The usefulness of a computer-aided diagnosis scheme for improving the performance of clinicians to diagnose non-mass lesions on breast ultrasonographic images

Mai Shibusawa, Ryohei Nakayama, Yuko Okanami, Yumi Kashikura, Nao Imai, Takashi Nakamura, Hiroko Kimura, Masako Yamashita, Noriko Hanamura, Tomoko Ogawa

Mai Shibusawa・Yuko Okanami・Nao Imai・Hiroko Kimura・Masako Yamashita・Noriko Hanamura・Tomoko Ogawa
Department of Breast Surgery, Mie University Hospital, 2-174, Edobashi, Tsu city, Mie
514-8507, Japan
Telephone: +081-59-231-5584
Fax: +081-59-231-5338
E-mail: shibu-m@clin.medic.mie-u.ac.jp

Ryohei Nakayama
Department of Electronic and Computer Engineering, Ritsumeikan University, 1-1-1, Noji-higashi, Kusatsu city, Shiga, 525-8577, Japan

Yumi Kashikura
Department of Senology, Tsukuba Medical Center Hospital, 1-3-1, Amakubo, Tsukuba city, Ibaraki, 305-8558, Japan

Takashi Nakamura
Department of Breast Surgery, Nabari City Hospital, 1-178, Yurigaokanishi, Nabari city, Mie, 518-0481, Japan

Abstract

Purpose The purpose of this study was to evaluate the usefulness of a computer-aided diagnosis (CAD) scheme for improving the performance of clinicians to diagnose non-mass lesions appearing as hypoechoic areas on breast ultrasonographic images.

Methods The database included 97 ultrasonographic images with hypoechoic areas：48 benign cases [benign lesion with benign mammary tissue or fibrocystic disease (n=20), fibroadenoma (n=11) and intraductal papilloma (n=17)] and 49 malignant cases [ductal carcinoma *in situ* (n=17) and invasive ductal carcinoma (n=32)]. Seven clinicians, three expert breast surgeons, and four general surgeons participated in the observer study. They were asked their confidence level concerning the possibility of malignancy in all 97 cases with and without the use of the CAD scheme. Receiver operating characteristic (ROC) analysis was performed to evaluate the usefulness of the CAD scheme.

Results The areas under the ROC curve (AUC) improved for all observers when they used the CAD scheme and increased from 0.649 to 0.783 (P =0.0167). Notably, the AUC for the general surgeon group increased from 0.625 to 0.793(P=0.045).

Conclusions This study showed that the performance of clinicians to diagnose of non-mass lesions appearing as hypoechoic areas on breast ultrasonographic images was improved by the use of a CAD scheme.

Keywords: computer-aided diagnosis・breast ultrasonography・breast non-mass lesions

Introduction

Breast ultrasonography is thought to be more useful than mammography for detecting small breast masses in dense breasts [1]. The combined approach of screening with ultrasonography and mammography is significantly more effective than mammography alone in the detection of small breast cancer in women with dense breast tissue [2]. Thus the introduction of ultrasonography to breast cancer screening would allow for the earlier detection and diagnosis of breast cancer and would decrease the level of associated mortality. However, the accuracy of detection and the diagnosis of lesions on breast ultrasonographic images depend on the skill and experience of the operator. Although the sensitivity would be improved by introducing breast ultrasonography to cancer screening, a high false-positive rate might result in an increase in the number of unnecessary detailed examinations (e.g., pathologic examinations such as fine-needle aspiration cytology, core needle biopsy and vacuum-assisted biopsy) [2, 3].

Computer-aided diagnosis (CAD) schemes can potentially make a differential diagnosis more accurate and less dependent on the skill of the observer. CAD schemes allow clinicians to refer to the results of a computer-based analysis in the differential diagnosis of unknown images [4]. CAD schemes has been shown to be useful for improving diagnostic accuracy in various image modalities. Nakayama et al. showed that a CAD scheme to histologically classify breast lesions improved the performance of radiologists in the differential diagnosis of clustered microcalcifications on mammograms [5]. Its efficacy has also been demonstrated in breast ultrasonography. Ying et al. showed that the performance of clinicians in the differential diagnosis of masses on breast ultrasonography was significantly improved when they could view the likelihood of different malignancies that had been determined by a CAD scheme [6]. Clinicians usually take the histological classification of a lesion into account in their differential diagnosis of lesions on breast ultrasonographic images. Kashikura et al. demonstrated that presenting the likelihood of a histological classification, as determined by a CAD scheme, improved the performance of clinicians in the differential diagnosis of breast masses detected by breast ultrasonography [7].

Breast ultrasonographic images often show both masses that are defined as solid lesions, which have an evaluable shape, margin and size, and non-mass lesions, which are difficult to identify as masses and in which the shape and margin are indistinct. It is more difficult to differentially diagnose non-mass lesions than mass lesions due to their indistinct shape and margin. Non-mass lesions appear as lesions that are mostly identified based on abnormalities of the ducts, hypoechoic areas in the mammary gland, architectural distortion, multiple small cysts, or echogenic foci without a hypoechoic area. Among these, hypoechoic areas in the mammary gland are most commonly encountered in clinical practice. In the present study, we investigated whether presenting the likelihood of the different histological classifications determined by a CAD scheme would improve the performance of clinicians in the differential diagnosis of non-mass lesions appearing as hypoechoic areas in the mammary gland on breast ultrasonographic images.

Materials and methods

Case selection

The database included 97 breast ultrasonographic images of non-mass lesions appearing as hypoechoic areas in the mammary gland, including 48 benign cases and 49 malignant cases. All of the cases were considered to correspond to the five histological classifications that are often encountered in clinical practice and which are sometimes considered to be non-mass lesions appearing as hypoechoic areas in the mammary gland. The classifications of the 48 benign cases were as follows: benign lesion with benign mammary tissue or fibrocystic disease (n=20), fibroadenoma (n=11), and intraductal papilloma (n=17). The classifications of the 49 malignant cases were as follows: ductal carcinoma *in situ* (n=17) and invasive ductal carcinoma (n=32). The images were acquired at Mie University Hospital in 2013-2014 using an ultrasound diagnostic system (APLIO XG SSA-790A, Toshiba Medical Systems Corp.). We selected images with hypoechoic areas in the mammary gland that were difficult to identify as masses. These hypoechoic areas were defined by the clinicians who usually engage in the diagnosis of the breast ultrasonographic images and who did not participate in this observer study.

All of the diagnoses were pathologically confirmed by core needle biopsy, vacuum-assisted biopsy, or excisional biopsy after the initial diagnosis by breast ultrasonography. Each of the benign cases selected had been reexamined at 6 -12 months after the initial diagnosis, at which time it was confirmed that the ultrasonographic images had not changed. Images that were captured after vacuum assisted biopsy or excisional biopsy, or after the administration of medication, were excluded due to the possible influence of the examinations or therapies.

CAD scheme

In the present study, the rectangular region of interest (ROI), which included an entire non-mass, was selected manually in ultrasonographic images, as shown in Fig.1a. The non-mass was automatically segmented with a level set method based on an active contour model that used the ROI as the initial contour of the non-mass region, as shown in Fig.1b.

In our previous study [8, 9], we selected nine objective features that could be used to histologically classify masses by taking into account the differences in the features on the ultrasonographic images of the different histological classifications. We then showed the statistical significance of the identification of the histological classification using these objective features. Thus, in the present study, we used the same objective features as those used in our previous study. These objective features were (1) the depth-width ratio, (2) the degree of indistinctness along the margin, (3) the homogeneity in the internal echoes, (4) the echo level of the internal echoes, (5) the echo level of the posterior echoes, (6) the circularity measure of the non-mass shape, (7) the polygon measure the non-mass shape, (8) the lobulated shape measure of the non-mass shape, and (9) the irregularity measure of the non-mass shape.

A classifier based on a *k*-nearest neighbor (*k*-NN) algorithm that included the nine objective features was employed to distinguish five different histological classifications of non-mass lesions. In the *k*-NN algorithm, a test case was classified by a majority vote of the histological classifications of its *k* neighbors in the feature space. A leave-one-out-by-patient test method was employed for the training and testing of the classifier. In this method, the training was carried out for all but one case from the database; the case that was not used for training was used for the testing of the trained classifier. This procedure was repeated until every patient case in our database had been used once.

Observer study

To evaluate the usefulness of the CAD scheme [8, 9], we conducted an observer study using 97 ultrasonographic images that included non-mass lesions appearing as hypoechoic areas in the mammary gland. Seven clinicians participated in the observer study, including three experienced breast surgeons who each had more than 5 years of experience in the diagnosis of breast images in a university hospital (Expert group) and four general surgeons who had 3-21 years of experience and who sometimes diagnosed breast images in a local hospital (General group).

In the observer study, an image was first displayed on a monitor without the CAD results. The observers were asked to indicate their confidence level with regard to the possibility of malignancy using a continuous rating scale of 0-1, where 0 was “definitely benign” and 1 was “definitely malignant”. After indicating their confidence level without viewing the CAD results, the observer could view the CAD results, which displayed the likelihood of the five histological classifications as a bar graph. The observers referred to the CAD results and indicated their confidence level again. In this manner, the observers indicated their confidence levels for all 97 cases.

We informed the observers that the purpose of this study was to evaluate whether the accuracy of the differential diagnosis of non-mass lesions appearing as hypoechoic areas in the mammary gland on breast ultrasonography could be improved by presenting the likelihood of the different histological classifications. The observers were also informed that there was no time limit for this study. We did not inform the observers of the details about the differential diagnosis ability of the CAD scheme, such as the sensitivity or the specificity levels in the diagnosis of the benign and malignant cases, or the accuracy levels of the histological classifications. Furthermore, we did not inform the observers about the number of cases that were histologically classified as benign or malignant.

Statistical analysis

We used receiver operating characteristic (ROC) analysis based on a sequential-test method to analyze the results of the observer study [10, 11]. The AUCs were obtained using the DBM MRMC software package (version 2.2) [12], which was developed by researchers at the University of Iowa and the University of Chicago. The significance of the differences in the AUCs of the observers with and without the use of the CAD scheme was tested using the Dorfman-Berbaum-Metz method [12]. P values of less than 0.05 were considered to indicate statistical significance.

Results

The specificity and the sensitivity of the CAD scheme in the diagnosis of the 97 cases used in the observer study were 89.5% (43/48) and 87.8% (43/49), respectively. The AUC of the CAD scheme was 0.93. The accuracy of the histological classifications was as follows: benign lesions (80.0% [16/20]), fibroadenoma (81.8% [9/11]), intraductal papilloma (70.6% [12/17]), ductal carcinoma *in situ* (82.3% [14/17]), and invasive ductal carcinoma (84.3% [27/32]).

Table 1 shows the average AUC for each observer, each observer group, and all of the observers with and without the use of the CAD scheme. In the general group, the AUC of Observer A increased significantly from 0.661 without the CAD scheme to 0.942 with the CAD scheme($P < 0.0001$), while the AUC of Observer B increased slightly from 0.617 without the CAD scheme to 0.652 with the CAD scheme($P=0.198$). The AUCs of the observers in the expert group showed less variation than the AUCs of the observers in the general group.

Figure 2 shows the average ROC curves for all of the observers in the diagnosis of benign and malignant non-mass lesions that appeared as hypoechoic areas in the mammary gland with and without the use of the CAD scheme. The average AUC for all observers increased from 0.649 without the CAD scheme to 0.783 with the CAD scheme ($P = 0.0167$). The diagnostic accuracy of all observers improved when they were presented with the likelihood of the histological classification.

Figure 3 shows the average ROC curves for the general group and the expert group with and without the use of the CAD scheme. The average AUCs for both groups were improved by the use of the CAD scheme. The average AUC of the general group increased from 0.625 without the CAD scheme to 0.793 with the CAD scheme ($P = 0.045$), with the difference being slightly significant. Although the average AUC of the expert group also increased from 0.681 without the CAD scheme to 0.769 with the CAD scheme ($P = 0.327$), the difference was not statistically significant.

Figure 4 shows the average of the beneficial or detrimental changes in the confidence levels of all of the observers for each of the cases with the use of the CAD scheme. We assumed that the CAD scheme had a beneficial effect on an observer’s diagnosis when the change in the average confidence levels with and without the use of the CAD scheme was >0.1. We assumed that the use of the CAD scheme had a detrimental effect when the change in the average confidence levels with and without the CAD scheme was <-0.1. The use of the CAD scheme was found to have a beneficial effect in 48 cases and a detrimental effect in five cases. The histological classifications in the beneficial cases were as follows: benign lesion (n=11), fibroadenoma (n=7), intraductal papilloma (n=8), ductal carcinoma *in situ* (n=9) and invasive ductal carcinoma (n=13). The histological classifications in the detrimental cases were as follows: ductal carcinoma *in situ* (n=1) and invasive ductal carcinoma (n=4).

Table 2 shows the numbers of beneficial and detrimental cases for each observer, each group, and all observers. The number of beneficial cases was much larger than the number of detrimental cases in all observers.

Figure 5 shows a case of ductal carcinoma *in situ* in which the use of the CAD scheme had a beneficial effect in all observers. The likelihood of the different histological classifications, as determined by the CAD scheme, was as follows: benign lesion (1%), fibroadenoma (18%), intraductal papilloma (2%), ductal carcinoma *in situ* (71%) and invasive ductal carcinoma (7%). With the use of the CAD scheme, the average confidence levels improved from 0.63 to 0.76 in the general group, from 0.65 to 0.79 in the expert group and from 0.64 to 0.77 in all observers.

Figure 6 shows a case of invasive ductal carcinoma in which the beneficial effect was most pronounced in the general group. The likelihood of the different histological classifications, as determined by the CAD scheme, was as follows: benign lesion (9%), fibroadenoma (9%), intraductal papilloma (8%), ductal carcinoma *in situ* (10%) and invasive ductal carcinoma (63%). With the use of the CAD scheme, the average confidence levels improved from 0.70 to 0.88 in the general group, whereas the confidence level of the expert group only showed a slight improvement from 0.87 to 0.93.

Figure 7 shows a case of invasive ductal carcinoma in which CAD had a detrimental effect on all observers. The likelihood of the different histological classifications, as determined by the CAD scheme, was as follows: benign lesion (7%), fibroadenoma (8%), intraductal papilloma (76%), ductal carcinoma *in situ* (7%) and invasive ductal carcinoma (2%). This detrimental effect was caused by an incorrect evaluation by the CAD scheme.

Discussion

In recent years, the improved resolution and image quality of ultrasonographic images have allowed for the increased detection of breast lesions of various morphologies. In clinical practice, clinicians encounter an increased number of lesions that are difficult to define as masses with apparent margins [13]. It is easy to perform whole breast screening for Asian women with small breasts and to observe their breast lesions in detail. Thus, there is a greater opportunity to identify lesions of various morphologies. According to the guidelines of the Japan Association of Breast and Thyroid Sonography [14], lesions on breast ultrasonography are divided into masses and non-mass lesions. A mass is defined as a solid lesion that is distinguishable from the surrounding tissue and that has an evaluable shape, margin, and size. Non-mass lesions appear as lesions that are difficult to identify as masses and that are mostly identified based on abnormalities of the ducts, hypoechoic areas in the mammary gland, architectural distortion, multiple small cysts or echogenic foci without a hypoechoic area. On the other hand, the lexicon of features on breast ultrasonography published by the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) [15] defines three types of lesions: masses, calcification and special cases, and does not include non-mass lesions. However, "non-mass-like enhancement" has been added to the lexicon of breast magnetic resonance imaging (MRI) [16]. Sotome et al. reported that 95% of non-mass lesions on breast ultrasonography appeared as non-mass-like enhancements on MRI [17]. We are therefore of the opinion that non-mass lesions should be added to the lexicon of ultrasonography. Benign mammary tissue or fibrocystic disease, fibroadenoma, intraductal papilloma, ductal carcinoma *in situ* and invasive ductal carcinoma sometimes appear as non-mass lesions, appearing as hypoechoic areas in the mammary gland on breast ultrasonography. It is more difficult to differentially diagnose non-mass lesions than mass lesions because their shape and margins are indistinct. We are of the opinion that a CAD scheme for non-mass lesions would be very useful for improving the diagnostic accuracy of breast ultrasonography.

In the observer study, the AUCs of all observers improved when they were presented with the likelihood of the different histological classifications. The average AUC of the general group showed greater improvement than that of the expert group. We therefore, considered that the general clinicians were more inclined to follow the CAD scheme than the experienced breast surgeons. Hence, the CAD scheme would be particularly useful for improving the diagnostic accuracy of general clinicians. The AUC of Observer A was higher than that of the CAD scheme. We considered that Observer A was able to understand the tendency of the CAD results and to use them well because he was the youngest doctor and was used to dealing with computers. The AUCs of the observers in the general group varied widely. The degrees of confidence and proficiency in the use of computers may have influenced their CAD performance. Although the average AUC in the expert group improved with the use of the CAD scheme, the difference was not statistically significant. The AUCs of each of the observers in the expert group showed less variation than the general group. Perhaps the experts did not accept the CAD results of the CAD scheme without questioning them.

Figure 5 shows a case of ductal carcinoma *in situ* in which the CAD scheme showed a beneficial effect in all observers. All of the observers rated the likelihood of malignancy as slightly high or uncertain without the use of the CAD scheme. However, the CAD scheme clearly indicated that ductal carcinoma *in situ* was the most likely of the five histological classifications. The CAD results allowed the observers to confidently rate the likelihood of malignancy as higher.

Figure 6 shows a case of invasive ductal carcinoma in which a particularly pronounced beneficial effect was observed in the general group. Although one of observers in the general group answered that this case might be benign without the use of the CAD scheme, his confidence level improved and he rated the case as malignant when he used the CAD scheme. The other observers in the general group also confidently rated the case as being more likely to be malignant with the use of the CAD scheme. Echogenic foci were present inside the lesion of this case. The experienced clinicians identified the echogenic foci and were able to easily determine that the lesion was malignant. The CAD scheme was helpful for the inexperienced general clinicians who could not identify the echogenic foci. In most cases, the CAD scheme helped to improve the diagnostic accuracy of the less experienced clinicians and the performance of experienced clinicians in cases in which it was difficult to decide whether a lesion was benign or malignant.

The CAD scheme was not effective in improving the AUCs in one case of ductal carcinoma *in situ* and four cases of invasive ductal carcinoma. The CAD scheme misdiagnosed four invasive ductal carcinomas as benign cases [fibroadenoma (n=3) and intraductal papilloma (n=1)]. Figure 7 shows a case that was indicated as being possibly malignant by all observers without the use of the CAD scheme. However, the CAD scheme incorrectly determined that intraductal papilloma was the most likely of the five histological classifications. This result had a detrimental effect on the confidence levels of the observers. Although the CAD scheme that was used in the present study had a high level of sensitivity, it showed the lowest level of sensitivity in the diagnosis of intraductal papilloma (70.6%). Intraductal papillomas are often difficult to differentially diagnose as malignant lesions, especially from ductal carcinoma *in situ*, because they are both intraductal lesions and because their imaging features are similar. It is also often difficult to assume the histological classification and to precisely differentiate benign lesions from malignant lesions.

Furthermore, the differential diagnosis of benign or malignant lesions by a CAD scheme using a continuous rating scale is associated with a limitation and a problem. In the present study, the numerical value that indicated the clinicians' confidence level changed easily on different occasions and was not replicable.

We are therefore of the opinion that, in addition to calculating the likelihood of the different histological classifications, the CAD scheme must be expanded to allow it to determine the BI-RADS category classifications. The use of the category classifications in the diagnosis of lesions would allow for a more uniform diagnosis and could be adapted for use in clinical practice. In a future study, we will develop a new CAD scheme to be used in the diagnosis of both mass and non-mass lesions, and which will also be able to diagnose breast lesions using the BI-RADS category classifications.

Conclusion

The ability to view the likelihood of the histological classifications of non-mass lesions appearing as hypoechoic areas in the mammary gland improved the performance of the clinicians in the present study. In clinical practice, the CAD scheme would be useful for improving diagnostic ability for non-mass lesions appearing as hypoechoic areas in the mammary gland on breast ultrasonographic images.

Conflicts of Interest

The authors declare that they have no conflicts of interest in association with the present study.

Ethical Standards

This study was performed at Mie University Hospital, and was approved by the Ethical Review Board of Mie University Hospital. Informed consent was obtained from all patients for being included in the study.

References

1) Crystal P, Strano SD, Shcharynski S, et al. Using sonography to screen women with mammographically dense breasts. AJR. 2003;181:177-82.
2) Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomized controlled trial. Lancet. 2015. doi: 10.1016/S0140-6736(15)00774-6.
3) Tohno E, Ueno E, Watanabe H. Ultrasound screening of breast cancer. Breast Cancer. 2009;16:18-22.
4) Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph. 2007;31:198-211.
5) Nakayama R, Namba K, Watanabe R, et al. Potential usefulness of presentation of histological classifications with computer-aided diagnosis (CAD) scheme in differential diagnosis of clustered microcalcifications on mammograms. Breast Imaging: 12th International Workshop, IWDM 2014, Gifu City, Japan, June 29 - July 2, 2014;8539:115-20.
6) Ying W, Hong W, Yanhui G, et al. Novel computer-aided diagnosis algorithms on ultrasound image: Effects on solid breast masses discrimination, J Digit Imaging. 2010;23:581-91.
7) Kashikura Y, Nakayama R, Hizukuri A, et al. Improved differential diagnosis of breast masses on ultrasonographic images with a computer-aided diagnosis scheme for determining histological classifications. Acad Radiol. 2013;20:471-7.
8) Hizukuri A, Nakayama R, Kashikura Y, et al. Computer-aided diagnosis scheme for determining histological classification of breast masses using clinical features and texture features. Proc. of the Forth International Workshop on Regional Innovation Studies - Biomedical Engineering-. 2012;83-6.
9) Hizukuri A, Nakayama R, Kashikura Y, et al. Computerized Determination Scheme for Histological Classification of Breast Mass Using Objective Features Corresponding to Clinicians' Subjective Impressions on Ultrasonographic Images. Journal of Digital Imaging. 2013;26: 958-70.
10) Kobyashi T, Xu XW, Macmahon H, et al. Effect of a computer-aided

diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. Radiology. 1996;199:843-8.

11) Uozumi T, Nakamura K, Watanabe H, et al. ROC analysis of detection of metastatic pulmonary nodules on digital chest radiographs with temporal subtraction. Acad Radiol. 2001;8:871-8.
12) Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generralization to the population of readers and patients with the jackknife method. Invest Radiol. 1992;27:723-31.
13) 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.
14) Japan Association of Breast and Thyroid Sonography (JABTS). Guideline for breast ultrasound: management and diagnosis, 3rd edition. Nankodo, Tokyo: 2014. p.63-91
15) American College of Radiology(ACR). Breast Imaging Reporting and Data System, Ultrasound. 5th edition. Reston, VA ; 2013. p.1-153.
16) American College of Radiology(ACR). Breast Imaging Reporting and Data System, Magnetic resonance imaging. 4th edition. Reston, VA ; 2003. p.1-114.
17) Sotome K, Yamamoto Y, Hirano A, et al. The role of contrast enhanced MRI in the diagnosis of non-mass image-forming lesions on breast ultrasonography. Breast Cancer. 2007;14:371-80.

Figure legend

Figure 1(a). The manual setting of the region of interest (ROI) in a case involving a benign lesion with benign mammary tissue or fibrocystic disease.

Figure 1(b). The automatic segmentation result of the non-mass lesion with a level set method in a case involving a benign lesion with benign mammary tissue or fibrocystic disease.

Figure 2. The average ROC curves for all observers in distinguishing between benign and malignant non-mass lesions with and without the use of the CAD scheme.
The average area under the ROC curve (AUC) improved from 0.649 to 0.783 with the use of the CAD scheme (P=0.0167).

Figure 3. The average ROC curves for the general and expert groups in distinguishing between benign and malignant non-mass lesions with and without the use of the CAD scheme. The AUC improved from 0.625 to 0.793 in the general group (P=0.045), and from 0.681 to 0.769 in the expert group (P=0.327).

Figure 4. The average of the beneficial or detrimental changes in the confidence levels with the CAD scheme in each of the cases for all observers. The vertical axis shows changes in the confidence levels before and after the use of the CAD scheme, the horizontal axis shows the case number. When the change in the average confidence levels with and without the use of the CAD scheme was > 0.1 or < -0.1, the CAD scheme was considered to have had a beneficial or detrimental effect, respectively.

Figure 5. A case of ductal carcinoma *in situ* in which the CAD scheme was effective for all observers.

Figure 6. A case of invasive ductal carcinoma in which the CAD scheme was only effective in the general group.

Figure 7. A case of invasive ductal carcinoma in which the CAD scheme was ineffective.

Figure1a

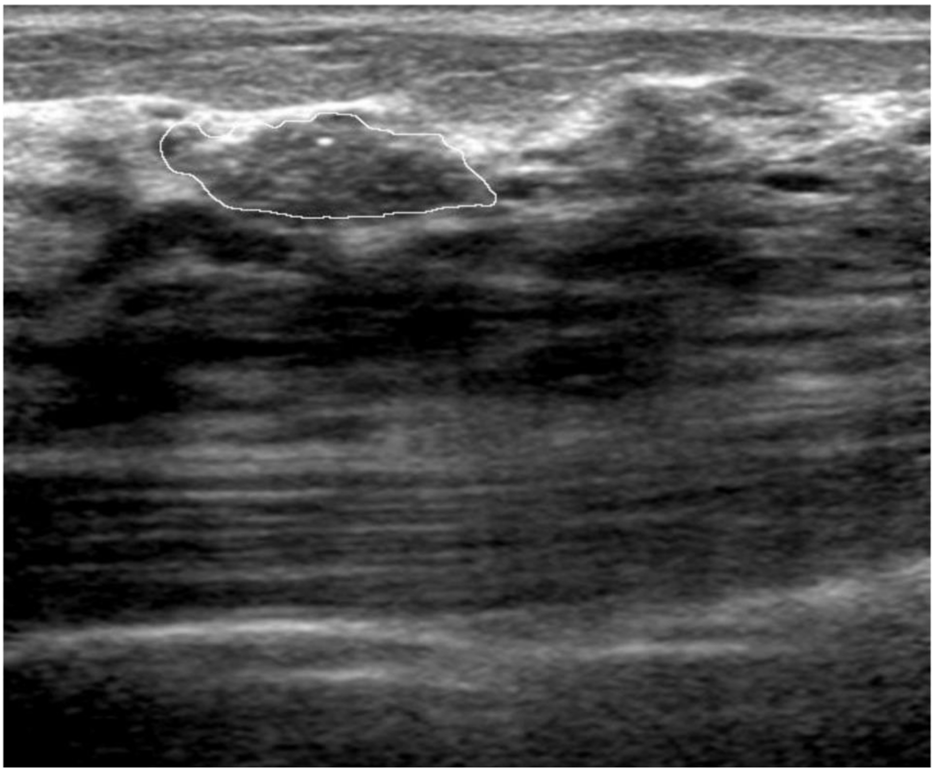


Figure1b

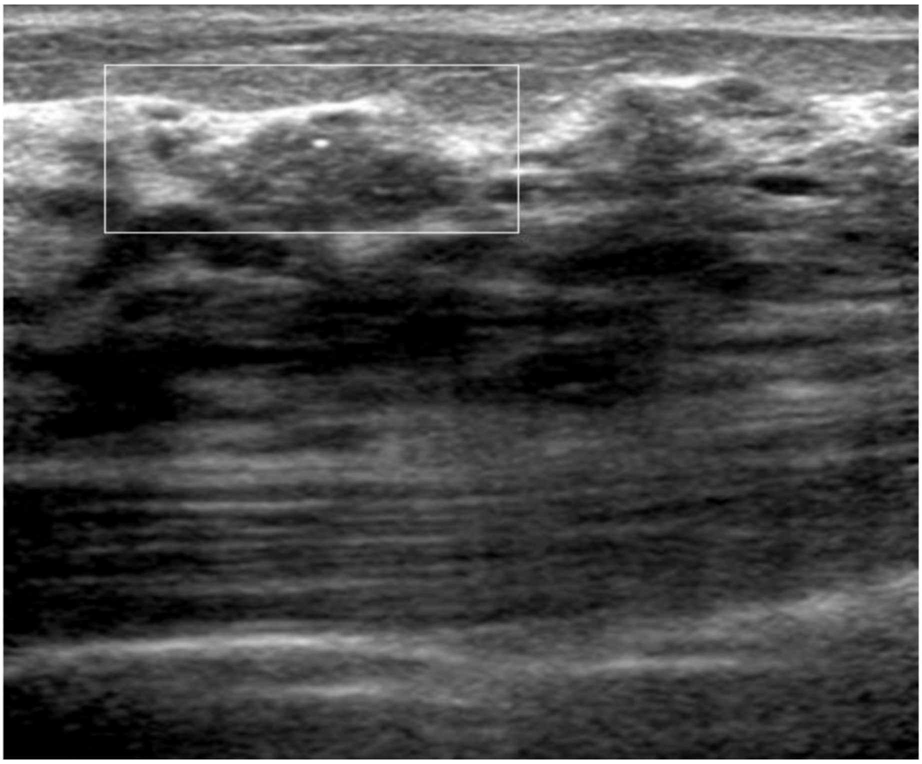


Figure2

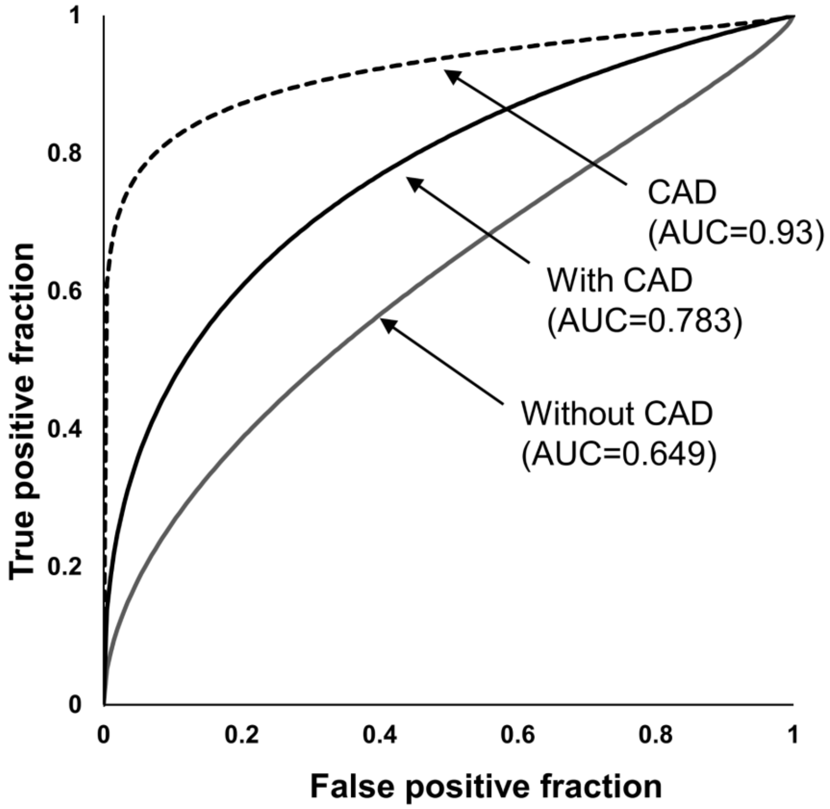


Figure3

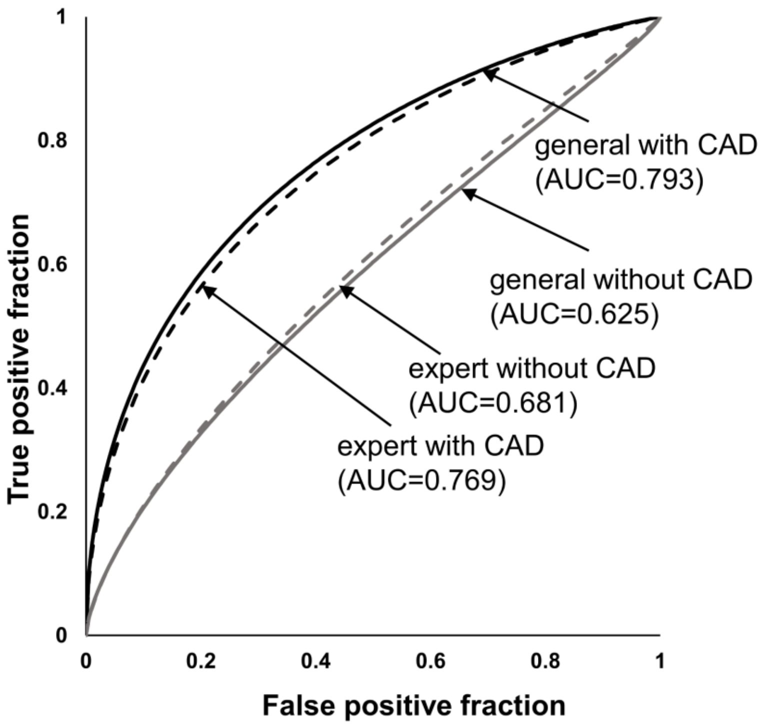


Figure4

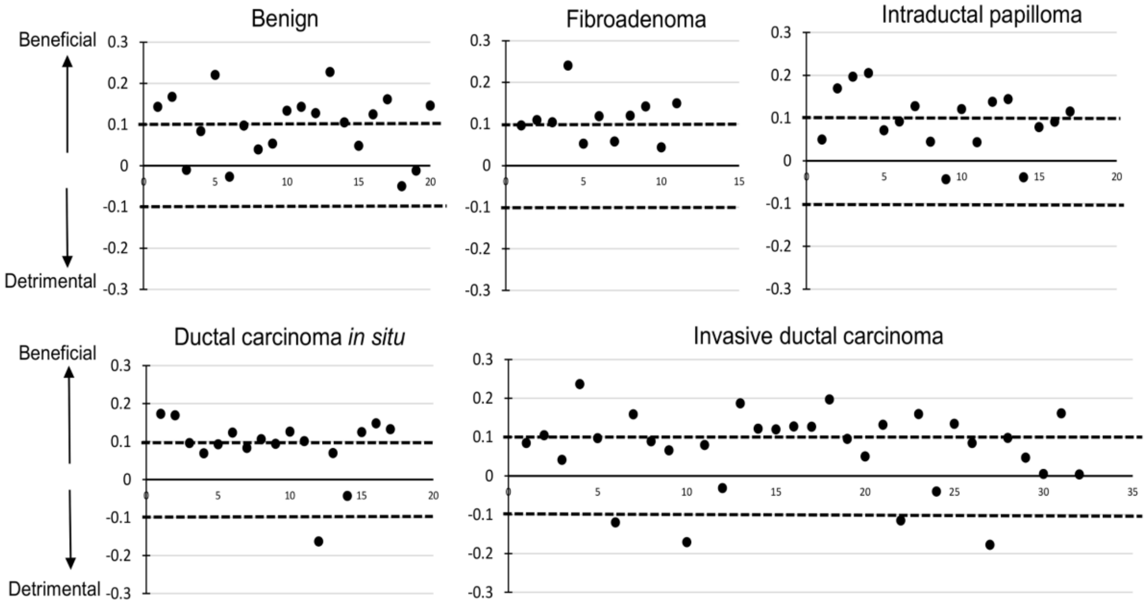


Figure5

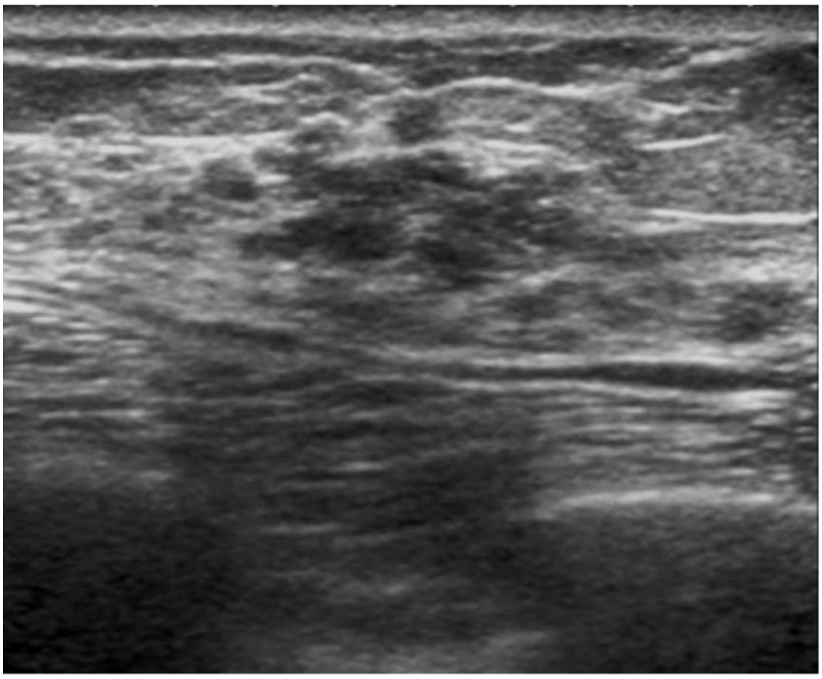


Figure6

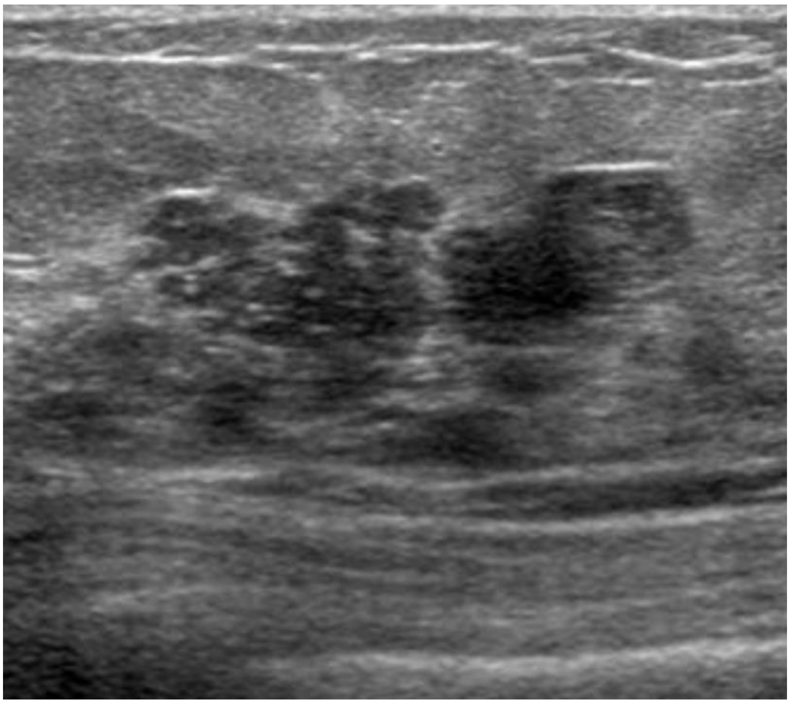
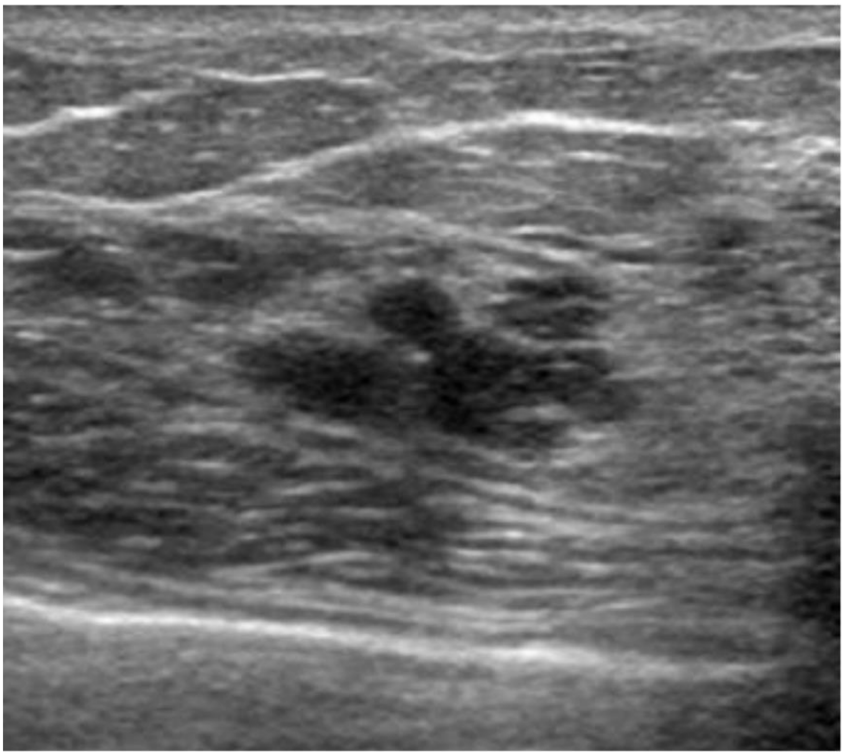


Figure7



	AUC		
	Without CAD	With CAD	P
General			
A	0.661	0.942	< 0.0001
B	0.617	0.652	0.198
C	0.627	0.832	< 0.0001
D	0.593	0.746	0.003
Mean	0.625	0.793	0.045
Expert			
E	0.742	0.756	0.264
F	0.603	0.828	< 0.0001
G	0.697	0.722	0.0503
Mean	0.681	0.769	0.327
All			
Mean	0.649	0.783	0.0167

Table 1. The AUCs for each observer, each observer group and all observers with and without CAD

	Beneficial	Detrimental
General		
A	67	7
B	7	1
C	48	9
D	40	4
Mean	56	4
Expert		
E	9	1
F	60	7
G	3	0
Mean	32	1
All		
Mean	58	5

Table 2. The numbers of beneficial and detrimental cases for each observer, each group and all observers