Original Article

Ultrasound-based prediction of preoperative core biopsy categories in solid breast tumor using machine learning

Ting Liang[1,2#], Junhui Shen[3#], Jiexin Wang[4#], Weilin Liao[5], Zhi Zhang[5], Juanjuan Liu[2], Zhanwu Feng[2], Shufang Pei[2], Kebing Liu[1]

[1]Department of Ultrasound, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, China; [2]Department of Ultrasound, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China; [3]Department of Rehabilitation Medicine, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China; [4]Department of Ultrasound, Affiliated Hospital of Guangdong Medical University, Zhanjiang, China; [5]School of Geography and Planning, Sun Yat-sen University, Guangzhou, China

Contributions: (I) Conception and design: T Liang, J Shen; (II) Administrative support: K Liu; (III) Provision of study materials or patients: S Pei; (IV) Collection and assembly of data: J Liu, Z Feng; (V) Data analysis and interpretation: W Liao, Z Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Kebing Liu, MD. Department of Ultrasound, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine, No. 16 Jichang Road, Guangzhou 510405, China. Email: liukebing@vip.163.com; Shufang Pei, MD. Department of Ultrasound, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, No. 106 Zhongshan Er Road, Guangzhou 510080, China. Email: peishufang2008@163.com.

Background: The American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) used with ultrasonography cannot guide the individual management of solid breast tumors, but preoperative core biopsy categories (CBCs) can. We aimed to use machine learning to analyze clinical and ultrasonic features for predicting CBCs and to aid in the development of a new ultrasound (US) imaging reporting system for solid tumors of the breast.

Methods: This retrospective study included women with solid breast tumors who underwent US-guided core needle biopsy from March 1, 2019, to December 31, 2019. All patients were randomly assigned to a training or validation cohort (7:3 ratio). CBC was predicted using 5 machine learning models: random forest (RF), support vector machine (SVM), k-nearest-neighbor (KNN), multilayer perceptron (MLP), and ridge regression (RR). In the validation cohort, the area under the curve (AUC) and accuracy were ascertained for every algorithm. Based on AUC values, the optimal algorithm was determined, and the features' importance was depicted.

Results: A total of 1,082 female patients were included (age range, 12–96 years; mean age ± standard deviation, 42.22±13.37 years). The proportion of the 4 CBCs was 4% (44/1,185) for the B1 group, 60% (714/1,185) for the B2 group, 5% (57/1,185) for the B3 group, and 31% (370/1,185) for the B5 group. In the validation cohort, AUCs of the optimal algorithm constructed RF were 0.78, 0.88, 0.64, and 0.92 for B1, B2, B3, and B5, respectively, with an accuracy of 0.82.

Conclusions: Machine learning could strongly predict CBC, particularly in B2 and B5 categories of solid breast tumors, with RF being the optimal machine learning model.

Keywords: Ultrasound (US); core biopsy categories (CBCs); solid breast tumor; machine learning; prediction

Submitted Aug 21, 2022. Accepted for publication Feb 23, 2023. Published online Mar 03, 2023.
doi: 10.21037/qims-22-877
View this article at: https://dx.doi.org/10.21037/qims-22-877

Introduction

Breast tumors are common across the globe and lower the quality of life of adult women. Histologically, breast tumors are classified as benign, malignant, or borderline. Each breast lesion can be treated using different methods, including such as biopsy, operation, and imaging follow-up (1,2).

Imaging follow-ups for breast tumors are the most common method and include magnetic resonance imaging (MRI), ultrasound (US), and mammography. These are considered important techniques for assessing breast lesions and preliminarily evaluating their pathological properties. However, these methods are not without disadvantages; mammography is not good at discovering a nodule in dense breast tissue (3), and MRI often has a high false-positive rate for malignant tumors and a high cost of operation (4). In contrast, US does not use radiation, costs less, and is widely available. Therefore, US is the preferred imaging method for evaluating breast nodules in Asian countries, where women tend to have denser breast tissue (5).

The American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) is the most commonly use classification tool in US. This system can predict the probability of malignancy and is suitable only for screening, nut it is not suitable for precise individual management for radiologists for two reasons. First, radiologists currently pay too much attention to probability in providing the BI-RADS category and often ignore the specific pathological type (e.g., mucinous cancer, sclerosing adenosis, or fibroadenoma). To some extent, the radiologists' diagnostic accuracy and understanding of the specific pathological type will inevitably decrease in the future. Second, lesions treated with neoadjuvant therapy are classified into BI-RADS category 6 for the entire treatment period. At the moment, BI-RADS category is not capable of reflecting dynamic changes. Therefore, radiologists must master the diagnosis of specific pathological types, which can sometimes provide the necessary guidance for individualized clinical management.

There are many types of solid tumors of the breast, and it is very difficult for radiologists to be proficient in the detailed characteristics of each type of tumor. Core biopsy categories (CBCs) help to alleviate this difficulty. According to core biopsy reporting guidelines, all breast histological types can be classified into 5 categories (B1–B5) (6). CBCs can become the basis for building a new radiological reporting system to help radiologists master pathological types. Thus, machine learning should be applied to conduct CBC multiclassification.

Machine learning with many algorithms can automatically analyze and obtain rules from existing data and then use the rules to predict unknown data. Machine learning can also complete classification tasks, including in the areas of radiology, critical care medicine, and cardiology (7-9). Recent studies have shown that machine learning can be beneficial to the diagnosis and treatment of cancer, including imaging examinations for evaluating breast cancer (9-12). Sutton *et al.* indicated that a machine learning classifier combining MRI radiomics with molecular subtypes could accurately assess pathologic complete response after post-neoadjuvant chemotherapy for breast cancer (13). Wu *et al.* reported that machine learning achieved a high level of differentiation between triple-negative and non-triple-negative subtypes of breast cancer (14). We aimed to use machine learning to analyze the clinical and ultrasonic features for predicting CBCs and to aid in the development of a new US imaging reporting system for solid tumors of the breast. We present the following article in accordance with the TRIPOD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-22-877/rc).

Methods

Participants

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board of Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences. The informed consent requirement was waived due to the retrospective nature of the study. We acquired details of the pathological characteristics of breast nodules from histological reports. From March 1, 2019, to December 31, 2019, 1,082 female patients were included in the study (age range, 12–96 years; mean age ± standard deviation, 42.22±13.37 years). A total of 1,185 nodules

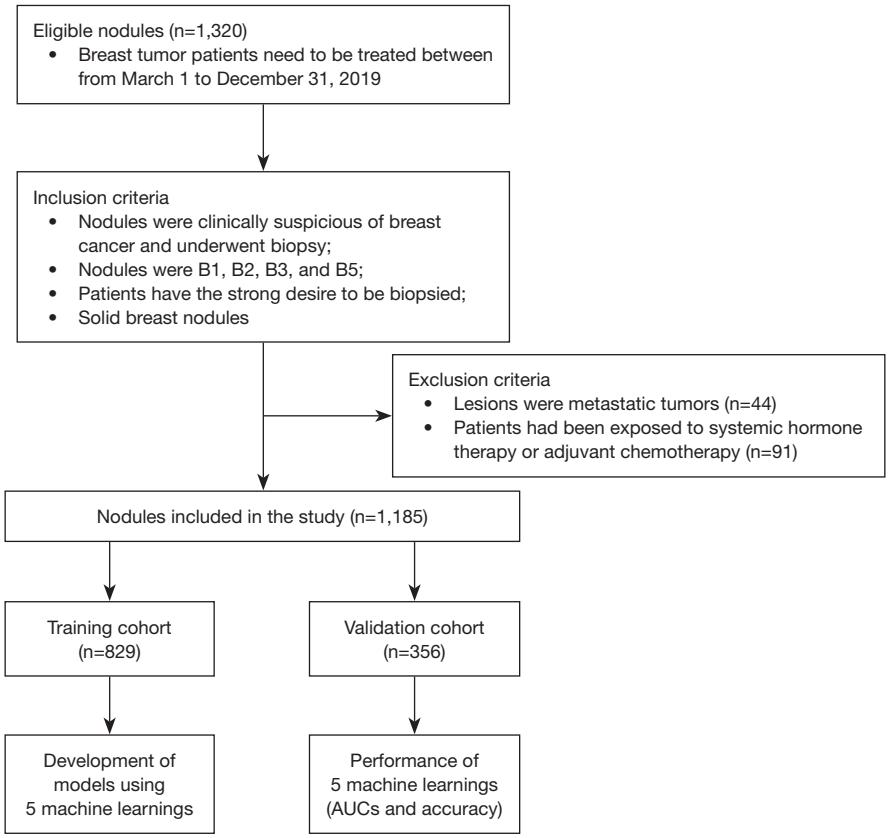


Figure 1 Flowchart of participant inclusion and exclusion. AUC, area under the curve.

(815 benign nodules and 370 malignant nodules) met the inclusion and exclusion criteria. All patients underwent US-guided core needle biopsies. Nodules were marked on the skin, and ultrasonic images were obtained in the electronic system.

The inclusion criteria were as follows: (I) nodules were clinically suspicious for breast cancer and underwent biopsy; (II) nodules were B1, B2, B3, or B5; (III) patients expressed a strong desire to be biopsied; and (IV) solid breast nodules were defined as having little or no anechoic components based on US images. The exclusion criteria were as follows: (I) lesions were metastatic tumors; (II) patients had undergone systemic hormone therapy or adjuvant chemotherapy; and (III) nodules were B4. (The B4 category was excluded, as it cannot provide a definite diagnosis or valuable information for clinical practice, and further processing is required to determine its pathological type.)

The workflow is shown in *Figure 1*.

Clinical characteristic acquisition, ultrasonic image acquisition, and interpretation

Clinical characteristics included age, height, weight, and body mass index (BMI). The images were obtained using a 14-MHz linear transducer (Toshiba Aplio 500, Toshiba, Tokyo, Japan). The nodules' images were acquired in a standard manner. All images included at least 2 orthogonal planes (radial and antiradial planes or transverse and longitudinal planes). According to the ACR BI-RADS fifth edition classification criteria and a previous study (15), all images were analyzed retrospectively by two breast radiologists (reader 1 with 10 years' experience and reader 2 with 5 years' experience). The radiologists strictly recorded 15 ultrasonic features, as follows: orientation, shape, echogenic pattern, margin, posterior features, calcifications, vascularity grade, vascularity distribution, background echotexture of parenchyma (BEP), anteroposterior thickness

of breast parenchyma (TBP), anteroposterior thickness ratio of breast parenchyma to tissue before pectoralis fascia (RPT), anteroposterior thickness ratio of breast parenchyma to mammary fat (RPF), lymph node metastasis, tumor size, and BI-RADS category. RPT and RPF were the adjusted parameters of TBP. They were acquired after TBP was corrected via thickness of tissue, pectoralis fascia, and fat. The detailed interpretations of all ultrasonic features are listed in Appendix 1. Two radiologists were blinded to the pathological characteristic but not to patient age. Intra- and interobserver agreements of 15 ultrasonic features were evaluated. For the records of each ultrasonic feature, any disagreements between the two radiologists were resolved by final consensus following discussion.

Core biopsy reporting categories (6)

According to histological examination, all lesions were classified into 5 categories based on clinical and ultrasonic characteristics:

(I) B1 lesion indicates normal tissue.
(II) B2 indicates benign lesion, including fibroadenomas, fibrocystic change, sclerosing adenosis, and duct ectasia as well as other nonparenchymal lesions such as abscesses and fat necrosis.
(III) B3 lesion indicates uncertain malignant potential, and this category mainly consists of lesions that may provide benign histology on core biopsy but either are known to show heterogeneity or to be associated with an increased risk of malignancy. The category includes atypical intraductal epithelial proliferations, flat epithelial atypia, lobular neoplasia, phyllodes tumors, papillary lesions, radial scars, mucocele-like lesions, and rare lesions.
(IV) B4 indicates suspicious nodules (technical problems may have led to crushed or poorly fixed cores that contain probable carcinoma that cannot provide a definitive diagnosis). Thus, B4 lesions were excluded in this study.
(V) B5 indicates malignant nodules.

Statistical analysis

Statistical analysis was performed with SPSS version 22.0 (IBM Corp., Armonk, NY, USA). The statistical significance levels were two-sided, and a P value <0.05 was deemed to be statistically significant.

Multiple comparisons of the B1, B2, B3, and B5 categories

Multiple differences in all characteristics were assessed among the 4 categories. Continuous variables were compared using the least significant difference, whereas categorical variables were compared using the Bonferroni test.

Machine learning in characteristics analysis

RStudio version 1.1.463 (RStudio, The R Foundation for Statistical Computing, Vienna, Austria) was used as R software. Because there are many machine learning models, we chose the more common ones to complete the classification. The machine learning models included were random forest (RF), support vector machine (SVM), k-nearest-neighbor (KNN), multilayer perceptron (MLP), and ridge regression (RR). In operations, all cases were randomly split into training and validation cohorts based on the machine learning model. We achieved normalization for each feature in the data preprocessing steps. For the hyperparameters, we applied a simple tuning process instead of a detailed one since this study did not focus on the optimal hyperparameters. The hyperparameters used for each classifier are shown in Appendix 2.

Performance of machine learning

The maps of distribution were depicted for variables. The table of the validation cohort was used to predict the biopsied categories, including area under the curve (AUC) and accuracy. The receiver operating characteristic (ROC) was provided for every algorithm. According to the highest AUC, we determined the optimal algorithm, whose weight map was provided to show the importance of ultrasonic and clinical features.

Results

Clinical and ultrasonic characteristics

In our study, the proportions of the 4 categories were 4% (44/1,185) for the B1 group, 60% (714/1,185) for the B2 group, 5% (57/1,185) for the B3 group, and 31% (370/1,185) for the B5 group. The baseline clinical and ultrasonic characteristics are listed in *Table 1*. There were statistically significant differences in 17 features (P<0.05) but not for echo pattern or BEP (P>0.05). The horizontal

Table 1 Baseline clinical and ultrasonic characteristics

Features	B1 (n=44)	B2 (n=714)	B3 (n=57)	B5 (n=370)	P
Age (years), mean ± SD	47.55±11.46	36.82±11.34	42.45±8.73	52.02±11.92	<0.001
Height (cm), mean ± SD	158.68±4.67	158.93±4.82	159.02±5.26	157.67±5.07	0.001
Weight (kg), mean ± SD	55.71±8.08	53.96±7.68	56.34±7.91	57.89±8.14	<0.001
BMI (kg/m^2), mean ± SD	22.12±2.98	21.37±2.95	22.31±3.25	23.28±3.09	<0.001
Echo pattern, n [%]					0.050
Hyperechoic	0 [0]	0 [0]	0 [0]	0 [0]	
Complex cystic and solid	0 [0]	5 [1]	1 [2]	2 [1]	
Hypoechoic	27 [61]	332 [46]	19 [33]	152 [41]	
Isoechoic	0 [0]	14 [2]	3 [5]	11 [3]	
Heterogeneous	17 [39]	363 [51]	34 [60]	205 [55]	
Shape, n [%]					<0.001
Oval	17 [39]	182 [25]	12 [21]	18 [5]	
Round	1 [2]	11 [2]	0 [0]	2 [1]	
Irregular	26 [59]	521 [73]	45 [79]	350 [95]	
Margin, n [%]					<0.001
Circumscribed	11 [25]	222 [31]	11 [19]	17 [5]	
Indistinct	17 [39]	115 [16]	14 [25]	47 [13]	
Angular	14 [32]	353 [49]	27 [47]	167 [45]	
Microlobulated	2 [5]	24 [3]	5 [9]	139 [38]	
Orientation, n [%]					<0.001
Parallel	37 [84]	648 [91]	47 [82]	267 [72]	
Not parallel	7 [16]	66 [9]	10 [18]	103 [28]	
Posterior feature, n [%]					<0.001
No posterior feature	15 [34]	85 [12]	9 [16]	20 [5]	
Enhancement sound	4 [9]	75 [11]	8 [14]	21 [6]	
Shadowing	13 [30]	161 [23]	10 [18]	63 [17]	
Combined pattern	12 [27]	393 [55]	30 [53]	266 [72]	
Calcification, n [%]					<0.001
Within a mass	4 [9]	72 [10]	6 [11]	119 [32]	
Outside of a mass	0 [0]	3 [0]	0 [0]	0 [0]	
Intraductal calcifications	0 [0]	1 [0]	0 [0]	1 [0.3]	
None	40 [91]	638 [89]	51 [89]	250 [68]	
Vascularity distribution, n [%]					<0.001
Absent	31 [70]	260 [36]	15 [26]	44 [12]	
Vessels in rim	6 [14]	128 [18]	15 [26]	35 [10]	
Internal	7 [16]	326 [46]	27 [48]	291 [78]	

Table 1 *(continued)*

Table 1 *(continued)*

Features	B1 (n=44)	B2 (n=714)	B3 (n=57)	B5 (n=370)	P
Vascularity grade, n [%]					<0.001
Grade I	31 [70]	260 [36]	14 [25]	45 [12]	
Grade II	8 [18]	271 [38]	28 [49]	107 [29]	
Grade III	4 [9]	126 [18]	13 [23]	140 [38]	
Grade IV	1 [2]	57 [8]	2 [4]	78 [21]	
Lymph node metastasis reported, n [%]					<0.001
No	43 [98]	689 [96]	55 [96]	255 [69]	
Yes	1 [2]	25 [4]	2 [4]	115 [31]	
BEP, n [%]					0.281
Homogenous	10 [23]	148 [21]	11[19]	60 [16]	
Inhomogeneous	34 [77]	566 [79]	46 [81]	310 [84]	
BI-RADS category, n [%]					<0.001
3	2 [5]	36 [5]	3 [5]	1 [0]	
4A	15 [34]	135 [19]	10 [18]	7 [2]	
4B	15 [34]	335 [47]	12 [21]	37 [10]	
4C	11 [25]	166 [23]	23 [40]	96 [26]	
5	1 [2]	42 [6]	9 [16]	229 [62]	
Tumor size (mm), mean ± SD	11.76±6.50	17.91±9.39	18.03±7.90	24.46±11.66	<0.001
TBP (mm), mean ± SD	7.73±3.21	8.12±2.93	7.75±2.49	9.06±4.47	<0.001
RPT, mean ± SD	0.51±0.14	0.52±0.14	0.52±0.13	0.48±0.21	0.004
RPF, mean ± SD	1.96±1.97	1.93±1.75	1.86±1.23	1.53±1.72	0.004

SD, standard deviation; BMI, body mass index; BEP, background echotexture of parenchyma; BI-RADS, Breast Imaging Reporting and Data System; TBP, anteroposterior thickness of breast parenchyma; RPT, thickness ratio of breast parenchyma to tissue before pectoralis fascia; RPF, thickness ratio of breast parenchyma to mammary fat.

bar graphs of categorical variables are shown in *Figure 2*, the histograms of continuous variables are shown in *Figure 3*, and multiple comparisons of clinical and ultrasonic characteristics are shown in Appendix 3.

Performance of machine learning

All nodules were split into a training cohort (829 nodules) and a validation cohort (356 nodules). In the validation cohort, the ROCs of 5 algorithms are plotted in *Figure 4*. The prediction results, AUC, and accuracy of the 5 algorithms are listed in *Table 2*.

According to the above results, we found that RF was the best algorithm for predicting CBCs. The AUCs were 0.78, 0.88, 0.64, and 0.92, for B1, B2, B3, and B5, respectively. In summary, RF showed the best predictive ability with an accuracy of 0.82. Since each rank variable was treated as an independent variable in the categorical variables, up to 45 features were included in RF. Because the top 10 variables play a major role in building the model, the weight map of these variables is shown in *Figure 5*.

Discussion

Our study demonstrated relationships between several clinical and almost all US features with biopsy categories of solid tumors of the breast. Based on clinical and US features, machine learning models were used to predict

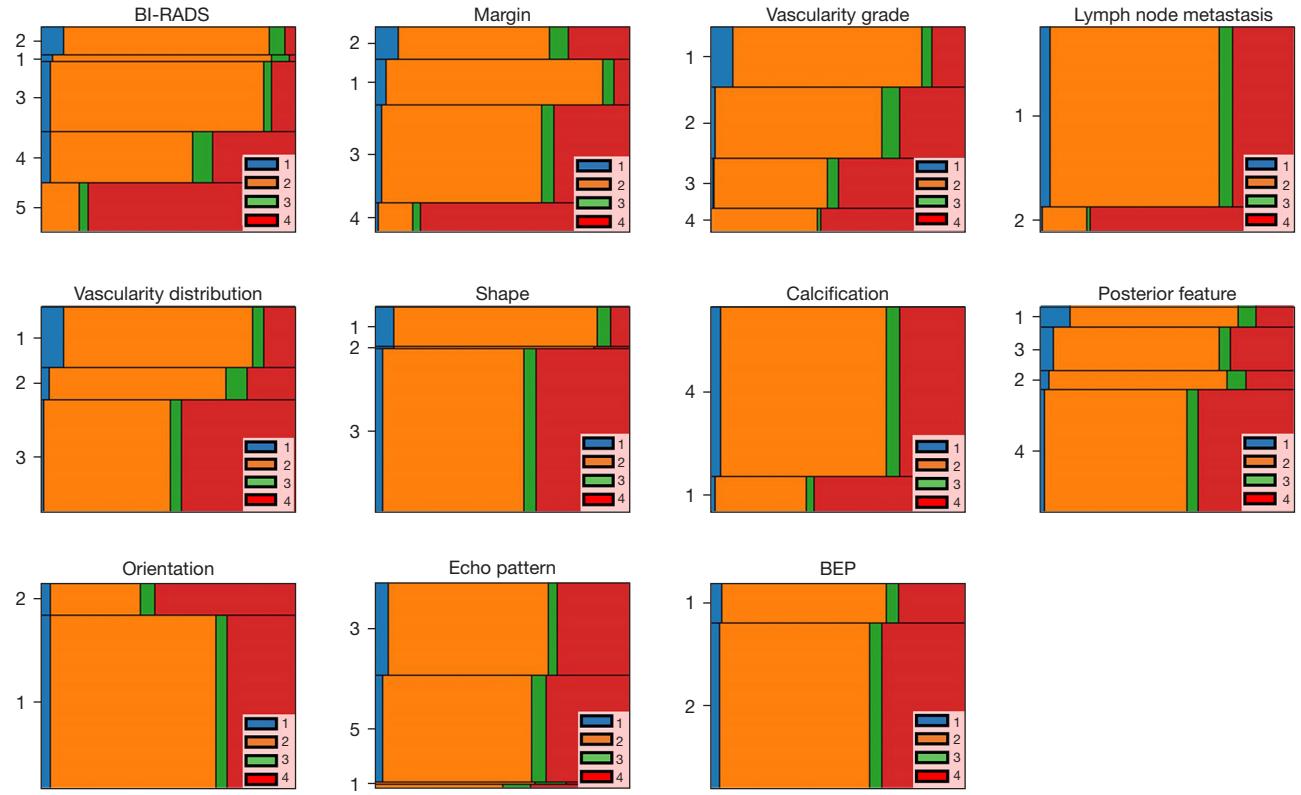


Figure 2 Horizontal bar graphs of the categorical variable. The ordinates represent the assignments of the subcategory variable, and the height of each assignment represents its proportion in all cases. The abscissa represents the preoperative biopsy categories, and different colours indicate different categories. Blue represents the B1 category, orange the B2 category, green the B3 category, and red the B5 category. BI-RADS, Breast Imaging Reporting and Data System; BEP, background echotexture of parenchyma.

CBC. This is in contrast to a previous study, in which MRI features of B2, B3, and B5 lesions were analyzed in only 61 cases and focused solely on the distribution of MRI features (16). Another study investigated several clinical and US features of 102 B3 lesions but did not analyze in detail the relationship between B3 lesions and each feature of the BI-RADS lexicon (17). To the best of our knowledge, there are no previous studies that have used machine learning models to predict CBC based on US features.

Ultrasonic terminologies of BI-RADS are used to depict the morphology and function of breast tumors. However, each feature can exist in both benign and malignant tumors. Hypoecho, calcification, and abundant vascularity have been reported in benign and malignant tumors (18-20); a similar phenomenon was found in our study (*Table 1*). Benign and malignant lesions cannot be accurately predicted with a single US feature or clinical feature. Therefore, a model that assesses the characteristics of breast lesions is needed.

Machine learning is beneficial for building a predictive model and plays an important role in radiological studies (9,21). In our study, we used 5 machine learning models to analyze US and clinical features and to conduct multiclassification for predicting CBC. Our study showed outstanding performance of the 5 machine learning models in predicting the B2 and B5 lesions, but they could only barely predict B1 or B3 lesions (*Table 2*). We tried to retrain the RF model by overrepresenting B1 and B3 within the same class. However, there was no significant improvement using this approach (Appendix 4).

Similar to previous ultrasonic studies (22,23), our study found that RF showed the best overall performance among the 5 machine learning models (*Table 2*). RF contains multiple decision trees, and the output category is determined by the mode of the category output by the individual tree. Therefore, RF is advantageous for conducting the classification. According to the weight map

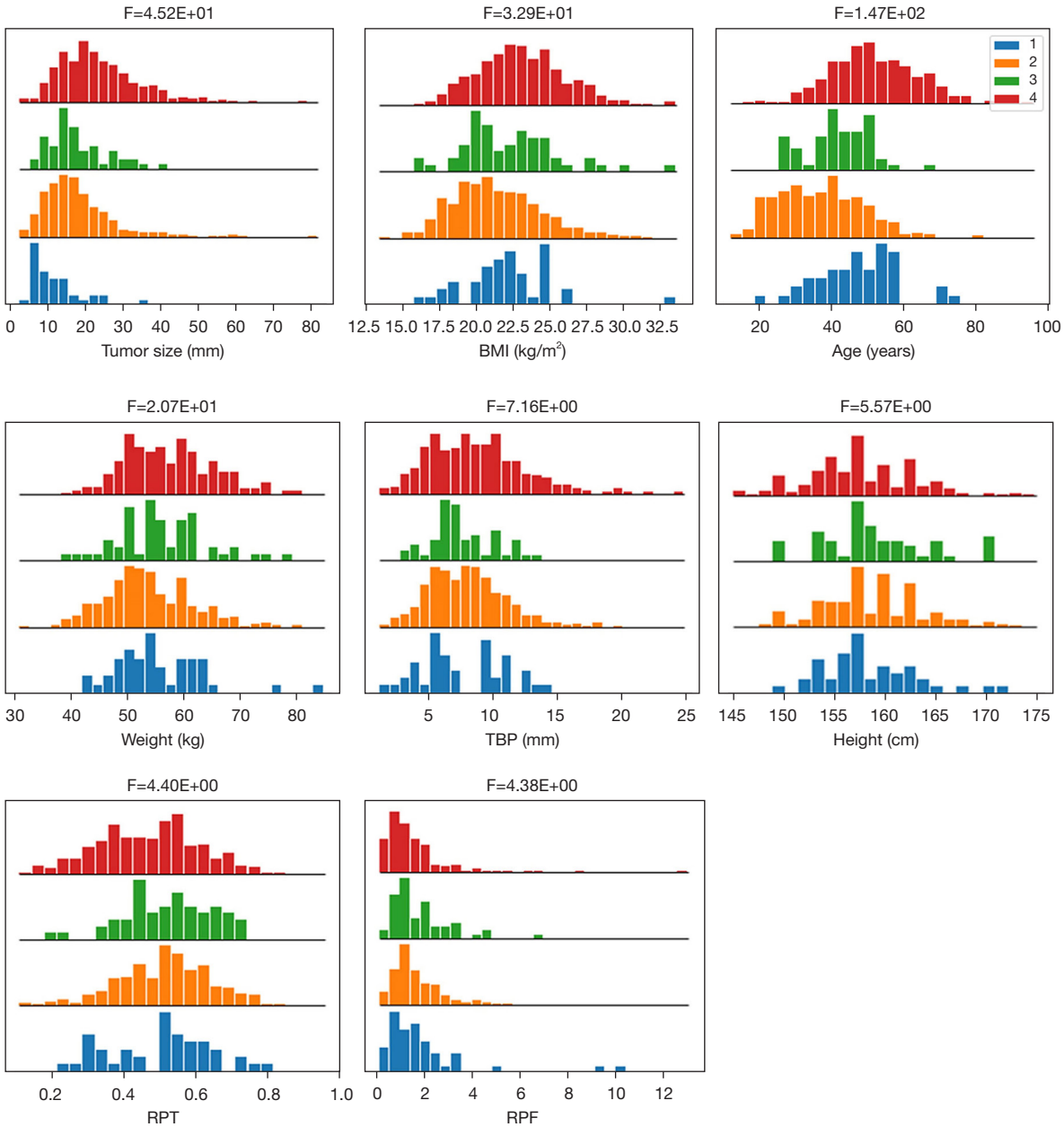


Figure 3 Histograms of each continuous variable. Blue represents the B1 category, orange the B2 category, green the B3 category, and red the B5 category. Different colors represent corresponding categories. BMI, body mass index; TBP, anteroposterior thickness of breast parenchyma; RPT, thickness ratio of breast parenchyma to tissue before pectoralis fascia; RPF, thickness ratio of breast parenchyma to mammary fat.

of RF, we found that the top 10 features were age, tumor size, BMI, microlobulated margin, TBP, RPF, RPT, weight, height, and no reported lymph node metastasis.

Age is not included in the risk stratification of ACR US BI-RADS, but it influences the doctor's determination. Increasing age is closely related to a higher incidence of breast cancer worldwide (24,25). Ordinarily, tumor size is not considered to be a risk factor in US BI-RADS. Nevertheless, based on the tumor biological behavior, the faster the tumor size increases, the higher the possibility of

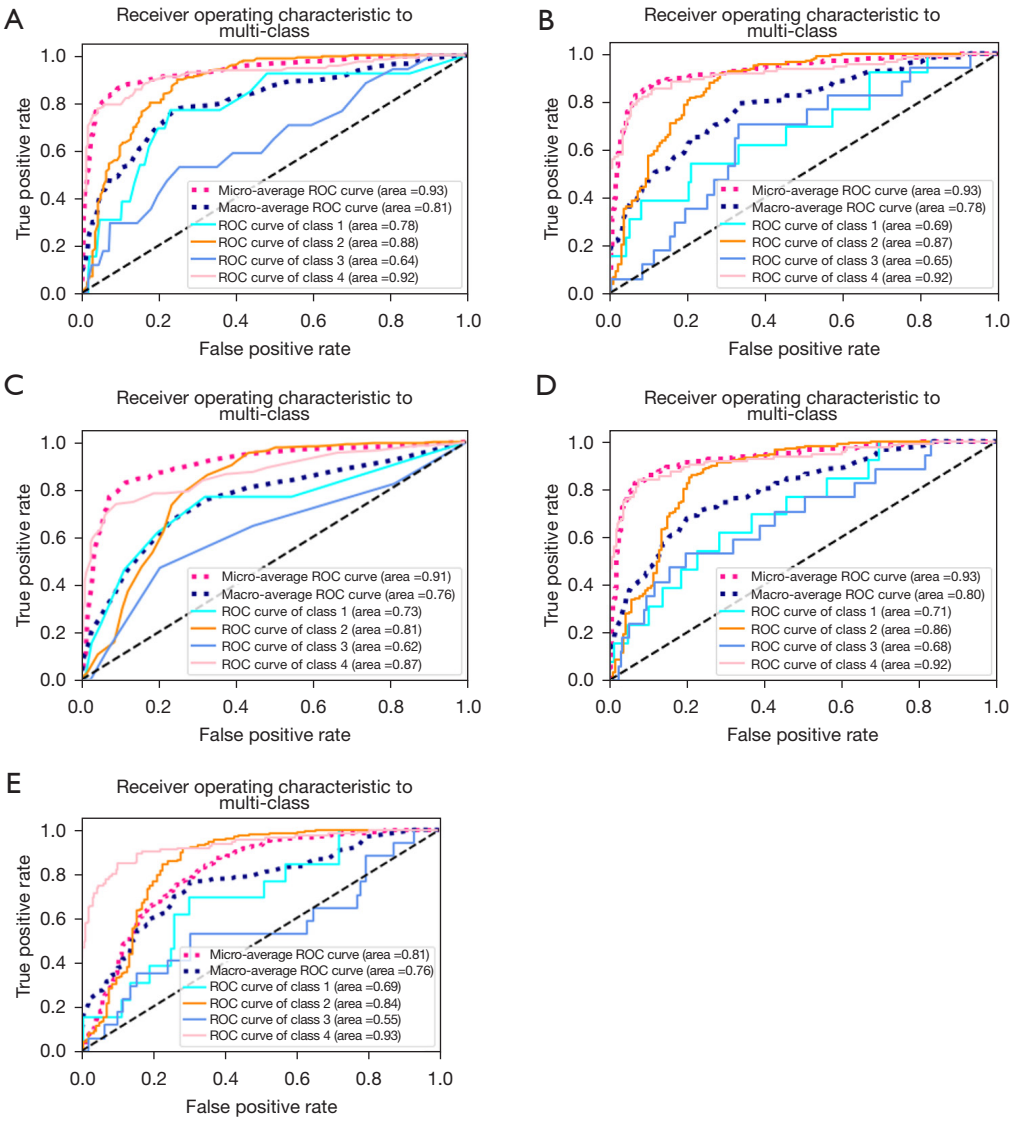


Figure 4 ROC curves of the 5 machine learning models. (A) RF; (B) SVM; (C) KNN; (D) MLP; (E) RR. Class 1: B1 category; class 2: B2 category; class 3: B3 category; class 4: B5 category. ROC, receiver operating characteristic; RF, random forest; SVM, support vector machine; KNN, k-nearest-neighbor; MLP, multilayer perceptron; RR, ridge regression.

malignancy. BMI showed a strong ability to predict CBC. BMI was highest in the B5 category, and a higher BMI indicated a greater likelihood for malignancy. It has long been recognized that overweight and obese individuals are at increased risk of postmenopausal breast cancer, notably hormone receptor-positive cancers (26). In our study, increased age and higher BMI were associated with a higher CBC.

Lesions with microlobulated margins are likely to be classified as malignant (27). The greater desmoplastic reaction leads to noncircumscribed margins, including microlobulated margins (28). As the parameters of breast parenchyma, TBP, RPF, and RPT performed well in our predictive task. The performance of TBP was very strong, and the TBP of the B5 category was the greatest. The interaction of tumor cells and host stromal cells is associated with breast cancer, and the parenchymal stromal cells play an important role in the formation and development of breast cancer, acting as a matrix to promote tumor growth (29,30). A small number of ultrasonic studies found that

Table 2 AUCs and accuracy of the 5 algorithms

Model	HIS	PRE					
		B1	B2	B3	B5	AUC	Accuracy
RF	B1	0	12	0	1	0.78	0.82
	B2	0	204	0	11	0.88	
	B3	0	16	0	1	0.64	
	B5	0	23	0	88	0.92	
SVM	B1	0	13	0	0	0.69	0.81
	B2	0	204	0	11	0.87	
	B3	0	15	0	2	0.65	
	B5	0	27	0	84	0.92	
KNN	B1	0	13	0	0	0.73	0.77
	B2	0	210	0	5	0.81	
	B3	0	16	0	1	0.62	
	B5	0	46	0	65	0.87	
MLP	B1	1	10	0	2	0.71	0.80
	B2	1	198	0	16	0.86	
	B3	1	15	0	1	0.68	
	B5	0	25	0	86	0.92	
RR	B1	0	10	0	3	0.69	0.81
	B2	0	203	0	12	0.84	
	B3	0	16	0	1	0.55	
	B5	0	24	0	87	0.93	

AUC, area under the curve; HIS, actual histological result; PRE, predicted histological result; RF, random forest; SVM, support vector machine; KNN, k-nearest-neighbor; MLP, multilayer perceptron; RR, ridge regression.

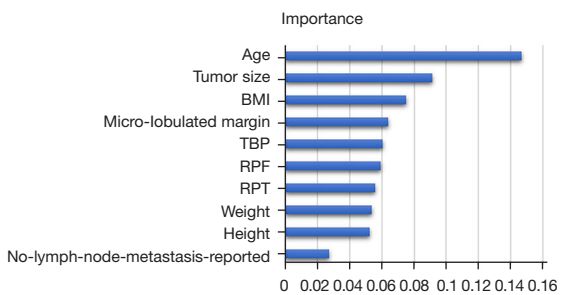


Figure 5 Weight map of the RF model. The weight of age was the highest, and the weights of the remaining features decrease sequentially. BMI, body mass index; TBP, anteroposterior thickness of breast parenchyma; RPT, thickness ratio of breast parenchyma to tissue before pectoralis fascia; RPF, thickness ratio of breast parenchyma to mammary fat; RF, random forest.

thick breast parenchyma promotes the occurrence of breast cancer (31-34). Similar to previous studies (35,36), our study found that the reported lymphatic metastasis was able to predict malignancy.

Our study was not without limitation. First, although RF was determined as the optimal machine learning for CBC, it missed 13 B5 and B2 lesions. As expert mathematicians further perfect the RF, we will try to use the latest RF algorithm to improve the accuracy of CBC prediction in the future. Second, the sample size was small, particularly in the B1 and B3 categories. A larger data set is needed to validate the robustness of RF and to build the application to automatically predict the preoperative biopsy category. Third, the study did not integrate other clinical risk factors

due to the inherent incompleteness of the retrospective study data. Prospective studies with complete data sets (i.e., those including family history, serological, and menopausal period) should be carried out. Last but not the least, the subjective analysis of some ultrasonic features led to inevitable deviation. Consequently, objective parameters (i.e., radiomics features and contrast-enhanced US) should be used to improve this model.

Conclusions

Based on clinical and ultrasonic features, we used machine learning models to predict the CBCs of solid breast tumors. Of the 5 models, RF was the optimal machine learning model, as it had the highest accuracy of 0.82. Especially for B2 and B5, RF performed well, and its AUCs were 0.88 and 0.92, respectively. In practice, our study can aid in clinically precise individualized management with BI-RADS and can also assist in developing a new US imaging reporting system for solid breast tumors.

Acknowledgments

We thank all the study participants and the medical teams who contributed to this study.
Funding: This work was supported by the Guangdong Medical Science and Technology Research Fund (Nos. A2022225, B2022089, and C2018001).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-22-877/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-877/coif). The authors have no 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). This study was approved by the institutional review board of Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences. The informed consent requirement was waived due to the nature of the retrospective study.

References

1. Neal L, Sandhu NP, Hieken TJ, Glazebrook KN, Mac Bride MB, Dilaveri CA, Wahner-Roedler DL, Ghosh K, Visscher DW. Diagnosis and management of benign, atypical, and indeterminate breast lesions detected on core needle biopsy. Mayo Clin Proc 2014;89:536-47.
2. Rungruang B, Kelley JL 3rd. Benign breast diseases: epidemiology, evaluation, and management. Clin Obstet Gynecol 2011;54:110-24.
3. Yap YS, Lu YS, Tamura K, Lee JE, Ko EY, Park YH, Cao AY, Lin CH, Toi M, Wu J, Lee SC. Insights Into Breast Cancer in the East vs the West: A Review. JAMA Oncol 2019;5:1489-96.
4. Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, et al. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Eur J Cancer 2010;46:1296-316.
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. Lee A, Anderson N, Carder P, Cooke J, Deb R, Ellis IO, Howe M, Jenkins JA, Knox F, Stephenson T. Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening. London: The Royal College of Pathologists, 2016.
7. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. J Am Coll Cardiol 2016;68:2287-95.
8. Nanayakkara S, Fogarty S, Tremeer M, Ross K, Richards B, Bergmeir C, Xu S, Stub D, Smith K, Tacey M, Liew D,

Pilcher D, Kaye DM. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. PLoS Med 2018;15:e1002709.
9. Zhang B, Tian J, Pei S, Chen Y, He X, Dong Y, Zhang L, Mo X, Huang W, Cong S, Zhang S. Machine Learning-Assisted System for Thyroid Nodule Diagnosis. Thyroid 2019;29:858-67.
10. Achilonu OJ, Fabian J, Bebington B, Singh E, Eijkemans MJC, Musenge E. Predicting Colorectal Cancer Recurrence and Patient Survival Using Supervised Machine Learning Approach: A South African Population-Based Study. Front Public Health 2021;9:694306.
11. Yu KH, Lee TM, Yen MH, Kou SC, Rosen B, Chiang JH, Kohane IS. Reproducible Machine Learning Methods for Lung Cancer Detection Using Computed Tomography Images: Algorithm Development and Validation. J Med Internet Res 2020;22:e16709.
12. Bitencourt AGV, Gibbs P, Rossi Saccarelli C, Daimiel I, Lo Gullo R, Fox MJ, Thakur S, Pinker K, Morris EA, Morrow M, Jochelson MS. MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer. EBioMedicine 2020;61:103042.
13. Sutton EJ, Onishi N, Fehr DA, Dashevsky BZ, Sadinski M, Pinker K, Martinez DF, Brogi E, Braunstein L, Razavi P, El-Tamer M, Sacchini V, Deasy JO, Morris EA, Veeraraghavan H. A machine learning model that classifies breast cancer pathologic complete response on MRI post-neoadjuvant chemotherapy. Breast Cancer Res 2020;22:57.
14. Wu T, Sultan LR, Tian J, Cary TW, Sehgal CM. Machine learning for diagnostic ultrasound of triple-negative breast cancer. Breast Cancer Res Treat 2019;173:365-73.
15. Lin X, Zhuang S, Yang S, Lai D, Chen M, Zhang J. Development and internal validation of a conventional ultrasound-based nomogram for predicting malignant nonmasslike breast lesions. Quant Imaging Med Surg 2022;12:5452-61.
16. Meucci R, Pistolese Chiara A, Perretta T, Vanni G, Portarena I, Manenti G, Ryan Colleen P, Castrignanò A, Di Stefano C, Ferrari D, Lamacchia F, Pellicciaro M, Materazzo M, Buonomo Oreste C. MR imaging-guided vacuum assisted breast biopsy: Radiological-pathological correlation and underestimation rate in pre-surgical assessment. Eur J Radiol Open 2020;7:100244.
17. Giuliani M, Rinaldi P, Rella R, D'Angelo A, Carlino G, Infante A, Romani M, Bufi E, Belli P, Manfredi R. A new risk stratification score for the management of ultrasound-detected B3 breast lesions. Breast J 2018;24:965-70.
18. Kim YR, Kim HS, Kim HW. Are Irregular Hypoechoic Breast Masses on Ultrasound Always Malignancies?: A Pictorial Essay. Korean J Radiol 2015;16:1266-75.
19. Grigoryev M, Thomas A, Plath L, Durmus T, Slowinski T, Diekmann F, Fischer T. Detection of microcalcifications in women with dense breasts and hypoechoic focal lesions: comparison of mammography and ultrasound. Ultraschall Med 2014;35:554-60.
20. Park AY, Kwon M, Woo OH, Cho KR, Park EK, Cha SH, Song SE, Lee JH, Cha J, Son GS, Seo BK. A Prospective Study on the Value of Ultrasound Microflow Assessment to Distinguish Malignant from Benign Solid Breast Masses: Association between Ultrasound Parameters and Histologic Microvessel Densities. Korean J Radiol 2019;20:759-72.
21. Lu CF, Hsu FT, Hsieh KL, Kao YJ, Cheng SJ, Hsu JB, Tsai PH, Chen RJ, Huang CC, Yen Y, Chen CY. Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas. Clin Cancer Res 2018;24:4429-36.
22. Destrempes F, Trop I, Allard L, Chayer B, Garcia-Duitama J, El Khoury M, Lalonde L, Cloutier G. Added Value of Quantitative Ultrasound and Machine Learning in BI-RADS 4-5 Assessment of Solid Breast Lesions. Ultrasound Med Biol 2020;46:436-44.
23. Wildeboer RR, Mannaerts CK, van Sloun RJG, Budäus L, Tilki D, Wijkstra H, Salomon G, Mischi M. Automated multiparametric localization of prostate cancer based on B-mode, shear-wave elastography, and contrast-enhanced ultrasound radiomics. Eur Radiol 2020;30:806-15.
24. Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, Chen WQ, Shao ZM, Goss PE. Breast cancer in China. Lancet Oncol 2014;15:e279-89.
25. 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 PPV(3) of ACR BI-RADS Ultrasound categories 4 and 5 in the diagnostic setting? Eur Radiol 2018;28:2492-8.
26. Jiralerspong S, Goodwin PJ. Obesity and Breast Cancer Prognosis: Evidence, Challenges, and Opportunities. J Clin Oncol 2016;34:4203-16.
27. Aho M, Irshad A, Ackerman SJ, Lewis M, Leddy R, Pope TL, Campbell AS, Cluver A, Wolf BJ, Cunningham JE. Correlation of sonographic features of invasive ductal mammary carcinoma with age, tumor grade, and hormone-receptor status. J Clin Ultrasound 2013;41:10-7.
28. Irshad A, Leddy R, Pisano E, Baker N, Lewis M, Ackerman S, Campbell A. Assessing the role of ultrasound

in predicting the biological behavior of breast cancer. AJR Am J Roentgenol 2013;200:284-90.
29. Fidler IJ, Poste G. The "seed and soil" hypothesis revisited. Lancet Oncol 2008;9:808.
30. Li H, Mendel KR, Lan L, Sheth D, Giger ML. Digital Mammography in Breast Cancer: Additive Value of Radiomics of Breast Parenchyma. Radiology 2019;291:15-20.
31. Hou XY, Niu HY, Huang XL, Gao Y. Correlation of Breast Ultrasound Classifications with Breast Cancer in Chinese Women. Ultrasound Med Biol 2016;42:2616-21.
32. Kim WH, Lee SH, Chang JM, Cho N, Moon WK. Background echotexture classification in breast ultrasound: inter-observer agreement study. Acta Radiol 2017;58:1427-33.
33. Chang RF, Hou YL, Lo CM, Huang CS, Chen JH, Kim WH, Chang JM, Bae MS, Moon WK. Quantitative analysis of breast echotexture patterns in automated breast ultrasound images. Med Phys 2015;42:4566-78.
34. Liu Z, Li Z, Qu J, Zhang R, Zhou X, Li L, Sun K, Tang Z, Jiang H, Li H, Xiong Q, Ding Y, Zhao X, Wang K, Liu Z, Tian J. Radiomics of Multiparametric MRI for Pretreatment Prediction of Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer: A Multicenter Study. Clin Cancer Res 2019;25:3538-47.
35. Jiang T, Su W, Zhao Y, Li Q, Huang P. Non-invasive prediction of lymph node status for patients with early-stage invasive breast cancer based on a morphological feature from ultrasound images. Quant Imaging Med Surg 2021;11:3399-407.
36. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, Mao R, Li F, Xiao Y, Wang Y, Hu Y, Yu J, Zhou J. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 2020;11:1236.

Cite this article as: Liang T, Shen J, Wang J, Liao W, Zhang Z, Liu J, Feng Z, Pei S, Liu K. Ultrasound-based prediction of preoperative core biopsy categories in solid breast tumor using machine learning. Quant Imaging Med Surg 2023;13(4):2634-2646. doi: 10.21037/qims-22-877

Supplementary

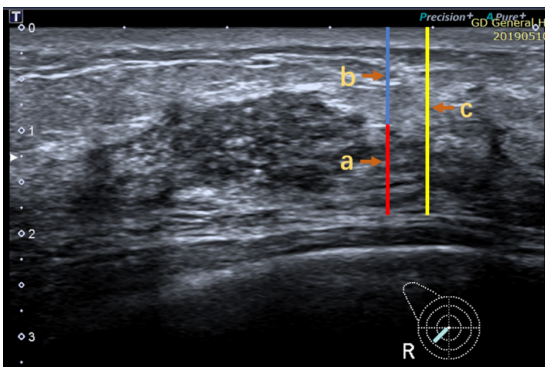
Appendix 1

According to the criteria of the ACR BI-RADS fifth edition and our previous studies (37,38), the detailed descriptions of imaging features are as follows.

Breast BEP

(I) BEP contains fat and fibro-glandular parenchyma (homogeneous and heterogeneous). Homogeneous echotexture consists of fat or fibroglandular homogeneity. Heterogeneity can be either focal or diffuse. The breast echotexture is characterized by multiple small areas of increased and decreased echogenicity and may be companied by shadowing.

(II) The qualitative parameters of breast parenchyma and fat adjacent to lesions were measured in the maximal diameter of the tumor's image. The mean diameter was based on 3 measurements. Each mean diameter should be obtained based on three measurements as follows.



RPT = a/c. RPT is the ratio (the largest thickness of TBP/thickness of pectoralis fascia). RPF = a/b. RPF is the ratio (the thickness of TBP/thickness of mammary fat).

Shape

(I) Oval: a mass that is elliptical or egg-shaped (may include 2 or 3 undulations; i.e., gently lobulated or macrolobulated).

(II) Round: a round mass is one that is spherical, ball-shaped, circular, or globular. It has an anteroposterior diameter equal to its transverse diameter; to qualify as a round mass, it must be circular in perpendicular projections.

(III) Irregular: this lesion shape is neither round nor oval.

Orientation

(I) Parallel: the long axis of the mass parallels the skin line. Masses that are only slightly obliquely oriented might be considered parallel.

(II) Not parallel: the long axis of the mass is not parallel to the skin line. The anterior-posterior or vertical dimension is greater than the transverse or horizontal dimension. These masses can also be obliquely oriented to the skin line. Round masses are not parallel in their orientation.

Margin

(I) Circumscribed: a circumscribed margin is one that is well defined, with an abrupt transition between the lesion and the

https://dx.doi.org/10.21037/qims-22-877

surrounding tissue. For a mass to be described as circumscribed at US, its entire margin must be sharply defined. Most circumscribed lesions have round or oval shapes.

(II) Indistinct: there is no clear demarcation of the entire margin or any portion of the margin from the surrounding tissue.

(III) Angular: some or all of the margin has sharp corners, often forming acute angles.

(IV) Microlobulated: the margin is characterized by short-cycle undulations, but the significant feature is that the margin of the mass is not circumscribed.

Echogenic pattern

(I) Hypoechoic: the term "hypoechoic" is defined relative to subcutaneous fat; hypoechoic masses, which are less echogenic than fat, are characterized by low-level echoes throughout.

(II) Isoechoic: isoechogenicity is defined as having the same echogenicity as subcutaneous fat.

(III) Hyperechoic: hyperechogenicity is defined as having increased echogenicity relative to fat or equal to fibroglandular tissue.

(IV) Complex cystic and solid: a complex mass contains both anechoic (cystic or fluid) and echogenic (solid) components.

(V) Heterogeneous: a mixture of echogenic patterns within a solid mass.

Posterior features

(I) No posterior features: no shadowing or enhancement is present deep to the mass; the echogenicity of the area immediately behind the mass is not different from that of the adjacent tissue at the same depth.

(II) Enhancement sound transmission: there is no impediment in the passage through the mass. Enhancement appears as a column that is more echogenic (whiter) deep to the mass.

(III) Shadowing: there is attenuation of the acoustic transmission. Sonographically, the area posterior to the mass appears darker.

(IV) Combined pattern: these are lesions with more than one pattern of posterior attenuation.

Calcifications

(I) Calcifications in a mass.

(II) Calcifications outside of a mass: calcifications situated in fat or fibro-glandular tissue.

(III) Intraductal calcifications.

(IV) No calcification.

Vascularity distribution

(I) Absent: cysts are the most common avascular lesions. Some solid masses also have little or no vascularity.

(II) Vessels in rim: the blood vessels may be marginal, forming part or all of a rim around a mass.

(III) Internal vascularity: blood vessels are present within the mass. Vessels may penetrate the margin of the mass or display an orderly or disorderly pattern within the mass.

Vascularity grade (39)

(I) Grade I: nonvascularity.

(II) Grade II: less than 1 mm in diameter.

(III) Grade III: a main vessel is visible in the area, and/or several small vessels are visible.

(IV) Grade IV: four or more vessels are visible.

https://dx.doi.org/10.21037/qims-22-877

Reported lymph node metastasis

(I) No: US doctor did not report lymph node metastasis.
(II) Yes: US doctor reported lymph node metastasis.

Tumor size

The largest diameter is the only feature used. The largest measurement should represent the longest axis of a lesion.

Appendix 2

The hyperparameters used for RF, SVM, KNN, MLP, and RR.

RF: n_estimators (200, test in [50, 100, 150, 200])
max_depth (5, test in [2, 3, 4, 5, 6])
n_features (none, since we want to test all features)

SVM: c (1, test in [0.1, 1, 10])
Kernel ('rbf', test in ['linear', 'poly', 'rbf', 'sigmoid'])
degree (3, test in [2, 3])

KNN: n_neighbors (40, test in [3, 5, 10, 20, 30, 40, 50])
weights ('uniform', test in ['uniform', 'distance'])

MLP: hidden_layer_sizes (128, test in [32, 64, 128])
activation ('relu', test in ['logistic', 'relu'])

RR: alpha (1.6, test in [0, 0.1, 0.2, …, 1.9, 2.0])

Parameters not mentioned above were set as the default values.

Appendix 3

Pairwise comparison of clinical and ultrasonic characteristics

Characteristics	HIS	B1	B2	B3	B5
Age	B1		*	*	*
	B2	*		*	*
	B3	*	*		*
	B5	*	*	*	
Height	B1				
	B2				*
	B3				
	B5		*		

(continued)

https://dx.doi.org/10.21037/qims-22-877

(continued)

Characteristics	HIS	B1	B2	B3	B5
Weight	B1				
	B2				*
	B3				
	B5		*		
BMI	B1		*		*
	B2	*		*	*
	B3		*		*
	B5	*	*	*	
Echo pattern	B1				
	B2				*
	B3				
	B5		*		
Shape	B1		*		*
	B2	*			*
	B3				
	B5	*	*		
Margin	B1				*
	B2				*
	B3				*
	B5	*	*	*	
Orientation	B1				
	B2				*
	B3				
	B5		*		
Posterior feature	B1				*
	B2				*
	B3				*
	B5	*	*	*	
Calcification	B1				*
	B2				*
	B3				*
	B5	*	*	*	

(continued)

https://dx.doi.org/10.21037/qims-22-877

(continued)

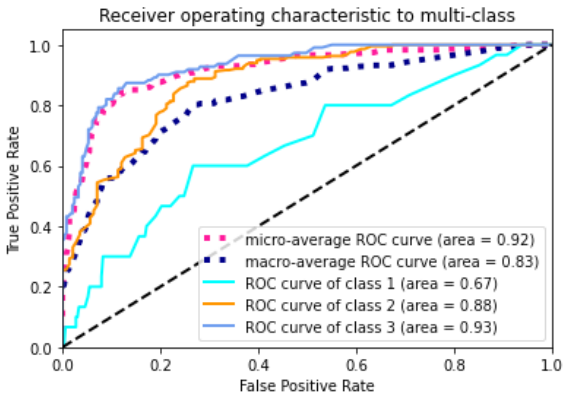
Characteristics	HIS	B1	B2	B3	B5
Vascularity distribution	B1		*	*	*
	B2	*		*	*
	B3	*	*		
	B5	*	*		
Vascularity grade	B1		*		*
	B2	*			*
	B3				*
	B5	*	*	*	
Lymph node metastasis reported	B1				*
	B2				*
	B3				*
	B5	*	*	*	
BEP	B1				
	B2				
	B3				
	B5				
BI-RADS category	B1		*		*
	B2	*			*
	B3				*
	B5	*	*	*	
Tumor size	B1		*	*	*
	B2	*			*
	B3	*	*		*
	B5	*	*	*	
TBP	B1				*
	B2				*
	B3				*
	B5	*	*	*	
RPT	B1				
	B2				*
	B3				
	B5		*		
RPF	B1				
	B2				*
	B3				
	B5		*		

*, P<0.05. HIS, actual histological result; BMI, body mass index; BEP, background echotexture of parenchyma; BI-RADS, Breast Imaging Reporting and Data System; TBP, anteroposterior thickness of breast parenchyma; RPT, thickness ratio of breast parenchyma to tissue before pectoralis fascia; RPF, thickness ratio of breast parenchyma to mammary fat.

https://dx.doi.org/10.21037/qims-22-877

Appendix 4

ROC curves, areas under the curve, and accuracy of the RF model with overrepresentation of B1 and B3 to the same class (class 1).



PRE: class 1 (B1 + B3), class 2 (B2), class 3 (B5), AUCs, accuracy.
HIS: class 1 (B1 + B3), [0, 26, 4, 0.67], 0.81
class 2 (B2) [0, 200, 15, 0.88]
class 3 (B5) [0, 24, 87, 0.93]

References

37. Liang T, Cong S, Yi Z, Liu J, Huang C, Shen J, Pei S, Chen G, Liu Z. Ultrasound-Based Nomogram for Distinguishing Malignant Tumors from Nodular Sclerosing Adenoses in Solid Breast Lesions. J Ultrasound Med 2021;40:2189-200.
38. Liang T, Shen J, Zhang S, Cong S, Liu J, Pei S, Shang S, Huang C. Using Ultrasound-Based Multilayer Perceptron to Differentiate Early Breast Mucinous Cancer and its Subtypes From Fibroadenoma. Front Oncol 2021;11:724656.
39. Adler DD, Carson PL, Rubin JM, Quinn-Reid D. Doppler ultrasound color flow imaging in the study of breast cancer: preliminary findings. Ultrasound Med Biol 1990;16:553-9.

https://dx.doi.org/10.21037/qims-22-877