

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)

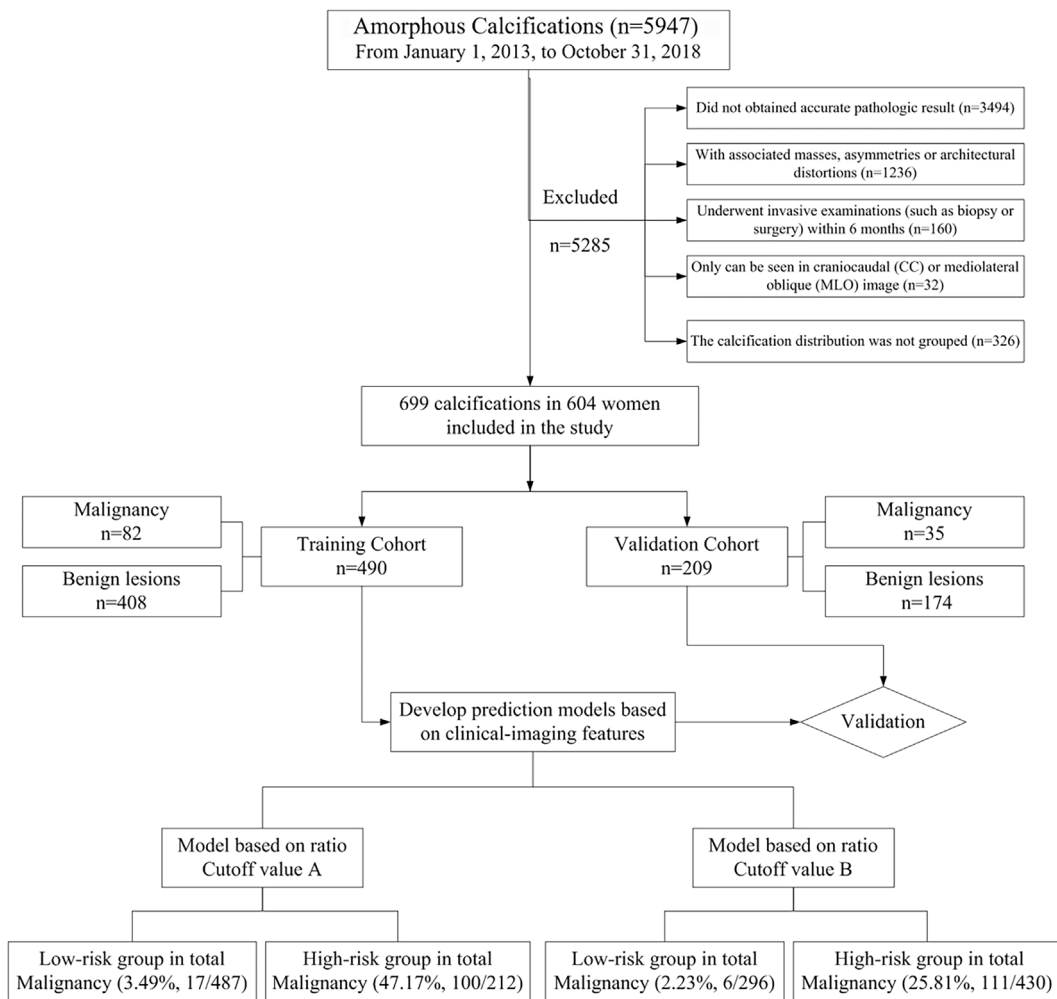


Figure S1 The flow of patients.