Original Article

The application of computer-aided diagnosis in Breast Imaging Reporting and Data System ultrasound training for residents—a randomized controlled study

Shuyi Lyu[1,2]^, Meiwu Zhang[1], Baisong Zhang[1], Libo Gao[1], Liu Yang[1], Susanna Guerrini[3], Eugene Ong[4], Yan Zhang[1,2]

[1]Department of Ultrasound, Ningbo No. 2 Hospital, Ningbo, China; [2]Department of Ultrasound, Zhenhai Hospital of Traditional Chinese Medicine, Ningbo, China; [3]Unit of Diagnostic Imaging, Department of Medical Sciences, Azienda Ospedaliero-Universitaria Senese, Siena, Italy; [4]Diagnostic Radiology, Mount Elizabeth Novena Hospital, Singapore, Singapore

Contributions: (I) Conception and design: S Lyu, M Zhang, Y Zhang; (II) Administrative support: Y Zhang, M Zhang; (III) Provision of study materials or patients: S Lyu, M Zhang, Y Zhang; (IV) Collection and assembly of data: B Zhang, L Gao, L Yang; (V) Data analysis and interpretation: S Lyu, B Zhang, L Gao, L Yang, S Guerrini, E Ong; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yan Zhang, BS. Department of Ultrasound, Ningbo No. 2 Hospital, No. 41, Northwest Street, Haishu District, Ningbo 315010, China; Department of Ultrasound, Zhenhai Hospital of Traditional Chinese Medicine, No. 51, Huancheng Street, Zhenhai District, Ningbo 315010, China. Email: zhangynb2@163.com.

Background: The consistency of Breast Imaging Reporting and Data System (BI-RADS) classification among experienced radiologists is different, which is difficult for inexperienced radiologists to master. This study aims to explore the value of computer-aided diagnosis (CAD) (AI-SONIC breast automatic detection system) in the BI-RADS training for residents.

Methods: A total of 12 residents who participated in the first year and the second year of standardized resident training in Ningbo No. 2 Hospital from May 2020 to May 2021 were randomly divided into 3 groups (Group 1, Group 2, Group 3) for BI-RADS training. They were asked to complete 2 tests and questionnaires at the beginning and end of the training. After the first test, the educational materials were given to the residents and reviewed during the breast imaging training month. Group 1 studied independently, Group 2 studied with CAD, and Group 3 was taught face-to-face by experts. The test scores and ultrasonographic descriptors of the residents were evaluated and compared with those of the radiology specialists. The trainees' confidence and recognition degree of CAD were investigated by questionnaire.

Results: There was no statistical significance in the scores of residents in the first test among the 3 groups ($P=0.637$). After training and learning, the scores of all 3 groups of residents were improved in the second test ($P=0.006$). Group 2 (52±7.30) and Group 3 (54±5.16) scored significantly higher than Group 1 (38±3.65). The consistency of ultrasonographic descriptors and final assessments between the residents and senior radiologists were improved ($\kappa 3 > \kappa 2 > \kappa 1$), with $\kappa 2$ and $\kappa 3 > 0.4$ (moderately consistent with experts), and $\kappa 1 = 0.225$ (fairly agreed with experts). The results of the questionnaire showed that the trainees had increased confidence in BI-RADS classification, especially Group 2 (1.5 to 3.5) and Group 3 (1.25 to 3.75). All trainees agreed that CAD was helpful for BI-RADS learning (Likert scale score: 4.75 out of 5) and were willing to use CAD as an aid (4.5, max. 5).

Conclusions: The AI-SONIC breast automatic detection system can help residents to quickly master BI-RADS, improve the consistency between residents and experts, and help to improve the confidence of residents in the classification of BI-RADS, which may have potential value in the BI-RADS training for radiology residents.

Trial Registration: Chinese Clinical Trial Registry (ChiCTR2400081672).

^ ORCID: 0000-0002-0126-7255.

Keywords: Ultrasound (US); Breast Imaging Reporting and Data System (BI-RADS); computer-aided diagnosis (CAD); training; resident

Submitted Nov 17, 2023. Accepted for publication Apr 09, 2024. Published online Apr 22, 2024.
doi: 10.21037/tcr-23-2122
View this article at: https://dx.doi.org/10.21037/tcr-23-2122

Introduction

The incidence of breast cancer surpassed that of lung cancer, becoming responsible for the highest number of new cases of malignant tumor worldwide and the most prevalent malignancy affecting women's health worldwide (1). Ultrasound (US) is an indispensable tool for breast imaging, complementing mammography and magnetic resonance imaging (MRI) (2-4). Moreover, it has the advantages of easy access, low cost, and no radiation risk. However, the low specificity and high interobserver variability of US remain problematic (5-7). Although the American College of Radiology (ACR) developed the Breast Imaging Reporting and Data System (BI-RADS) to standardize description of images and interpretation of reports (8), interobserver agreement ranges from poor to moderate among radiologists (9-11), and is especially impacted by residents without extensive training in breast US. This variation in lesion description and final classification has a serious impact on patient management. False positive results lead to additional imaging or invasive biopsies, thus increasing medical costs, adding to the patient's mental burden, and even causing anxiety.

Highlight box

Key findings

- The AI-SONIC breast automatic detection system is very helpful for resident physicians to master breast Breast Imaging Reporting and Data System (BI-RADS).

What is known and what is new?

- Many previous studies have validated the use of computer-aided diagnosis (CAD) to improve the accuracy of ultrasound (US) diagnosis of breast masses.
- We found that CAD can help residents better and faster master breast BI-RADS classification.

What is the implication, and what should change now?

- Standardized training in breast US is beneficial, and CAD is an effective tool for residents to learn the classification of US BI-RADS.

Many previous studies have validated the use of computer-aided diagnosis (CAD) to improve the accuracy of US diagnosis of thyroid nodules, breast masses, lung masses, etc. (12-14). Although the accuracy of BC detection varies from CAD to CAD, we suspect that this may be related to the difference in the database trained earlier, the difference in the delineation of interested experts, etc. But the application of CAD in BI-RADS training has not been reported. AI-SONIC Breast automatic detection system (Zhejiang Demetics Medical Technology Co., Zhejiang, China) is a CAD system independently developed by a Chinese company to automatically detect breast masses and identify the benign and malignant properties, with powerful automatic outlining, feature extraction, and the ability to differentiate benignancy and malignancy (15). Moreover, it can quantitatively display descriptions of several aspects involved in BI-RADS classification. A study has demonstrated a high degree of consistency between the system and senior radiologists. In addition, it can also reduce the biopsy rate of BI-RADS type 4 nodules (16).

The purpose of this study was to explore the feasibility of CAD application in US BI-RADS training for residents. We present this article in accordance with the CONSORT reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/rc).

Methods

Participants

A total of 12 residents in their first and second year of standardized training for residents in Ningbo No.2 Hospital from May 2020 to May 2021 were selected as the research participants, including 2 males and 10 females (25±2.5 years old). We designed a three-arm randomized trial. The 12 residents were randomly divided into three groups (Group 1, Group 2, Group 3) by drawing lots, and allocation ratio is 1:1:1. There were no statistical differences in age, gender, grade, educational background and major among the three groups (P value >0.05). For details, as shown in *Table 1*. All

Table 1 Basic information of three groups

Group	Number	Age (years)	Gender	Grade	Educational background	Major
Group 1	1	25	Male	2nd	Bachelor	Ultrasound
	2	26	Female	1st	Bachelor	Ultrasound
	3	25	Female	1st	Bachelor	Ultrasound
	4	24	Female	1st	Bachelor	Ultrasound
Group 2	5	25	Female	1st	Bachelor	Ultrasound
	6	24	Male	2nd	Bachelor	Ultrasound
	7	25	Female	1st	Bachelor	Ultrasound
	8	26	Female	1st	Bachelor	Ultrasound
Group 3	9	24	Female	1st	Bachelor	Ultrasound
	10	25	Female	1st	Bachelor	Ultrasound
	11	25	Female	2nd	Bachelor	Ultrasound
	12	26	Female	1st	Bachelor	Ultrasound

Group 1, independent study; Group 2, CAD-assisted study; Group 3, face-to-face teaching by experts. CAD, computer-aided diagnosis.

residents were aware of the study and agreed to participate. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ningbo No. 2 Hospital Ethics Committee (No. YJ-NBEY-KY-2024-022-01). Informed consent was obtained from the patients.

Research methods

Establishment of breast disease database

In the early stage, we established a database of breast diseases containing 200 pathologically confirmed or long-term follow-up verified cases. These medical records were all taken by the same breast US expert (Dr. A has 20 years of experience in breast US) on the same US instrument (Philips EPIQ7C 5–12 MHz). And divided into four sets: training set, learning set, test set A and test set B, each containing 50 cases. The training set includes images of typical ultrasonic features and some special cases described in ACR BI-RADS, such as retention cysts, postoperative scars, and images after neoadjuvant chemotherapy, etc., to help trainees master the ultrasonic image features of typical cases and understand the ultrasonic image manifestations of some special cases. The cases in the self-study set all included expert US descriptions of the masses, BI-RADS classification, and pathological findings. Trainees can refer to the US reports of experts and study against the pathological results. Both test set A and test set B were surgical cases, so BI-RADS were classified from grade 3 to grade 5. It is used to test the training effect of trainees. All cases have a complete medical history and excellent image quality, including video clips. All cases were reviewed and assigned a BI-RADS classification by 2 senior radiologists (A and B) dedicated to breast imaging. Dr. A and Dr. B have 20 and 15 years of experience in breast US, respectively.

Theoretical training

The cases in the training set were used as theoretical training materials. The training lasted for one day, during which the trainees were free to ask questions to the teaching experts. The first test and questionnaire were conducted on the day at the end of the training, and the second test and questionnaire were conducted one month after the training. The questionnaire included students' grade, experience of breast scan, knowledge of CAD, level of trust and concern, learning status of breast study month, and confidence in BI-RADS classification. The Likert scale is divided into 5 levels, with the lowest score being 1 and the highest score being 5.

Test methods

All residents were required to describe and classify lesions in accordance with the BI-RADS dictionary (17): investigating the consistency of the description from 6 dimensions: shape, orientation, margin, echo pattern, posterior features, and calcifications.

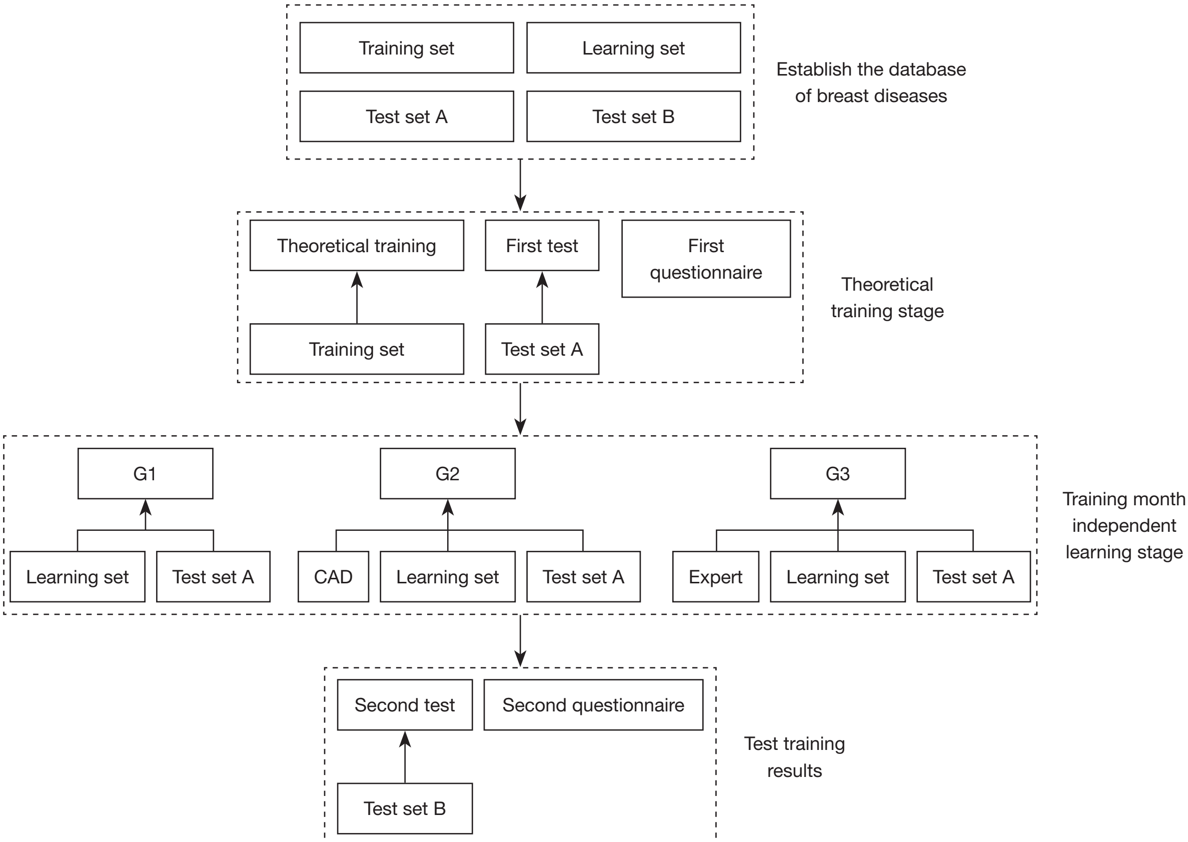

Figure 1 Training flow chart. G1, Group 1, independent study; G2, Group 2, CAD-assisted study; G3, Group 3, face-to-face teaching by experts. CAD, computer-aided diagnosis.

The classification of BI-RADS was then recorded and the corresponding results were recovered, followed by the feedback of the diagnosis and pathological results of the experts. By comparing the residents' findings with the description and classification of breast mass by experts, the consistencies between the residents and the experts were calculated. After the first test, 100 cases from test set A and the learning set were distributed to the residents, who reviewed the cases in the following month, each resident also receives an ACR BIRADS atlas. Group 1 was self-taught based on the US description of the mass and the BI-RADS classification combined with the pathological results, for 1 hour a day, 5 days a week, for a total of 20 hours. Group 2 was assisted by CAD to learn the 100 cases by themselves, referring to the description provided by CAD and combining with the pathological results, and was also given 20 hours. Group 3 received face-to-face reviews of each case by the experts, lasted about 3 hours. In the remaining 17 hours, residents learned according to expert explanation, US description, and pathological results. At the end of the training, a second test was conducted using the cases within test set B. Each test contained 50 questions, with 2 marks for each question consistent with the experts' classification results, and no marks for inconsistency. The detailed training process was shown in *Figure 1*.

AI-SONIC breast automatic detection system

The AI-SONIC breast automatic detection system can automatically label, process, analyze, and quantitatively describe the 6 above mentioned features of breast nodules, as well as provide BI-RADS classification and benign and malignant probability values (probability values range from 0 to 1, with 0 to 0.5 representing 'benign' and 0.6–1 representing a higher likelihood of 'malignant'). The most important feature of the AI-SONIC breast automatic detection system is that it objectively quantifies the characteristics of the BI-RADS dictionary, quantitatively describes the degree of breast lesions, and defines the

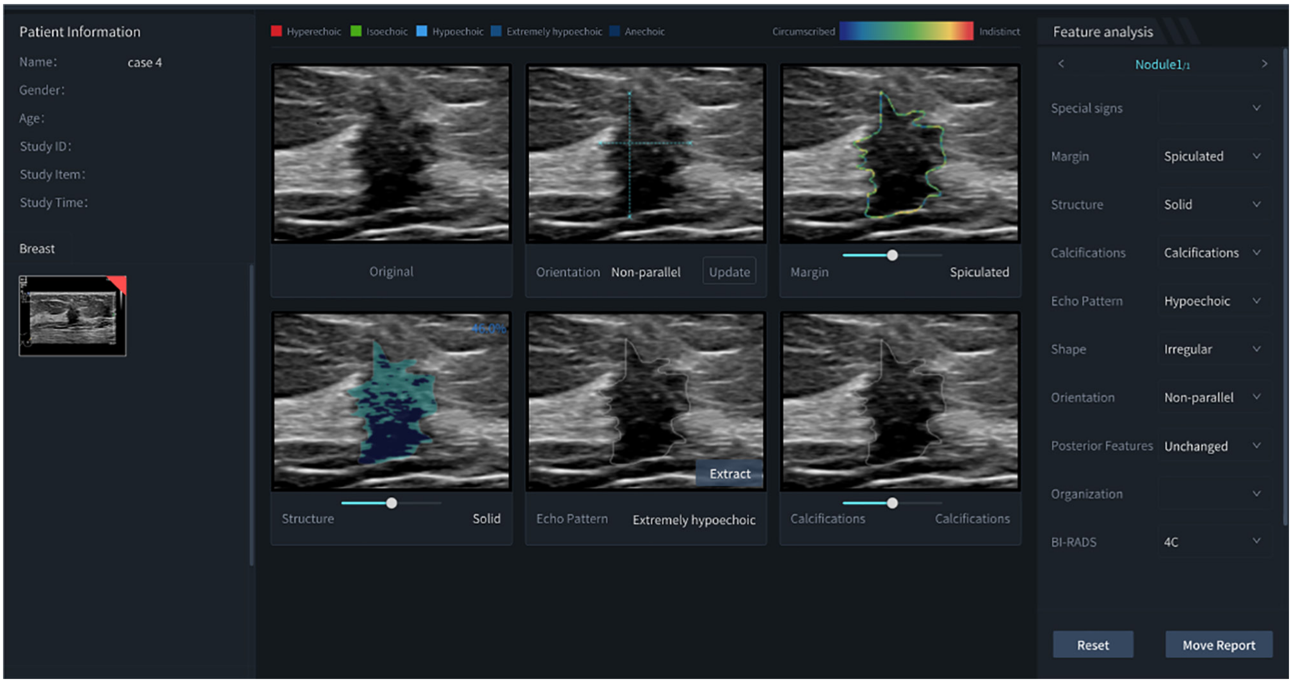

Figure 2 AI-SONIC breast automatic detection interface. Representative image of setting the ROI for AI-SONIC breast automatic detection system in a 53-year-old woman with a diagnosis of cancer in her right breast. The ROI was set automatically along the margin of the breast mass for analysis. After the ROI was set, the ultrasonographic features were automatically analyzed by AI-SONIC, and a final assessment was automatically visualized. The right side of the figure shows the details of the 6 important characteristics specified in the BI-RADS dictionary (add ref to reference). ROI, region of interest; BI-RADS, Breast Imaging Reporting and Data System.

characteristics of masses in precise score ranges (*Figure 2*). For each trait score range, there is a corresponding presentation pattern from the BI-RADS dictionary. Thus, radiologists can accurately classify each feature in breast US images. Prior to the training, 2 senior radiologists had reviewed CAD descriptions of the cases in the database and their consistency with the pathological findings.

Statistical analysis

The software SPSS 22.0 (IBM Corp., Armonk, NY, USA) was used for statistical analysis. Continuous variables were described using mean ± standard deviation (SD). One-way analysis of variance (ANOVA) was used to assess the difference between the 3 groups of residents. Kappa statistics were used for the consistency of US descriptors and final assessments between the residents and experts. Less than 0 indicated poor agreement; 0.00–0.20 indicated slight agreement; 0.21–0.40 indicated fair agreement; 0.41–0.60 indicated moderate agreement; 0.61–0.80 indicated substantial agreement; 0.81–1.00 indicated almost perfect agreement. The Wilcoxon rank sum test was used to compare the scores of the 2 questionnaires. A P value <0.05 was considered to indicated statistical significance.

Results

Pathology and final assessment for BI-RADS classifications of breast masses

The 2 tests included a total of 100 cases, of which 54 were benign and 46 were malignant. In the first test, 28 were benign and 22 were malignant, whereas 26 were benign and 24 were malignant in the second test. The final assessment for BI-RADS classifications by 2 experts were as follows: 32 cases of category 3 with 0 cases of malignancy; 26 cases of category 4a with 6 cases of malignancy; 13 cases of category 4b with 11 cases of malignancy; 22 cases of category 4c with 22 cases of malignancy; 7 cases of category 5 with 7 cases of malignancy. The consistency (κ =0.625–0.850) between the two experts and CAD in ultrasonic description and BI-RADS classification were both greater than 0.6. For

Table 2 Scores of residents' tests

Test	Group 1	Group 2	Group 3	P
The first test	30±5.89	28±6.32	32±5.16	0.63
The second test	38±3.65	52±7.30	54±5.16	0.006

Data are presented as mean ± standard deviation. P, comparison between the 3 groups in test. Group 1, independent study; Group 2, CAD-assisted study; Group 3, face-to-face teaching by experts. CAD, computer-aided diagnosis.

Table 3 Agreement for ultrasonographic descriptors and final assessments between the residents and senior radiologists

Ultrasound	Group 1, κ (95% CI)		Group 2, κ (95% CI)		Group 3, κ (95% CI)	
	The first test	The second test	The first test	The second test	The first test	The second test
Final assessment	0.125 (–0.030, 0.285)	0.225 (0.052, 0.389)	0.100 (–0.048, 0.252)	0.412 (0.232, 0.582)	0.150 (–0.008, 0.308)	0.425 (0.232, 0.600)
Features						
Shape	0.249 (0.039, 0.462)	0.371 (0.155, 0.580)	0.280 (0.074, 0.486)	0.460 (0.268, 0.663)	0.249 (0.039, 0.462)	0.520 (0.308, 0.696)
Orientation	0.240 (–0.027, 0.483)	0.400 (0.149, 0.661)	0.277 (–0.004, 0.538)	0.520 (0.271, 0.722)	0.204 (–0.074, 0.461)	0.560 (0.319, 0.785)
Margin	0.189 (–0.023, 0.396)	0.249 (0.002, 0.427)	0.158 (–0.053, 0.360)	0.430 (0.219, 0.634)	0.220 (0.002, 0.427)	0.460 (0.268, 0.663)
Echo pattern	0.250 (0.074, 0.423)	0.275 (0.105, 0.449)	0.225 (0.052, 0.389)	0.400 (0.232, 0.572)	0.200 (0.045, 0.371)	0.425 (0.232, 0.600)
Posterior features	0.129 (–0.034, 0.288)	0.232 (0.051, 0.409)	0.154 (–0.029, 0.324)	0.385 (0.196, 0.562)	0.206 (0.026, 0.387)	0.436 (0.244, 0.611)
Calcifications	0.280 (0.074, 0.486)	0.309 (0.082, 0.510)	0.189 (–0.023, 0.396)	0.490 (0.300, 0.671)	0.139 (–0.026, 0.252)	0.340 (0.126, 0.536)

Group 1, independent study; Group 2, CAD-assisted study; Group 3, face-to-face teaching by experts. CI, confidence interval; CAD, computer-aided diagnosis.

inconsistent cases, the two experts determined the final description and classification after consultation. The specific training process was shown in *Figure 1*.

Residents' BI-RADS classification scores

The basic information of the three groups was no difference, as shown in *Table 1*. There was no statistical significance in the scores of residents in the first test among the 3 groups (P=0.63). A 1 month after training, the results of the second test in the 3 groups were all improved compared with the previous test, and the differences were statistically significant (P=0.006). The scores of Group 2 and Group 3 were significantly higher than those of Group 1, and the difference were statistically significant (P<0.05). The detailed results are shown in *Table 2*.

Interobserver agreement on ultrasonographic descriptors and final classifications

A summary of the interobserver variability in US features and final assessments between the residents and senior radiologists is presented in *Table 3*. The consistency of final assessments was improved (κ3 > κ2 > κ1), κ2 and κ3 were both >0.4 (moderately consistent with experts), and κ1 =0.225 (fairly consistent with experts). Overall, the interobserver agreement in the first test ranged from 'slight' to 'fair' in terms of shape (κ =0.249–0.280), orientation (κ =0.204–0.277), margin (κ =0.158–0.220), echo pattern (κ =0.200–0.250), posterior features (κ =0.129–0.206), and calcifications (κ =0.139–0.280). However, the results of the interobserver variability were much different in the second test. The interobserver agreement between Group

Table 4 The pre-training questionnaire

No.	Questions	Reply
1	How many years have you been a resident physician?	1/2 years
2	Have you ever had any experience with breast ultrasound? (1= no experience, 5= very experienced)	1/2/3/4/5
3	How do you know about breast CAD? (1= don't know, 5= very know)	1/2/3/4/5
4	How much do you trust breast CAD? (1= distrust, 5= very trust)	1/2/3/4/5
5	Do you think breast CAD is helpful in learning BI-RADS classification? (1= not helpful, 5= very helpful)	1/2/3/4/5
6	Are you worried that CAD will replace doctors in the future? (1= not worried, 5= very worried)	1/2/3/4/5
7	Do you have confidence in the BI-RADS classification? (1= no confidence, 5= very confident)	1/2/3/4/5

CAD, computer-aided diagnosis; BI-RADS, Breast Imaging Reporting and Data System.

1 and the specialists was fair for shape, orientation, margin, echo pattern, posterior features, and calcifications (κ =0.371, 0.400, 0.249, 0.275, 0.232, and 0.309, respectively). Meanwhile, the interobserver agreement between Group 2 and the specialists was fair for posterior features and echo pattern (κ =0.385 and 0.400), and moderate for shape, orientation, margin, and calcifications (κ =0.460, 0.520, 0.430, and 0.490, respectively) in the second test. For Group 3 and the specialists, the inter-observer agreement was fair for calcifications (κ =0.340), and moderate for shape, orientation, margin, echo pattern, and posterior features (κ =0.520, 0.560, 0.460, 0.425, and 0.436, respectively) in the second test.

Questionnaire results

The questionnaires were scored by Likert scale (18). After the breast imaging training, the questionnaire results showed that the confidence of all trainees in BI-RADS assessments was improved, and improvements were more marked in Group 2 and Group 3, with Group 2 having their confidence index increased from 1.5 to 3.5 and Group 3 from 1.25 to 3.75. All trainees agreed that CAD was helpful in learning BI-RADS (4.75 out of 5), and were willing to use CAD as an assistant (4.5 out of 5). None of them worried that they would be replaced by CAD in the future (1.25 out of 5). The pre- and post-training questionnaires as shown in *Tables 4,5*.

Discussion

Previous studies have generally reported that CAD systems can improve the diagnostic performance of breast US (19-21). In this study, we innovatively applied CAD to the training of residents in breast BI-RADS learning. A total of 12 residents were randomly divided into 3 groups: Group 1 (independent study), Group 2 (CAD-assisted study), and Group 3 (face-to-face teaching by experts). The study period was 1 month, after which time we explored the effects of these 3 different educational approaches on residents' BI-RADS classification learning. We found that through training and learning, the second test scores increased in all 3 groups. The agreement between residents and experts was improved both in ultrasonographic descriptors and final assessments. However, the teaching effect of Group 2 and Group 3 was significantly better than that of Group 1. Additionally, the confidence of the trainees in BI-RADS classification was enhanced, with more evident improvements in Group 2 and Group 3. As previously reported in the literature (22), a standardized training program in breast US used to train residents proved beneficial.

Although BI-RADS is widely used to describe breast lesions, operator dependency and interobserver variability can lead to inconsistencies in diagnosis among practitioners (23-25). It was reported that the specificity of residents was significantly lower than that of senior radiologists in the assessment of breast lesions using the BI-RADS dictionary. CAD system has recently been applied to overcome observer variability in breast US, as well as to improve diagnostic performance (13,14). The AI-SONIC breast automated detection system allows for a quantitative representation of the radiologist's subjective description of the BI-RADS dictionary features, defining each feature with a precise range of scores and giving a final classification of the mass. The characteristic description of breast mass by beginners is subjective, which leads to low specificity of final classification. While previous study evaluating

Table 5 The post-training questionnaire

No.	Questions	Reply
1	How many years have you been a resident physician?	1/2 years
2	How many breast cases have you studied in the past month?	≤20/21–40/41–60/61–79/≥80
3	Have you studied the pattern of breast cases in the past month?	Self-taught/by CAD/by expert
4	Students learning with CAD please answer the following questions:	
	How much do you know about CAD through CAD practice? (1= don't know, 5= very know)	1/2/3/4/5
	How reliable do you think CAD is? (1= unreliable, 5= very reliable)	1/2/3/4/5
	Do you think breast CAD is helpful in learning BI-RADS classification? (1= not helpful, 5= very helpful)	1/2/3/4/5
5	Are you willing to CAD assist learning? (1= unwilling, 5= very willing)	1/2/3/4/5
6	Are you worried that CAD will replace doctors in the future? (1= not worried, 5= very worried)	1/2/3/4/5
7	Do you have confidence in the BI-RADS classification? (1= no confidence, 5= very confident)	1/2/3/4/5

CAD, computer-aided diagnosis; BI-RADS, Breast Imaging Reporting and Data System.

training has reported immediate results, the long-term effects of training have not been well documented (22). *Table 3* showed that the "posterior feature" scores of the three groups were relatively low. This may indicate that residents have difficulty grasping concepts in this area, and in future training, we will improve this training to improve their understanding. In our study, Group 2 learned BI-RADS with the help of CAD within 1 month after the first test, and the system presented US features to the trainees, which provided a useful reference for residents to learn the images on a case-by-case basis, allowing for a rule-based final classification, which, together with immediate pathology feedback, impressed the trainees and led to good results and basic mastery of BI-RADS classification. The Group 2 trainees were moderately consistent with experts (κ =0.412) and their scores (52±7.30) in the second test were significantly higher than those of Group 1 (38±3.65) yet slightly lower than those of Group 3 (54±5.16). The effect of CAD system-assisted teaching is close to that of face-to-face teaching by experts. The reason for this/our research outcome could possibly be the fact that the massive breast US images in the early model training period of the CAD system were manually outlined by breast US experts then the deep convolutional neural network was used to automatically outline lumps and conduct feature description and classification after unsupervised self-learning. To some extent, it reflects the opinions of experts, so the consistency between the system and experts in the description and classification of breast masses is very high (16). At present, the clinical workload of doctors with teaching duties (i.e., clinical teachers) in domestic hospitals is heavy, and there is no full-time teacher to teach residents face-to-face throughout the whole training session. The emergence of a trained CAD system may alleviate this situation. When learning the US BI-RADS, the residents can import the ultrasonic images into the CAD system, and the students can learn by referring to the quantitative score and classification results of the CAD description for masses, thus enabling themselves truly understand and flexibly use the BI-RADS classification. Further, students prefer this kind of teaching method due to its simplicity and remarkable accessibility. CAD systems may be potentially valuable in training inexperienced residents.

The low signal-to-noise ratio of breast US images, the utter commonness of homogeneous or heterogeneous images, and the easy confusion of benign and malignant tumors (26-28), inflammation and adenopathy are all challenges for relatively inexperienced residents. The pre-training questionnaire showed that students in all 3 groups had low confidence in mastering BI-BADS (1.25–1.5 points out of 5). The results of the questionnaire after training showed that all trainees had raised their confidence in mastering BI-RADS, with a more pronounced/marked increase in Group 2 (from 1.5 to 3.5) and Group 3 (from 1.25 to 3.75). All students found CAD helpful for breast BI-RADS learning (4.75 out of 5), and were willing to use CAD to assist their learning (4.5 out of 5). Hence, the application of CAD in the training of residents in breast

US can increase the self-confidence of the trainees, help residents to master the classification of BI-RADS in a short time, optimize the learning process, and reduce the burden on clinical teachers (29).

Video clips recommended by experts were used in all breast cancer cases in this study, which provides a more similar scenario to daily practice (9). Combined with pathological diagnosis, trainees can receive immediate feedback after observing the US images, which is thereby conducive to the training effect. In the future, we will continue to expand the cases, build a large database of breast US video clips, store US images from breast cancer patients and their clinical information, and help experienced radiologists to generate teaching documents for teaching and training. Due to the small number of trainees in this study, the fact that the teaching effect was only evaluated 1 month after the training (no follow-up on the long-term outcomes of the training), and all the surgical cases used for testing (with BI-RADS classifications above category 3), it was difficult to establish a perfect/thorough teaching evaluation because only BI-RADS classifications for category 3–5 masses by the trainees were evaluated. The CAD system also has certain limitations, with only 1 section being classified, which may lead to misdiagnosis since different focal sections can infer/deduce totally opposite diagnosis. For better repeatability, we used automatic diagnosis with maximum cross section.

Conclusions

Standardized training in breast US is beneficial, and CAD is an effective tool for residents to learn the classification of US BI-RADS, which is helpful for residents to truly understand and flexibly apply BI-RADS classification, improves the consistency between residents and experts, and boosts the confidence of residents in BI-RADS classification, thus facilitating the rapid mastery of BI-RADS classification for residents.

Acknowledgments

Funding: This work was supported by Ningbo Clinical Research Center for Medical Imaging (No. 2021L003), Provincial and Municipal Co-construction Key Discipline for Medical Imaging (No. 2022-S02), and Zhejiang Medical and Health Science and Technology Project (Nos. 2020KY837, 2021KY294, and 2023KY1098).

Footnote

Reporting Checklist: The authors have completed the CONSORT reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/rc

Trial Protocol: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/tp

Data Sharing Statement: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/dss

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2122/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). The study was approved by Ningbo No. 2 Hospital Ethics Committee (No. YJ-NBEY-KY-2024-022-01). Informed consent was obtained from the patients.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
2. Secco GM, Gutierrez PA, Secco VL, et al. Is breast ultrasound a good alternative to magnetic resonance

imaging for evaluating implant integrity? Radiologia (Engl Ed) 2022;64 Suppl 1:20-7.
3. Bader W, Vogel-Minea CM, Blohmer JU, et al. Best Practice Guideline - DEGUM Recommendations on Breast Ultrasound. Ultraschall Med 2022;43:570-82.
4. Yilmaz Y, Sezgin Bener G, Atahan K, et al. Can Magnetic Resonance Imaging Replace Mammography and Ultrasonography for the Detection of Breast Lesions? Cureus 2020;12:e8087.
5. Park HJ, Kim SM, La Yun B, et al. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of breast masses on ultrasound: Added value for the inexperienced breast radiologist. Medicine (Baltimore) 2019;98:e14146.
6. Elverici E, Zengin B, Nurdan Barca A, et al. Interobserver and Intraobserver Agreement of Sonographic BIRADS Lexicon in the Assessment of Breast Masses. Iran J Radiol 2013;10:122-7.
7. Park CS, Kim SH, Jung NY, et al. Interobserver variability of ultrasound elastography and the ultrasound BI-RADS lexicon of breast lesions. Breast Cancer 2015;22:153-60.
8. American College of Radiology. Breast imaging reporting and data system, 5th ed. Reston, VA, USA: American College of Radiology, 2013.
9. Youk JH, Jung I, Yoon JH, et al. Comparison of Inter-Observer Variability and Diagnostic Performance of the Fifth Edition of BI-RADS for Breast Ultrasound of Static versus Video Images. Ultrasound Med Biol 2016;42:2083-8.
10. Lee YJ, Choi SY, Kim KS, et al. Variability in Observer Performance Between Faculty Members and Residents Using Breast Imaging Reporting and Data System (BI-RADS)-Ultrasound, Fifth Edition (2013). Iran J Radiol 2016;13:e28281.
11. Schwab F, Redling K, Siebert M, et al. Inter- and Intra-Observer Agreement in Ultrasound BI-RADS Classification and Real-Time Elastography Tsukuba Score Assessment of Breast Lesions. Ultrasound Med Biol 2016;42:2622-9.
12. Yang Y, Hu Y, Shen S, et al. 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. Quant Imaging Med Surg 2021;11:3005-17.
13. Suzuki K. A review of computer-aided diagnosis in thoracic and colonic imaging. Quant Imaging Med Surg 2012;2:163-76.
14. Liu C, Chen S, Yang Y, et al. The value of the computer-aided diagnosis system for thyroid lesions based on computed tomography images. Quant Imaging Med Surg 2019;9:642-53.
15. Mai W, Zhou M, Li J, et al. The value of the Demetics ultrasound-assisted diagnosis system in the differential diagnosis of benign from malignant thyroid nodules and analysis of the influencing factors. Eur Radiol 2021;31:7936-44.
16. Lyu SY, Zhang Y, Zhang MW, et al. Diagnostic value of artificial intelligence automatic detection systems for breast BI-RADS 4 nodules. World J Clin Cases 2022;10:518-27.
17. Spak DA, Plaxco JS, Santiago L, et al. BI-RADS(®) fifth edition: A summary of changes. Diagn Interv Imaging 2017;98:179-90.
18. Vonk J. What's not to like about Likert? Developing a nonverbal animal preference scale (NAPS). Am J Primatol 2022;84:e23364.
19. Chan HP, Samala RK, Hadjiiski LM. CAD and AI for breast cancer-recent development and challenges. Br J Radiol 2020;93:20190580.
20. Le EPV, Wang Y, Huang Y, et al. Artificial intelligence in breast imaging. Clin Radiol 2019;74:357-66.
21. Xiao M, Zhao C, Zhu Q, et al. An investigation of the classification accuracy of a deep learning framework-based computer-aided diagnosis system in different pathological types of breast lesions. J Thorac Dis 2019;11:5023-31.
22. Yoon JH, Lee HS, Kim YM, et al. Effect of training on ultrasonography (US) BI-RADS features for radiology residents: a multicenter study comparing performances after training. Eur Radiol 2019;29:4468-76.
23. Abdullah N, Mesurolle B, El-Khoury M, et al. Breast imaging reporting and data system lexicon for US: interobserver agreement for assessment of breast masses. Radiology 2009;252:665-72.
24. Pesce K, Tajerian M, Chico MJ, et al. Interobserver and intraobserver variability in determining breast density according to the fifth edition of the BI-RADS® Atlas. Radiologia (Engl Ed) 2020;62:481-6.
25. Dobrosavljevic A, Milosevic Z, Plesinac S, et al. Interobserver variability and positive predictive value for ultrasonographic BI-RADS categories requiring pathohistological evaluation. Eur J Gynaecol Oncol 2016;37:95-9.
26. Ferré R, Paré M, Smith L, et al. Retroareolar Carcinomas in Breast Ultrasound: Pearls and Pitfalls. Cancers (Basel) 2016;9:1.
27. Wang ZL, Li N, Li M, et al. Non-mass-like lesions on breast ultrasound: classification and correlation with

histology. Radiol Med 2015;120:905-10.
28. Elfgen C, Varga Z, Reeve K, et al. The impact of distinct triple-negative breast cancer subtypes on misdiagnosis and diagnostic delay. Breast Cancer Res Treat 2019;177:67-75.
29. Chen Z, Zhang Y, Yan Z, et al. Artificial intelligence assisted display in thoracic surgery: development and possibilities. J Thorac Dis 2021;13:6994-7005.

Cite this article as: Lyu S, Zhang M, Zhang B, Gao L, Yang L, Guerrini S, Ong E, Zhang Y. The application of computer-aided diagnosis in Breast Imaging Reporting and Data System ultrasound training for residents—a randomized controlled study. Transl Cancer Res 2024;13(4):1969-1979. doi: 10.21037/tcr-23-2122

