A new nomogram for predicting the malignant diagnosis of Breast Imaging Reporting and Data System (BI-RADS) ultrasonography category 4A lesions in women with dense breast tissue in the diagnostic setting

Yaping Yang[1][#], Yue Hu[1][#], Shiyu Shen[1], Xiaofang Jiang[1], Ran Gu[1], Hongli Wang[1], Fengtao Liu[1], Jingsi Mei[1], Jing Liang[1], Haixia Jia[2], Qiang Liu[1], Chang Gong[1]

[1]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China; [2]Department of Breast Surgery, Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

[#]These authors contributed equally to this work.

Correspondence to: Chang Gong; Qiang Liu. Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Yanjiang West Road 107#, Yuexiu District, Guangzhou 510120, China. Email: gchang@mail.sysu.edu.cn or changgong282@163.com; victorlq@hotmail.com.

Background: Biopsy has been recommended for Breast Imaging Reporting and Data System (BI-RADS) category 4 lesions. However, the malignancy rate of category 4A lesions is very low (2-10%). Therefore, most biopsies of category 4A lesions are benign, and the results will generally cause additional health care costs and patient anxiety.

Methods: A prediction model was developed based on an analysis of 418 BI-RADS ultrasonography (US) category 4A patients at Sun Yat-sen Memorial Hospital. Univariate and multivariate logistic regression analyses were applied to identify significant variables for inclusion in the final nomogram. The predictive accuracy and discriminative ability were evaluated using the concordance index (C-index) and calibration curves. An independent cohort of 97 patients from the Second Affiliated Hospital of Guangzhou Medical University was used for external validation.

Results: The independent risk factors from the multivariate analysis for the training cohort were family history of breast cancer (OR =4.588, P=0.004), US features [margin (OR =2.916, P=0.019), shape (irregular *vs.* oval, OR =2.474, P=0.044; round *vs.* oval, OR =1.935, P=0.276), parallel orientation *vs.* not parallel (OR =2.204, P=0.040)], low suspicious lymph nodes (OR =7.664, P=0.019), and suspicious calcifications on mammography (MG) (OR =6.736, P=0.001). The C-index was good in the training [0.813, 95% confidence interval (95% CI), 0.733 to 0.893] and validation cohorts (0.765, 95% CI, 0.584 to 0.946). The calibration curves showed optimal agreement between the nomogram prediction and actual observations for the probability of malignancy. Also, the cutoff score was set to 100 for discriminating high and low risk. The model performed well in discerning different risk groups.

Conclusions: We developed a well-discriminated and calibrated nomogram to predict the malignancy of BI-RADS US category 4A lesions in dense breast tissue, which may help clinicians identify patients at lower or higher risk.

Keywords: Breast; ultrasonography (US); nomogram; diagnosis; predictive value of tests

Introduction

The performance of mammography (MG) is markedly lower in women with dense breast tissue (1). The incidence of dense breast tissue is higher in Asians than in White and other races (83.22% *vs.* 61.37%) (2). Ultrasonography (US) plays an indispensable role in dense breast examinations in both the screening and diagnostic settings (3-5).

A chapter on the US was added to the fourth edition of the Breast Imaging Reporting and Data System (BI-RADS) atlas by the American College of Radiology (ACR) in 2003 (6). BI-RADS category 4 lesions have a high likelihood of cancer diagnosis (2–95%). In the fifth edition of BI-RADS updated in 2013, category 4 was further divided into three subcategories: 4A, 4B, and 4C (7). Biopsy has been the recommended management approach for all 3 subcategories because of the suspicion of malignancy. However, the malignancy rate of category 4A lesions is rather low (2–10%), and the proportion of category 4A lesions was 44.5% amongst category 4-5 lesions (8). The average reported rate of BI-RADS category 4A lesions on the diagnostic US in our center is only 6.7% and is much lower for patients under 35 years old (2.7%) (8). Since most biopsies of category 4A lesions are benign, the exclusive use of biopsy can cause unnecessary health care costs and patient anxiety.

Although the ACR BI-RADS guideline has significantly improved the management of breast abnormalities, radiologists still make the final BI-RADS category assessment by weighing up the individual imaging features intuitively in general clinical practice. As such, an objective approach combining individual imaging features and patient information is required to reduce the number of unnecessary biopsies, especially for low suspicion BI-RADS category 4A lesions.

This study aimed to establish a nomogram, incorporating US and MG features and patient clinical information, to quantitatively predict the individual likelihood of malignancy of BI-RADS US category 4A lesions in women with dense breast tissue in the diagnostic setting. We also aimed to identify patients at lower risk of malignancy and eventually prevent unnecessary biopsies.

Methods

Patients

This study's training set data was compiled from 418 patients who met the inclusion criteria and were diagnosed and treated at Sun Yat-sen memorial hospital from January 2014 to June 2017. The inclusion criteria included female patients aged 18 or older who underwent whole breast examinations (US and MG examinations) in the diagnostic setting after experiencing symptoms such as palpable abnormalities, breast pain, and nipple discharge. All breasts were assessed as category 4A by radiologists using the same BI-RADS-US 2013 edition criteria and had dense tissue on MG images (breast composition category c or d according to the 2013 edition of the BI-RADS-MG criteria). All patients had a pathological diagnosis by breast biopsy and/ or surgical excision. To test the nomograms' generalizability, an external validation cohort of 97 patients provided by the Second Affiliated Hospital of Guangzhou Medical University between January 2015 and December 2017 was used. Only 1 4A lesion, which was the most likely to be malignant, was selected from 1 patient. All patients were fully informed about the overall variables in the developed nomogram. Ethical approval was provided by Sun Yat-sen Memorial Hospital Ethics Committee (SYSEC-KY-KS-2019-151). As this was a retrospective study, informed consent from the study participants was exempted.

A standardized data form was created to retrieve all relevant information, including demographic data (age, family history of breast cancer, history of benign breast biopsy), clinical examination of US imaging features (lesion size, margin, shape, orientation, lymph nodes, vascularity, echo, posterior features, calcification, and multiple solid lesions), and MG imaging features (i.e., BI-RADS assessment, suspicious mass/calcifications, architectural distortion, and asymmetry).

US and MG

Four dedicated breast radiologists with >5 years experience (RG, FTL, HLW, JSM)in our center performed the breast US examinations on patients using a high-frequency broad bandwidth linear array transducer equipped with a high-resolution US unit (S2000/S1000 and 18L6 transducer with a center frequency of 15 MHz; Siemens Medical Solutions, Erlangen, Germany). Imaging quality-related parameters were set according to the ACR guidelines (7,9). US was performed and interpreted by a single breast radiologist simultaneously according to the 2013 ACR BI-RADS-US criteria. Lesions were assigned as category 4A if the likelihood of malignancy was assessed to be low (2–10%). Such cases usually had a single suspicious mass with features such as no circumscribed margin and irregular shape or nonparallel orientation. Some cases had low suspicious

features of axillary lymph nodes (cortical thickness of 2.3–3 mm and short axis length less than 1 cm) (10-12). The BI-RADS-US assignment was independent of the MG findings, which is usual practice in US departments in China. US imaging features of patients from the training and validation cohorts were reviewed by another independent radiologist (YYP). The radiologist remained blinded to the patients' MG imaging and pathological information.

MG was performed by a technician with 10 years of experience. Each examination was carried out with the Digital MG machine (Planmed Nuance Excel; Planmed, Helsinki, Finland). Two dedicated radiologists, YH and SYS, with >5 years experience in breast imaging diagnosis, reviewed all MG images from the training and validation cohorts according to the 2013 ACR BI-RADS-MG criteria independently, with patients' US imaging and pathological information unknown beforehand.

According to the second edition of the BI-RADS-US guidelines, patients with category ≥4A lesions are recommended to undergo biopsy. However, the final decision is generally made by patients and referring physicians. We chose to perform a US-guided biopsy and/or surgical excision in our center to obtain the pathological diagnosis. In this study, the US-guided breast biopsy was conducted with a core instrument (with a 14- or 16-gauge needle) or vacuum-assisted biopsy machine (with an 8- or 11-gauge needle). More than 4 tissue samples were obtained and placed in formalin solution, then processed for histopathology by standard procedures. Open breast surgery was suggested for patients with indefinite or malignant histological results. The pathological results were classified as benign or malignant according to the 2013 ACR BI-RADS (7). Atypical ductal/lobular hyperplasia, peripheral duct papilloma and phyllodes tumor results were defined as benign.

Statistical analysis

The training set and validation set results were compared using the χ^2 test or Fisher's exact test. Risk factors were identified by univariate analysis based on the χ^2 test or Fisher's exact test, and the statistically significant variables (P<0.05) were included in the final models by applying multivariate logistic regression to the training set. A nomogram was formulated based on the multivariate analysis results and by using the rms package in R version 3.6.3. The nomogram was then subjected to the bootstrap resample method of 1,000 times for both the internal

validation of the training cohort and the external validation of the validation cohort. Bootstrap resampling for internal validation was performed to reduce the model's over-fitting bias and obtain more reliable prediction accuracy. External validation was also performed with an independent cohort of 97 patients from the Second Affiliated Hospital of Guangzhou Medical University between 2014 and 2017. The nomogram's predictive accuracy and discriminative ability were determined with the concordance index (C-index) and calibration curves. The C-index is a generalization of the area under the receiver operating characteristic (ROC) curve (AUC) (13). A larger C-index corresponds to more accuracy in prediction. Calibration curves were plotted to assess the calibration of the nomogram. The risk score for each patient was calculated from the regression coefficients, and the numerical variables of risk scores were divided into low- or high-risk groups according to the threshold determined by the maximal Youden index for predicting the malignancy rates. Then, we divided patients into a high-risk group or low-risk group according to the cutoff point. Histological type, molecular subtype, and the ki67 status of the malignant lesions were compared between the low-risk and high-risk groups using the χ^2 test or Fisher's exact test. The statistical analyses and modeling were performed using STATA (version 13; Stata Co., College Station, TX) and R software packages. All statistical tests were two-sided, with statistical significance defined as P<0.05.

Results

Study population

The training group comprised 418 female patients (median age, 44 years; range, 19–84 years) with BI-RADS-US category 4A lesions and dense breasts treated at Sun Yat-sen memorial hospital. Among them, 42 patients (10%) had malignant histology, while 376 patients (90%) had benign histology. An independent validation dataset was composed of 97 female patients (median age, 43 years; range, 24–64 years) diagnosed with BI-RADS-US category 4A lesions and dense breasts at the Second Affiliated Hospital of Guangzhou Medical University. A total of 10 patients (10.3%) presented with malignant histology, while 87 patients (89.7%) had benign histology in the validation population. Demographic data and MG and US imaging features of patients in the training and validation cohorts are listed in *Table 1*. There were no differences in the

Table 1 Demographic and imaging features of patients in the training and validation cohorts

Characteristics	Training data		Validation data		Total		P
	N	%	N	%	N	%	
Median [range]	44 [19–84]		43 [24–64]		43 [19–84]		
Age							0.217
≤35	51	12.2	18	18.6	69	13.4	
36–50	285	68.2	59	60.8	344	66.8	
>50	82	19.6	20	20.6	102	19.8	
Family history of breast cancer							0.062
No	393	94.0	86	88.7	479	93.0	
Yes	25	6.0	11	11.3	36	7.0	
History of benign breast biopsy							0.496
No	406	97.1	96	99.0	502	97.5	
Yes	12	2.9	1	1.0	13	2.5	
Histological finding							0.939
Benign	376	90.0	87	89.7	463	89.9	
Malignant	42	10.0	10	10.3	52	10.1	
Lymph node metastasis	5	1.2	1	1.0	6	1.2	0.890
Physical examinations							0.396
Breast pain	167	40.0	48	49.5	215	41.7	
Nipple discharge	40	9.6	8	8.2	48	9.3	
Palpable breast lesions	178	42.6	35	36.1	213	41.4	
Others	33	7.9	6	6.2	39	7.6	
U_Lesion size (mm)							0.342
<20	317	76.8	70	72.2	387	75.9	
≥20	96	23.2	27	27.8	123	24.1	
U_Margin							0.066
Circumscribed	237	56.7	45	46.4	282	54.8	
Not circumscribed	181	43.3	52	53.6	233	45.2	
U_Shape							0.94
Oval	221	52.9	53	54.6	274	53.2	
Round	47	11.2	10	10.3	57	11.1	
Irregular	150	35.9	34	35.1	184	35.7	
U_Orientation							0.990
Parallel	310	74.2	72	74.2	382	74.2	
Not parallel	108	25.8	25	25.8	133	25.8	
U_Lymph nodes							0.058
No/normal	412	98.6	92	94.8	504	97.9	
Low suspicious	6	1.4	5	5.2	11	2.1	

Table 1 (*continued*)

Table 1 (*continued*)

Characteristics	Training data		Validation data		Total		P
	N	%	N	%	N	%	
U_Internal vascularity							0.084
Yes	308	73.7	63	64.9	371	72.0	
No	110	26.3	34	35.1	144	28.0	
U_Echo pattern							0.892
Homogenous	345	82.9	81	83.5	426	83.0	
Heterogeneous	71	17.1	16	16.5	87	17.0	
U_Posterior features							0.879
No shadowing	398	95.2	92	94.8	490	95.1	
Shadowing	20	4.8	5	5.2	25	4.9	
U_Calcifications							0.858
No	367	88.0	86	88.7	453	88.1	
Yes	50	12.0	11	11.3	61	11.9	
U_Multiple solid lesions							0.526
No	257	61.5	63	64.9	320	62.1	
Yes	161	38.5	34	35.1	195	37.9	
M_density of mass							0.281
Equal density	326	78.0	80	82.5	406	78.8	
High density	92	22.0	17	17.5	109	21.2	
BI-RADS_M							0.169
1–3	305	73.0	64	66.0	369	71.7	
4 and above	113	27.0	33	34.0	146	28.3	
M_ Calcifications							0.633
No/normal	401	95.9	92	94.8	493	95.7	
Suspicious	17	4.1	5	5.2	22	4.3	
Location of suspicious calcifications							0.685
In the mass	12	70.6	4	80.0	16	72.7	
Outside the mass	5	29.4	1	20.0	6	27.3	
M_Architectural distortion							0.345
No	408	97.6	93	95.9	501	97.3	
Yes	10	2.4	4	4.1	14	2.7	
M_Asymmetry							0.576
No	390	93.3	92	94.8	482	93.6	
Yes	28	6.7	5	5.2	33	6.4	

U, ultrasonography; M, mammography; BI-RADS, Breast Imaging Reporting and Data System.

3010

Yang et al. A model to predict malignancy of BI-RADS US 4A lesions

demographic and imaging features between the training data and validation data.

Analysis of histological findings in the training set

In the training set, risk factors were determined by univariate analysis with the χ^2 test or Fisher's exact test for different demographic and imaging features. These variables comprised age, family history of breast cancer, history of benign breast biopsy, clinical examination of the mammary gland, US imaging features (lesion size, margin, shape, orientation, lymph nodes, vascularity, echo, posterior features, calcification, and multiple solid lesions), and MG imaging features (BI-RADS assessment, mass, calcifications, architectural distortion, and asymmetry) (*Table 2*). Among the malignant patients, the rate of family history of breast cancer was much higher than that of benign patients (19.0% *vs.* 4.5%, P=0.001). In the malignant patients, the lesion margin was predominantly not circumscribed (81.0% *vs.* 39.1%, P<0.001), the shapes of lesions were significantly more irregular (61.9% *vs.* 33.0%, P<0.001), and the orientation was significantly nonparallel (52.4% *vs.* 22.9%, P<0.001). Moreover, the rate of low suspicious lymph nodes was higher in malignant patients (7.1% *vs.* 0.8%, P=0.009). Nine lesions (21.4%) in the 42 malignant patients were found to be suspicious for MG calcifications, while in benign patients, the rate was 2.1% (P<0.001).

Development of the prediction model

The final univariate and multivariate logistic regression model variables were composed of family history of breast cancer, US imaging features (i.e., margin, shape, orientation, lymph nodes), and calcifications on MG imaging (*Table 3*). A nomogram incorporating all significant predictive factors was established (*Figure 1*). The nomogram identified low suspicious lymph nodes on US and suspicious calcifications on MG as the most significant factors contributing to prediction, followed by a family history of breast cancer. Other US imaging features (i.e., margin, shape, orientation) had a moderate effect on prediction.

Validation and calibration of the nomogram

In the primary dataset, Harrell's C-index produced by the prediction model was 0.813 (95% CI, 0.733 to 0.893) for the established nomogram. In the validation set, the nomogram's C-index for the prediction of malignant lesions

was 0.765 (95% CI, 0.584 to 0.946). The calibration plots presented excellent agreement in the training cohort and good agreement in the validation cohort between the nomogram prediction and the real observation made for the histological findings (*Figure 2A,B*).

Performance of the nomogram in stratifying the risk of patients

We assigned risk scores for each risk factor based on the regression coefficients from the final model that represented the histological findings' prediction (*Table 4*). The overall risk scores of all patients ranged between 0–304, and we chose 100 as the cutoff score with the maximal Youden index for predicting the malignancy rates. Then, patients were divided into a high-risk group and a low-risk group according to the cutoff point, and the AUC was 0.744 (95% CI, 0.664 to 0.824). In the low-risk group, 95.4% had benign pathology (395/414), while only 4.6% (19/414) had malignant pathology. 67.3% (68/101) had benign pathology in the high-risk group, and 32.7% (33/101) had malignant pathology.

In total, 52 breast lesions were malignant in the training and validation cohorts, and 50 of the 52 breast lesions underwent immunohistochemical testing. Histological type, molecular subtype, and the ki67 status of malignant lesions were compared between the low-risk and high-risk groups. There were no significant differences between the low-risk and high-risk groups in these clinicopathological characteristics (Table S1).

Discussion

Higher breast density reduces the sensitivity of MG and affects the results (14,15). Moreover, women with denser tissue, not fattier tissue, have a higher risk of breast cancer (16). The proportion of dense breast lesions in China is much higher than in any other country (17). The US has become an indispensable breast examination method for patients with dense breast lesions. In our center, patients with dense breasts undergo both MG and US examinations. The ACR BI-RADS guideline recommends biopsy for patients with category 4 lesions, causing unnecessary health care costs and patient anxiety, despite the very low malignancy rate of category 4A lesions (2–10%). Here, we developed and validated a nomogram based on demographic data and clinical examination images for predicting the malignant diagnosis of BI-RADS US category 4A lesions in

Table 2 Analysis of the histological findings in training set in US 4A patients

Characteristics	Benign		Malignant		Total		P
	N	%	N	%	N	%	
N	376		42		418		
Age							0.569
≤35	48	12.8	3	7.1	51	12.2	
36–50	255	67.8	30	71.4	285	68.2	
>50	73	19.4	9	21.4	82	19.6	
Family history of breast cancer							0.001
Yes	17	4.5	8	19.0	25	6.0	
No	359	95.5	34	81.0	393	94.0	
History of benign breast biopsy							0.841
No	365	97.1	41	97.6	406	97.1	
Yes	11	2.9	1	2.4	12	2.9	
Physical examinations							0.268
Breast pain	151	40.24	16	38.1	167	40.0	
Nipple discharge	33	8.8	7	16.7	40	9.6	
Palpable breast lesions	164	43.6	14	33.3	178	42.6	
Others	28	7.4	5	11.9	33	7.9	
U_Lesion size (mm)							0.169
<20	282	75.8	35	85.4	317	76.8	
≥20	90	24.2	6	14.6	96	23.2	
U_Margin							<0.001
Circumscribed	229	60.9	8	19.0	237	56.7	
Not circumscribed	147	39.1	34	81.0	181	43.3	
U_Shape							<0.001
Oval	212	56.4	9	21.4	221	52.9	
Round	40	10.6	7	16.7	47	11.2	
Irregular	124	33.0	26	61.9	150	35.9	
U_ Orientation							<0.001
Parallel	290	77.1	20	47.6	310	74.2	
Not parallel	86	22.9	22	52.4	108	25.8	
U_Lymph_nodes							0.009
Low suspicious	3	0.8	3	7.1	6	1.4	
No/normal	373	99.2	39	92.9	412	98.6	

Table 2 (*continued*)

3012

Yang et al. A model to predict malignancy of BI-RADS US 4A lesions

Table 2 (*continued*)

Characteristics	Benign		Malignant		Total		P
	N	%	N	%	N	%	
U_Internal vascularity							0.145
Yes	95	25.3	15	35.7	110	26.3	
No	281	74.7	27	64.3	308	73.7	
U_Echo pattern							0.381
Homogenous	313	83.5	32	78.0	345	82.9	
Heterogeneous	62	16.5	9	22.0	71	17.1	
U_Posterior features							0.442
No shadowing	357	94.9	41	97.6	398	95.2	
Shadowing	19	5.1	1	2.4	20	4.8	
U_Calcifications							0.325
No	332	88.5	35	83.3	367	88.0	
Yes	43	11.5	7	16.7	50	12.0	
U_Multiple solid lesions							0.345
No	234	62.2	23	54.8	257	61.5	
Yes	142	37.8	19	45.2	161	38.5	
M_density of mass							0.545
Equal density	296	78.7	30	71.5	326	78.0	
High density	80	21.3	12	28.6	92	22.0	
BI-RADS_M							0.182
1–3	278	73.9	27	64.3	305	73.0	
4 and above	98	26.1	15	35.7	113	27.0	
M_Calcifications							<0.001
No/normal	368	97.9	33	78.6	401	95.9	
Suspicious	8	2.1	9	21.4	17	4.1	
Location of suspicious calcifications							0.620
In the mass	5	62.5	7	77.8	12	70.6	
Outside the mass	3	37.5	2	22.2	5	29.4	
M_Architectural distortion							0.996
No	367	97.6	41	97.6	408	97.6	
Yes	9	2.4	1	2.4	10	2.4	
M_Asymmetry							0.903
No	351	93.4	39	92.9	390	93.3	
Yes	25	6.6	3	7.1	28	6.7	

U, ultrasonography; M, mammography; BI-RADS, Breast Imaging Reporting and Data System.

Table 3 Univariate and multivariate analyses of the training set for histological findings

Variable	Univariate analysis			Multivariate analysis		
	OR	95% CI	P	OR	95% CI	P
Family history of breast cancer						
No	Ref			Ref		
Yes	4.969	1.998–12.356	0.001	4.588	1.614–13.043	0.004
U_Margin						
Circumscribed	Ref			Ref		
Not circumscribed	6.621	2.982–14.699	<0.001	2.916	1.191–7.140	0.019
U_Shape						
Oval	Ref			Ref		
Round	4.122	1.451–11.708	0.008	1.935	0.591–6.337	0.276
Irregular	4.939	2.242–10.880	<0.001	2.474	1.023–5.981	0.044
U_ Orientation						
Parallel	Ref			Ref		
Not parallel	3.709	1.933–7.117	<0.001	2.204	1.038–4.682	0.040
U_Lymph_nodes						
Low suspicious	Ref			Ref		
No/normal	9.564	1.867–49.007	0.007	7.664	1.392–42.200	0.019
M_Calcifications						
No/normal	Ref			Ref		
Suspicious	12.545	4.538–34.679	<0.001	6.736	2.169–20.919	0.001

U, ultrasonography; M, mammography; 95% CI, 95% confidence interval.

women with dense breast tissue.

Patients with dense breast tissue and US BI-RADS 4A lesions are difficult to manage, and risk stratification is also challenging. For these cases, our nomogram incorporates 6 factors, composed of a family history of breast cancer, 4 US imaging features (margin, shape, orientation, lymph nodes), and calcifications on MG imaging. Harrell's C-index was used to predict the malignancy rate, which was 0.813 (95% CI, 0.733 to 0.893) in the training set and 0.765 (95% CI, 0.584 to 0.946) in the validation set.

Our nomogram score's cutoff point was 100, which divided the US of category 4A patients into two groups: high risk and low risk. In the low-risk group, 95.4% (395/414) had benign pathology, and 4.6% had malignant pathology (19/414), while in the high-risk group, 67.3% (68/101) had benign pathology and 32.7% (33/101) had malignant pathology. The prediction of our nomogram was

more accurate than the ACR BI-RADS category 4A lesions and other studies. In terms of the strain elastography features for BI-RADS 4A lesions, the malignancy rate was found to be 4.5% (4/88) for soft elastography and 12.6% (21/167) for not soft elastography (18). Menezes *et al.* used optoacoustic US to help downgrade benign masses classified as BI-RADS 4with the US, with 11.1% (3/27) BI-RADS 4b being correctly downgraded to BI-RADS 3 or 2, and 28.6% (2/7) malignant masses classified as BI-RASD 4a being incorrectly downgraded (19). Our nomogram of the low-risk group was able to help 95.4% of patients avoid a biopsy, based only on a few essential factors, i.e., family history of breast cancer, US imaging features, and calcifications on MG imaging. A study by Patterson *et al.* showed that breast cancer risk was higher in women aged >40 with BI-RADS 4A US lesions; however, this was not observed in our study (20).

3014

Yang et al. A model to predict malignancy of BI-RADS US 4A lesions
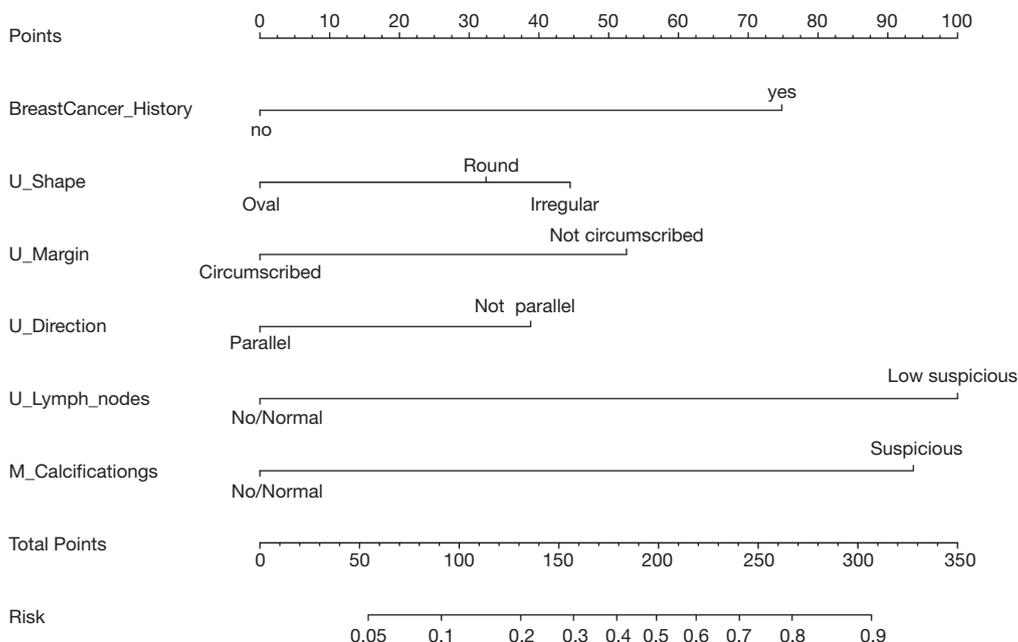
Figure 1 Prognostic nomogram for predicting patients' histological findings with Breast Imaging Reporting and Data System (BI-RADS) ultrasonography category 4A lesions in dense breast tissue. The prediction model was developed with the training dataset, which incorporated the family history of breast cancer, ultrasonography imaging features (shape, margin, orientation, and low suspicious lymph nodes), suspicious calcifications on MG imaging, and presented as a nomogram. The sum of the scores is located on the Total Points axis and corresponds to the malignancy risk of 4A lesions.
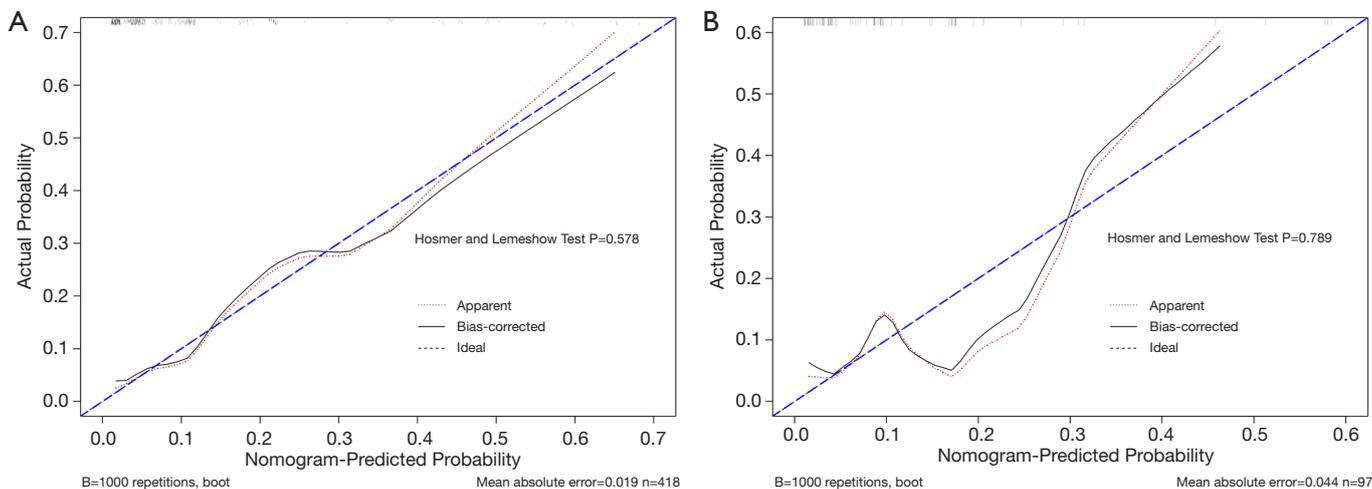
Figure 2 Calibration plots for histological findings between the nomogram prediction model and real observations in the (A) training cohort and (B) validation cohort. The calibration plots presented excellent agreement in the training cohort and good agreement in the validation cohort between the nomogram prediction and real observations made for the histological findings. The Hosmer and Lemeshow Test were 0.578 for the training cohort and 0.789 for the validation cohort.

Table 4 Point assignment and prediction scores for histological findings

Variable	Score
Family history of breast cancer	
No	0
Yes	74.8
U_Margin	
Circumscribed	0
Not circumscribed	52.6
U_Shape	
Oval	0
Round	33.4
Irregular	44.5
U_Orientation	
Parallel	0
Not parallel	38.8
U_Lymph_nodes	
Low suspicious	0
No/normal	100
M_Calcifications	
No/normal	0
Suspicious	93.7

U, ultrasonography; M, mammography.

In the multivariate analyses, 6 variables were found to be independent predictors for malignant diagnosis, including a family history of breast cancer, US imaging features (i.e., margin, shape, orientation, lymph nodes), and calcifications on MG imaging. Benndorf *et al.* reported that breast cancer's family history was an independent risk factor (odds ratio: 5.53) in the setting of an MG BI-RADS 3 assessment (21). Other previous studies have shown that the US features that most likely predict a malignant diagnosis are irregular shapes, nonparallel orientations, and no circumscribed margin (22-25). Our results were consistent with these findings, suggesting that other US features, such as internal echo, posterior features, calcification, and vascularity, cannot be used as significant predictors for malignancy (24,25). The US is a useful imaging modality to evaluate axillary nodes using morphological criteria, such as cortical thickening, hilar effacement, and nonhilar cortical blood flow. It can be highly specific in the identification of metastases (12). Although our data indicated that the rate of lymph node metastasis gradually increased as the BI-RADS US category steadily increased, 12.3% (9/73) of patients with a category 4A malignant lesion also presented a positive axillary status (8). In this study, the US lymph node status had the most important contribution in predicting malignancy in the nomogram, followed by suspicious calcifications on MG. Other MG features did not show distinct predictive value in this study. Our study, focused on women with dense breast tissue, showed that the US's role in dense breast tissue is more important, and noncalcified lesions can be obscured by normal dense tissue, which has also been reported in previous studies (3,4,26).

According to the second edition of the BI-RADS-US guidelines, patients with category ≥4A lesions are suggested to undergo biopsy. However, physicians in China may give other medical advice based on their own experience after clinical consultations and physical and imaging examinations. The rate of biopsy for BI-RADS-US 4A lesions may be less than 70–80%. Our nomogram provides a convenient and feasible tool to predict the malignancy risk of patients. For the high-risk group of patients, active biopsy or surgery is recommended, while for patients in the low-risk group who are not willing to undergo biopsy, intensive follow-up can be suggested.

There are several primary limitations in this study. First, this was a retrospective study design which is limited by the deficient collection of risk factors related to breast cancer in patients, such as menstruation, fertility, and smoking, amongst others. Second, we did not have the patients' genetic mutation (BRCA 1/2, etc.) data. This information is relevant to the dense breast tissue population and can indicate a higher likelihood of malignancy. Lastly, our nomogram did not include MRI imaging because the examination is expensive and inaccessible. The use of MRI is optional in patients with dense breast tissue, but its use is not universally recommended. Moreover, the determination of BI-RADS-US category 4A lesions is an uncertain process dependent on the institution and radiologist training. Therefore, cohorts of 4A patients may look different between different institutions/countries.

In future studies, we will carry out prospective multicenter radiomics investigations combining the risk factors associated with breast cancer in patients with radiomics data to establish a more sensitive and specific breast cancer prediction model.

In conclusion, we have developed a well-discriminated and calibrated nomogram which provides an individual

prediction of malignancy of BI-RADS US category 4A lesions in dense breast tissue based entirely upon the family history of breast cancer, US imaging features (i.e., margin, shape, and orientation), low suspicious lymph nodes on the US, and suspicious calcifications on MG. Our nomogram may help clinicians identify patients with a lower risk of malignancy, thus preventing unnecessary and expensive biopsies.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/qims-20-1203). The authors have no conflicts of interest to declare.

Ethical Statement: Ethical approval was provided by Sun Yat-sen Memorial Hospital Ethics Committee (SYSEC-KY-KS-2019-151). As this was a retrospective study, informed consent from the study participants was exempted.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, White E. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081-7.
2. del Carmen MG, Halpern EF, Kopans DB, Moy B, Moore RH, Goss PE, Hughes KS. Mammographic breast density and race. AJR Am J Roentgenol 2007;188:1147-50.
3. 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.
4. Ohuchi N, Suzuki A, Sobue T, Kawai M, Yamamoto S, Zheng YF, Shiono YN, Saito H, Kuriyama S, Tohno E, Endo T, Fukao A, Tsuji I, Yamaguchi T, Ohashi Y, Fukuda M, Ishida T. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. Lancet 2016;387:341-8.
5. Shen S, Zhou Y, Xu Y, Zhang B, Duan X, Huang R, Li B, Shi Y, Shao Z, Liao H, Jiang J, Shen N, Zhang J, Yu C, Jiang H, Li S, Han S, Ma J, Sun Q. A multi-centre randomised trial comparing ultrasound vs mammography for screening breast cancer in high-risk Chinese women. Br J Cancer 2015;112:998-1004.
6. Mendelson E, Baum J, Berg W, Merritt C, Rubin E. Breast imaging reporting and data system, BI-RADS: ultrasound. Reston VA: American College of Radiology 2003.
7. Sickles E, D'Orsi C, Bassett L, Appleton C, Berg W, Burnside E. ACR BI-RADS® Atlas Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology 2013.
8. Hu Y, Yang Y, Gu R, Jin L, Shen S, Liu F, Wang H, Mei J, Jiang X, Liu Q, Su F. Does patient age affect the PPV3 of ACR BI-RADS Ultrasound categories 4 and 5 in the diagnostic setting? Eur Radiol 2018;28:2492-8.
9. American College of Radiology. ACR practice parameter for theperformance of a breast ultrasound examination. Available online: https://www.acr.org/-/media/ACR/Files/Practice-Parameters/USBreast.pdf?la=en. Res.38-2016.

Accessed Oct 20, 2020.

10. Lee B, Lim AK, Krell J, Satchithananda K, Coombes RC, Lewis JS, Stebbing J. The efficacy of axillary ultrasound in the detection of nodal metastasis in breast cancer. AJR Am J Roentgenol 2013;200:W314-20.

11. Deurloo EE, Tanis PJ, Gilhuijs KG, Muller SH, Kröger R, Peterse JL, Rutgers EJ, Valdés Olmos R, Schultze Kool LJ. Reduction in the number of sentinel lymph node procedures by preoperative ultrasonography of the axilla in breast cancer. Eur J Cancer 2003;39:1068-73.

12. Ecanow JS, Abe H, Newstead GM, Ecanow DB, Jeske JM. Axillary staging of breast cancer: what the radiologist should know. Radiographics 2013;33:1589-612.

13. Harrell FE, Jr. Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer -Verlag New York; 2001.

14. Philpotts LE. Density Variation among Mammographic Modalities Will Likely Impact Imaging Management and Risk Stratification. Radiology 2019;291:328-9.

15. Evans DG, van Veen EM, Howell A, Astley S. Heritability of mammographic breast density. Quant Imaging Med Surg 2020;10:2387-91.

16. Boyd NF, Rommens JM, Vogt K, Lee V, Hopper JL, Yaffe MJ, Paterson AD. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 2005;6:798-808.

17. Yang Y, Liu J, Gu R, Hu Y, Liu F, Yun M, Xiao Q, Wu M, Liu Q, Su F. Influence of factors on mammographic density in premenopausal Chinese women. Eur J Cancer Prev 2016;25:306-11.

18. Koh J, Kim EK, Kim MJ, Yoon JH, Park VY, Moon HJ. Role of elastography for downgrading BI-RADS category 4a breast lesions according to risk factors. Acta Radiol 2019;60:278-85.

19. Menezes GLG, Pijnappel RM, Meeuwis C, Bisschops R, Veltman J, Lavin PT, van de Vijver MJ, Mann RM. Downgrading of Breast Masses Suspicious for Cancer by Using Optoacoustic Breast Imaging. Radiology 2018;288:355-65.

20. Patterson SK, Neal CH, Jeffries DO, Joe A, Klein K, Bailey J, Pinsky R, Paramagul C, Watcharotone K. Outcomes of solid palpable masses assessed as BI-RADS 3 or 4A: a retrospective review. Breast Cancer Res Treat 2014;147:311-6.

21. Benndorf M, Wu Y, Burnside ES. A history of breast cancer and older age allow risk stratification of mammographic BI-RADS 3 ratings in the diagnostic setting. Clin Imaging 2016;40:200-4.

22. Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, Jackson VP, Sayre JW, Bassett LW. Benign versus malignant solid breast masses: US differentiation. Radiology 1999;213:889-94.

23. Hong AS, Rosen EL, Soo MS, Baker JA. BI-RADS for sonography: positive and negative predictive values of sonographic features. AJR Am J Roentgenol 2005;184:1260-5.

24. Bouzghar G, Levenback BJ, Sultan LR, Venkatesh SS, Cwanger A, Conant EF, Sehgal CM. Bayesian probability of malignancy with BI-RADS sonographic features. J Ultrasound Med 2014;33:641-8.

25. Elverici E, Barca AN, Aktas H, Ozsoy A, Zengin B, Cavusoglu M, Araz L. Nonpalpable BI-RADS 4 breast lesions: sonographic findings and pathology correlation. Diagn Interv Radiol 2015;21:189-94.

26. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of Breast Cancer Screening: Systematic Review to Update the 2009 US Preventive Services Task Force Recommendation. Ann Intern Med 2016;164:256-67.

Table S1 Comparison of clinicopathological characteristics of malignant lesions between the low-risk and high-risk groups

	Low risk, n (%)	High risk, n (%)	P value
Molecular subtype			
ER+ & HER2-	14 (36.8)	24 (63.2)	0.939
ER+ & HER2+	1 (25.0)	3 (75.0)	
ER- & HER2+	0 (0)	1 (100)	
Triple negative	3 (42.9)	4 (57.1)	
Histological type			
IDC	9 (34.6)	17 (65.4)	0.956
ILC	0 (0)	1 (100)	
DCIS	6 (40.0)	9 (60.0)	
Others	4 (40.0)	6 (60.0)	
Ki67			
<15	13 (34.2)	25 (65.8)	0.735
≥15	5 (41.7)	7 (58.3)	

IDC, invasive ductal carcinoma; DCIS, ductal carcinoma *in situ*; ILC, invasive lobular carcinoma; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.