son EJ, Gweon HM, Kim H, Park YJ, Kim JA. Comparison of Strain and Shear Wave Elastography for the Differentiation of Benign From Malignant Breast Lesions, Combined With B-mode Ultrasonography: Qualitative and Quantitative Assessments. Ultrasound in medicine & biology 2014;**40**(10):2336-44 doi: 10.1016/j.ultrasmedbio.2014.05.020[published Online First: Epub Date]|.
17. Alcock C, Layer GT. Predicting occult malignancy in nipple discharge. ANZ journal of surgery 2010;**80**(9):646-9 doi: 10.1111/j.1445-2197.2010.05270.x[published Online First: Epub Date]|.
18. Sanders MA, Roland L, Sahoo S. Clinical implications of subcategorizing BI-RADS 4 breast lesions associated with microcalcification: a radiology-pathology correlation study. Breast J 2010;**16**(1):28-31 doi: 10.1111/j.1524-4741.2009.00863.x[published Online First: Epub Date]|.
19. Itoh A, Ueno E, Tohno E, et al. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006;**239**(2):341-50 doi: 10.1148/radiol.2391041676[published Online First: Epub Date]|.
20. Wu S. Medlical Statistics. BeiJing: People's Medical Publishing House, 2002:164-78.
21. Li G, Li DW, Fang YX, et al. Performance of shear wave elastography for differentiation of benign and malignant solid breast masses. PloS one 2013;**8**(10):e76322 doi: 10.1371/journal.pone.0076322[published Online First: Epub Date]|.
22. Munoz Carrasco R, Alvarez Benito M, Rivin del Campo E. Value of mammography and breast ultrasound in male patients with nipple discharge. European journal of radiology 2013;**82**(3):478-84 doi: 10.1016/j.ejrad.2012.10.019[published Online First: Epub Date]|.
23. Hamed ST, Abdo MH, Ahmed HH. Breast discharge: ultrasound and Doppler evaluation. Journal of the Egyptian National Cancer Institute 2008;**20**(3):262-70
24. Parajuly SS, Lan PY, Yun MB, Gang YZ, Hua Z. Diagnostic potential of strain ratio measurement and a 5 point scoring method for detection of breast cancer: Chinese experience. Asian Pacific journal of cancer prevention : APJCP 2012;**13**(4):1447-52
25. Sabel MS, Helvie MA, Breslin T, et al. Is duct excision still necessary for all cases of suspicious nipple discharge? Breast J 2012;**18**(2):157-62 doi: 10.1111/j.1524-4741.2011.01207.x[published Online First: Epub Date]|.
26. Choi JJ, Kang BJ, Kim SH, et al. Role of sonographic elastography in the differential diagnosis of papillary lesions in breast. Japanese journal of radiology 2012;**30**(5):422-9 doi: 10.1007/s11604-012-0070-y[published Online First: Epub Date]|.

Figure Lengend

Figure 1. Flowchart of this study

* We used the Breast Imaging Reporting and Data System (BI-RADS) to classified the breast lesions in B-mode ultrasound. In the BI-RADS, number 3 refers to probably benign, number 4 refers to suspicious malignancy, number 5 refers to highly suggestive of malignancy. In the level 4 category, a subclassification is used: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy.

Figure 2. The image of a 45-years old patient with invasive ductal carcinoma

The left picture shows a B-mode ultrasonography locating the region of interest (ROI). The right picture shows the elastography of the breast lesion, with the blue color indicates a harder tissue and the red color indicated a softer tissue.

Figure 3. The image of a 41-years old patient with fibroadenoma

A B-mode ultrasonography at the left of the picture shows a location of the region of interest (ROI). An elastography picture at the right shows the elasticity of the fibroadenoma.

Figure 4. The receiver-operating-characteristics (ROC) curve of elasticity score, VTQm and VTQc

VTQm, virtual touch^{TM} quantification at margin of a lesion. VTQc, virtual touch^{TM} quantification at center of a lesion. The figure shows the area under the curve (AUC), specificity and sensitivity of Elasticity score, VTQm and VTQc. A cut-off point of 2.50 was selected for elasticity score yielding a sensitivity of 0.872 (95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. A cut-off point of 2.860 m/s was selected for VTQm yielding a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of

3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857.

Table 1. The distribution of benign and malignant lesions classified by microdochectomy

Diagnosis after biopsy	Number of lesions (%)
Benign	326
No pathological lesions	167 (51.2)
Fibrocystic breast disease	82(25.1)
Fibroadenoma	60 (18.4)
Atypical ductal hyperlasia	11(3.5)
Intraductal pailloma	6(1.8)
Maglinant	78
Invasive ductal carcinoma	61(78.2)
Ductal carcinoma in situ	6(7.7)
Paillary carcinoma	7(9.0)
Lobular carcinoma	4(5.1)

Measurements	Correctly positive	Correctly negative	Sensitivity (%)	Specificity (%)	PPV(%)	NPV(%)
Elasticity (n=404)	68/78	240/326	87.2 (79.5 to 94.0)	73.6 (66.0 to 80.5)	28.5 (22.8 to 34.1)	94.9 (91.5 to 98.4)
VTQm (n=404)	57/78	242/326	73.1(62.8 to 82.1)	74.2(67.3 to 80.5)	27.6 (21.8 to 33.3)	91.9 (87.8 to 95.9)
VTQc (n=404)	57/78	227/326	73.1(62.8 to 82.1)	69.8(61.6 to 76.1)	60.7 (51.7 to 69.8)	96.6 (94.5 to 98.7)
Synthesized score (n=404)	70/78	234/326	90.0 (83.3 to 96.2)	72.0 (66.9 to 77.0)	43.5 (35.8 to 51.1)	96.7 (94.5 to 99.0)

Table 2. Predictive value of the synthesized quantitative measurement

Abbreviations: SE, systematic error. CI, confidence interval. PPV, positive predictive value. NPV, negative predictive value. VTQm, virtual touch^{TM} quantification at the margin of a lesion. VTQc, virtual touch^{TM} quantification at the center of a lesion.

*Compared to Synthesized score. The comparison was made using area under the curve(AUC).

Table 3. The BI-RADS classification and the synthesized measures

BI-RADS*	Elasticity score	VTQm	VTQc	Synthesized score	
	Mean(sd)	Mean(sd)	Mean(sd)	Benign(n, %)	Malignant(n, %)
3(n=293)	1.98(0.64)	2.64(0.51)	2.86(0.55)	207(85.2)	86(53.4)
4A(n=46)	2.84(0.68)	2.85(0.76)	3.02(0.76)	25(10.3)	21(13.1)
4B(n=23)	3.91(0.29)	3.24(0.86)	3.48(0.78)	4(1.6)	19(11.8)
4C(n=16)	3.71(0.61)	2.90(0.70)	3.34(0.84)	5(2.1)	11(6.8)
5(n=26)	4.00(0.81)	3.31(0.62)	3.71(0.75)	2(0.8)	24(14.9)

Abbreviations: BI-RADS: the Breast Imaging Reporting and Data System. VTQm, virtual touch^{TM} quantification at margin of a lesion. VTQc, virtual touch^{TM} quantification at center of a lesion. SD, standardized deviation. N, number of counts.

* In the BI-RADS, number 3 refers to probably benign, number 4 refers to suspicious malignancy, number 5 refers to highly suggestive of malignancy. In the 4 category, a subclassification is used: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy.

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	1
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	5
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	5
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	5
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Yes
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	5
Test methods	7	The reference standard and its rationale.	5
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	5-8
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	5-8
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	6, 8
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	6, 8
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	9-10
	13	Methods for calculating test reproducibility, if done.	9
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	10
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	10
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	10, figure 1
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	10
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	10
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	10, table 2
	20	Any adverse events from performing the index tests or the reference standard.	10
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	11-12, tabl e 2
	22	How indeterminate results, missing data and outliers of the index tests were handled.	7, 11-12
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	No
	24	Estimates of test reproducibility, if done.	No
DISCUSSION	25	Discuss the clinical applicability of the study findings.	12-13

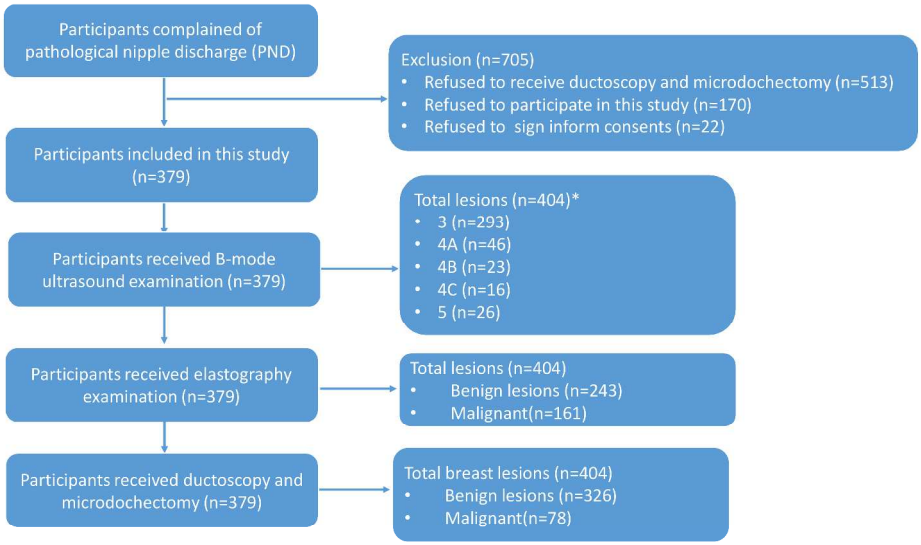

Figure 1. Flowchart of this study

* We used the Breast Imaging Reporting and Data System (BI-RADS) to classified the breast lesions in B-mode ultrasound. In the BI-RADS, number 3 refers to probably benign, number 4 refers to suspicious malignancy, number 5 refers to highly suggestive of malignancy. In the level 4 category, a subclassification is used: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy.

1188x839mm (150 x 150 DPI)

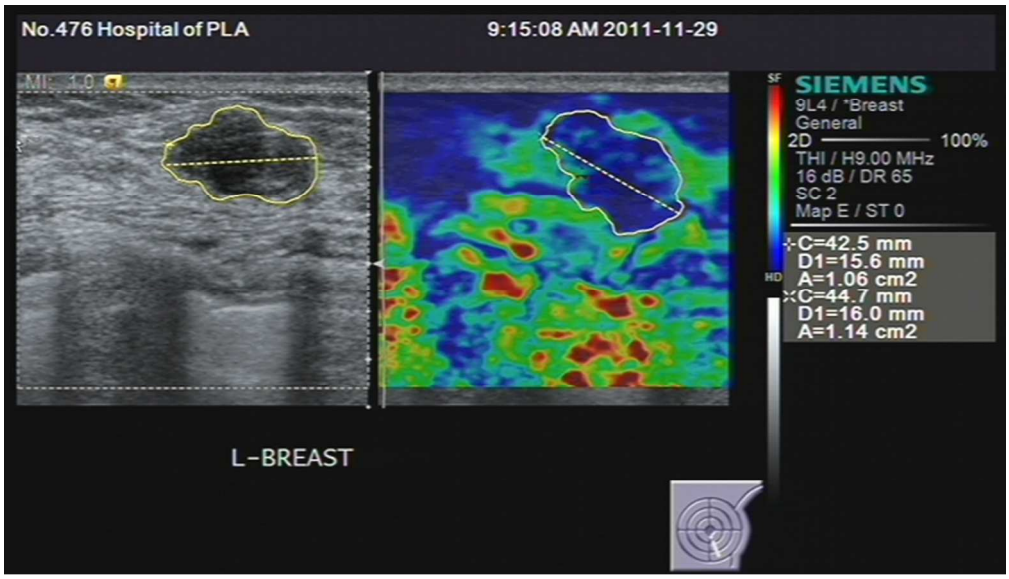

Figure 2. The image of a 45-years old patient with invasive ductal carcinoma
The left picture shows a B-mode ultrasonography locating the region of interest (ROI). The right picture shows the elastography of the breast lesion, with the blue color indicates a harder tissue and the red color indicated a softer tissue.

56x32mm (600 x 600 DPI)

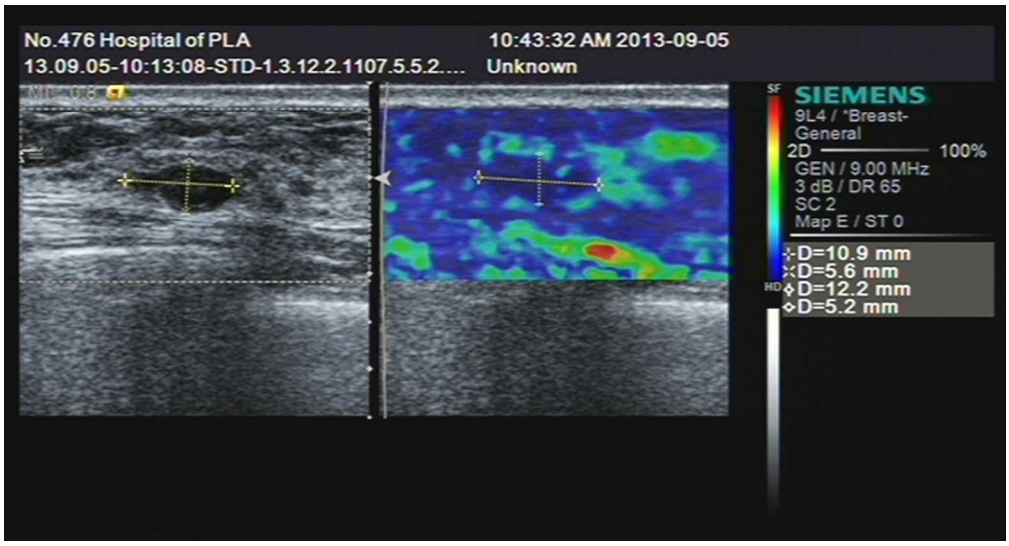

Figure 3. The image of a 41-years old patient with fibroadenoma
A B-mode ultrasonography at the left of the picture shows a location of the region of interest (ROI). An elastography picture at the right shows the elasticity of the fibroadenoma.

53x28mm (600 x 600 DPI)

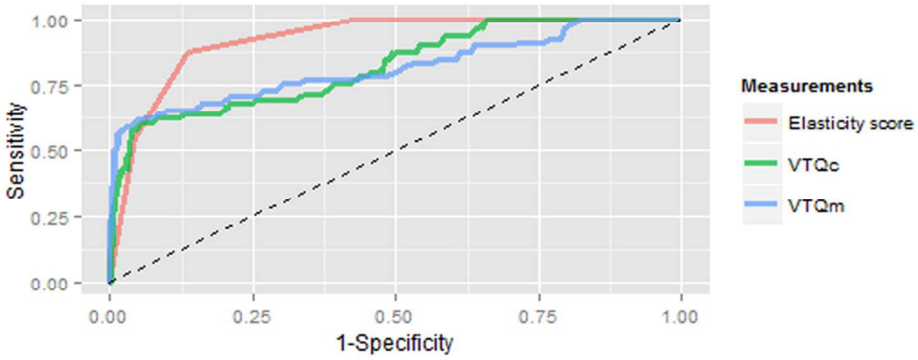

Figure 4. The receiver-operating-characteristics (ROC) curve of elasticity score, VTQm and VTQc
VTQm, virtual touchTM quantification at margin of a lesion. VTQc, virtual touchTM quantification at center of a lesion. The figure shows the area under the curve (AUC), specificity and sensitivity of Elasticity score, VTQm and VTQc. A cut-off point of 2.50 was selected for elasticity score yielding a sensitivity of 0.872 (95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. A cut-off point of 2.860 m/s was selected for VTQm yielding a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of 3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857.

44x19mm (600 x 600 DPI)

BMJ Open

Diagnostic accuracy of shear wave elastography for prediction of breast malignancy in patients with pathological nipple discharge

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-008848.R2
Article Type:	Research
Date Submitted by the Author:	13-Oct-2015
Complete List of Authors:	Guo, xiao-bo; 476 Clinical Department of Fuzhou General Hospital of Nanjing Military Region, Department of Ultrasonography Liu, Ying; 476 Clinical Department of Fuzhou General Hospital of Nanjing Military Region, Department of Ultrasonography Li, Wanhu; Shandong Cancer Hospital and Institute, Department of Radiology
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Radiology and imaging
Keywords:	Breast tumours < ONCOLOGY, ONCOLOGY, Adult oncology < ONCOLOGY

Diagnostic accuracy of shear wave elastography for prediction of breast malignancy in patients with pathological nipple discharge

Xiaobo Guo[1] Ying Liu[1] Wanhu Li[2]

1. Department of Ultrasonography, 476 Clinical Department of Fuzhou General Hospital of Nanjing Military Region, Fuzhou 350002, China

2. Department of Radiology, Shandong Cancer Hospital and Institute, Jinan 250117, China

Email address: Xiaobo Guo: 782605810@qq.com; Ying Liu: 1637824435@qq.com; Wanhu Li: lwhvzg@126.com.

Corresponding author: Xiaobo Guo, phone number：0531-86833526, E-mail: 782605810@qq.com.

ABSTRACT

Objectives: Pathological nipple discharge (PND) may indicate malignant breast lesions, the role of shear wave elastography (SWE) in predicting these malignant lesions were not evaluated. We aim to evaluate the diagnostic value of SWE for this condition.

Design: Prospective diagnostic accuracy study comparing a combination of qualitative and quantitative measurements of SWE (index test) to a ductoscopy and microdochectomy for histological diagnosis (reference test).

Setting: Fuzhou general hospital of Nanjing military command.

Participants: A total of 379 patients with PND were finally included from January, 2011 to March 2014, after we screened 1084 possible candidates. All participants were evaluated through a SWE with qualitative parameters generated by Virtual Touch^{TM} tissue imaging (VTI) and quantitative parameters generated by Virtual Touch^{TM} tissue quantification (VTQ). All the patients were consented to receive a ductoscopy and microdochectomy for histological diagnosis, and the results were set as a reference test.

Outcome measures: Sensitivity and specificity of the combined VTI and VTQ of the SWE for detection of malignancy in patients with PND.

Results: The 379 participants presented with 404 lesions. The results of pathological examination showed that 326(80.7%) of the 404 lesions were benign and the other 78 (19.3%) were malignant. An area under the curve (AUC) of elasticity score, VTQm and VTQc were 0.872, 0.825, 0.857 respectively, with the corresponding cut-off point as 2.50, 2.860m/s and 3.015m/s respectively. After a combination of these measurements, the sensitivity, specificity, positive and negative predictive value (PPV and NPV) were 89.7%, 72.1%, 43.5% and 96.7%, respectively. The sensitivity

analysis showed 82.0% of the sensitivity and 96.8% of the specificity, in which patients with no pathological findings in ductoscopy were excluded.

Conclusion: Ultrasonography elastography is sensitive for PND patients and could be used as a triage test before ductoscopy examination. Studies for further improvement of diagnostic sensitivity are warranted.

Key words: Shear wave elastography; Pathological nipple discharge; Microdochectomy; Breast Imaging- Reporting and Data System.

Strengths and limitations of this study

- Diagnostic accuracy of shear wave elastography (SWE) for detecting malignancy of PND patients is rarely studied.
- For the first time, this study tested diagnostic accuracy of a synthesized measurement of qualitative and quantitative measures of SWE for detection of malignancy in PND patients.
- Limitations include that the weight of each measurement in the synthesized score was assigned evenly and the surgeon was not blinded.

Introduction

In 2014, the American cancer society reported 235,030 new cases of breast cancer and 40,430 deaths because of it[1]. The prevalence rate of breast cancer is the highest in all types of malignant carcinoma in Chinese women [2]. The cases of Chinese women with breast cancer account for 12.2% of all newly diagnosed breast cancer around the world and 9.6% of deaths from breast cancer [2].

Pathological nipple discharge (PND) is the third most common complaint of patients prompting referral to doctors for breast diseases[3]. PND is believed to be an indicator for breast carcinoma, especially when the nipple discharge is bloody[3]. So mastectomies were carried out unnecessarily without obtaining a histopathological diagnosis before the year 1950. Although mastectomies are not necessarily needed nowadays, patients with PND would be normally suggested to take further examinations to rule out breast malignancy. These examinations involve mammography, ultrasonography, ductoscopy with or without galactography, as well as cytological tests. Mammography and B-mode ultrasonography are used for screening patients with high risk of breast carcinoma, but they are not sensitive enough for detecting malignancy in PND patients, since small lesions may be occult with these 2 modalities[4]. Ductoscopy and microdochectomy are the gold standard for diagnosis of malignancy in PND patients[5 6]. In clinical practice patients with PND will conventionally be suggested for a ductoscopy, but ductoscopy and microdochectomy are invasive examinations with high cost, so some patients will refuse ductoscopy and turn to other non-invasive tests instead. Cytological tests could only be used as supportive evidence[7]. Therefore, a triage test is needed before the PND patients are suggested to take the ductoscopy and microdochectomy examination.

Ultrasonographic elastography works through the assumption that tissue compression produces strain within the tissue, and that the strain is smaller in firm tissue than in soft tissue. So we could evaluate the firmness of the tissue to predict malignancy through measuring tissue strain ratio, which is usually measured as elasticity[8]. Recently, several studies indicated that ultrasonographic elastography may be a useful tool for differentiating benign and malignant masses. The sensitivity of predicting malignancy in patients complaining breast lesions ranges from 80% to 98%, while the specificity ranges from 66% to 84%[9-14]. Recently, a shear wave elastography (SWE) is reported to be less operator-dependent and predictive of ductal or intraductal carcinoma[8 15 16]. A five-level elasticity score is normally used for qualitative assessment of the images of elastography, while a quantitative measurement of elastography is also becoming popular; a combination of these ultrasonic measurements seems to be beneficial for improving diagnostic accuracy of breast malignancies[8 16]. We noticed that the diagnostic value of SWE to detect malignancy in patients with PND has rarely been studied, especially with a combination of qualitative and quantitative measurements of SWE. So we conducted a study aiming to find out the sensitivity and specificity accuracy of shear wave elastography (SWE) for predicting malignancy in patients with PND.

Methods and materials

This study is designed for evaluating diagnostic accuracy of shear wave elastography (SWE) for detection of malignancy in patients with PND. The study protocol is approved by the ethical review board in Fuzhou general hospital of Nanjing military command (No. FZJQ2011018), in which the study was conducted and the data were collected. Data collection was planned before the index and reference test.

Patients

Patients with PND symptoms less than a year duration were recruited from outpatient settings in the Fuzhou general hospital of Nanjing military command. The PND was defined as nipple discharge in women who stopped breastfeeding and were not being pregnant[17]. The patients were screened for eligibility according to the following inclusion criteria: (1) with an age over 18 years; (2) with or without palpable breast masses; (3) agree to participate in this study; (4) providing a written consent for further examination, biopsy and use of the study data. We excluded the participants who (1) refused to receive a ductoscopy and microdochectomy or a biopsy; (2) had serious damage in internal organs that may bias the results; (3) could not receive local anesthesia because of allergy; (4) were taking or have taken chemotherapy or radiotherapy for malignant breast tumors.

Ultrasonographic imaging

The eligible participants were asked to take ultrasonographic examinations before receiving ductoscopy and microdochectomy. The ultrasonographic examinations included a B-mode ultrasound (BUS) and an elastography with acoustic radiation force Impulse (ARFI), which were performed with a Siemens ACUSON S2000 ultrasonographic system (Siemens Ltd., China, product standard: YZB/USA 3876-2010). One operator (with 10 years of experience in ultrasonography imaging) ran the BUS and the ARFI to get the BUS and elastography images, using a superficial probe (9L4, Siemens Ltd., China) with a frequency of 4 to 9 Hz.

The BUS images of the breast lesions were obtained and categorized according to Breast Imaging Reporting and Data System (BI-RADS). In the BI-RADS, number 1 refers to negative findings; number 2 refers to benign findings; number 3 refers to probably benign findings; number 4 refers

to suspicious malignancy; number 5 is highly suggestive of malignancy. Moreover, we applied a subclassification scheme for BI-RADS 4: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy [18]. Two radiologists (XBG and YL who are with at least 5 years of experience in ultrasonography) independently assessed the BUS images in the following aspects: the shape of the mass, depth, orientation, margin, boundary of the lesions, echo pattern, posterior acoustic features, as well as surrounding tissue changes. All the detected masses were divided into two types (solid and cystic) according to the presence of a cystic portion. In cases involving a solid mass in a dilated duct (intraductal) or cyst (intracystic) with a predominant cystic portion, the radiologists considered the cases as cystic type lesions.

After the BUS examination, the operator switched the ultrasonic mode to the virtual touch^{TM} quantification (VTQ). The patients were told to hold the breath, when the operator detected the inside and margin of the lesion, with a previous set region of interest (ROI) in a fixed region of 5×5cm. The operator recorded the value of VTQ (m/s) and the depth of the ROI (cm) from skin for each measurement, in a total of 5 measurements. We calculated the mean value of the 5 measurements as the final VTQ score. Additionally, the value of VTQ was measured at both the margin (VTQm) and center (VTQc) of the breast lesions. Measurable shear wave velocity (SWE) ranges from 0 to 9 m/s, a value of ROI exceed the range would be shown as “X.XX m/s” on the screen. The participants with this “X.XX m/s” value were excluded.

After finishing examination with the VTQ mode, the operators switched to virtual touch^{TM} imaging (VTI). ARFI sequences were generated to evaluate the elasticity scores of the breast lesions in VTI mode. The VTI images were evaluated and categorized through a five-level scoring

system[19]. An elasticity score of 1 indicates strain in the entire hypoechoic lesion (the entire lesion is equally shaded in green); score 2 indicates that strain is not seen in part of the hypoechoic lesion (the lesion is shown as a mosaic of green and blue); score 3 indicates that strain is in the peripheral areas only (peripheral areas in green); score 4 indicates no strain in the entire hypoechoic area (lesion is shown in blue); score 5 indicates no strain in the entire hypoechoic lesion or the surrounding area(both the entire hypoechoic lesion and its surrounding area are shown in blue). A higher elasticity score indicates a stiffer tissue. For cystic-type lesions, we analyzed the elasticity scores in solid areas. After the B-mode ultrasonography and ARFI elastography examination, 2 radiologists independently reviewed the images and the measurements without awareness of the medical history of the participants. Any discrepancy happened was solved by discussion.

Microdochectomy and pathological examination

After finishing the examination of BUS imaging and elastography, the participants received a ductoscopy examination. Microdochectomy was performed to get a biopsy sample if the ductoscopy indicated a breast lesion. The result of microdochectomy and pathological examination was set as a reference test (the golden standard). The patients were asked to stay in supine position with the arm preferably extending on a side board. After skin disinfection, the fluid-producing ducts were identified by massaging the breast towards the nipple. After dilating the target ducts with lacrimal dilators, a surgeon (with 5 years of experience in microdochectomy) introduced endoscope (a fiberoptic ductoscopy, 0.8mm of sheath, 80mm of length; manufactured in Denzlingen, Germany) through the ducts. The surgeon was not blinded from the results of the ultrasonographic examinations. After inserting the ductoscopy, the surgeon injected

saline into the ducts to keep them expanded. The ductoscopy was advanced through the ducts under direct vision from the endoscope camera, all tiers of branches were examined until the endoscope could not be advanced further or an obstructing lesion was detected. When an obstructing lesion was found, the image of this lesion would be taken using an endoscope camera, moreover, an endoscopy guided biopsy would be performed as scheduled. After biopsy and image taken, the patient would be referred to a surgeon for further treatment advice.

After detection of ductal lesions, methylene blue was injected into the discharged mammary duct. Local anesthesia was performed and then an incision around the areola was implemented. The pathological duct was separated according to tissue with methylene-blue stain, and the duct was then dissected and excised. The operator then pulled the clip out of duct together with the scope after the excision was finished. The biopsy samples over 1 mm^3 were fixed using formalin after sample collection and were sent for histologic examination.

Statistical analysis

Statistical analysis was performed to specify the diagnostic sensitivity and specificity of a synthesized measurement of VTI (elasticity score) and VTQ measurements. First, we would assess the interobserver variability between the two radiologists (XBG and YL) using the Cohen's kappa test. The value of kappa over 0.75 was considered as excellent, 0.40 to 0.75 as fair to good, and below 0.40 as poor. Second, we calculated the sensitivity, specificity and the area under the curve (AUC) of elasticity score, VTQm and VTQc using receiver operating characteristics (ROC) curve. A valued of AUC of 0.5 was considered as no diagnostic value; 0.5 to 0.7, low diagnostic value; 0.7 to 0.9, mediate diagnostic value; 0.9 to 1.0, high diagnostic value[20]. We calculated the cut-off point of the elasticity score, VTQm and VTQc and gave the sensitivity and specificity respectively,

based on both the highest sum of sensitivity and specificity and a relatively high sensitivity of the measurements. According to the cut-off points, we synthesized the measurements by the following formula: the synthesized score=elasticity score + VTQm + VTQc. A synthesized score of 1 was given, if a value over the cut-off point was observed in a breast lesion; otherwise, a score of -1 was be given. A breast lesion with a final score below 0 was classified as benign, otherwise it was categorized as malignant. Third, after differentiating benign and malignant lesions using the synthesized score, we evaluated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The diagnostic value of this synthesized measurement was compared with the biopsy results using the McNemar's test. We also compared the BI-RADS with the synthesized measurement using the chi-square test. All statistical analysis was performed using Statistical Package for the Social Sciences (SPSS, Version 20.0, Armonk, NY: IBM Corp.) and R software (www.r-project.org, version 3.1.1).

Results

A total of 379 patients with PND were included in this study, after 1084 possible candidates were screened and 705 of them were excluded. The 379 participants presented with 404 lesions, while 25 participants showed breast lesions on both sides. The process of eligibility screening and a brief description of the results of the BUS, elastography, ductoscopy and microdochectomy examinations were shown in figure 1. The age of the included participants ranged from 25 to 67 years. The result of pathological examination showed that 326(80.7%) of the 404 lesions were benign and 78 (19.3%) were malignant. Table 1 shows the types of the benign and malignant lesions. Figure 2 and 3 show examples of invasive ductal carcinoma and fibroadenoma diagnosed with B-mode ultrasonography and elastography. Of all the breast lesions, 191 (47.2%) were

palpable masses; 127 of them were benign and 64 were malignant. A total of 59 participants complained of bloody discharge in 62 breast lesions(15.3%); 9 of them were benign and 53 were malignant. The size of the lesions ranged from 3.2mm to 38.7mm. The benign lesions showed a mean size of 8.16mm, while the malignant lesions showed a mean size of 15.58mm. No patients with SWE exceeding the range were found. Moderate or severe pain (with a score of visual analog scale over 50 mm) in the breast was reported by 133 of the 227 participants who received ductoscopy. Other adverse events were not reported. The agreement between the two radiologists for ultrasound elastography was good (kappa=0.76, p<0.01).

The diagnostic cut-off points of the elasticity score, VTQm and VTQc

We calculated the cut-off points for elasticity scores, VTQm and VTQc. The results showed that a cut-off point of 2.50 for elasticity score yielded a sensitivity of 0.872(95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. We also calculated a cut-off point of 2.860 m/s for VTQm with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of 3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857. Table 2 showed raw numbers and diagnostic estimates of the elasticity score, VTQm and VTQc. Figure 4 shows the ROC curve of each measurement.

Diagnostic accuracy of a synthesized measurement of the elasticity score, VTQm and VTQc

We used a synthesized measurement of both quantitative and qualitative parameters to predict malignant breast lesions. The measurement categorized 243 (60.1%) breast lesions as benign and 161 (39.9%) as malignant. Table 3 shows a comparison of BI-RADS with the synthesized

measurement for elastography, and it showed that using BI-RADS diagnosed significantly less participants with malignancy than the synthesized measurement (X-squared = 72.035, p<0.001). The biopsy results showed that 326 (80.6%) breast lesions were classified as benign while 78 (19.4%) were malignant. We used the McNemar's test to compare the synthesized measurement with the biopsy results, the result showed that significantly more malignant lesions were found by elastography with the synthesized measurement (McNemar's chi-squared = 67.919, p<0.001). The sensitivity, specificity, positive and negative predictive value (PPV and NPV) were 0.897, 0.721, 0.435 and 0.967, respectively. Table 2 shows the raw numbers and diagnostic estimates of the synthesized measurement. In the sensitivity analysis, we excluded 152 participants who were found with no pathological changes by ductoscopy and included 227 participants. These participants presented with 237 breast lesions. The measurement of synthesized score categorized 168 (70.9%) breast lesions as benign and 69 (29.1%) as malignant. Results of the reference test showed that 159(67.1%) of the 237 lesions were benign and the other 78 (32.9%) were malignant. The sensitivity and specificity were 0.820 and 0.968, respectively.

Discussion

The results of our study showed that a combination of qualitative and quantitative measurements for SWE elastography is sensitive in predicting malignancy in PND patients, which is consistent with several recent studies testing the diagnostic value of elastography in breast masses[8 16 21]. However, several studies reported that ultrasonography was not predictive to malignancy in PND patients[22-24]. We discussed our major findings and the remaining questions below.

Predicting malignancy in patients with PND is still challenging, Sabel and colleagues found that 95%

of the PND lesions were benign, which need not be excised. However, most patients will choose duct excision for palliative purposes[25]. In the screen phase of our study, 53 patients refuse to take ductoscopy or duct excision and therefore were excluded. The patients refuse to take the examinations for that no evidence of malignancy was found and they were afraid of the contraindications of ductoscopy or duct excision. Therefore, a non-invasive method with accurate diagnosis is needed by patients with PND. Alcock and Layer found no predictive value of ductoscopy and ultrasonography in differentiating benign and malignant breast lesions in PND patients. The authors of this study suggested that ductal surgery was the only reliable way of providing a diagnosis[17]. In our study, we found the SWE elastography with a sensitivity of 90% for screening malignancy in PND patients. This result indicated that SWE elastography could be used as a helpful triage test before suggesting PND patients to ductoscopy examinations. And compared with B-mode ultrasound, SWE elastography is with lower false negative risk, according to the comparison of the BI-RADS and synthesized measurements in our study. And therefore SWE elastography is of better diagnostic performance than B-mode ultrasonography. So our study, for the first time, revealed that ultrasonographic elastography is valuable in predicting malignant lesions in patients with PND.

The size of the breast mass might be an important factor in differentiating benign and malignant lesions. We found that PND patients with malignant lesions presented a bigger size of the breast mass (malignant versus benign, 15.58 mm versus 8.15 mm), which might be easier to be detected by B-mode ultrasonography and elastography. This finding is consistent with other studies using elastography to detect breast lesions[8 21]. However, whether the size of the breast mass could be used as an independent indicator for predicting breast malignancy needs more studies.

It is interesting to find out that 513 patients with PND refused to take ductoscopy examination in our study. Absence of these patients might contribute to a high incidence of malignant breast lesions, and thus made the diagnostic value of elastography overestimated. The reason may be that patients with PND worried about post-operative pain and procedure cost, so they turned to other non-invasive tests. When the non-invasive tests showed no positive findings, they would refuse to take ductoscopy, which is the reason that we carried out this study to find out a reliable and pain-free test for them.

The synthesized score we used in this study is similar to a published method for classification of benign and malignant breast lesions, except that we did not include the BI-RADS[8]. The BI-RADS is a classification system, which is in part coincident with the elasticity score[19]. We chose elasticity score instead of the BI-RADS into the synthesized score. The elasticity score is a qualitative measurement, which was treated as continuous data in several studies[8 21 26]. We also treated the elasticity score as a continuous data, so that we could calculate a cut-off point and include it into the synthesized score. The way that we handled the elasticity score may bring bias into our results. However, neither did we find a better way to solve the problem, nor did the previous studies. Therefore, in future studies, a qualitative measurement should be developed, which differentiate benign and malignant breast lesions more clearly than the elasticity score.

In this study, we incorporated both VTQ at the margin (VTQm) and in the center (VTQc) into the synthesized score. To the best of our knowledge, we are the first to use these two quantitative measurements, instead, previous studies used VTQc alone[8 9 21]. Inclusion of both VTQm and VTQc could increase the understanding of the breast lesions, since they were reported to be different[19 21]. The results of our study indicated that both VTQm and VTQc were significantly

higher in malignant lesions than benign ones. So we might have assigned more weight in the VTQ measurement, and therefore made breast lesions with higher VTQ easier to be malignant in the synthesized score.

We have several limitations in this study. First, a total of 152 participants were not found of pathological change after ductoscopy examination, and thus were not performed microdochectomy to get a biopsy result. These participants were classified as with benign lesions without solid evidence. We might have missed the malignant lesions, since ductoscopy could only reach a depth within 2cm[5]. To ensure the reliability of our study, we ran a sensitivity analysis, in which all these participants were classified as with malignant lesions. And we found the result was consistent. Moreover, we found that the incidence of malignancy in PND patients is 19.3% in our study, while a reported incidence ranged from 4% to 29%[27-30]. This supported that the 152 participants were classified as with benign lesions. Second, the weight of elasticity score, VTQm and VTQc were assigned evenly, and therefore made the value of VTQ a more important factor than elasticity score in predicting malignancy in PND patients. Third, the surgeon who performed the ductoscopy was not blinded from the ultrasonographic results, which may introduce performance bias.

In conclusion, ultrasonography elastography is sensitive for PND patients and could be considered as a triage test before ductoscopy examination. However, a study for further improvement of diagnostic sensitivity is warranted by assigning proper weights to the qualitative and quantitative measurements for elastography.

Authors' contributions

XBG and YL contributed to the design and manuscript writing. XBG and WHL conceptualized the study

and data analysis. YL assisted with data collection. All authors read and approved the final manuscript.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments: This study was financially supported by the Key Project of science and technology of Fujian Province of China (No. 2009Y0043). We would like to thank all the participants who participated in this study.

Data sharing

No additional data available.

Reference

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: A Cancer Journal for Clinicians 2014;**64**(1):9-29 doi: 10.3322/caac.21208[published Online First: Epub Date]|.
2. Fan L, Strasser-Weippl K, Li J-J, et al. Breast cancer in China. The Lancet Oncology 2014;**15**(7):e279-e89
3. Chen L, Zhou WB, Zhao Y, et al. Bloody nipple discharge is a predictor of breast cancer risk: a meta-analysis. Breast cancer research and treatment 2012;**132**(1):9-14 doi: 10.1007/s10549-011-1787-5[published Online First: Epub Date]|.
4. Jagmohan P, Pool FJ, Putti TC, Wong J. Papillary lesions of the breast: imaging findings and diagnostic challenges. Diagnostic and interventional radiology (Ankara, Turkey) 2013;**19**(6):471-8 doi: 10.5152/dir.2013.13041[published Online First: Epub Date]|.
5. Denewer A, El-Etribi K, Nada N, El-Metwally M. The role and limitations of mammary ductoscope in management of pathologic nipple discharge. Breast J 2008;**14**(5):442-9 doi: 10.1111/j.1524-4741.2008.00620.x[published Online First: Epub Date]|.
6. Makita M, Akiyama F, Gomi N, Iwase T. Mammary ductoscopy and watchful follow-up substitute microdochectomy in patients with bloody nipple discharge. Breast cancer (Tokyo, Japan) 2014 doi: 10.1007/s12282-014-0561-z[published Online First: Epub Date]|.
7. Kooistra BW, Wauters C, van de Ven S, Strobbe L. The diagnostic value of nipple discharge cytology in 618 consecutive patients. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2009;**35**(6):573-7 doi: 10.1016/j.ejso.2008.09.009[published Online First: Epub Date]|.
8. Kim YS, Park JG, Kim BS, Lee CH, Ryu DW. Diagnostic value of elastography using acoustic radiation force impulse imaging and strain ratio for breast tumors. J Breast Cancer 2014;**17**(1):76-82 doi: 10.4048/jbc.2014.17.1.76[published Online First: Epub Date]|.
9. Tozaki M, Isobe S, Sakamoto M. Combination of elastography and tissue quantification using the acoustic radiation force impulse (ARFI) technology for differential diagnosis of breast masses. Japanese journal of radiology 2012;**30**(8):659-70 doi: 10.1007/s11604-012-0106-3[published

Online First: Epub Date]|.

10. Tozaki M, Isobe S, Fukuma E. Preliminary study of ultrasonographic tissue quantification of the breast using the acoustic radiation force impulse (ARFI) technology. European journal of radiology 2011;**80**(2):e182-7 doi: 10.1016/j.ejrad.2011.05.020[published Online First: Epub Date]|.
11. Meng W, Zhang G, Wu C, Wu G, Song Y, Lu Z. Preliminary results of acoustic radiation force impulse (ARFI) ultrasound imaging of breast lesions. Ultrasound in medicine & biology 2011;**37**(9):1436-43 doi: 10.1016/j.ultrasmedbio.2011.05.022[published Online First: Epub Date]|.
12. Taylor K, O'Keeffe S, Britton PD, et al. Ultrasound elastography as an adjuvant to conventional ultrasound in the preoperative assessment of axillary lymph nodes in suspected breast cancer: a pilot study. Clinical radiology 2011;**66**(11):1064-71 doi: 10.1016/j.crad.2011.05.015[published Online First: Epub Date]|.
13. Sadigh G, Carlos RC, Neal CH, Dwamena BA. Accuracy of quantitative ultrasound elastography for differentiation of malignant and benign breast abnormalities: a meta-analysis. Breast cancer research and treatment 2012;**134**(3):923-31 doi: 10.1007/s10549-012-2020-x[published Online First: Epub Date]|.
14. Gong X, Xu Q, Xu Z, Xiong P, Yan W, Chen Y. Real-time elastography for the differentiation of benign and malignant breast lesions: a meta-analysis. Breast cancer research and treatment 2011;**130**(1):11-8 doi: 10.1007/s10549-011-1745-2[published Online First: Epub Date]|.
15. Kılıç F, Ustabaşıoğlu FE, Samancı C, et al. Ductal Carcinoma In Situ Detected by Shear Wave Elastography within a Fibroadenoma. Journal of breast cancer 2014;**17**(2):180-3 doi: 25013441[published Online First: Epub Date]|.
16. Youk JH, Son EJ, Gweon HM, Kim H, Park YJ, Kim JA. Comparison of Strain and Shear Wave Elastography for the Differentiation of Benign From Malignant Breast Lesions, Combined With B-mode Ultrasonography: Qualitative and Quantitative Assessments. Ultrasound in medicine & biology 2014;**40**(10):2336-44 doi: 10.1016/j.ultrasmedbio.2014.05.020[published Online First: Epub Date]|.
17. Alcock C, Layer GT. Predicting occult malignancy in nipple discharge. ANZ journal of surgery 2010;**80**(9):646-9 doi: 10.1111/j.1445-2197.2010.05270.x[published Online First: Epub Date]|.
18. Sanders MA, Roland L, Sahoo S. Clinical implications of subcategorizing BI-RADS 4 breast lesions associated with microcalcification: a radiology-pathology correlation study. Breast J 2010;**16**(1):28-31 doi: 10.1111/j.1524-4741.2009.00863.x[published Online First: Epub Date]|.
19. Itoh A, Ueno E, Tohno E, et al. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006;**239**(2):341-50 doi: 10.1148/radiol.2391041676[published Online First: Epub Date]|.
20. Wu S. Medlical Statistics. BeiJing: People's Medical Publishing House, 2002:164-78.
21. Li G, Li DW, Fang YX, et al. Performance of shear wave elastography for differentiation of benign and malignant solid breast masses. PloS one 2013;**8**(10):e76322 doi: 10.1371/journal.pone.0076322[published Online First: Epub Date]|.
22. Munoz Carrasco R, Alvarez Benito M, Rivin del Campo E. Value of mammography and breast ultrasound in male patients with nipple discharge. European journal of radiology

2013;**82**(3):478-84 doi: 10.1016/j.ejrad.2012.10.019[published Online First: Epub Date]|.

23. Hamed ST, Abdo MH, Ahmed HH. Breast discharge: ultrasound and Doppler evaluation. Journal of the Egyptian National Cancer Institute 2008;**20**(3):262-70
24. Parajuly SS, Lan PY, Yun MB, Gang YZ, Hua Z. Diagnostic potential of strain ratio measurement and a 5 point scoring method for detection of breast cancer: Chinese experience. Asian Pacific journal of cancer prevention : APJCP 2012;**13**(4):1447-52
25. Sabel MS, Helvie MA, Breslin T, et al. Is duct excision still necessary for all cases of suspicious nipple discharge? Breast J 2012;**18**(2):157-62 doi: 10.1111/j.1524-4741.2011.01207.x[published Online First: Epub Date]|.
26. Choi JJ, Kang BJ, Kim SH, et al. Role of sonographic elastography in the differential diagnosis of papillary lesions in breast. Japanese journal of radiology 2012;**30**(5):422-9 doi: 10.1007/s11604-012-0070-y[published Online First: Epub Date]|.
27. Seltzer MH. Breast complaints, biopsies, and cancer correlated with age in 10,000 consecutive new surgical referrals. Breast J 2004;**10**(2):111-7
28. Lau S, Kuchenmeister I, Stachs A, Gerber B, Krause A, Reimer T. Pathologic nipple discharge: surgery is imperative in postmenopausal women. Annals of surgical oncology 2005;**12**(7):546-51 doi: 10.1245/aso.2005.04.013[published Online First: Epub Date]|.
29. Vargas HI, Vargas MP, Eldrageely K, Gonzalez KD, Khalkhali I. Outcomes of clinical and surgical assessment of women with pathological nipple discharge. The American surgeon 2006;**72**(2):124-8
30. Morrogh M, Morris EA, Liberman L, Borgen PI, King TA. The predictive value of ductography and magnetic resonance imaging in the management of nipple discharge. Annals of surgical oncology 2007;**14**(12):3369-77 doi: 10.1245/s10434-007-9530-5[published Online First: Epub Date]|.

Figure Lengend

Figure 1. Flowchart of this study

* There were 379 participants presenting 404 lesions. In the results, we count the number of lesions instead of the patients. The index test is shear wave elastography on the basis of B-mode ultrasound. The reference test is ductoscopy and subsequent microdochectomy if pathological changes was found in ductoscopy.

Figure 2. The image of a 45-years old patient with invasive ductal carcinoma

The left picture shows a B-mode ultrasonography locating the region of interest (ROI). The right picture shows the elastography of the breast lesion, with the blue color indicates a harder tissue and the red color indicated a softer tissue.

Figure 3. The image of a 41-years old patient with fibroadenoma

A B-mode ultrasonography at the left of the picture shows a location of the region of interest (ROI). An elastography picture at the right shows the elasticity of the fibroadenoma.

Figure 4. The receiver-operating-characteristics (ROC) curve of elasticity score, VTQm and VTQc

VTQm, virtual touch^{TM} quantification at margin of a lesion. VTQc, virtual touch^{TM} quantification at center of a lesion. The figure shows the area under the curve (AUC), specificity and sensitivity of Elasticity score, VTQm and VTQc. A cut-off point of 2.50 was selected for elasticity score yielding a sensitivity of 0.872 (95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. A cut-off point of 2.860 m/s was selected for VTQm yielding a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of 3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857.

Table 1. The distribution of benign and malignant lesions classified by microdochectomy

Diagnosis after biopsy	Number of lesions (%)
Benign	326
No pathological lesions	167 (51.2)
Fibrocystic breast disease	82(25.1)
Fibroadenoma	60 (18.4)
Atypical ductal hyperlasia	11(3.5)
Intraductal pailloma	6(1.8)
Maglinant	78
Invasive ductal carcinoma	61(78.2)
Ductal carcinoma in situ	6(7.7)
Paillary carcinoma	7(9.0)
Lobular carcinoma	4(5.1)

Measurements	Correctly positive	Correctly negative	Sensitivity (%)	Specificity (%)	PPV(%)	NPV(%)
Synthesized score (n=404)	70/78	235/326	89.7 (83.3 to 96.2)	72.1 (66.9 to 77.0)	43.5 (35.8 to 51.1)	96.7 (94.5 to 99.0)

Table 2. Predictive value of the synthesized quantitative measurement

Abbreviations: SE, systematic error. CI, confidence interval. PPV, positive predictive value. NPV, negative predictive value.

Table 3. The BI-RADS classification and the synthesized measures

BI-RADS*	Elasticity score	VTQm	VTQc	Synthesized score	
	Mean(sd)	Mean(sd)	Mean(sd)	Benign(n, %)	Malignant(n, %)
3(n=293)	1.98(0.64)	2.64(0.51)	2.86(0.55)	207(85.2)	86(53.4)
4A(n=46)	2.84(0.68)	2.85(0.76)	3.02(0.76)	25(10.3)	21(13.1)
4B(n=23)	3.91(0.29)	3.24(0.86)	3.48(0.78)	4(1.6)	19(11.8)
4C(n=16)	3.71(0.61)	2.90(0.70)	3.34(0.84)	5(2.1)	11(6.8)
5(n=26)	4.00(0.81)	3.31(0.62)	3.71(0.75)	2(0.8)	24(14.9)

Abbreviations: BI-RADS: the Breast Imaging Reporting and Data System. VTQm, virtual touch^{TM} quantification at margin of a lesion. VTQc, virtual touch^{TM} quantification at center of a lesion. SD, standardized deviation. N, number of counts.

* In the BI-RADS, number 3 refers to probably benign, number 4 refers to suspicious malignancy, number 5 refers to highly suggestive of malignancy. In the 4 category, a subclassification is used: (1) 4A: low suspicious for malignancy; (2) 4B: intermediate suspicious for malignancy; (3) 4C: moderate concern, but not classic for malignancy.

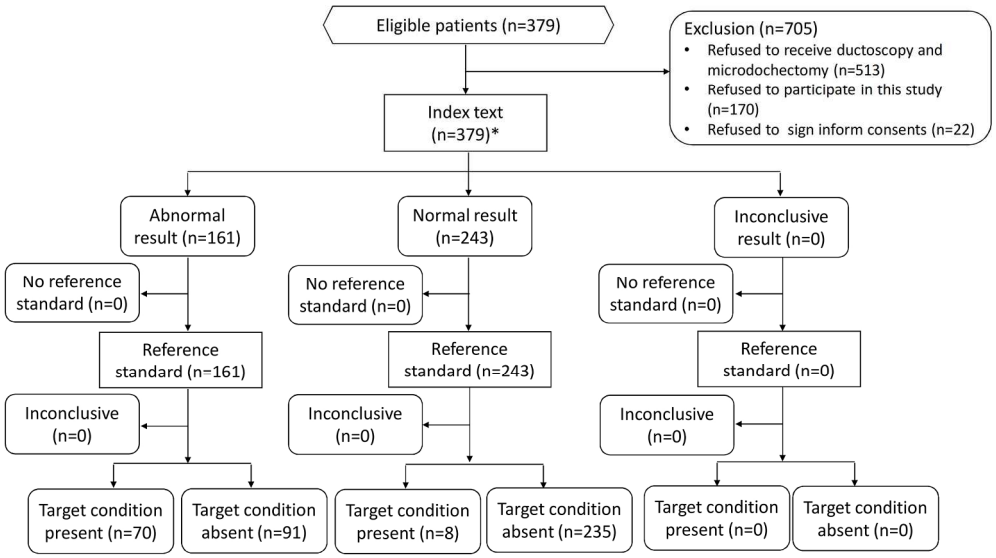

Figure 1. Flowchart of this study

* There were 379 participants presenting 404 lesions. In the results, we count the number of lesions instead of the patients. The index test is shear wave elastography on the basis of B-mode ultrasound. The reference test is ductoscopy and subsequent microdochectomy if pathological changes was found in ductoscopy.

199x115mm (300 x 300 DPI)

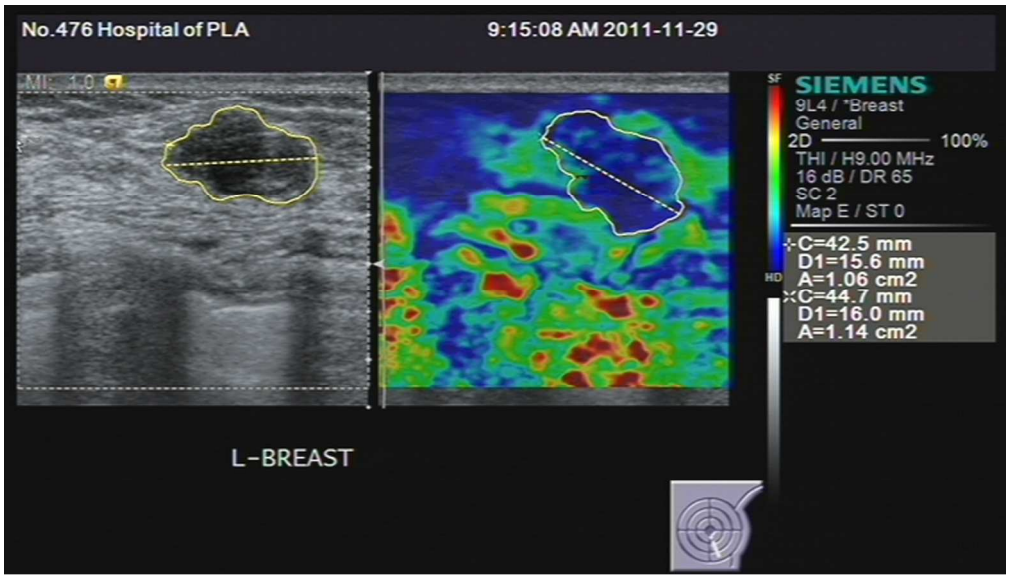

Figure 2. The image of a 45-years old patient with invasive ductal carcinoma
The left picture shows a B-mode ultrasonography locating the region of interest (ROI). The right picture shows the elastography of the breast lesion, with the blue color indicates a harder tissue and the red color indicated a softer tissue.

250x142mm (300 x 300 DPI)

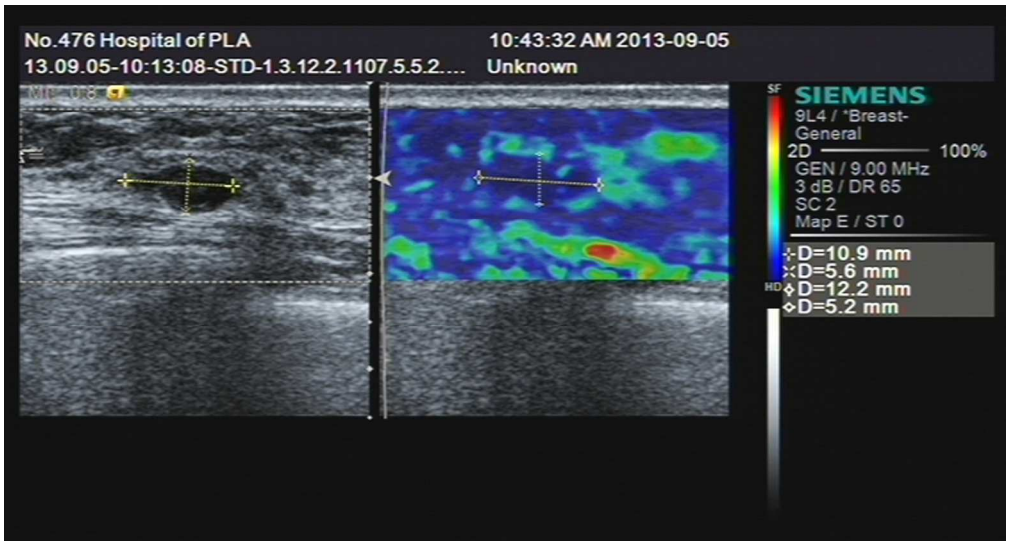

Figure 3. The image of a 41-years old patient with fibroadenoma
A B-mode ultrasonography at the left of the picture shows a location of the region of interest (ROI). An elastography picture at the right shows the elasticity of the fibroadenoma.

199x107mm (300 x 300 DPI)

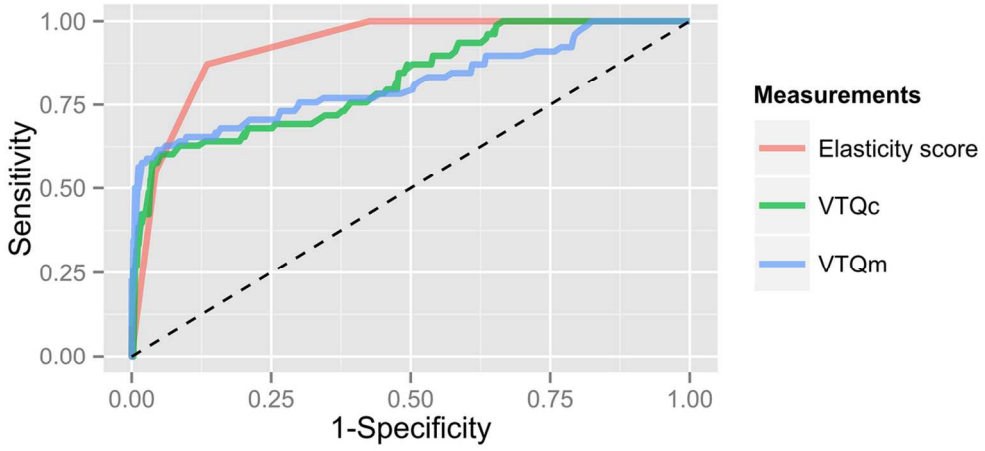

Figure 4. The receiver-operating-characteristics (ROC) curve of elasticity score, VTQm and VTQc
VTQm, virtual touchTM quantification at margin of a lesion. VTQc, virtual touchTM quantification at center of a lesion. The figure shows the area under the curve (AUC), specificity and sensitivity of Elasticity score, VTQm and VTQc. A cut-off point of 2.50 was selected for elasticity score yielding a sensitivity of 0.872 (95%CI, 0.795 to 0.940), a specificity of 0.736 (95%CI, 0.660 to 0.805) and an area under the curve of 0.854. A cut-off point of 2.860 m/s was selected for VTQm yielding a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.742 (95%CI, 0.673 to 0.805) and an AUC of 0.825. Additionally, a cut-off point of 3.015 m/s was calculated for VTQc with a sensitivity of 0.731 (95%CI, 0.628 to 0.821), a specificity of 0.698 (95%CI, 0.616 to 0.761) and an AUC of 0.857.
115x53mm (300 x 300 DPI)

STARD checklist for reporting of studies of diagnostic accuracy

(version January 2003)

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	1
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	5
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	5
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	5
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Yes
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	5
Test methods	7	The reference standard and its rationale.	5
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	5-8
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	5-8
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	6, 8
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	6, 8
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	9-10
	13	Methods for calculating test reproducibility, if done.	9
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	10
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	10
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	10, figure 1
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	10
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	10
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	10, table 2
	20	Any adverse events from performing the index tests or the reference standard.	10
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	11-12, tabl e 2
	22	How indeterminate results, missing data and outliers of the index tests were handled.	7, 11-12
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	No
	24	Estimates of test reproducibility, if done.	No
DISCUSSION	25	Discuss the clinical applicability of the study findings.	12-13