Original Article

Development and validation of a nomogram for discriminating between benign and malignant breast masses by conventional ultrasound and dual-mode elastography: a multicenter study

Keen Yang[1#], Xiuqin Ye[1#], Hongtian Tian[1], Qiaoying Li[2], Qinghua Liu[3], Jingjing Li[3], Jinhan Guo[4], Jinfeng Xu[1], Fajin Dong[1]

[1]Department of Ultrasound, Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University; The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, China; [2]Department of Ultrasound, Tangdu Hospital, Fourth Military Medical University, Xi'an, China; [3]Department of Ultrasound, Rizhao People's Hospital, Rizhao, China; [4]Department of Ultrasound, Longhua Branch of Shenzhen People's Hospital, Shenzhen, China

Contributions: (I) Conception and design: K Yang, X Ye, J Xu, F Dong; (II) Administrative support: J Xu, F Dong; (III) Provision of study materials or patients: H Tian, Q Li, Q Liu, J Li, J Guo; (IV) Collection and assembly of data: K Yang, X Ye, H Tian, Q Li, Q Liu, J Li, J Guo; (V) Data analysis and interpretation: K Yang, X Ye; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jinfeng Xu. Department of Ultrasound, Shenzhen People's Hospital, 1017 Dongmen North Road, Luohu District, Shenzhen 518020, China. Email: xujinfeng@yahoo.com; Fajin Dong. Department of Ultrasound, Shenzhen People's Hospital, 1017 Dongmen North Road, Luohu District, Shenzhen 518020, China. Email: dongfajin@szhospital.com.

Background: This study developed and validated an ultrasound nomogram based on conventional ultrasound and dual-mode elastography to differentiate breast masses.

Methods: The data of 234 patients were collected before they underwent breast mass puncture or surgery at 4 different centers between 2016 and 2021. Patients were divided into 5 datasets: internal validation and development sets from the same hospital, and external validation sets from the 3 other hospitals. In the development cohort, age and 294 different ultrasound and elastography features were obtained from ultrasound images. Univariate logistic regression and least absolute shrinkage and selection operator (LASSO) regression were used for data reduction and visualization. Multivariable logistic regression analysis was used to develop the prediction model and ultrasound nomogram. Receiver operating characteristic (ROC) curve analysis, calibration curves, integrated discrimination improvement, and the net reclassification index were used to evaluate nomogram performance; decision curve analysis (DCA) and clinical impact curves were used to estimate clinical usefulness.

Results: In the development cohort, margin, posterior features, shape, vascularity, (the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (the mean shear wave elastography value of the breast mass)—shell mean/A mean1.5(E), (the ratio of strain elastography of adipose tissue near a breast mass) divided by [the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass)]—B/A'1.5 were selected as predictors in multivariable logistic regression analysis, comprising Model 1. Among the 5 cohorts, Model 1 performed best, with areas under the curve (AUC) of 0.92, 0.84, 0.87, 0.93, and 0.89, respectively. The AUCs were 0.90, 0.82, 0.83, 0.91, and 0.85, respectively, in Model 2 (margin + posterior features + shape + vascularity) and 0.80, 0.76, 0.77, 0.87, and 0.80, respectively, in Model 3 [shell mean/A mean1.5(E) + B/A'1.5].

Conclusions: Our ultrasound nomograms facilitate exposure to the features and visualization of breast cancer. Shell mean/A mean1.5(E), B/A'1.5 integrated with margin, posterior features, shape, and vascularity are superior at identifying breast cancer, and are worthy of further clinical investigation.

Keywords: Breast cancer; elastography; nomogram; prediction model; ultrasound

Submitted Mar 12, 2022. Accepted for publication Dec 07, 2022. Published online Jan 11, 2023.
doi: 10.21037/qims-22-237
View this article at: https://dx.doi.org/10.21037/qims-22-237

Introduction

Among cancers affecting females, breast cancer ranks first in incidence (31%) and second in mortality (15%) (1). Breast cancer mortality has decreased as treatments have improved (2). An effective population-based screening study showed that early detection of breast cancer can lead to reduced local and distant recurrence rates, as well as better 5-year survival rate (3). Breast masses can be diagnosed via ultrasound, mammography, and magnetic resonance imaging (4,5). Although each of these modalities has its advantages, ultrasound is particularly attractive due to its lack of radiation, high patient acceptance, and ability to detect lesions in dense breasts that would otherwise be obscured on mammography (6,7). However, the main drawbacks of ultrasound are the vast number of false positives identified and the low specificity of biopsies. In previous studies, the specificity of ultrasound has ranged from 60% to 75% (8,9).

Elastography, a new technique, can monitor the response of tissue to acoustic energy (10,11). As a breast mass progresses, it can affect the stiffness of the tissue, providing new imaging targets to assess biological features of the disease (12). By detecting the stiffness of the mass and its surrounding tissue, strain elastography (SE) and shear wave elastography (SWE) can be used to differentiate between benign and malignant breast masses, improving the efficiency and accuracy of breast cancer detection (13-15).

The "shell" is defined as the tissue surrounding breast masses. Previous studies have investigated the importance of different shell sizes (e.g., 1, 2, and 3 mm) in distinguishing benign from malignant breast masses (16-20). However, the prediction value of different shear modulus values combined with SE and conventional ultrasound in differentiating benign from malignant breast masses has not been explored. Furthermore, some radiologists regard elastography as an unreliable method for which they are unsure how to interpret results; consequently, they are hesitant to use elastography in regular practice to diagnose breast tumors (21).

To overcome these barriers to using elastography, the primary aim of the present study was to develop and validate an ultrasound nomogram based on conventional ultrasound and dual-mode elastography for the diagnosis of breast masses. A secondary aim was to validate our findings using various validation sets. We present the following article in accordance with the TRIPOD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-22-237/rc).

Methods

Patients

This retrospective study included patients from 4 different hospitals in 4 distinct regions of China. The study period was from 2016 to 2021.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the Regional Ethics Board of Shenzhen People's Hospital. All patients provided informed consent.

Initially, 499 patients met the inclusion criteria and were enrolled in the study. Specifically, to be eligible for inclusion, patients had to have had: (I) an ultrasound-detected mass; (II) a solid or nearly solid (<20% cystic) mass; (III) a mass between 5 and 50 mm in size; (IV) at least a 3-mm tissue margin surrounding the lesions; and (V) surgery or biopsy within a week after ultrasound detection of the mass.

Patients were excluded if they had no pathology results (n=15), lacked clinical information (n=6), had unsatisfactory elastography images (n=7), had received neoadjuvant chemotherapy (n=3), and/or had missing values for elastography data (n=234), leaving 234 patients for analysis. These 234 patients were separated into 5 datasets according to hospital: the internal validation and development sets came from the same hospital, whereas the external validation sets came from the 3 other hospitals.

Patient inclusion and exclusion pathways are shown in Figure S1.

Image acquisition

For image acquisition, patients were placed in the supine position with the breast fully exposed. Sonographers noted the size, shape, margin, echogenicity, posterior features, and vascularity of the suspected mass using a Mindray Resona 7 (L11-3U probe) (Mindray, Shenzhen, China).

Conventional ultrasound features are based on the American College of Radiology descriptions (22). In the present study, the final conventional ultrasound features considered included the following: shape (regular =0; irregular =1); margin (distinct =0, indistinct =1); echo (saccadic echo =0, iso-echo =1, heterogeneous echo =2, hypo-echo =3); posterior features (enhancement =0; no posterior acoustic features =1; shadowing =2); and vascularity (no =0; little =1, rich =2).

During SWE measurements, probes were kept perpendicular to the skin and the skin was touched gently with as little additional pressure as possible. A shear wave quality mode quality control chart (QCC) was used to assess the reliability of the SWE image. When 95% of the QCC was filled with green (indicating high quality of the SWE image) and the largest diameter was located, the image was stored for further analysis. Then, the breast mass was segmented and the edge was outlined manually. The tissue surrounding the edge of the breast mass was automatically circled to obtain Young's modulus (E), shear wave velocity (Cs), and shear modulus (G) data for –0.5-, 0.5-, 1.0-, 1.5-, 2.0-, 2.5-, and 3.0-mm shells.

During SE measurements, probes were kept perpendicular to the skin and the skin was gently touched with as little additional pressure as possible. Images were saved after 3 seconds of image stabilization. For image analysis, the breast mass was segmented and the edge was outlined manually, and, as a reference point, the adipose tissue was manually delineated at the same depth as the mass. The tissue surrounding the edge of the breast mass was automatically circled to obtain the strain ratios for –0.5-, 0.5-, 1.0-, 1.5-, and 2.0-mm shells.

In summary, we gathered E, Cs, and G data for –0.5-, 0.5-, 1.0-, 1.5-, 2.0-, 2.5-, and 3.0-mm shells (i.e., tissue surrounding the breast mass), A (mass), and A' (shell + mass), as well as for –0.5-, 0.5-, 1.0-, 1.5-, and 2.0-mm shells, A, B (adipose tissue near the lesions), and A' on SE. In all, 287 ultrasound elastography features were included for analysis.

The image acquisition process is shown in *Figure 1*.

In this study, ultrasound images were captured, stored, and analyzed by sonographers with 5 years of experience in breast ultrasound at 4 centers. The senior sonographer-F.D, with 15 years of breast ultrasound experience, then assessed image quality and analysis findings. To reduce the impact of subjective factors, a single-blinded approach was adopted, meaning that until the 5 sonographers had acquired and analyzed the images or evaluated them, they were unaware of the patients' pathology findings and other imaging data.

Statistical analysis

Normally distributed continuous variables were presented as the mean ± SD, whereas categorical data were presented as numbers and percentages.

Statistical analyses were performed using R (R Foundation for Statistical Computing, Vienna, Austria). Univariable logistic regression was used to select conventional ultrasound features. The R packages "glmnet", "rms", "pROC", "ggplot2", "rmda", and "nricens" were used for least absolute shrinkage and selection operator (LASSO) regression (23), to draw nomograms, receiver operating characteristic (ROC) curves, decision curve analysis (DCA), and clinical impact curves, as well as to calculate the net reclassification index (NRI), respectively. It is important to note that for SWE and SE features, LASSO regression was performed and the variable with the highest lamda value was chosen for the multivariable logistic regression analysis.

Multivariate logistic regression was used in the development cohort to calculate the prediction model. To provide physicians with a quantitative and visual tool to predict the incidence of individual breast cancer, we built an ultrasound nomogram based on multivariate logistic analysis in the development cohort (24).

Prediction performance of the nomogram

First, we used data from the 4 validation cohorts to draw calibration curves, and the calibration performance of the nomogram was assessed using the unreliability test. Second, the area under the curve (AUC), integrated discrimination improvement (IDI) (25), and NRI were used to evaluate the model in the development and validation cohorts.

Clinical use

DCA can obtain one-to-one correspondence between threshold probability and net benefit value, and can also

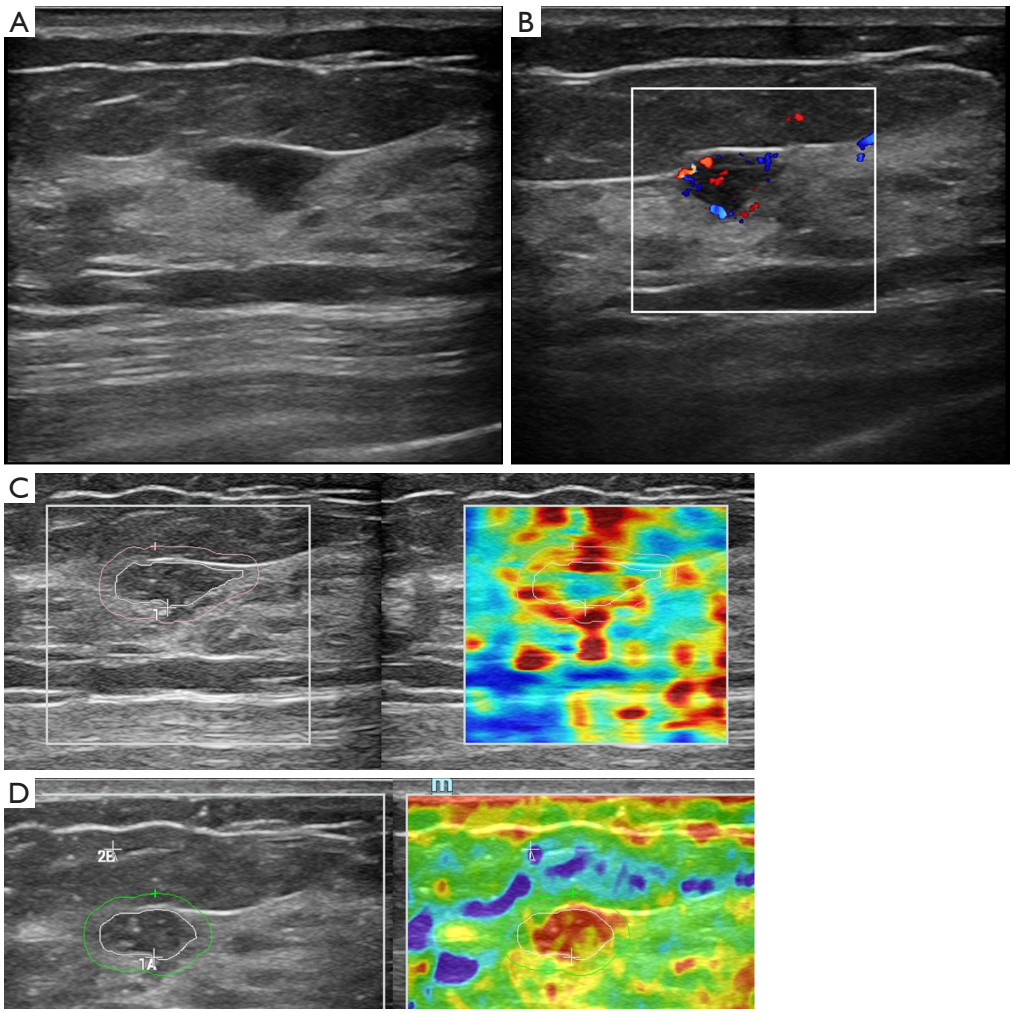

Figure 1 A 59-year-old woman with invasive ductal carcinoma. (A,B) There is a mass at 10 o'clock in the right breast, with a length of approximately 12 mm, irregular shape, indistinct margin, heterogeneous echo, no posterior acoustic features, and rich vascularity within the mass. (C) The value of the shell mean/A mean1.5(E) is 1.16; the area between the red and white lines is the shell, which is 1.5 mm. (D) The value of B/A'1.5 is 3.35; the area between the green and white lines is the shell, which is 1.5 mm. According to the nomogram, the probability of breast cancer is 0.88. shell mean/A mean1.5(E): (in Young's modulus, the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (in Young's modulus, the mean shear wave elastography value of the breast mass). B/A'1.5: (the ratio of strain elastography of adipose tissue near a breast mass) divided by (the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass).

draw this relationship as a line graph (26). DCA and clinical impact curves were used to evaluate the additional diagnostic value of dual-mode elastography.

Results

Patient characteristics

The study initially included a total of 499 breast mass samples. After application of the exclusion criteria, 234 patients were retained the data analysis.

The development cohort consisted of 116 patients (mean age 44.9±13.7 years) and 116 lesions (58 benign, 58 malignant). The Breast Imaging-Reporting and Data System (BI-RADS) was used to classify benign and malignant breast masses in the development cohort. Of the benign breast masses, 35 (60.3%) were BI-RADS 3, 11

(19.0%) were BI-RADS 4A, 8 (13.8%) were BI-RADS 4B, and 4 (6.9%) were BI-RADS 4C. Of the malignant breast masses 1 (1.7%) was BI-RADS 3, 8 (13.8%) were BI-RADS 4A, 8 (13.8%) were BI-RADS 4B, 40 (69.0%) were BI-RADS 4C, and 1 (1.7%) was BI-RADS 5.

The internal validation cohort consisted of 24 patients (mean age 46.2±13.8 years) and 24 lesions (13 benign, 11 malignant). There were 3 external validation cohorts. Cohort 1 consisted of 35 patients (mean age 40.7±11.7 years) and 35 lesions (19 benign, 16 malignant); Cohort 2 consisted of 29 patients (mean 44.3±12.4 years) and 29 lesions (16 benign, 13 malignant); Cohort 3 consisted of 30 patients (mean age 46.3±11.2 years) and 30 lesions (14 benign, 16 malignant).

Patient characteristics in the development and validation cohorts are provided in *Table 1*.

Selection of ultrasound features

According to univariable logistic regression, posterior features [if=2, odds ratio (OR) 22.65, 95% confidence interval (CI): 7.62–77.88], shape (if=1, OR 11.36, 95% CI: 4.27–36.27), margin (if=1, OR 26.70, 95% CI 10.40–76.95), vascularity (if=2, OR 10.33, 95% CI: 3.50–4.65) are appropriate predictors because their OR values are higher than that of age (OR 1.09, 95% CI: 1.05–1.14; *Table 1*). Based on the results of LASSO regression, 264 SWE features were reduced to 7 (Figure S2A,S2B) and 23 SE features were reduced to 1 (Figure S2C,S2D) in the development cohort. Finally, in Young's modulus, (the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (the mean shear wave elastography value of the breast mass)-shell mean/A mean1.5(E), (the ratio of strain elastography of adipose tissue near a breast mass) divided by [the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass)]—B/A'1.5, posterior features, shape, margin, and vascularity were used to construct Model 1, the nomogram (*Table 2*).

Ultrasound nomogram and prediction models

The diagnostic performance of the different variables used in Model 1 is presented in the Table S1. Model 1 included all the selected features [i.e., margin + posterior features + shape + vascularity + shell mean/A mean1.5(E) + B/A'1.5]; the nomogram according to Model 1 is plotted in *Figure 2*. To better compare the diagnostic value of dual-mode elastography imaging in breast masses, we built another 2 models: Model 2 contained margin + posterior features + shape + vascularity and Model 3 contained shell mean/A mean1.5(E) + B/A'1.5.

Model validation

Using data from the development and validation cohorts, we performed ROC curve analysis for the 3 models (*Table 2*; *Figure 3*) and used the validation data to draw calibration curves (*Figure 4*). *Table 2* presents the sensitivity and specificity of cut-off values determined using each of the 3 models in all 5 cohorts (development, internal validation, and external Cohorts 1, 2, and 3). Across the 5 cohorts, Model 1 had the best AUC compared to Model 2 and Model 3.

Across the 5 cohorts, Model 1 obviously improved the reclassification performance compared with Models 2 and 3, with the results presented in *Table 3*.

Prediction performance of the nomogram

As shown in the DCA (*Figure 5*), although the threshold probability is 0.4, all models performed better than all-negative samples or all-positive samples. This also reveals that, when the high risk threshold above 0.6, the Model 1 had the highest net benefit among the 3 models. In addition, the clinical impact curve of Model 1 performed best.

Discussion

We have developed and validated a diagnostic nomogram based on conventional ultrasound and dual-mode elastography ultrasound to differentiate between benign and malignant breast masses. The nomogram combines 2 elastography features with 4 conventional ultrasound features. The accuracy of differentiating between benign and malignant breast masses is greatly improved by dual-mode elastography ultrasound.

In this study, 264 SWE features were reduced to 7 features and 23 SE features were reduced to 1. The number of survivals and deaths should be more than 10-fold the number of variables used to generate the nomogram in order to lower the predicted error in the prediction probability to <10% (24). In particular, our case load was too small to build a 12-variable logistics model because the number of case samples needed to be more than 10-fold the number of control samples (24). So, we only chose

Table 1 Patient characteristics in the development and validation cohorts

Characteristics	Development cohort				Validation cohorts							
					Internal		External 1		External 2		External 3	
	Benign (n=58)	Malignant (n=58)	P value	OR (95% CI)	Benign (n=13)	Malignant (n=11)	Benign (n=19)	Malignant (n=16)	Benign (n=16)	Malignant (n=13)	Benign (n=14)	Malignant (n=16)
Age (years)[a]	38.53±12.18	51.33±12.22		1.09 (1.05–1.14)	39.62±10.21	54.00±13.90	35.68±9.24	46.69±11.75	41.94±12.57	47.23±11.97	41.86±11.04	50.19±10.13
Size of mass (mm)[a]	16.71±6.70	18.14±7.35	0.27		18.92±8.77	25.45±11.63	17.01±7.50	21.58±21.58	16.26±8.10	24.20±10.14	18.41±6.87	21.10±8.31
Shape[b]			<0.01									
Regular	30 (85.7)	5 (14.3)			3 (100.0)	0 (0)	10 (83.3)	2 (16.7)	7 (87.5)	1 (12.5)	8 (88.9)	1 (11.1)
Irregular	28 (34.6)	53 (65.4)		11.36 (4.27–36.27)	10 (47.6)	11 (52.4)	9 (39.1)	14 (60.9)	9 (42.9)	12 (57.1)	6 (28.6)	15 (71.4)
Margin[b]			<0.01									
Distinct	50 (82.0)	11 (18.0)			8 (88.9)	1 (11.1)	11 (73.3)	4 (26.7)	10 (100.0)	0	9 (100.0)	0
Indistinct	8 (14.5)	47 (85.5)		26.70 (10.40–76.95)	5 (33.3)	10 (66.7)	8 (40.0)	12 (60.0)	6 (31.6)	13 (68.4)	5 (23.8)	16 (76.2)
Echo[b]			<0.01									
Saccadic echo	2 (3.5)	0			2 (15.4)	0	3 (15.8)	0	2 (12.5)	1 (7.7)	2 (14.3)	0
Iso-echo	20 (34.5)	3 (5.2)		8.64×10^{5} (2.36×10^{-64}–NA)	7 (53.8)	3 (27.3)	6 (31.6)	2 (12.5)	5 (31.3)	3 (23.1)	6 (42.9)	3 (18.8)
Heterogeneous echo	17 (29.3)	17 (29.3)		5.76×10^{6} (4.90×10^{-64}–NA)	1 (7.7)	3 (27.3)	8 (42.1)	5 (31.3)	6 (37.5)	2 (15.4)	4 (28.6)	4 (25.0)
Hypo-echo	19 (32.8)	38 (65.5)		1.15×10^{7} (8.98×10^{-64}–NA)	3 (23.1)	5 (45.5)	2 (10.5)	9 (56.3)	3 (18.8)	7 (53.8)	2 (14.3)	9 (56.3)
Posterior features[b]			<0.01									
Enhancement	37 (84.1)	7 (15.9)			2 (22.2)	7 (77.8)	16 (57.1)	12 (42.9)	14 (77.8)	4 (22.2)	13 (48.1)	14 (51.9)
No posterior acoustic features	14 (40.0)	21 (60.0)		7.93 (2.88–24.14)								
Shadowing	7 (18.9)	30 (81.1)		22.65 (7.62–77.88)	11 (73.3)	4 (26.7)	3 (42.9)	4 (57.1)	2 (18.2)	9 (81.8)	1 (33.3)	2 (66.7)
Vascularity[b]			0									
None	22 (66.7)	11 (33.3)			8 (72.7)	3 (27.3)	15 (78.9)	4 (21.1)	11 (25.0)	3 (75.0)	6 (85.7)	1 (14.3)
A little	30 (65.2)	16 (34.8)		1.07 (0.42–2.79)								
Rich	6 (16.2)	31 (83.8)		10.33 (3.50–4.65)	5 (38.5)	8 (61.5)	4 (25.0)	12 (75.0)	5 (33.3)	10 (66.7)	8 (37.8)	15 (65.2)
Pathology												
Intraductal papilloma	3 (5.2)				1 (7.7)		1 (5.3)				1 (7.1)	
Atypical hyperplasia	3 (5.2)				0		1 (5.3)		1 (6.25)		0	
Mammary duct expansion	1 (1.7)				2 (15.4)		0				0	
Inflammation	1 (1.7)				1 (7.7)		0		1 (6.25)		1 (7.1)	
Adenosis	10 (17.2)				2 (15.4)		5 (26.3)		4 (25.0)		2 (14.3)	
Fibroadenoma	40 (70.0)				7 (53.8)		12 (63.2)		10 (62.5)		10 (71.4)	
Ductal carcinoma *in situ*		1 (1.7)				2 (18.2)		0		1 (7.7)		2 (12.5)
Invasive ductal carcinoma		53 (91.4)				7 (63.6)		14 (87.5)		12 (92.3)		14 (87.5)
Lymphoma		0				0		1 (6.3)				0
Mucinous carcinoma		2 (3.4)				0		0				0
Breast invasive lobular carcinoma		2 (3.4)				2 (18.2)		1 (6.3)				0

[a], continuous variables are presented as the mean ± SD and were compared using t-tests. [b], categorical variables are presented as n (%) and were compared using chi-squared tests. SD, standard deviation; NA, missing value.

Table 2 Diagnostic performance of multivariable combined diagnosis

Model	Cut-off	Sensitivity (%)	Specificity (%)	AUC	95% CI
Development cohort					
Model 1	0.51	87.9	86.2	0.92	0.87–0.97
Model 2	0.65	81.0	89.7	0.90	0.84–0.96
Model 3	0.62	63.7	91.4	0.80	0.71–0.88
Internal validation cohort					
Model 1	0.73	90.9	76.9	0.84	0.66–1.00
Model 2	0.48	100.0	61.5	0.82	0.65–0.99
Model 3	0.50	100.0	53.8	0.76	0.55–0.96
External validation cohort 1					
Model 1	0.44	87.5	68.4	0.87	0.75–0.98
Model 2	0.48	87.5	63.2	0.83	0.70–0.96
Model 3	0.55	68.8	89.5	0.77	0.60–0.95
External validation cohort 2					
Model 1	0.62	100.0	75.0	0.93	0.84–1.00
Model 2	0.78	92.3	87.5	0.91	0.80–1.00
Model 3	0.64	92.3	81.3	0.87	0.74–1.00
External validation cohort 3					
Model 1	0.70	100.0	71.4	0.89	0.77–1.00
Model 2	0.80	93.8	78.6	0.85	0.69–1.00
Model 3	0.42	87.5	71.4	0.80	0.63–0.97

Model 1 consisted of margin, posterior features, shape, vascularity, shell mean/A mean1.5(E), and B/A'1.5. Model 2 consisted of margin, posterior features, shape, and vascularity. Model 3 consisted of shell mean/A mean1.5(E) and B/A'1.5. shell mean/A mean1.5(E): (in Young's modulus, the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (in Young's modulus, the mean shear wave elastography value of the breast mass). B/A'1.5: (the ratio of strain elastography of adipose tissue near a breast mass) divided by (the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass). AUC, area under the curve; CI, confidence interval.

the maximum coefficient correlation variable in the shear wave elastography and the strain elastography into the multivariable logistic model based on the results of LASSO regression. Surprisingly, Model 1 outperformed Models 2 and 3 in terms of AUC, sensitivity, and specificity across all 5 data cohorts. This demonstrates that the ability of Model 1 to predict breast cancer is credible, reducing biopsies and the physical and mental toll on patients.

Studies have shown that elastography plays an important role in discriminating breast masses (27-29). Zhang *et al.* (20) found that G and Cs parameters of SWE can guide the differentiation of benign and malignant breast masses. Xie *et al.* (16) showed that the shell of maximum elastography value (Emax) can help to identify breast masses. Xu *et al.* (18) reported that within a shell range of 1–3 mm, the 2.5-mm Emax shell performs best in the differential diagnosis of non-mass lesions in the breast. Yang *et al.* (17) found that among the 1-, 2-, and 3-mm shells, the 3-mm shell combination of Emax and minimum elastography value (Emin) differentiated breast masses most effectively. However, to the best of our knowledge, the proportions of the maximum, minimum, mean, and standard deviation of different shells and lesions in E, Cs, and G, such as (the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (the mean shear wave elastography value of the breast mass), have not been

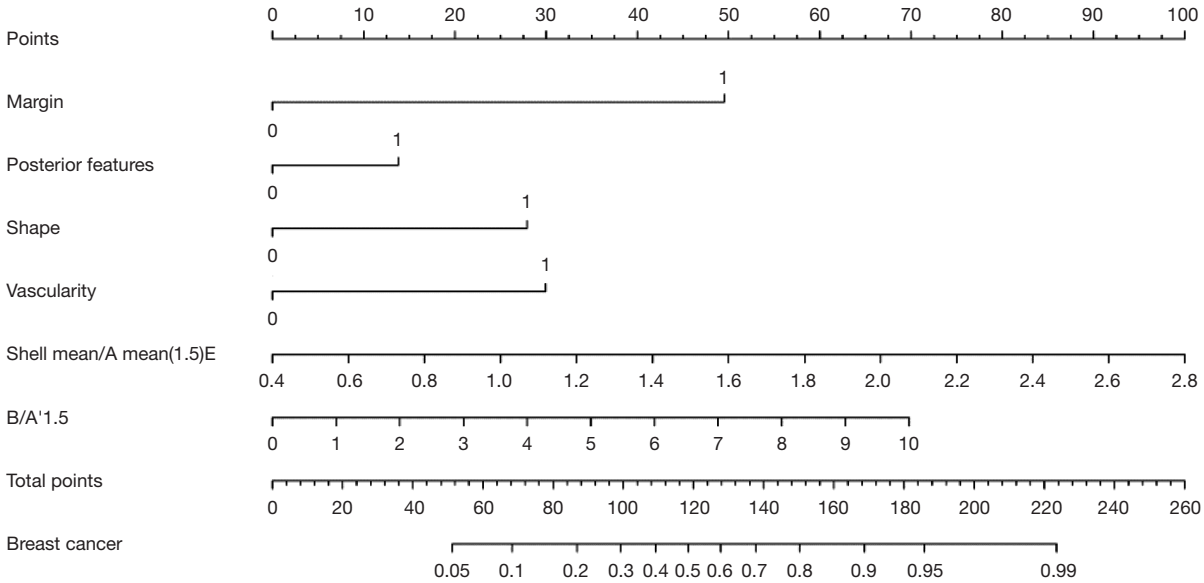

Figure 2 Breast cancer rate nomogram. The ultrasound nomogram was developed in the development cohort, with margin, posterior features, shape, vascularity, shell mean/A mean1.5(E), and B/A'1.5. shell mean/A mean1.5(E): (in Young's modulus, the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (in Young's modulus, the mean shear wave elastography value of the breast mass). B/A'1.5: (the ratio of strain elastography of adipose tissue near a breast mass) divided by (the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass).

studied yet. We examined the significance of these factors in the ultrasound detection of breast cancer, and found that the combination of shell mean/A mean1.5(E) and B/A'1.5 performed best in the diagnosis of breast masses.

In terms of using conventional ultrasound to diagnose breast masses, many studies have used BI-RADS (30-32); however, the AUCs in those studies ranged from approximately 0.79 to 0.83, which is not that different from the AUC of Model 2 in the present study. In the present study, the conventional ultrasound features were derived from the fifth edition of the BI-RADS lexicon, so Model 2 (margin, posterior features, shape, and vascularity) can represent the results of BI-RADS classification to a certain extent. Furthermore, the diagnostic AUC value of the BI-RADS classification is 0.90, lower than that of the nomogram derived from the present study. In addition, it has been shown that for BI-RADS categories, the highest concordance was typically found in the benign group, with generally lower concordance in the malignant and suspicious groups (33). Therefore, we used conventional ultrasound features and color Doppler features instead of the BI-RADS classification to construct our model.

Age is known to play a significant role in predicting breast cancer (34). However, according to univariate logistic regression analysis, the OR for age is 1.091, which is much lower than the ORs for margin, shape, vascularity, and posterior features. Therefore, age was not included in the multivariate logistic regression analysis in this study. A possible explanation for this strange phenomenon may be that Shenzhen is a young one, and most of the patients treated at the Shenzhen People's Hospital are young. Another reason could be that our case load was too small, indicating that the study should be repeated with a larger sample size.

In the development cohort, Model 1 reduced the rate of misdiagnosis of benign and malignant breast masses, with an estimated net gain of 3.85–34.08% of patients compared with Model 2. This finding implies that dual-mode elastography can help with breast cancer diagnosis. Moreover, compared with Model 3, Model 1 lowered the rate of incorrectly diagnosed benign and malignant breast masses, with an estimated net gain of 35.08–75.26% of patients. We used the IDI to reflect the overall improvement of the model (25). In the development cohort, Model 1 had a better prediction ability than Model 2 based on the IDI of Model 1 versus Model 2, which was 4.04% (P=0.05). All these results were well validated in all the validation cohorts, which supports the wide applicability of

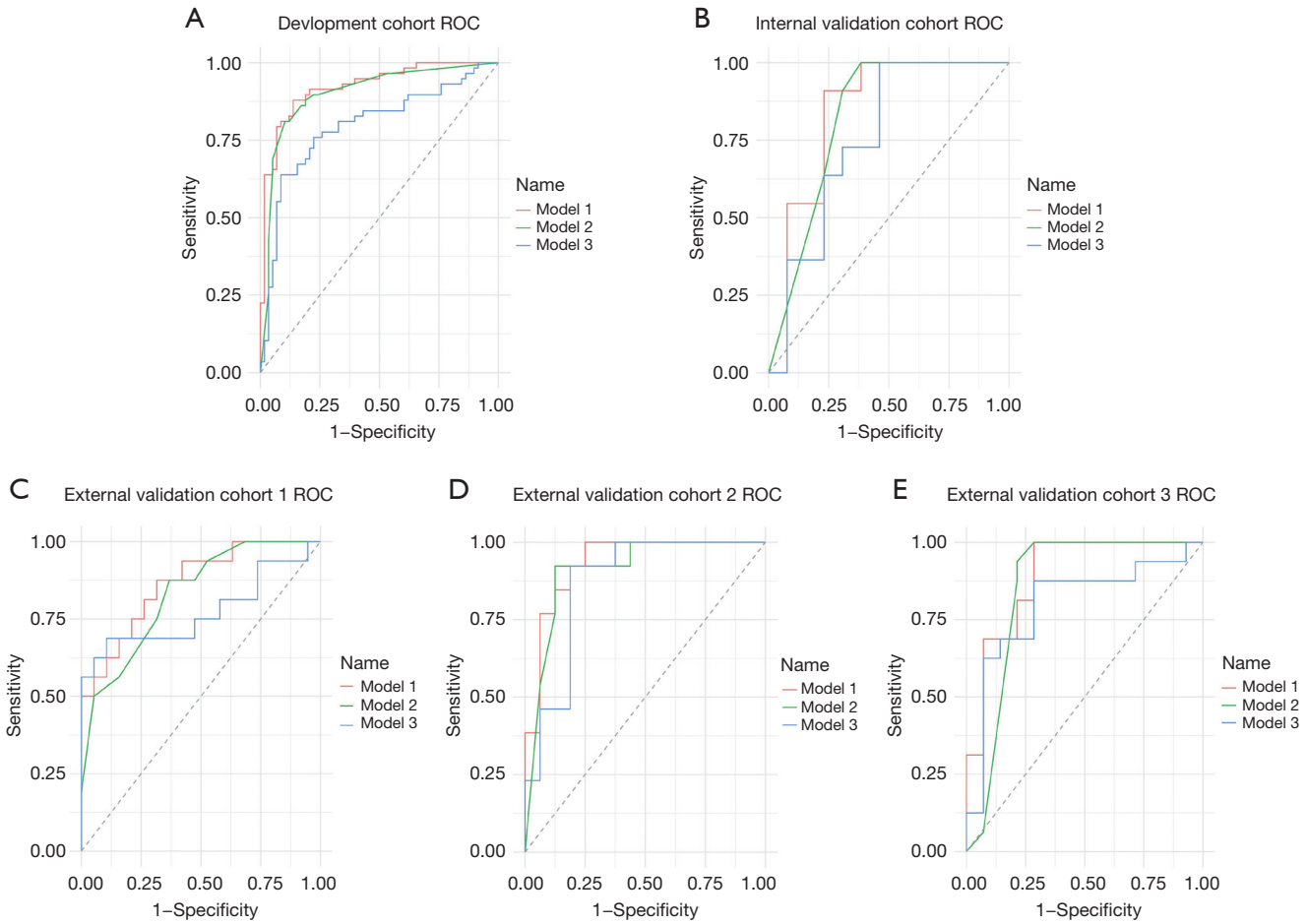

Figure 3 ROC curves of the (A) development, (B) internal validation, and (C-E) external validation cohorts (C, Cohort 1; D, Cohort 2; E, Cohort 3). Red, green, and blue lines indicate Models 1, 2, and 3, respectively. Model 1 consisted of margin, posterior features, shape, vascularity, shell mean/A mean1.5(E), and B/A'1.5. Model 2 consisted of margin, posterior features, shape, and vascularity. Model 3 consisted of shell mean/A mean1.5(E) and B/A'1.5. shell mean/A mean1.5(E): (in Young's modulus, the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (in Young's modulus, the mean shear wave elastography value of the breast mass). B/A'1.5: (the ratio of strain elastography of adipose tissue near a breast mass) divided by (the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass). ROC, receiver operating characteristic.

our results.

To assess clinical value, we used DCA and clinical impact curves (26) because the aforementioned results were explored with statistical rigor. The results of these analyses showed that the net benefit rate of Model 1 was higher than those of Models 2 and 3, suggesting that Model 1 has better clinical utility and is worthy of consideration for clinical use.

This study has several limitations. First, retrospective study designs have inherent limitations, such as selection bias and recall bias. In addition, our case load was too small. Finally, shell functionality was limited to 1 instrument, which is not conducive to the generalization of the results.

Conclusions

Our ultrasound nomograms provide exposure to the features and visualization of breast cancer. Dual-mode elastography plays an important role in the diagnosis of breast cancer; in particular, shell mean/A mean1.5(E), B/A'1.5 integrated with margin, posterior features, shape, and vascularity may be more helpful in differentiating breast

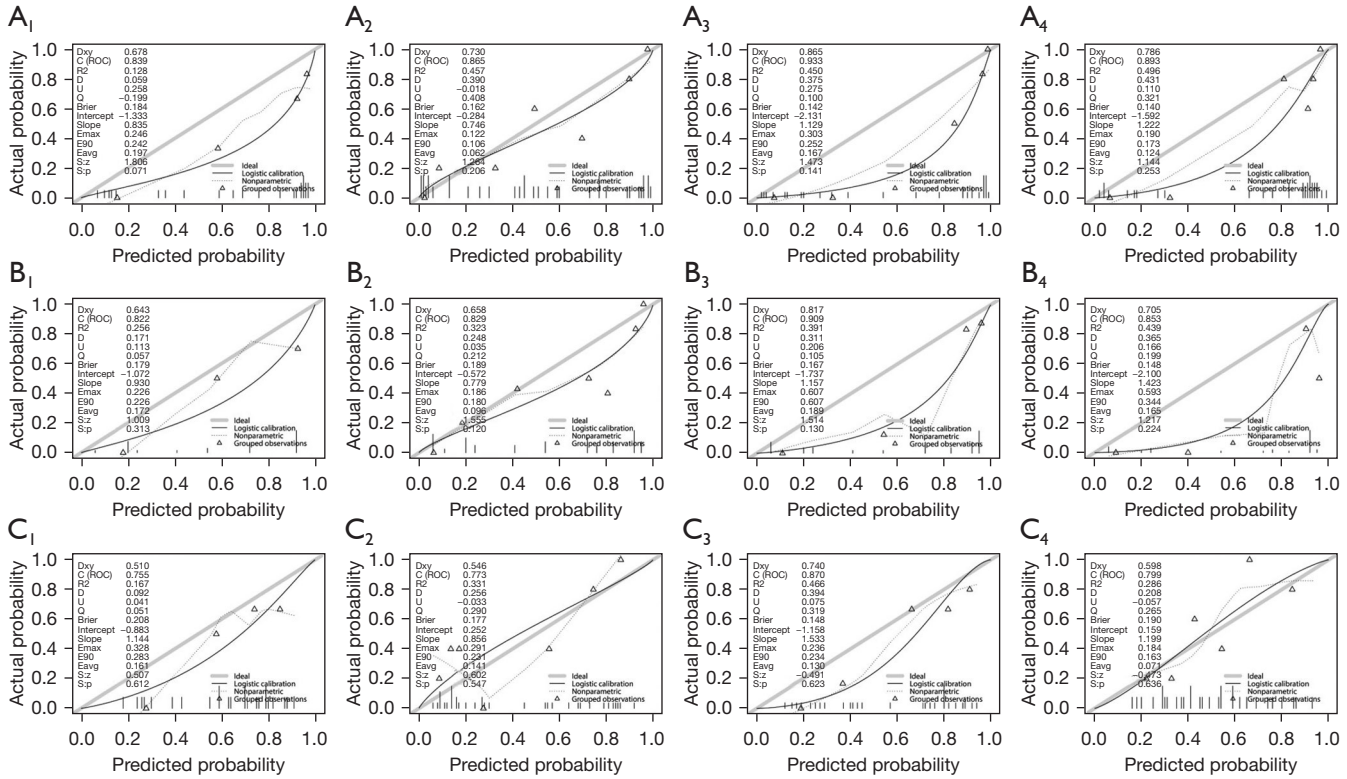

Figure 4 Calibration curves in the internal validation cohort (A1,B1,C1) and external validation cohorts 1 (A2,B2,C2), 2 (A3,B3,C3), and 3 (A4,B4,C4). (A-C) Calibration curves for Models 1 (A), 2 (B), and 3 (C). P>0.05 as the standard, all the models mentioned above passes unreliability test.

Table 3 The net reclassification index and integrated discrimination improvement in the five cohorts

Cohort	NRI (%) categorical (95% CI)	P value	NRI (%) continuous (95% CI)	P value	IDI, % (95% CI)	P value
Development cohort						
Model 1 *vs.* Model 2	18.97 (3.85, 34.08)	0.01	75.86 (42.19, 109.54)	<0.01	4.04 (0.04, 8.01)	0.05
Model 1 *vs.* Model 3	55.17 (35.08, 75.26)	<0.01	131.03 (103.57, 158.50)	<0.01	30.42 (22.02, 38.83)	<0.01
Internal validation cohort						
Model 1 *vs.* Model 2	30.77 (–2.16, 63.70)	0.07	74.13 (10.16, 138.1)	0.02	4.29 (–3.92, 12.5)	0.31
Model 1 *vs.* Model 3	30.77 (–2.16, 63.70)	0.07	104.90 (42.03, 167.76)	<0.01	20.83 (7.03, 34.63)	<0.01
External validation cohort 1						
Model 1 *vs.* Model 2	14.80 (–13.07, 42.67)	0.30	95.40 (37.02, 153.77)	<0.01	6.51 (–2.85, 15.87)	0.17
Model 1 *vs.* Model 3	8.22 (–26.34, 42.79)	0.64	78.29 (19.68, 136.89)	0.01	10.73 (–3.76, 25.21)	0.15
External validation cohort 2						
Model 1 *vs.* Model 2	25.00 (–2.39, 52.39)	0.07	122.12 (68.24, 175.99)	<0.01	8.76 (0.78, 16.74)	0.03
Model 1 *vs.* Model 3	25.00 (–11.75, 61.75)	0.18	97.12 (40.52, 153.71)	<0.01	17.65 (2.62, 32.68)	0.02

Table 3 (*continued*)

Table 3 (*continued*)

Cohort	NRI (%) categorical (95% CI)	P value	NRI (%) continuous (95% CI)	P value	IDI, % (95% CI)	P value
External validation cohort 3						
Model 1 *vs.* Model 2	28.57 (4.91, 52.24)	0.02	83.93 (23.04, 144.82)	0.01	2.8 (–2.39, 7.98)	0.29
Model 1 *vs.* Model 3	55.36 (18.93, 91.78)	<0.01	114.29 (62.44, 166.13)	<0.01	30.01 (12.97, 47.06)	<0.01

NRI categorical refers to the net reclassification index (NRI) of clinically meaningful categories (the thresholds are 0.2 and 0.4). For NRI continuous, no classifications are considered. Model 1 consisted of margin, posterior features, shape, vascularity, shell mean/A mean1.5(E), and B/A'1.5. Model 2 consisted of margin, posterior features, shape, and vascularity. Model 3 consisted of shell mean/A mean1.5(E) and B/A'1.5. shell mean/A mean1.5(E): (in Young's modulus, the mean shear wave elastography value of 1.5 mm surrounding tissues in a breast mass) divided by (in Young's modulus, the mean shear wave elastography value of the breast mass). B/A'1.5: (the ratio of strain elastography of adipose tissue near a breast mass) divided by (the ratio of strain elastography of (the breast mass adds the 1.5 mm surrounding tissues in the breast mass). CI, confidence interval; IDI, integrated discrimination improvement.

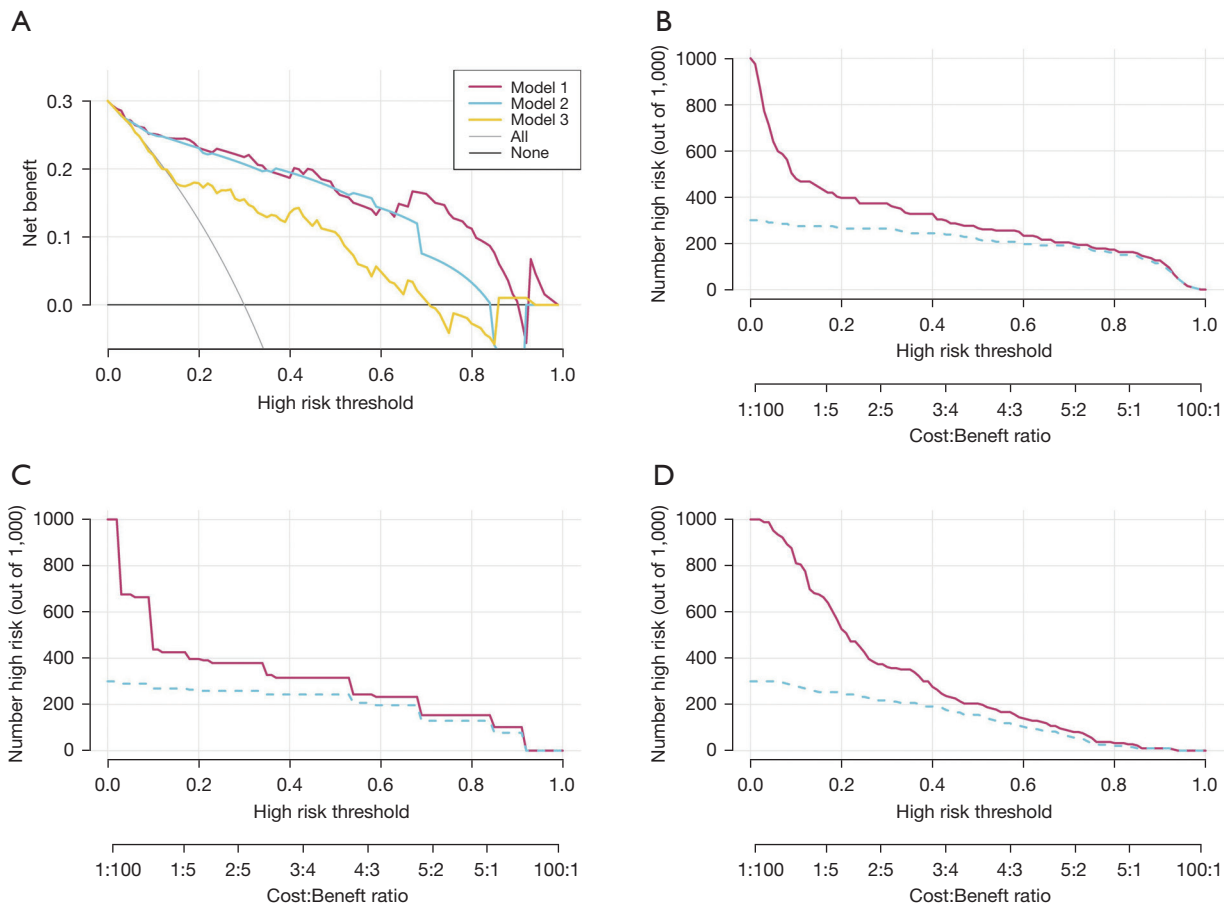

Figure 5 DCA (A) and clinical impact curves of Models 1 (B), 2 (C), and 3 (D) in the development cohort. In the DCA, the *y*-axis shows the net benefit and the *x*-axis shows the threshold probability. The horizontal black line indicates that all samples were negative, none were treated, and the net benefit was 0. The slanting black line means that all the samples are positive, all are treated. Red, blue, and yellow lines indicate Models 1, 2, and 3, respectively. On the clinical impact curves, red represents the number of people classified as positive by the model under each threshold probability, and the blue represents the number of people who are true positives under each threshold probability. DCA, decision curve analysis.

masses and are worthy of consideration for clinical use.

Acknowledgments

The authors are grateful for the support and help of our team in the Ultrasound Department, Shenzhen People's Hospital.
Funding: This project was supported by the Commission of Science and Technology of Shenzhen (No. GJHZ20200731095401004).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-22-237/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-237/coif). All authors reported that this work was supported by the Commission of Science and Technology of Shenzhen (No. GJHZ20200731095401004). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Regional Ethics Board of Shenzhen People's Hospital and informed consent was provided by all patients.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022;72:7-33.
2. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69:438-51.
3. Cedolini C, Bertozzi S, Londero AP, Bernardi S, Seriau L, Concina S, Cattin F, Risaliti A. Type of breast cancer diagnosis, screening, and survival. Clin Breast Cancer 2014;14:235-40.
4. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, Böhm-Vélez M, Mahoney MC, Evans WP 3rd, Larsen LH, Morton MJ, Mendelson EB, Farria DM, Cormack JB, Marques HS, Adams A, Yeh NM, Gabrielli G; . Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA 2012;307:1394-404.
5. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast cancer. Lancet 2005;365:1727-41.
6. Berg WA, Bandos AI, Mendelson EB, Lehrer D, Jong RA, Pisano ED. Ultrasound as the Primary Screening Test for Breast Cancer: Analysis From ACRIN 6666. J Natl Cancer Inst 2015;108:djv367.
7. Geisel J, Raghu M, Hooley R. The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound. Semin Ultrasound CT MR 2018;39:25-34.
8. Yu P, Lei J, Xu B, Wang R, Shen Z, Tian J. Correlation Between (18)F-FDG PET/CT Findings and BI-RADS Assessment Using Ultrasound in the Evaluation of Breast Lesions: A Multicenter Study. Acad Radiol 2020;27:682-8.
9. Huang S, Ye X, Yang K, Tian H, Ding Z, Chen J, Xu J, Dong F. The significance of dual-mode elastography in the diagnosis of breast lesions by physicians with different levels of experience. Quant Imaging Med Surg 2022;12:1438-49.
10. Zhou J, Zhan W, Dong Y, Yang Z, Zhou C. Stiffness of the surrounding tissue of breast lesions evaluated by ultrasound elastography. Eur Radiol 2014;24:1659-67.
11. Evans A, Whelehan P, Thomson K, McLean D, Brauer K, Purdie C, Baker L, Jordan L, Rauchhaus P, Thompson A. Invasive breast cancer: relationship between shear-wave elastographic findings and histologic prognostic factors. Radiology 2012;263:673-7.
12. Ozturk A, Grajo JR, Dhyani M, Anthony BW, Samir AE. Principles of ultrasound elastography. Abdom Radiol (NY) 2018;43:773-85.
13. Liu G, Zhang MK, He Y, Liu Y, Li XR, Wang ZL. BI-RADS 4 breast lesions: could multi-mode ultrasound be helpful for their diagnosis? Gland Surg 2019;8:258-70.

14. Goddi A, Bonardi M, Alessi S. Breast elastography: A literature review. J Ultrasound 2012;15:192-8.
15. Kapetas P, Clauser P, Woitek R, Wengert GJ, Lazar M, Pinker K, Helbich TH, Baltzer PAT. Quantitative Multiparametric Breast Ultrasound: Application of Contrast-Enhanced Ultrasound and Elastography Leads to an Improved Differentiation of Benign and Malignant Lesions. Invest Radiol 2019;54:257-64.
16. Xie X, Zhang Q, Liu S, Ma Y, Liu Y, Xu M, Xu B. Value of quantitative sound touch elastography of tissues around breast lesions in the evaluation of malignancy. Clin Radiol 2021;76:79.e21-8.
17. Yang H, Xu Y, Zhao Y, Yin J, Chen Z, Huang P. The role of tissue elasticity in the differential diagnosis of benign and malignant breast lesions using shear wave elastography. BMC Cancer 2020;20:930.
18. Xu P, Wu M, Yang M, Xiao J, Ruan ZM, Wu LY. Evaluation of internal and shell stiffness in the differential diagnosis of breast non-mass lesions by shear wave elastography. World J Clin Cases 2020;8:2510-9.
19. Huang L, Ma M, Du Z, Liu Z, Gong X. Quantitative evaluation of tissue stiffness around lesion by sound touch elastography in the diagnosis of benign and malignant breast lesions. PLoS One 2019;14:e0219943.
20. Zhang L, Xu J, Wu H, Liang W, Ye X, Tian H, Dong F. Screening Breast Lesions Using Shear Modulus and Its 1-mm Shell in Sound Touch Elastography. Ultrasound Med Biol 2019;45:710-9.
21. Mesurolle B, El Khoury M, Chammings F, Zhang M, Sun S. Breast sonoelastography: Now and in the future. Diagn Interv Imaging 2019;100:567-77.
22. Magny SJ, Shikhman R, Keppke AL. Breast Imaging Reporting and Data System. In: StatPearls. edn. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.; 2022.
23. Cai W, van der Laan M. Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator. Int J Biostat 2020. [Epub ahead of print]. doi: 10.1515/ijb-2017-0070.
24. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26:1364-70.
25. Hayashi K, Eguchi S. The power-integrated discriminant improvement: An accurate measure of the incremental predictive value of additional biomarkers. Stat Med 2019;38:2589-604.
26. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med 2014;160:122-31.
27. Itoh A, Ueno E, Tohno E, Kamma H, Takahashi H, Shiina T, Yamakawa M, Matsumura T. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006;239:341-50.
28. Athanasiou A, Tardivon A, Tanter M, Sigal-Zafrani B, Bercoff J, Deffieux T, Gennisson JL, Fink M, Neuenschwander S. Breast lesions: quantitative elastography with supersonic shear imaging--preliminary results. Radiology 2010;256:297-303.
29. Berg WA, Mendelson EB, Cosgrove DO, Doré CJ, Gay J, Henry JP, Cohen-Bacrie C. Quantitative Maximum Shear-Wave Stiffness of Breast Masses as a Predictor of Histopathologic Severity. AJR Am J Roentgenol 2015;205:448-55.
30. Xiao X, Dong L, Jiang Q, Guan X, Wu H, Luo B. Incorporating Contrast-Enhanced Ultrasound into the BI-RADS Scoring System Improves Accuracy in Breast Tumor Diagnosis: A Preliminary Study in China. Ultrasound Med Biol 2016;42:2630-8.
31. Luo WQ, Huang QX, Huang XW, Hu HT, Zeng FQ, Wang W. Predicting Breast Cancer in Breast Imaging Reporting and Data System (BI-RADS) Ultrasound Category 4 or 5 Lesions: A Nomogram Combining Radiomics and BI-RADS. Sci Rep 2019;9:11921.
32. Lee HJ, Kim EK, Kim MJ, Youk JH, Lee JY, Kang DR, Oh KK. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. Eur J Radiol 2008;65:293-8.
33. Wen W, Liu J, Wang J, Jiang H, Peng Y. A National Chinese Survey on Ultrasound Feature Interpretation and Risk Assessment of Breast Masses Under ACR BI-RADS. Cancer Manag Res 2021;13:9107-15.
34. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015;65:5-29.

Cite this article as: Yang K, Ye X, Tian H, Li Q, Liu Q, Li J, Guo J, Xu J, Dong F. Development and validation of a nomogram for discriminating between benign and malignant breast masses by conventional ultrasound and dual-mode elastography: a multicenter study. Quant Imaging Med Surg 2023;13(2):865-877. doi: 10.21037/qims-22-237

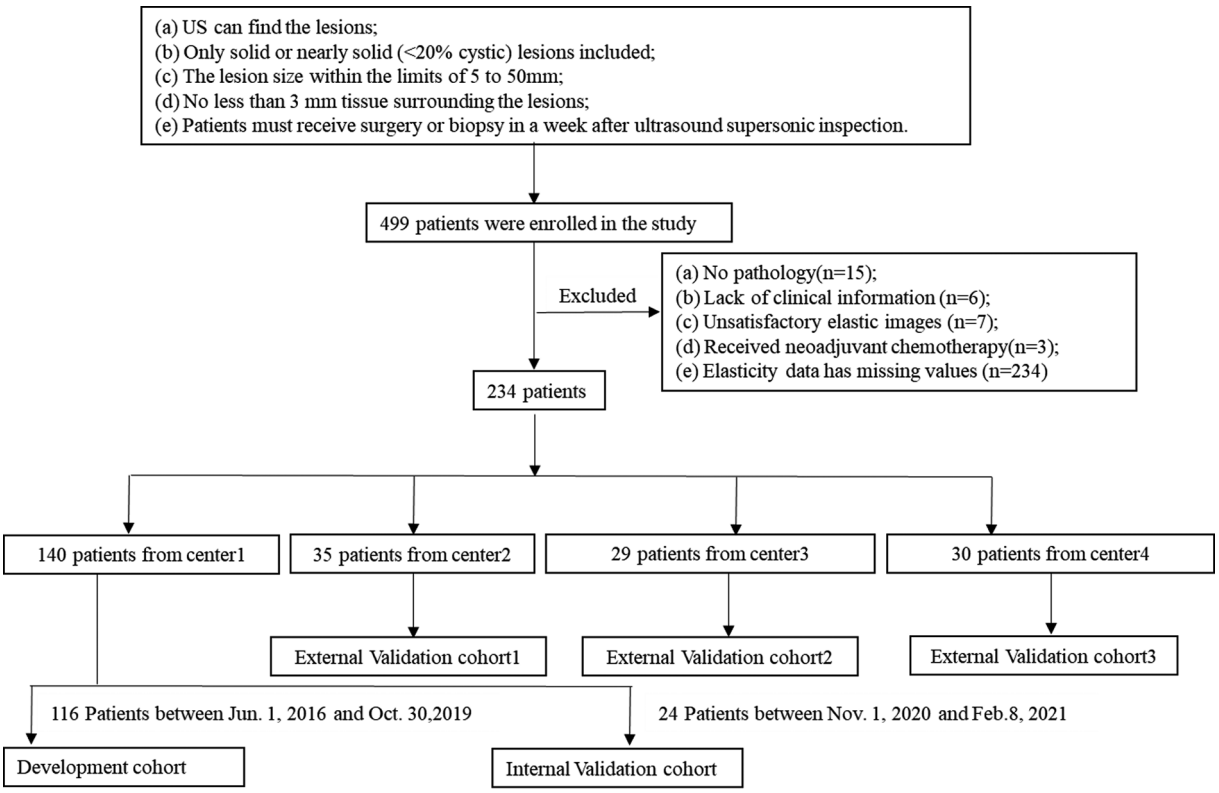

Figure S1 Flowchart of patient selection. US, ultrasound.

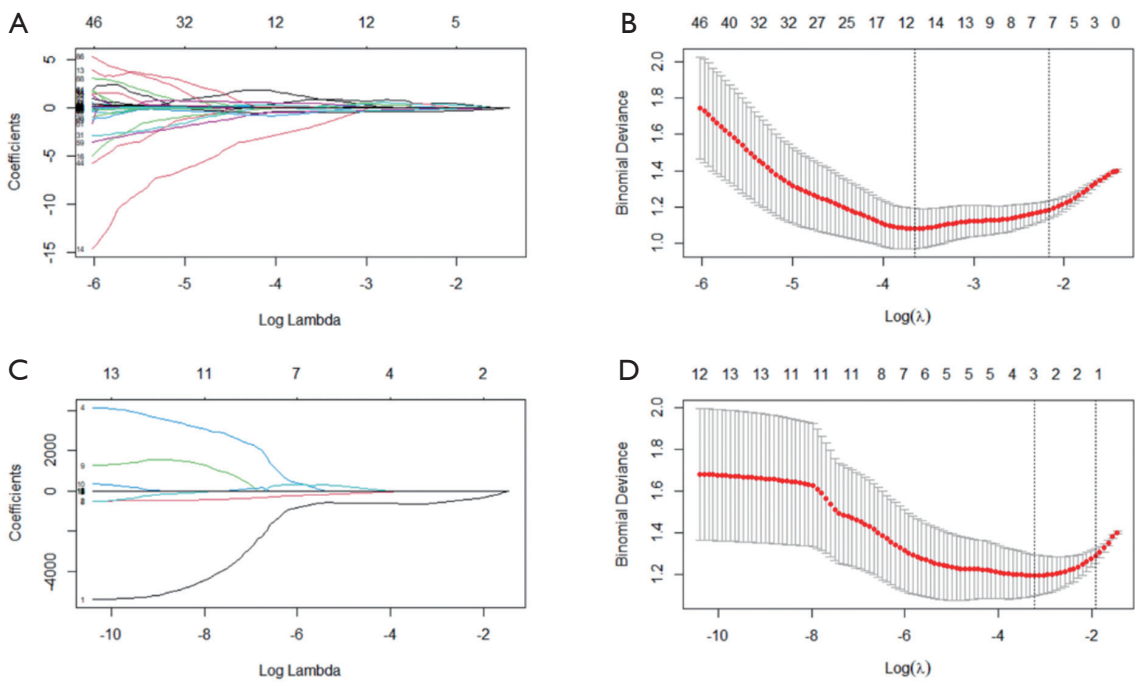

Figure S2 LASSO binary logistic regression results of shear wave elastography (A) and strain elastography (C) characteristics. The optimal penalization coefficient, λ, of shear wave elastography (B) and strain elastography (D) characteristics was selected by repeated procession of the 10-fold cross validation. LASSO, least absolute shrinkage and selection operator.

https://dx.doi.org/10.21037/qims-22-237

Table S1 Diagnostic performance of multivariable diagnosis in the development cohort

Variable	Odds ratio (95% CI)	Wald	P value
Shape =1	3.393 (0.846–15.452)	1.68	0.093
Margin =1	8.742 (2.561–33.294)	3.36	0.001
Posterior features =2	1.828 (0.445–7.289)	0.86	0.389
Vascularity =2	3.705 (1.041–14.333)	1.99	0.047
Shell mean/A mean(1.5)E	6.194 (0.888–57.836)	1.71	0.087
B/A'1.5	1.357 (0.901–2.113)	1.44	0.151

https://dx.doi.org/10.21037/qims-22-237