



Comprehensive quantitative malignant risk prediction of pure grouped amorphous calcifications: clinico-mammographic nomogram

Lijuan Shen^{1,2#^}, Tingting Jiang^{1,3#^}, Pengzhou Tang^{1,3}, Huijuan Ge^{3,4^}, Chao You^{1,3^}, Weijun Peng^{1,3^}

¹Department of Radiology, Fudan University Shanghai Cancer Center, Shanghai, China; ²Department of Nuclear Medicine, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ³Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China; ⁴Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China

Contributions: (I) Conception and design: C You, W Peng, L Shen; (II) Administrative support: W Peng, C You; (III) Provision of study materials or patients: C You, W Peng, L Shen; (IV) Collection and assembly of data: L Shen, T Jiang, P Tang, H Ge; (V) Data analysis and interpretation: L Shen, T Jiang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Weijun Peng. Department of Radiology, Fudan University Shanghai Cancer Center, 270 Dongan Road, Xuhui District, Shanghai 200032, China. Email: cjr.pengweijun@vip.163.com; Chao You. Department of Radiology, Fudan University Shanghai Cancer Center, 270 Dongan Road, Xuhui District, Shanghai 200032, China. Email: youchao8888@aliyun.com.

Background: Pure grouped amorphous calcifications are classified as Breast Imaging Reporting and Data System (BI-RADS) category 4B suspicious calcifications and recommended for biopsy. However, the biopsies often reveal benign findings, especially in screening mammograms (92.4–97.2%).

Methods: Mammograms of 699 pure grouped amorphous calcifications with final pathological results were analyzed in this retrospective study. The maximum span (MS) of the group of calcifications and the MS of the parallel/vertical direction of the mammary duct (MPS/MVS) were measured, and the MPS to MVS ratio was calculated. Based on the MS and ratio, 2 prediction nomograms with other clinic-mammographic features were developed. The discrimination performance of the models was assessed and compared by the area under the receiver operating characteristic curve (AUC). Scatterplots were created to determine the cutoff values with fewer misdiagnoses of malignant calcifications and fewer false positives.

Results: Ultimately, 2 prediction models were successfully developed based on the 4 risk factors of age, purpose of the mammogram, whether multiple or single calcifications, and the MS [odds ratio (OR) =1.06, P=0.02]/ratio (OR =6.05, P<0.001). Both models had good discrimination. The ratio model performed better than the MS model in the training cohort (AUC of 0.875 and 0.834, respectively, P=0.003) and validation cohort (AUC 0.908 and 0.867, respectively, P=0.047). For the group with probably benign calcifications (as detected by the ratio nomogram), the malignancy rates were 2.7% [95% confidence interval (CI): 1.00% to 6.53%] and 1.19% (95% CI: 0.06% to 7.37%) in the training and validation cohorts, respectively, and 44.12% and 47.70% of benign biopsies were detected in the training and validation cohorts, respectively.

Conclusions: The clinico-mammographic quantitative malignancy risk prediction nomogram showed favorable discrimination and calibration performance. The ratio model showed better diagnostic efficiency than the MS model, and identified >40% of benign biopsies.

Keywords: Breast; mammography; amorphous calcification; predictive value of tests; nomogram

^ ORCID: Lijuan Shen, 0000-0003-4250-834X; Tingting Jiang, 0000-0002-0366-3606; Huijuan Ge, 0000-0003-1059-4212; Chao You, 0000-0001-9420-1919; Weijun Peng, 0000-0002-7254-140X.

Submitted Aug 17, 2021. Accepted for publication Feb 03, 2022.

doi: 10.21037/qims-21-797

View this article at: <https://dx.doi.org/10.21037/qims-21-797>

Introduction

Mammography is widely accepted as the most reliable technique for assessing breast calcifications. Microcalcifications are associated with approximately 55% of non-palpable breast carcinomas and 85–95% of ductal carcinomas in situ (DCIS) which are detected by screening mammography (1,2). The 5th Breast Imaging Reporting and Data System (BI-RADS) (3) provides standardized descriptions for the morphology and distribution of calcifications, which should both be considered when analyzing and planning therapeutic management (4–10). Grouped amorphous calcifications are the most common type of suspicious calcifications. For this type, the positive predictive value (PPV) is approximately 20%, the category is assessed as 4B, and a biopsy is recommended (4).

Despite the widely known benefits of the early detection of breast cancer, concerns continue to be raised regarding unnecessary biopsies and surgeries on benign microcalcifications (11–13). In recent years, researchers have debated the acceptable risk and the benefit balance on grouped amorphous calcifications due to the low PPV, especially in screening mammography, which has a PPV of 2.8–7.6% (4–7). These studies have suggested that the results of biopsies of grouped amorphous calcifications are often benign. Many factors affect the PPV of calcifications, including the distribution, morphology, patient age/menopausal status, whether the mammogram was for screening or diagnostic purposes, and history of breast or ovarian cancer (2–10,14,15). We also found that the malignancy rate of grouped amorphous calcifications was lower than other calcification distributions in our previous study (16). However, to date, very few studies have focused on the assessment of pure grouped amorphous calcifications.

This study sought to qualitatively predict the malignancy rates of grouped amorphous calcifications using user-friendly nomograms, and identify calcifications that are probably benign. We present the following article in accordance with the STARD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-797/rc>).

Methods

Participants

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of Fudan University Shanghai Cancer Center, and individual consent for this retrospective analysis was waived. We consecutively reviewed all amorphous calcifications (n=5,974) on both screening and diagnostic full-field digital mammograms at our institution with high-resolution (5-megapixel) monitors (BARCO MDNG -5121) between January 2013 and December 2018. The clinical and pathological information was also reviewed on the electronic medical record system. At our institution, all grouped amorphous calcifications are assessed as BI-RADS 4B and recommended for biopsy.

The exclusion criteria for this study were as follows: (I) accurate pathologic results were not obtained (n=3,494, most of these calcifications had a diffused or scattered distribution); (II) associated masses, suspicious asymmetries, or architectural distortions were found (n=1,236); (III) the patient had undergone invasive breast surgery within 6 months of the mammogram (n=160); (IV) the calcifications could only be observed in 1 view (n=32); (V) the calcification distribution was not grouped (n=326). [Figure S1](#) shows the flow of patients throughout the study. Ultimately, 699 pure grouped amorphous calcifications from 604 women were included in this study.

Data acquisition

All anonymous full-field digital mammograms, demographics, and clinical features were reviewed and evaluated by 2 experienced breast radiologists (with 11 and 15 years of experience, respectively, who were blinded to the pathologic results) based on 5th edition of BI-RADS (the only widely used structured guidelines for breast imaging throughout the world). The breast composition, whether multiple calcifications were present

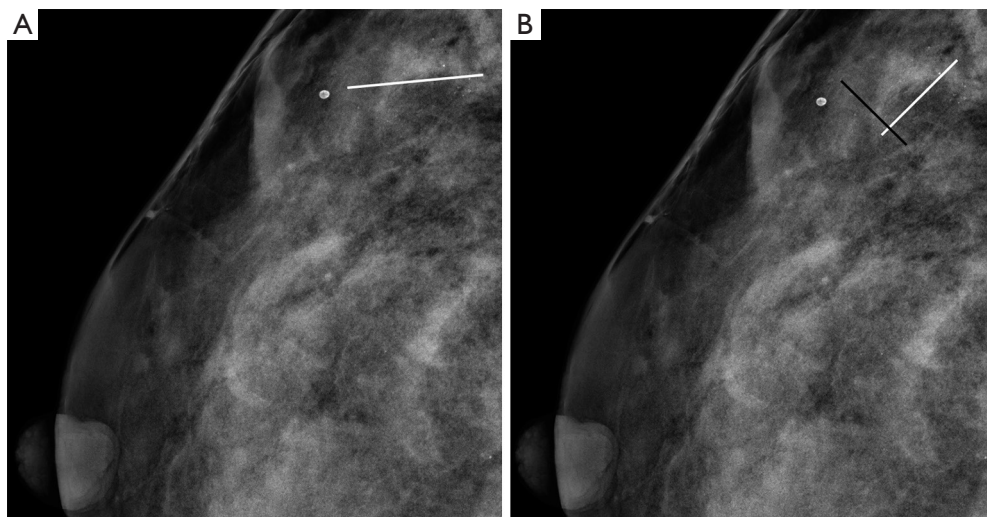


Figure 1 Illustration of the measurements. MS (A, white line), MPS (B, white line), and MVS (B, black line). MS, maximum span of the calcifications; MPS, maximum span of the calcifications in the parallel direction of the mammary duct; MVS, maximum span of the calcifications in the vertical direction of MPS.

(classified as follows: single; multiple unilateral; multiple bilateral), side, location, and BI-RADS classification were recorded. The maximum span (MS) of the calcifications (4), the MS of the calcifications in the parallel direction of the mammary duct (parallel direction from the breast nipple toward the chest wall (17), (abbreviated as MPS), and the MS of the calcifications in the vertical direction of the MPS (abbreviated as MVS) were measured (*Figure 1*), and the ratio of the MPS to the MVS (abbreviated as ratio) were calculated. The MS was recorded as the larger of the measurements on the 2 views of the mammograms, while the MPS, MVS, and ratio were recorded as the measurements on the view with the larger ratio. At first, the 2 radiologists measured 60 cases each, and an absolute agreement 2-way random intraclass correlation coefficient (ICC) was used for the repeatability analysis of the MS and ratio. Next, the 2 radiologists assessed all the included calcifications together. If there was any disagreement, another more senior radiologist (26 years' experience) was invited to evaluate and discuss the issue with the 2 radiologists, before forming a final conclusion.

Case demographics and clinical features, including age, menopausal status, a personal or family history of breast, ovarian cancer, or other malignant tumor, the purpose of the mammogram (classified as follows: 0: screening mammogram, 1: diagnostic examination), and the histopathologic features from the electronic medical records were also documented. The final pathologic results were

categorized as benign or malignant. Any high-risk lesions that could not be clearly identified as benign or malignant were excluded, as stipulated by exclusion criterion (I).

Statistical analysis

The programs of R software (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria, <http://www.r-project.org>) and SPSS 22.0 (SPSS Inc., IBM Corp., Chicago, IL, USA) were used for statistical analysis. Cases were divided into training and validation cohorts at a ratio of 7:3 using the “caret” package based on stratified randomization with fixed benign and malignant proportions. Mann-Whitney U tests, χ^2 tests, or Fisher's exact tests were used to compare all clinical and mammographic features as appropriate. A 2-sided $P < 0.05$ indicated a statistically significant difference. A stepwise multivariate binary logistic regression analysis was conducted to select risk factors and develop prediction models using the training cohort. To prevent overfitting, the ratio and MS were separately introduced to develop 2 prediction models. The probability of malignancy (prob-score) was calculated for each group of calcifications. The discrimination performance of the models was assessed and compared by the area under the receiver operating characteristic curve (AUC) with the 95% confidence interval (CI). Calibration curves and the Hosmer-Lemeshow goodness-of-fit test were used to assess the calibration performance and the overall fit of the

Table 1 Multivariate logistic regression analysis of the 2 models

Features	MS model		Ratio model	
	P value	OR (95% CI)	P value	OR (95% CI)
MS	0.022*	1.06 (1.01, 1.13)	–	–
Ratio	–	–	<0.001*	6.05 (3.29, 11.11)
Purpose	<0.001*	15.72 (8.64, 28.62)	<0.001*	16.98 (8.88, 32.47)
Quantity				
Single	–	–	–	–
Multiple unilateral	0.256	0.60 (0.25, 1.44)	0.629	0.80 (0.33, 1.95)
Multiple bilateral	0.005*	0.37 (0.19, 0.75)	0.009*	0.37 (0.17, 0.78)
Age groups				
<40 years	0.003*	0.19 (0.06, 0.56)	0.007*	0.21 (0.07, 0.65)
≥40 and <50 years	0.002*	0.23 (0.09, 0.58)	0.007*	0.26 (0.10, 0.70)
≥50 and <60 years	0.022*	0.31 (0.11, 0.84)	0.021*	0.28 (0.10, 0.82)
≥60 years	–	–	–	–

*, P<0.05. MS model, the prediction model based on the factor of MS (maximum span of the group of calcifications); Ratio model, the prediction model based on the factor of ratio (the ratio of the maximum span of the group of calcifications in the parallel direction to the maximum span in the vertical direction of the mammary duct). OR, odds ratio; CI, confidence interval.

models. A P value >0.05 indicated non-significant deviance from the theoretical perfect calibration. A cutoff value A with the highest Youden index was found. A cutoff value B with fewer misdiagnoses of malignant calcifications and fewer false positives was also found by creating a scatterplot with the prob-scores and actual pathologic results (malignant or benign). We divided the calcifications into probably benign and malignant groups based on the above 2 cutoff values.

Results

Included cases and clinicopathologic characteristics

The final study included 699 grouped amorphous calcifications from 604 women, including 582 benign and 117 malignant lesions (Tables S1,S2). The time between the mammography and biopsy or surgical excision was <30 days. The cases had a mean age of 46.54±8.79 years (24–86 years), and most were premenopausal (74.2%, 523/699; Table S2). There were 490 groups of calcifications in the training cohort, and 209 in the validation cohort, and there was no difference in the pathological results (P=0.922) or clinical or mammographic features (P=0.134–0.978) between the 2 cohorts.

Development and validation of the prediction models

The MS and ratio had good repeatability, the ICC was 0.94 (95% CI: 0.87 to 0.97) for the MS, and 0.92 (95% CI: 0.86 to 0.95) for the ratio. The MS and ratio differed significantly between the benign calcifications (MS 9.45±5.29 mm and ratio 1.06±0.38) and malignant calcifications (MS 11.41±4.64 mm and ratio 1.52±0.64) (P<0.001). Additionally, the MPS, whether multiple calcifications were present, purpose of the mammogram (screening or diagnostic), menopausal status, age, and age group differed significantly between the benign and malignant calcification groups. Some risk factors, such as whether the calcifications were increased findings compared to the previous examination, could not be included in the statistical analyses due to the small amount of positive data. Finally, the MS/ratio models were separately developed with the 4 risk factors of the MS/ratio, whether multiple calcifications were present, purpose of the mammogram, and age group (LR =118.69, R²=0.362/LR =152.48, R²=0.450), and the prob-score/prob-score1 was calculated. The odds ratios (ORs) with 95% CIs and the p values for the included risk factors are detailed in Table 1. The MS model had AUCs of 0.834 (95% CI: 0.782 to 0.887) and 0.867 (95% CI: 0.791 to 0.944), while the ratio model had

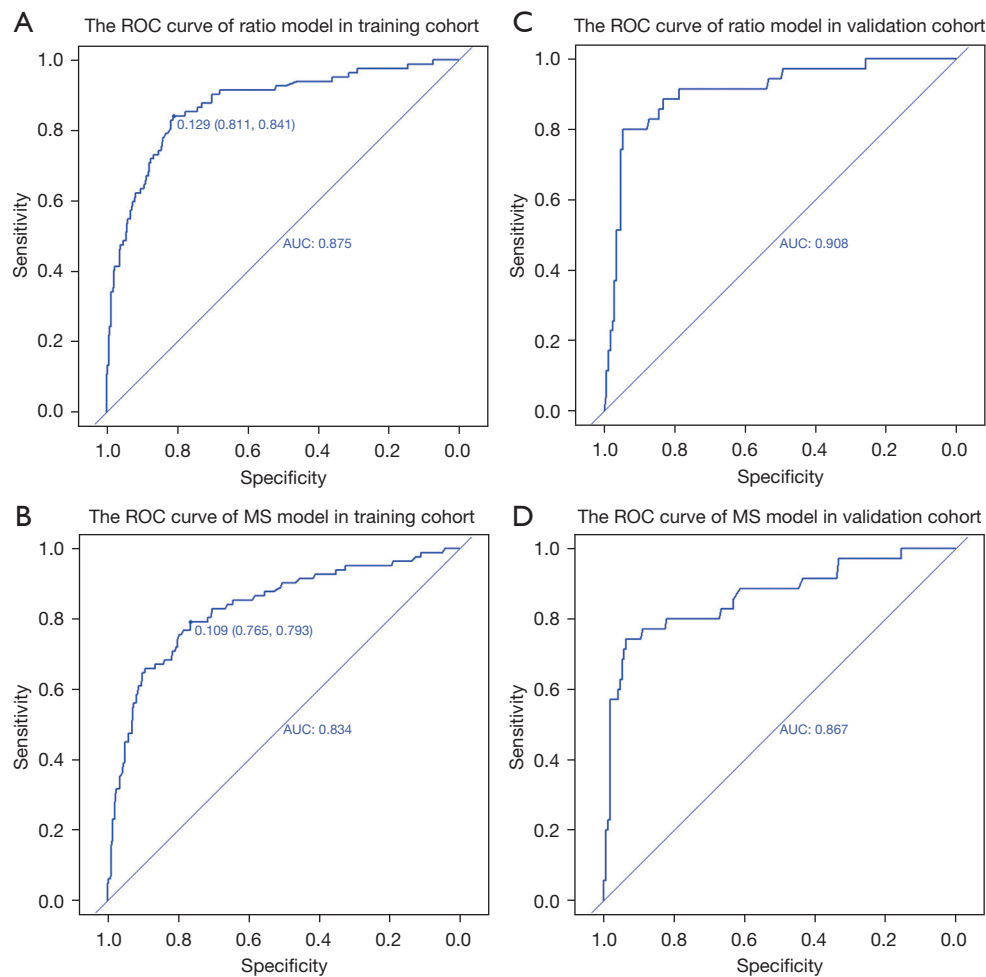


Figure 2 The ROC curves. The AUCs of the ratio model (A,C) showed better performance than the MS model (B,D). The cutoff value A with the highest Youden index was 0.129 (A) in the ratio model with a sensitivity of 0.841 and specificity of 0.811. The cutoff value was 0.109 (C) in the MS model with a sensitivity of 0.793 and specificity of 0.765. ROC, receiver operating characteristic; AUC, area under the ROC curve; MS, maximum span of the calcifications.

AUCs of 0.875 (95% CI: 0.830 to 0.920) and 0.908 (95% CI: 0.851 to 0.965) for the training and validation cohorts, respectively (Figure 2). The Delong test confirmed that the ratio model performed better than the MS model in both the training ($P=0.003$) and validation ($P=0.047$) cohorts. The accuracy, sensitivity, specificity, PPV, and negative predictive value (NPV) of the ratio model were all higher than those of the MS model (see Table 2). The Hosmer-Lemeshow test ($P>0.05$) and calibration curves (Figure 3) revealed that the ratio and MS model had good calibration.

The cutoff value A with the highest Youden index was 0.129 in the ratio model (Figure 2A), and 0.109 in the MS model (Figure 2B). The cutoff value B with fewer

misdiagnoses of malignant calcifications was 0.05 in both the ratio (Figure 4A,4B) and MS models (Figure 4C,4D). The cutoff value B had lower accuracy, specificity and PPV, but higher sensitivity and NPV than the cutoff value A (Table 3). More benign lesions were identified by the cutoff value B of prob-score1 than that of prob-score both in the training and validation cohorts, with only 1 more misclassified malignant lesion in the training cohort. In relation to the malignancy rate of the probably benign groups in the 2 models divided by the cutoff value B (Table 4), the malignancy rate of the probably benign group was lower in the ratio model than the MS model (Table 4). In relation to the 699 cases, the malignancy rate of

Table 2 Accuracy and predictive results of the 2 prediction models

Prediction Performance	Training Cohort (490, 82/408) [†]		Validation Cohort (209, 35/174) [†]	
	MS model	Ratio model	MS model	Ratio model
AUC	0.834	0.875	0.867	0.908
95% CI	0.782, 0.887	0.830, 0.920	0.791, 0.944	0.851, 0.965
Accuracy (%)	76.5	81.4	74.2	81.8
Sensitivity (%)	79.3	84.1	80.0	88.6
Specificity (%)	76.0	80.9	73.0	80.5
PPV (%)	39.9	46.9	37.3	47.7
NPV (%)	96.5	97.8	94.8	97.2

[†], numbers of (total, malignant/benign lesions). MS model, the prediction model based on the factor of MS (maximum span of the group of calcifications); Ratio model, the prediction model based on the factor of ratio (the ratio of the maximum span of the group of calcifications in the parallel direction to the maximum span in the vertical direction of the mammary duct). AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

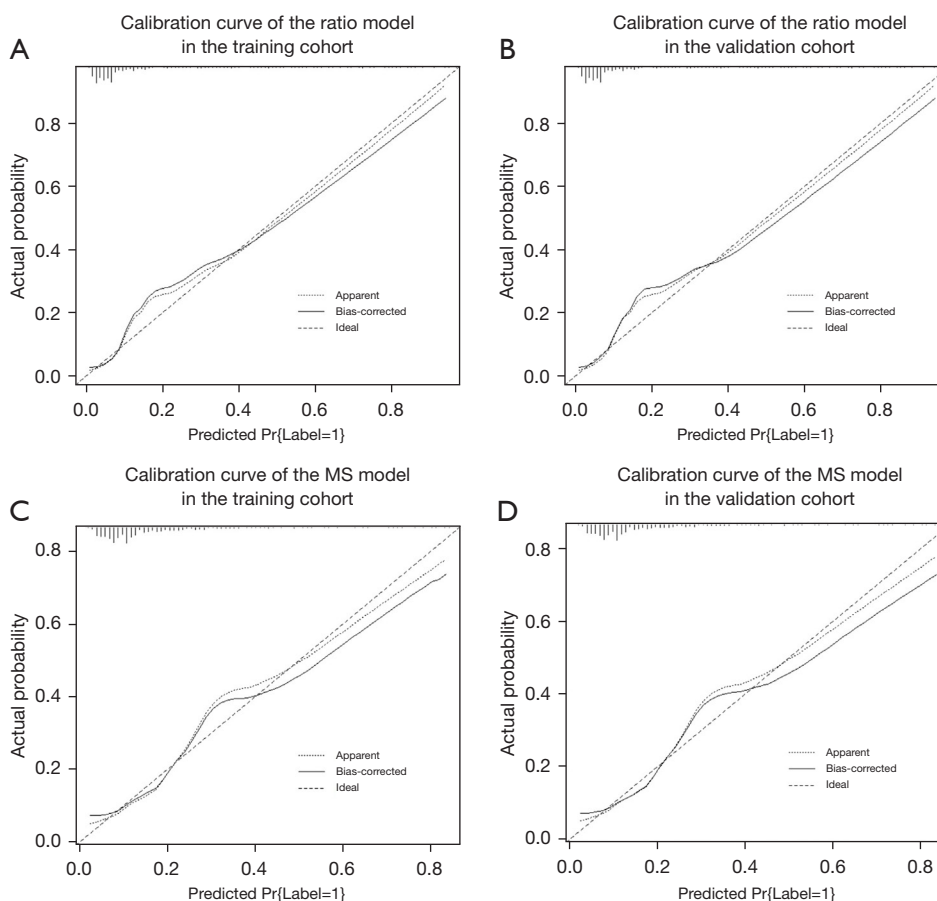


Figure 3 Calibration curves. Calibration curves of the ratio model (A,B) and MS model (C,D) describe the consistency between the predicted results of the malignancy and the actual pathologic results. The 45° straight line indicates the ideal prediction. A smaller distance between the curves and the line indicates the higher prediction accuracy of the nomogram. MS, maximum span of the calcifications.

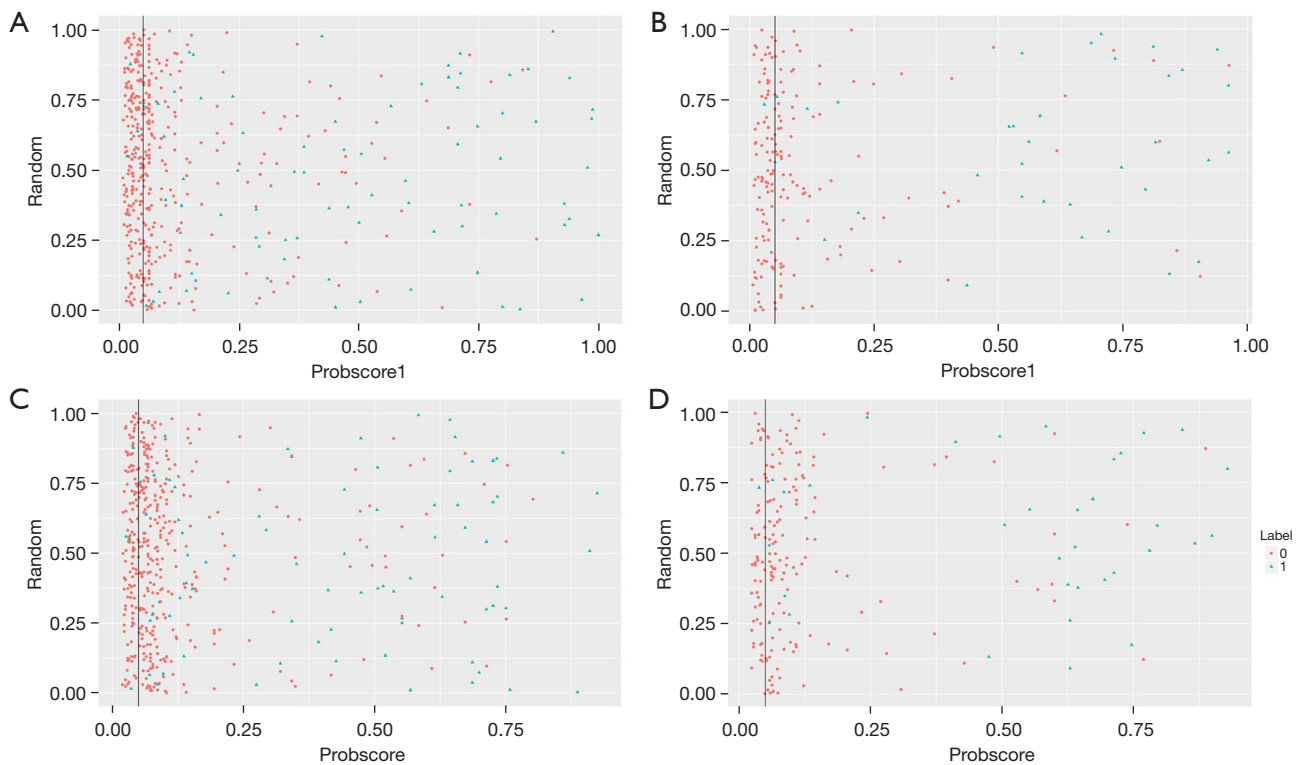


Figure 4 Scatterplots. The red circles represent benign lesions, and the blue triangles represent malignant lesions. The black vertical line indicates the cutoff value B (0.05). The prob-scores below 0.05 were allocated to the probably benign group. More benign lesions were allocated to the benign group in the ratio model (prob-score1) than the MS (maximum span of the calcifications) model (prob-score) in both the training cohort (A,C) and validation cohort (B,D), with only 1 more misclassified malignant lesion in the training cohort.

Table 3 Accuracy and predictive results of the 2 cutoff values of our models

Prediction performance	Training cohort (490, 82/408) [§]				Validation cohort (209, 35/174) [§]			
	Cutoff value A		Cutoff value B		Cutoff value A		Cutoff value B	
	MS	Ratio	MS	Ratio	MS	Ratio	MS	Ratio
Malignancy misdiagnosed [†]	17 (20.7%)	13 (15.9%)	4 (4.9%)	5 (6.1%)	7 (20.0%)	4 (11.4%)	1 (2.9%)	1 (2.9%)
Benign lesions avoided biopsy [‡]	310 (80.9%)	330 (76.0%)	115 (28.2%)	180 (44.1%)	127 (73.0%)	140 (80.5%)	48 (27.5%)	83 (47.7%)
Accuracy (%)	76.5	81.4	39.4	52.4	74.2	81.8	39.2	56.0
Sensitivity (%)	79.3	84.1	95.1	93.9	80.0	88.6	97.1	97.1
Specificity (%)	76.0	80.9	28.2	44.1	73.0	80.5	27.6	47.7
PPV (%)	39.9	46.9	21.0	25.2	37.3	47.7	21.3	27.2
NPV (%)	96.5	97.8	96.6	100.0	94.8	97.2	98.0	98.8

[†], number (percentage of misdiagnosed malignancy in total malignancy); [‡], number (percentage of avoided biopsy benign lesions in total benign lesions); [§], numbers of (total, malignant/benign lesions); MS, the prediction model based on the factor of MS (maximum span of the group of calcifications); Ratio, the prediction model based on the factor of ratio (the ratio of the maximum span of the group of calcifications in the parallel direction to the maximum span in the vertical direction of the mammary duct). PPV, positive predictive value; NPV, negative predictive value.

Table 4 Malignancy rate of the probably benign groups divided by the cutoff value B

Cohort	Model	Total No.	Malignancy		
			No.	Rate (%)	95% CI (%)
Training cohort (490, 82/408) [†]	MS model	119	4	3.36	1.08; 8.90
	Ratio model	185	5	2.70	1.00; 6.53
Validation cohort (209, 35/174) [†]	MS model	49	1	2.04	0.11; 12.24
	Ratio model	84	1	1.19	0.06; 7.37

[†], numbers of (total, malignancy/ benign lesions). MS model: the prediction model based on the factor of MS (maximum span of the group of calcifications); Ratio model: the prediction model based on the factor of ratio (the ratio of the maximum span of the group of calcifications in the parallel direction to the maximum span in the vertical direction of the mammary duct). CI, confidence interval.

the probably benign group divided by the cutoff value B was 2.23% (95% CI: 0.91% to 5.03%) in the ratio model, which was also lower than that in the MS model (2.98%, 95% CI: 1.10% to 7.18%). Additionally, 45.19% (95% CI: 41.11% to 49.34%) of the biopsies of benign lesions in the ratio model were detected (44.12% in the training cohort and 47.70 in the validation cohort), which was more than that of the MS model (28.01%, 95% CI: 24.43% to 31.88%). Of the 6 misdiagnosed malignant cases, 5 were DCIS, and all 6 cases were estrogen receptor/progesterone receptor (ER/PR) positive. Next, the predictive nomogram (*Figure 5*) was successfully plotted based on the ratio model.

Discussion

In this study, we primarily investigated the diagnostic value of the range spans of groups of calcifications as quantitative factors to establish predictive nomograms with improved differentiating ability. We first proposed the concept of the ratio of the MS in the parallel direction to the MS in the vertical direction of the mammary duct of the grouped amorphous calcifications. The ICCs confirmed the good repeatability of the MS and ratio. Additionally, both the MS and ratio were meaningful in the differential diagnosis of benign and malignant calcifications. We successfully developed 2 prediction models based on the MS and ratio, both of which had good discrimination and calibration performance. The ratio model showed better discrimination performance than the MS model, and identified a probably benign group with a malignancy rate <3%.

Numerous studies (4,7-10,14-16,18-20) have shown that each suspicious descriptor and the combined descriptors of the 5th BI-RADS have predictive value in stratifying the risk of malignancy for microcalcifications detected on mammograms. However, few studies have focused on pure

grouped amorphous calcifications, which is a controversial calcification type due to the high false-positive rate of biopsies. To our knowledge, this study had the largest sample size of 699 grouped amorphous calcifications. There is consensus that >90% of biopsies of screening-detected grouped amorphous calcifications reveal benign pathological results, and more nuanced management is needed to reduce the number of biopsies of benign calcifications (4-7).

Farshid *et al.* (21) found that the mammographic extent of microcalcifications >15 mm was associated with malignancy, but there is no recommendation about the range measurement of calcifications in clinical practice. We referred to the recommendations for pulmonary nodule measurements, which suggest using the average of the long- and short-axis spans for assessment as it relates better to tumor volume, and recommend manual measurements as they are convenient for extensive application in clinical practice (22). As the PPV of the segmental distribution descriptor is significantly higher than the grouped distribution (3), we manually measured the MPS, MVS, and MS, and calculated the ratio of the MPS to MVS. We hypothesized that the ratio would be better related to the segmental distribution tendency of the calcifications than the MS, and may be a risk factor for malignancy. Our hypothesis was supported by results showing that the ratio model had better discrimination performance than the MS model. However, further research involving more radiologists and institutions needs to be conducted to confirm this finding and its repeatability.

Age and menopausal status have always been important risk factors for the benign and malignant calcifications of the breast, which have been distinguished in BI-RADS category 4 calcifications (14) and amorphous calcifications (4). In this study, age group and menopausal status were also found to be malignancy risk factors, and elderly women

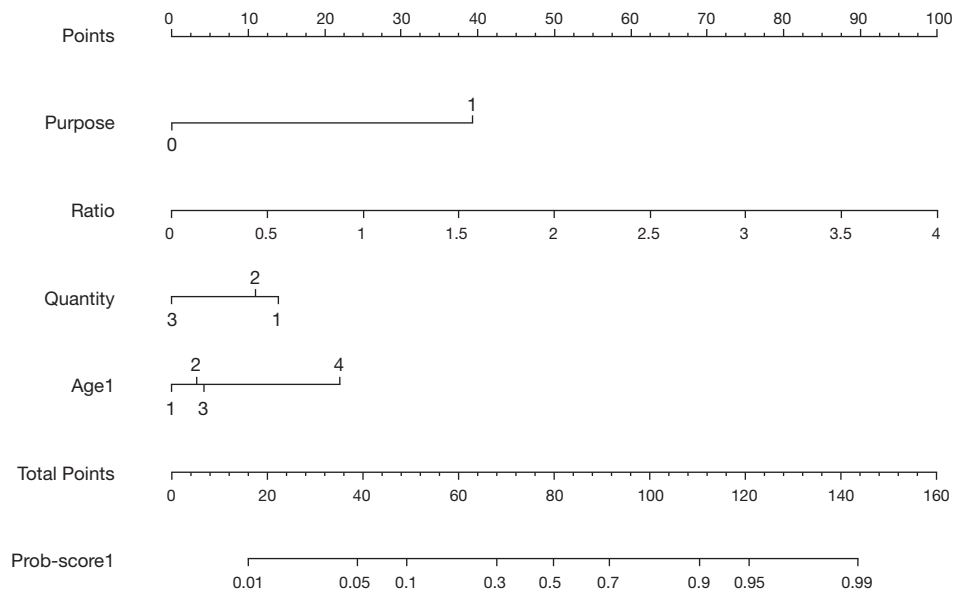


Figure 5 The prediction nomogram based on the ratio model for prob-score1 (the probability of malignancy). Every risk factor corresponds to “Points”, and “Total Points” corresponds to “prob-score1”. In “purpose”, 0 represents the screening mammogram, and 1 represents the diagnostic examination. “Ratio” represents the ratio of the maximum spans in the parallel direction to the vertical direction of the mammary duct of the calcifications. “Quantity” represents the quantity of the groups of calcifications, 1 represents single group of calcifications, 2 represents multiple unilateral groups of calcifications, and 3 represents multiple bilateral groups of calcifications. “Age1” represents the age groups, 1 represents <40 years old, 2 represents ≥ 40 and <50 years old, 3 represents ≥ 50 and <60 years old, and 4 represents ≥ 60 years old.

(aged ≥ 60 years) had a higher malignancy rate than younger women, which is consistent with the literature on amorphous calcifications (4,23).

A nomogram is a visualization statistical tool based on logistic regression that can take multiple risk factors into account, and thus provide more comprehensive predictions (24). Nomograms have been successfully employed to predict lymph nodes metastasis, disease-free survival, and the lymphedema risk of breast cancer patients (25-29). In this study, we developed a user-friendly nomogram based on the ratio model, which displayed excellent discrimination performance in the differential diagnosis of benign and malignant calcifications. This non-invasive method only relied on routine clinical information and simple measurements from mammograms, and thus could be easily used in clinical applications. With the cutoff value A, which corresponded to the highest Youden index, we were able to obtain higher accuracy, specificity, PPV, and more benign lesions avoided biopsy, but we did not think that the percentage of misdiagnosed malignancies was clinically acceptable. Thus, we found the cutoff value B, and the malignancy rate for the detected probably benign group was close to 2%, which is the upper malignancy

likelihood limit for BI-RADS category 3 recommended for follow-up management. We believe that an initial short interval follow-up examination period of 6 months can be recommended. With the cutoff value B, >40% of benign lesions were detected.

Our study had several limitations. First, it was a retrospective single-center study without external validation from other hospitals. Second, we only included suspicious calcifications obtained from pathological results, and did not separately study screening and diagnostic mammograms. This may have introduced bias to the results, as the percentages of screening and diagnostic mammograms are likely to vary greatly among institutions. Thus, prospective multicenter large-scale research studies that focus on screening mammograms including follow-up patients need to be conducted to enable further generalization. Third, the manual assessment of the mammographic features and the measurement of the spans would introduce subjective interpretation among doctors. Future work should seek to include a combination of computer-aided detection and deep learning to improve reproducibility. Finally, the PPV of the probably benign group was slightly higher than 2%, and 6 malignancies were misdiagnosed in our study.

Thus, future work should seek to include a combination of computer-aided detection radiomics and deep learning to improve differential diagnosis performance (30,31).

Conclusions

We developed a clinic-mammographic nomogram for individual quantitative malignancy risk prediction of pure grouped amorphous calcifications, which showed a favorable discrimination performance. The range span measurement of calcifications is meaningful in predicting malignancy probability, and the ratio showed better diagnostic performance than the MS model. With a cutoff value of 0.05, the ratio model-based nomogram may help clinicians to identify grouped amorphous calcifications that are probably to be benign. More than 40% of biopsies could be detected with <3% misdiagnoses of malignancy. Further studies need to be conducted to validate these results.

Acknowledgments

The authors would like to thank Caiyue Ren and Tingdan Hu for their assistance in the data analysis.

Funding: This work was supported by The National Natural Science Foundation of China (No. 81901703 to CY; No. 61731008 to WP), the Shanghai “Rising Stars of Medical Talent” Youth Development Program, the Medical Imaging Practitioner Program SHWRS (No. [2020]087 to CY.), and the Shanghai Municipal Health Planning Commission Youth Project (No. 20184Y0010 to CY).

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-21-797/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-797/coif>). The authors report that this work was supported by The National Natural Science Foundation of China (No. 81901703 to CY; No. 61731008 to WP), the Shanghai “Rising Stars of Medical Talent” Youth Development Program, the Medical Imaging Practitioner Program SHWRS (No. [2020]087 to CY), and the Shanghai Municipal Health Planning Commission Youth Project (No. 20184Y0010 to CY). The authors have no other conflicts of

interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work, including ensuring that any questions related to the accuracy or integrity of any part of the work have been 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 Institutional Review Board of Fudan University Shanghai Cancer Center, and individual consent for this retrospective analysis was waived.

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. de Roos MA, van der Vegt B, de Vries J, Wesseling J, de Bock GH. Pathological and biological differences between screen-detected and interval ductal carcinoma in situ of the breast. *Ann Surg Oncol* 2007;14:2097-104.
2. Mordang JJ, Gubern-Mérida A, Bria A, Tortorella F, Mann RM, Broeders MJM, den Heeten GJ, Karssemeijer N. The importance of early detection of calcifications associated with breast cancer in screening. *Breast Cancer Res Treat* 2018;167:451-8.
3. American College of Radiology. *Breast Imaging Reporting And Data System*, 5th edn. Reston, VA: American College of Radiology, 2013.
4. Oligane HC, Berg WA, Bandos AI, Chen SS, Sohrabi S, Anello M, Zuley ML. Grouped Amorphous Calcifications at Mammography: Frequently Atypical but Rarely Associated with Aggressive Malignancy. *Radiology* 2018;288:671-9.
5. Kim SY, Kim HY, Kim EK, Kim MJ, Moon HJ, Yoon JH. Evaluation of malignancy risk stratification of microcalcifications detected on mammography: a study based on the 5th edition of BI-RADS. *Ann Surg Oncol* 2015;22:2895-901.
6. Iwase M, Tsunoda H, Nakayama K, Morishita E, Hayashi N, Suzuki K, Yamauchi H. Overcalling low-risk findings:

- grouped amorphous calcifications found at screening mammography associated with minimal cancer risk. *Breast Cancer* 2017;24:579-84.
7. Moy L. Should We Continue to Biopsy All Amorphous Calcifications? *Radiology* 2018;288:680-1.
 8. Burnside ES, Ochsner JE, Fowler KJ, Fine JP, Salkowski LR, Rubin DL, Sisney GA. Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy. *Radiology* 2007;242:388-95.
 9. Kaltenbach B, Brandebusch V, Möbus V, Mall G, Falk S, van den Bergh M, Chevalier F, Müller-Schimpfle M. A matrix of morphology and distribution of calcifications in the breast: Analysis of 849 vacuum-assisted biopsies. *Eur J Radiol* 2017;86:221-6.
 10. Kim J, Kim EK, Kim MJ, Moon HJ, Yoon JH. "Category 4A" microcalcifications: how should this subcategory be applied to microcalcifications seen on mammography? *Acta Radiol* 2018;59:147-53.
 11. Tina Shih YC, Dong W, Xu Y, Shen Y. Assessing the Cost-Effectiveness of Updated Breast Cancer Screening Guidelines for Average-Risk Women. *Value Health* 2019;22:185-93.
 12. McPherson K, Al Waheidi S. The overestimation and the inappropriate promotion of the benefits of mammographic screening in breast cancer research and interventions in the Gaza Strip. *Lancet* 2018;391 Suppl 2:S21.
 13. Oeffinger KC, Fontham ET, Etzioni R, Herzig A, Michaelson JS, Shih YC, Walter LC, Church TR, Flowers CR, LaMonte SJ, Wolf AM, DeSantis C, Lortet-Tieulent J, Andrews K, Manassaram-Baptiste D, Saslow D, Smith RA, Brawley OW, Wender R; American Cancer Society. Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. *JAMA* 2015;314:1599-614.
 14. Lei C, Wei W, Liu Z, Xiong Q, Yang C, Yang M, Zhang L, Zhu T, Zhuang X, Liu C, Liu Z, Tian J, Wang K. Mammography-based radiomic analysis for predicting benign BI-RADS category 4 calcifications. *Eur J Radiol* 2019;121:108711.
 15. Berg WA, Arnoldus CL, Teferra E, Bhargavan M. Biopsy of amorphous breast calcifications: pathologic outcome and yield at stereotactic biopsy. *Radiology* 2001;221:495-503.
 16. Shen L, Ma X, Jiang T, Shen X, Yang W, You C, Peng W. Malignancy Risk Stratification Prediction of Amorphous Calcifications Based on Clinical and Mammographic Features. *Cancer Manag Res* 2021;13:235-45.
 17. Love SM, Barsky SH. Anatomy of the nipple and breast ducts revisited. *Cancer* 2004;101:1947-57.
 18. Shin HJ, Kim HH, Ko MS, Kim HJ, Moon JH, Son BH, Ahn SH. BI-RADS descriptors for mammographically detected microcalcifications verified by histopathology after needle-localized open breast biopsy. *AJR Am J Roentgenol* 2010;195:1466-71.
 19. Mazari FAK, Sharma N, Dodwell D, Horgan K. Human Epidermal Growth Factor 2-positive Breast Cancer with Mammographic Microcalcification: Relationship to Pathologic Complete Response after Neoadjuvant Chemotherapy. *Radiology* 2018;288:366-74.
 20. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ; National Evaluation Team for Breast Cancer Screening in Netherlands Study Group (NETB). Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. *Eur J Cancer* 2013;49:3517-25.
 21. Farshid G, Sullivan T, Downey P, Gill PG, Pieterse S. Independent predictors of breast malignancy in screen-detected microcalcifications: biopsy results in 2545 cases. *Br J Cancer* 2011;105:1669-75.
 22. Bankier AA, MacMahon H, Goo JM, Rubin GD, Schaefer-Prokop CM, Naidich DP. Recommendations for Measuring Pulmonary Nodules at CT: A Statement from the Fleischner Society. *Radiology* 2017;285:584-600.
 23. Ferreira VCCS, Etchebehere ECSC, Bevilacqua JLB, de Barros N. Suspicious amorphous microcalcifications detected on full-field digital mammography: correlation with histopathology. *Radiol Bras* 2018;51:87-94.
 24. Kattan MW. Judging new markers by their ability to improve predictive accuracy. *J Natl Cancer Inst* 2003;95:634-5.
 25. Sadeghi M, Alamdaran SA, Daneshpajouhnejad P, Layegh P, Afzalaghaee M, Boroumand F, Shafaat O, Hashemi SK. A logistic regression nomogram to predict axillary lymph node metastasis in early invasive breast cancer patients. *Breast J* 2019;25:769-71.
 26. Gooch JC, Schnabel F, Chun J, Pirraglia E, Troxel AB, Guth A, Shapiro R, Axelrod D, Roses D. A Nomogram to Predict Factors Associated with Lymph Node Metastasis in Ductal Carcinoma In Situ with Microinvasion. *Ann Surg Oncol* 2019;26:4302-9.
 27. Pan X, Yang W, Chen Y, Tong L, Li C, Li H. Nomogram for predicting the overall survival of patients with inflammatory breast cancer: A SEER-based study. *Breast* 2019;47:56-61.
 28. Xie X, Xiong Z, Li X, Huang X, Ye F, Tang H, Xie X. Nomogram to Predict Internal Mammary Lymph Nodes

- Metastasis in Patients With Breast Cancer. *Front Oncol* 2019;9:1193.
29. Han L, Zhu Y, Liu Z, Yu T, He C, Jiang W, Kan Y, Dong D, Tian J, Luo Y. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol* 2019;29:3820-9.
30. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. *Radiology* 2019;292:60-6.
31. Martin-Noguerol T, Luna A. External validation of AI algorithms in breast radiology: the last healthcare security checkpoint? *Quant Imaging Med Surg* 2021;11:2888-92.

Cite this article as: Shen L, Jiang T, Tang P, Ge H, You C, Peng W. Comprehensive quantitative malignant risk prediction of pure grouped amorphous calcifications: clinico-mammographic nomogram. *Quant Imaging Med Surg* 2022;12(5):2672-2683. doi: 10.21037/qims-21-797

Table S1 Detailed pathological results

Pathological result	Total cohort % (number, number of total)	Training cohort % (number, number of total)	Validation cohort % (number, number of total)
Malignant lesions	16.7 (117/699)	16.7 (82/490)	16.7 (35/209)
DCIS	44.4 (52/117)	42.6 (35/82)	48.6 (17/35)
IDC	50.4 (59/117)	50.0 (41/82)	51.4 (18/35)
Also had DCIS	49.2 (29/59)	46.3 (19/41)	55.6 (10/18)
Others	5.1 (6/117)	7.3 (6/82)	0.0 (0/35)
Benign lesions	83.3 (582/699)	83.3 (408/490)	83.3 (174/209)
Breast hyperplasia	59.5 (346/582)	60.3 (246/408)	57.5 (100/174)
Breast hyperplasia with fibroadenoma forming tendency	15.8 (92/582)	16.4 (67/408)	14.4 (25/174)
High risk lesions	15.8 (92/582)	14.2 (58/408)	19.5 (34/174)
Fibroadenoma	7.0 (41/582)	7.6 (31/408)	5.7 (10/174)
Others	1.9 (11/582)	1.5 (6/408)	2.9 (5/174)

% (number, number of total). DCIS, breast ductal carcinoma in situ. IDC, invasive ductal carcinoma. Other malignant lesions included Invasive lobular carcinoma, mucinous breast carcinoma and solid papillary intraductal carcinoma. High risk lesions included atypical ductal hyperplasia, intraductal papillomatosis, atypical lobular hyperplasia and radial scar. Other benign lesion included calcium salt deposition and inflammation.

Table S2 Comparison of malignancy rate in different subgroups

		Total No.	Malignancy No.	Malignancy rate % (95% CI)
BI-RADS 4A		564	48	8.51 (6.40, 11.20)
BI-RADS 4B		135	69	51.11 (42.40, 59.75)
Purpose of mammogram	Screening	575	45	7.83 (5.82, 10.41)
	Diagnostic exam	124	72	58.06 (48.86, 66.76)
Quantity	Single	372	86	23.12 (19.00, 27.81)
	Multiple unilateral	112	15	13.39 (7.94, 21.44)
	Multiple bilateral	215	16	7.44 (4, 45, 12.02)
Age groups	1. <40	143	26	18.18 (12.43, 25.69)
	2. ≥40 and <50	340	44	12.94 (9.75, 17.09)
	3. ≥50 and <60	161	29	18.01 (12.58, 25.01)
	4. ≥60	55	17	30.91 (19.52, 44.97)
Menopausal Status	Premenopausal	523	72	13.77 (10.99, 17.09)
	Postmenopausal	176	45	25.57 (19.44, 32.79)

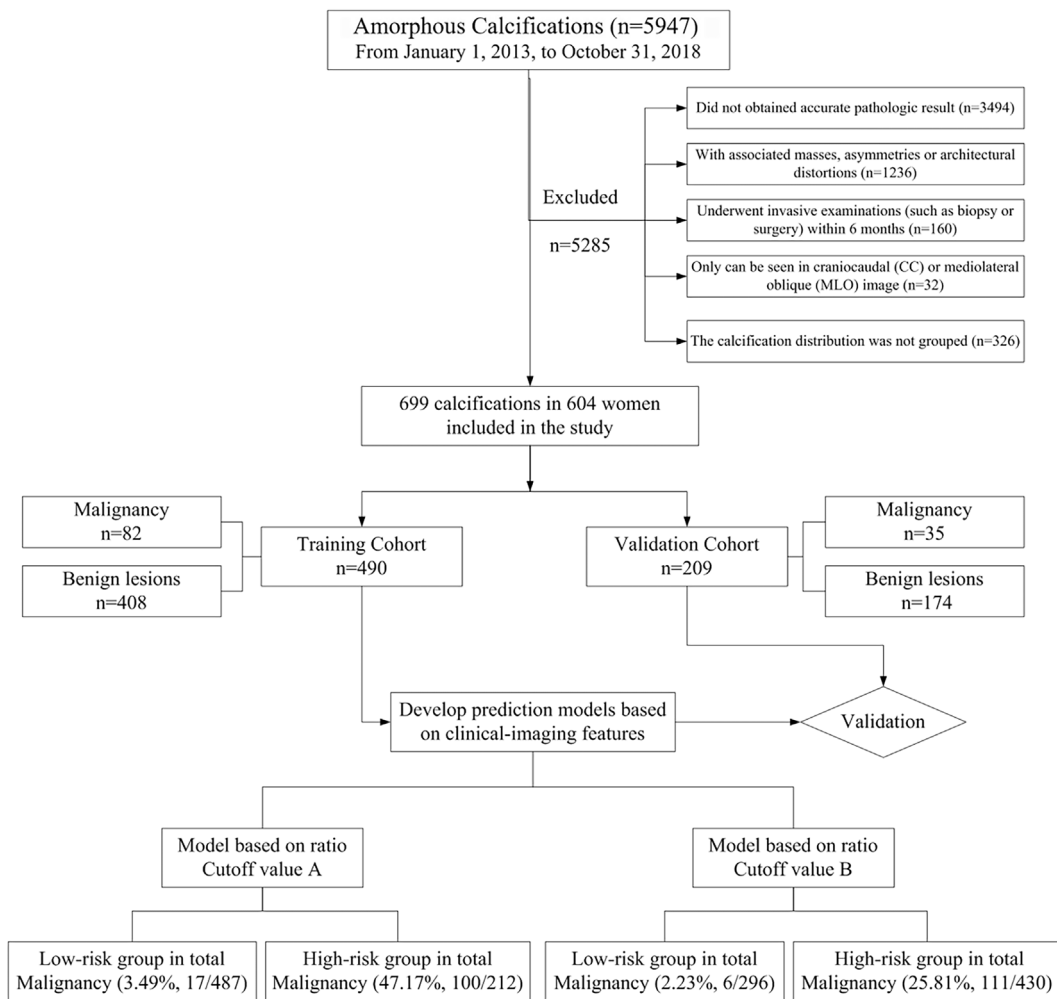


Figure S1 The flow of patients.