European Radiology (2019) 29:2518–2525
https://doi.org/10.1007/s00330-018-5886-0

BREAST

Decrease in interpretation time for both novice and experienced readers using a concurrent computer-aided detection system for digital breast tomosynthesis

Eun Young Chae [1] · Hak Hee Kim [1] · Ji-wook Jeong [2] · Seung-Hoon Chae [2] · Sooyeul Lee [2] · Young-Wook Choi [3]

Received: 21 September 2018 / Accepted: 13 November 2018 / Published online: 13 December 2018

Abstract

Objectives To compare the diagnostic performance and interpretation time of digital breast tomosynthesis (DBT) for both novice and experienced readers with and without using a computer-aided detection (CAD) system for concurrent read.

Methods CAD system was developed for concurrent read in DBT interpretation. In this observer performance study, we used an enriched sample of 100 DBT cases including 70 with and 30 without breast cancers. Image interpretation was performed by four radiologists with different experience levels (two experienced and two novice). Each reader completed two reading sessions (at a minimum 2-month interval), once with and once without CAD. Three different rating scales were used to record each reader's interpretation. Reader performance with and without CAD was reported and compared for each radiologist. Reading time for each case was also recorded.

Results Average area under the receiver operating characteristic curve values for BI-RADS scale on using CAD were 0.778 and 0.776 without using CAD, demonstrating no statistically significant differences. Results were consistent when the probability of malignancy and percentage probability of malignancy scales were used. Reading times per case were 72.07 s and 62.03 s (SD, 37.54 s vs 34.38 s) without and with CAD, respectively. The average difference in reading time on using CAD was a statistically significant decrease of 10.04 ± 1.85 s, providing 14% decrease in time. The time-reducing effect was consistently observed in both novice and experienced readers.

Conclusion DBT combined with CAD reduced interpretation time without diagnostic performance loss to novice and experienced readers.

Key Points

- *The use of a concurrent DBT-CAD system shortened interpretation time.*
- *The shortened interpretation time with DBT-CAD did not come at a cost to diagnostic performance to novice or experienced readers.*
- *The concurrent DBT-CAD system improved the efficiency of DBT interpretation.*

Keywords Digital breast tomosynthesis · Computer-assisted diagnosis · Breast cancer

✉ Hak Hee Kim
hhkim@amc.seoul.kr

[1] Department of Radiology, Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-Gil, Songpa-Gu, Seoul 05505, Republic of Korea

[2] Medical Imaging Research Section, Electronics & Telecommunications Research Institute, 218 Gajeongno, Yuseong-gu, Daejeon 34129, Republic of Korea

[3] Advanced Medical Device Research Division, Korea Electrotechnology Research Institute, 111 Hanggaul-ro, Sangnok-gu, Ansan-si, Gyeonggi-do 15588, Republic of Korea

Abbreviations

CAD	Computer-aided detection
DBT	Digital breast tomosynthesis
FFDM	Full-field digital mammography

Introduction

Although there are obvious limitations, mammography is the only proven screening modality that has been shown to reduce breast cancer mortality in randomized trials [1, 2]. A limitation of full-field digital mammography (FFDM) is the superimposition of breast tissue, which can obscure cancers or make normal structures appear suspicious [3, 4]. Therefore, the sensitivity and specificity of screening mammography are limited in women with dense breast tissues [5–7]. Digital breast tomosynthesis (DBT) may be an effective screening tool for these women. By reducing the effect of tissue superimposition, DBT allows better visualization of malignancy and also facilitates discrimination between normal tissue structures and breast cancers. Compared with FFDM alone, DBT in combination with FFDM has shown increased sensitivity, reduced recall rates, and fewer false positives in breast cancer detection [8–11].

Interpretation of DBT studies may be challenging for radiologists partially because it involves the analysis of a series of thin slices through the breast, which requires longer reading time than that associated with FFDM [9, 12]. To address the challenges of reviewing DBT studies, computer-aided detection (CAD) software may help radiologists detect cancers and improve efficiency in the DBT interpretation workflow. However, to date, few studies have investigated the effect of using CAD systems for DBT [13, 14]. We hypothesized that the potential benefit of a CAD system is to reduce the overall interpretation time while maintaining better diagnostic performance of radiologists compared with FFDM. To the best of our knowledge, there have been no published reports regarding the investigation of DBT-CAD systems according to experience levels of radiologists. In this study, we aimed to develop and validate the new DBT-CAD system. For this, we evaluated a new DBT-CAD system in an observer performance study and compared the diagnostic accuracy of DBT interpretation and reading time for novice and experienced readers with and without the CAD system.

Materials and methods

Data set

This retrospective study was approved by the institutional review board, and the informed consent requirement was waived. Between March 2016 and June 2016, a total of 100 consecutive patients underwent DBT examinations at our hospital. For studying observer performance, these 100 consecutive cases who underwent screening ($N = 14$) or diagnostic ($N = 86$) DBT examinations constituted our study population. This case series comprised 70 biopsy-proven breast cancers, and 30 non-cancer cases consisting of 17 benign and 13 negative cases. The reference standard was biopsy-proven histopathology ($N = 76$) or clinical imaging follow-up for at least 1 year ($N = 24$). DBT examinations were performed bilaterally ($N = 60$) or unilaterally ($N = 40$) using a commercially available device (SenoClaire, GE Healthcare). Craniocaudal (CC) and mediolateral oblique (MLO) views of each breast were obtained. The system acquired nine projections along a -12.5 to + 12.5 degree arc using a step-and-shoot technique. Images were reconstructed by iterative reconstruction (ASiR, Adaptive Statistical Iterative Reconstruction, GE Healthcare).

Computer-aided detection system

Lesion detection step

In this study, we used a newDBT-CAD system, which was developed for this study. First, 3D DBT images were softly binarized using a customized Sigmoidal function to better highlight breast lesions compared with the neighboring fibroglandular breast tissues; subsequently, an edge detection step using a Canny-edge detection algorithm was applied. The 3D Hough transform technique generated a list of Hough spheres by detecting locally spherical edges around a breast lesion, after which the Hough spheres were converted into circumscribing parallelepiped cubes to effectively cover the entire range of examined spheres.

The Hough transform [15, 16] is widely used in the detection of various types of figures, for example, by exploiting the duality between points on a curve and its parameters. The resulting Hough cube is considered as a candidate for the 3D lesion image and labeled normal or lesion whether or not there is an overlap with the ground truth.

False-positive (FP) reduction step

We performed an automatic FP reduction of a list of mass candidate cubes, extracted from Hough transform-based lesion detection step (B) using a convolutional neural networks (CNN). In this study, we used a well-known CNN model ResNet-50 without modifying the network architecture [17]. In this study, the ResNet-50 convolutional neural network is used because of its computational simplicity, better performance in a wide range of machine learning processes, and it especially works well on unsupervised dataset.

For the train database of images, a rescaled thumbnail consisting of 128×128 pixels in grayscale space was created from each parallelepiped candidate cube. After data augmentation, such as the typical horizontal flipping and rotation processes for solving the class imbalance problem of our dataset, input images were labeled false-positive and true-positive regardless of an overlap with the ground truth. These images were again partitioned into a training set (60%), a testing set (20%), and a validation set (20%). We also used the k-fold cross-validation to estimate classification accuracy on a relatively small data set. In this study, we adopted $k = 5$ for simplicity.

To enhance the detection accuracy of breast tumor mass, we applied transfer learning techniques with an additional dataset consisting of artificially generated thumbnail images (128×128

pixels), containing mass-like 2D objects and various noise and background gray levels. For first-stage transfer learning, we loaded pre-trained weights for the ResNet-50 model on ImageNet database and subsequently performed second-stage transfer learning on mass-like 2D object dataset described above. To enhance its training efficiency, the intermediate layers above the first input layer were frozen except for the last six convolutional layers. For the final-stage FP reduction step, the resulting confidence scores for thumbnail images were examined to merge the bounding boxes of mass candidates.

The ResNet-50 model in this study were trained on the 393 DBT image volumes of 200 breasts containing 181 tumor masses, which produces an FP rate of 3.42 per DBT volume at a sensitivity of 90% [18].

Image interpretation

Image interpretation was performed by four radiologists with different experience levels in breast imaging and DBT. Two of the radiologists (readers 1 and 2) were classified as experienced readers, and the remaining two (readers 3 and 4) were novice readers. Two novice readers were still in the training of breast imaging after completing a radiology residency program. They had a year of experience in breast radiology. Experienced readers were fellowship-trained, dedicated breast radiologists with 14 and 6 years of clinical experience in breast imaging, reading at least 5000 mammograms each year. They had at least 6 years of experience in DBT and had participated in prior studies involving the interpretation of DBT images [19, 20]. We used a viewer system to review DBT images. Each reader completed two reading sessions, once with and once without CAD. The predetermined minimum interval between the reading sessions was 2 months to reduce any possible bias caused by the possibility of readers remembering what they had previously read. Each case was de-identified and randomized between the reading sessions. Readers were blinded to the readings of other radiologists, clinical information, and histological diagnosis.

The readers were asked to detect the presence or absence of abnormalities in each reading session. Three different scales were used to record each reader's interpretation: modified BI-RADS (BI-RADS scale without category 0), probability of malignancy scale (1, definitely not cancer; 2, almost certainly not cancer; 3, probably not cancer; 4, possibly cancer; 5, probably cancer; 6, almost certainly cancer; 7, definitely cancer), and percentage probability of malignancy scale (0–100%, i.e., each reader's subjective assessment of the probability that malignancy was positive). These scales were used to calculate diagnostic performance and derive receiver operating characteristic (ROC) curves. They also recorded the reading time with a stopwatch in each case, measured in seconds. Reading time only included the time to review DBT images and did not include the time to record interpretation results.

Data analysis

The primary endpoint was the overall diagnostic performance based on the area under the ROC curve (AUC) using three different scales. The 95% confidence intervals (CIs) of various AUC scores were determined by bootstrapping. Secondary endpoints included sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) according to the modified BI-RADS scale. Modified BI-RADS scale was dichotomized to compute standard operating metrics (e.g., sensitivity, specificity, PPV, and NPV), with a BI-RADS score of 4 or higher that indicates a positive test result. The McNemar test was used to compare diagnostic performance, and CIs of these values were based on a robust variance estimator. The reading times with or without CAD were compared using paired *t* test. Generalized estimating equations with logit link were used to account for the correlation in interpretation times due to repeat reads of an image within and across readers. Statistical analyses were

Table 1 Area under the curve (AUC) for novice and experienced readers with and without computer-aided detection (CAD)

Scales	Experience level	Average AUC (95% CI)		Difference (standard error)	*p* value
		Without CAD	With CAD		
BI-RADS	Novice	0.752 (0.688–0.816)	0.758 (0.694–0.822)	0.006 (0.021)	0.784
	Experienced	0.800 (0.738–0.861)	0.796 (0.735–0.858)	-0.003 (0.019)	0.866
	All	0.776 (0.732–0.821)	0.778 (0.733–0.822)	0.001 (0.014)	0.924
Probability of malignancy	Novice	0.772 (0.709–0.834)	0.766 (0.700–0.832)	-0.005 (0.020)	0.787
	Experienced	0.808 (0.746–0.871)	0.818 (0.758–0.878)	0.010 (0.018)	0.593
	All	0.788 (0.744–0.833)	0.791 (0.746–0.836)	0.006 (0.013)	0.844
Percentage probability of malignancy	Novice	0.779 (0.718–0.840)	0.784 (0.722–0.847)	0.005 (0.018)	0.771
	Experienced	0.812 (0.750–0.873)	0.828 (0.768–0.887)	0.016 (0.019)	0.391
	All	0.794 (0.750–0.838)	0.806 (0.762–0.849)	0.012 (0.013)	0.376

CI confidence interval

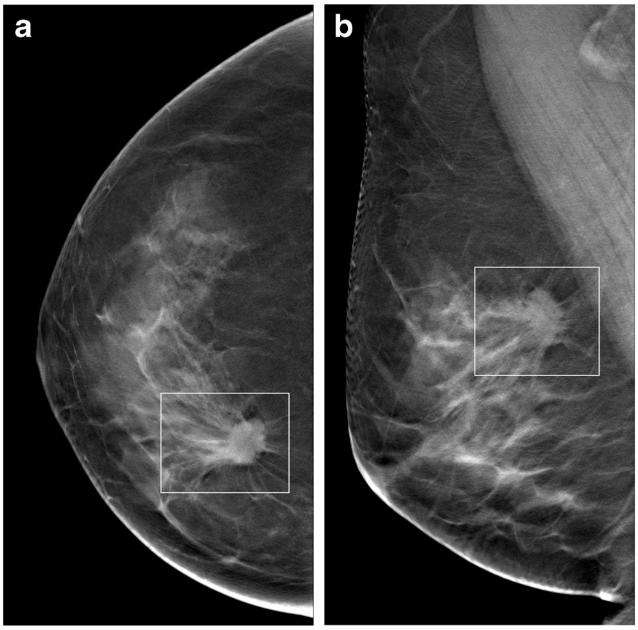

Fig. 1 Digital breast tomosynthesis images in a 55-year-old woman with an invasive ductal carcinoma show a spiculated mass in craniocaudal (**a**) and mediolateral oblique (**b**) views, with superimposed CAD markers (squares)

performed using the SPSS software (version 23.0, Statistical Package for the Social Sciences, Chicago, IL). A p value of < 0.05 was considered to be statistically significant.

Results

Of the 100 women (age range, 39–81 years; mean age, 52.5 years) selected for this study, 70 had biopsy-proven breast cancer. The majority of patients had heterogeneously dense ($n = 48$) or extremely dense ($n = 34$) breast tissue, while 15 patients had scattered areas of fibroglandular density and three patients had almost entirely fatty breasts. Nine of 100 women had undergone previous breast operation. Among these 70 breast cancer patients, the mean tumor size was 23.2 ± 15.3 mm (range, 6–78 mm). The tumors were predominantly invasive ductal carcinomas (59/70; 84.3%); others included 6 ductal carcinomas in situ, 2 invasive lobular carcinomas, 2 mucinous carcinomas, and 1 metaplastic carcinoma. Of 70 breast cancers, 51 lesions appeared as a mass on DBT, 10 as mass with calcification, 4 as calcification without an accompanying mass, 1 as an architectural distortion, and 1 as an asymmetry. The remaining 3 lesions were invisible at DBT due to overlapping dense breast tissue.

Table 1 summarizes the AUCs by scale and expertise (Figs. 1 and 2). Results across interpretative scales were remarkably consistent. For the BI-RADS scale, the average AUCs across readers were 0.778 with CAD (95% CI, 0.733–0.822) and 0.776 without CAD (95% CI, 0.732–0.821). The difference in AUC with and without CAD was not statistically significant, and the results were consistent with the results obtained on using the probability of malignancy and percentage probability of malignancy scales.

Table 2 summarizes the reader-specific AUCs for the probability of malignancy scale, and Fig. 3a and b shows these curves. Although there were variations in the AUC values between the novice and experienced readers, results obtained with CAD were consistent in all readers. Almost identical results were observed for the BI-RADS and percentage probability of malignancy scales (image not shown).

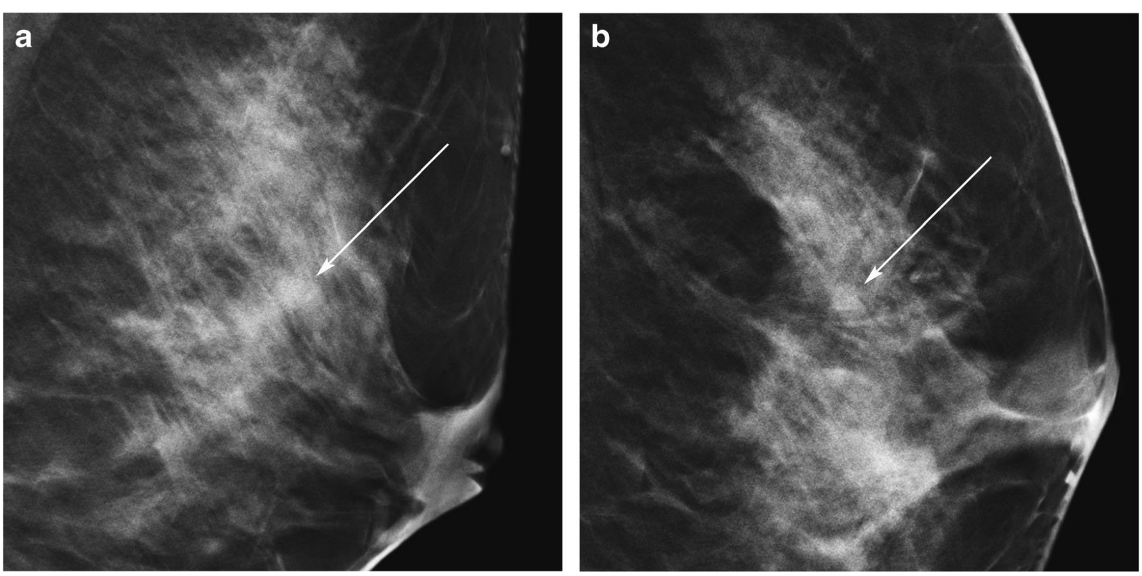

Fig. 2 A 51-year-old woman with a 7-mm invasive ductal carcinoma that was not detected by CAD. Arrows in mediolateral oblique (**a**) and craniocaudal (**b**) images indicate the location of the lesion, not CAD markers

Table 2 AUC for each reader with and without CAD using each scale

Scales	Reader no.		Average AUC (95% CI)		Difference (standard error)	*p* value
			Without CAD	With CAD		
BI-RADS	Novice	1	0.718 (0.621–0.815)	0.755 (0.661–0.848)	0.037(0.026)	0.158
		2	0.785 (0.702–0.868)	0.764 (0.674–0.853)	-0.021 (0.034)	0.530
	Experienced	3	0.787 (0.698–0.875)	0.801 (0.716–0.887)	0.015 (0.025)	0.566
		4	0.802 (0.716–0.888)	0.770 (0.671–0.860)	-0.031 (0.028)	0.257
Probability of malignancy	Novice	1	0.769 (0.680–0.857)	0.787 (0.698–0.876)	0.018 (0.021)	0.378
		2	0.791 (0.706–0.876)	0.771 (0.680–0.863)	-0.020 (0.031)	0.527
	Experienced	3	0.803 (0.709–0.898)	0.840 (0.756–0.924)	0.036 (0.028)	0.189
		4	0.809 (0.723–0.895)	0.785 (0.695–0.874)	-0.024 (0.026)	0.341
Percentage probability of malignancy	Novice	1	0.769 (0.681–0.858)	0.790 (0.702–0.877)	0.021 (0.018)	0.263
		2	0.794 (0.711–0.878)	0.778 (0.686–0.870)	-0.016 (0.032)	0.615
	Experienced	3	0.816 (0.725–0.907)	0.852 (0.770–0.933)	0.036 (0.027)	0.185
		4	0.801 (0.715–0.887)	0.791 (0.699–0.883)	-0.010 (0.027)	0.708

CI confidence interval

The average reading time of radiologists to review a DBT examination per case was 72.07 s (SD, 37.54 s) in the absence of CAD and 62.03 s (SD, 34.38 s) in the presence of CAD (Table 3). The average difference in reading time was a statistically significant decrease by 10.04 ± 1.85 s (standard error; $p < 0.001$) on using CAD, providing 14% decrease in time. The effect of reducing the reading time was consistently shown to the novice ($p = 0.007$) and experienced readers ($p < 0.001$).

Summary measures of diagnostic accuracy are presented in Table 4 for the modified BI-RADS scale (dichotomized at category 3 or less vs 4 or higher). The differences in sensitivity, specificity, PPV and NPV with and without the use of CAD were not statistically significant, and the results were consistent in both the novice and experienced readers.

Discussion

This reader study evaluated a new DBT-CAD system to compare the diagnostic performance and interpretation time with and without CAD. The main finding of this study is that the

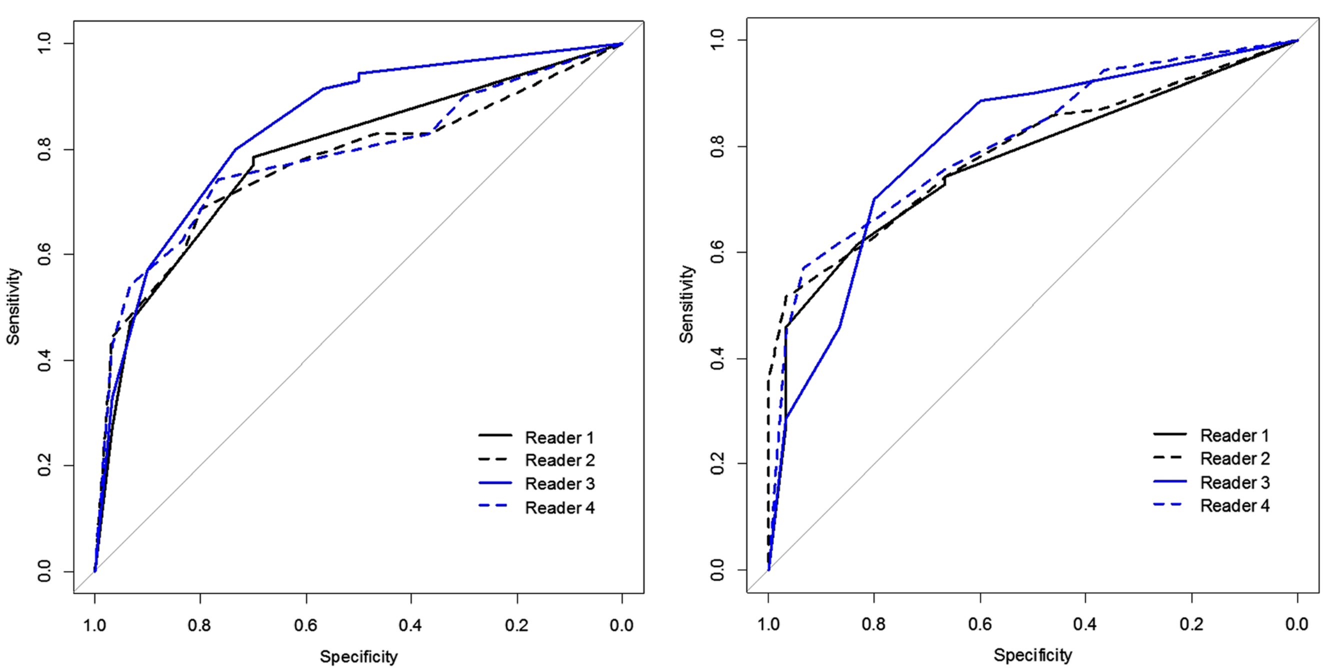

Fig. 3 Receiver operating characteristic curves with (**a**) and without (**b**) concurrent computer-aided detection for the probability of malignancy scale

Table 3 Average reading time (seconds) with and without CAD per case

Reader	Reading time (SD)		Difference (standard error)	*p* value
	Without CAD	With CAD		
Novice	70.83 (45.56)	62.06 (42.04)	8.77 (3.19)	0.007
Experienced	73.31 (27.32)	62.01 (24.53)	11.31 (1.88)	< 0.001
All	72.07 (37.54)	62.03 (34.38)	10.04 (1.85)	< 0.001

SD standard deviation

use of concurrent CAD system for DBT does reduce interpretation time significantly without a loss of diagnostic performance in cancer detection. This observation was consistent for both novice and experienced readers.

Because there are more images to review, it is not surprising that the mean time to interpret DBT images increases. One may reconstruct DBT data to read by selecting different slice thicknesses. For example, if a 1-mm thickness is chosen for a 50-mm compressed breast, there will be 50 slices for the radiologist to review. Reconstructing the image to a thickness of 0.5 mm will result in 100 images to be reviewed. Studies have shown that addition of DBT to FFDM results in a twofold increase in time to interpret images compared with the use of FFDM alone [9, 12]. This additional interpretation time is a substantial consideration for radiologists who read large volume of screening examinations.

In this study, we demonstrated that the reading time to review a DBT examination was significantly reduced on using CAD system, with an average difference of 10.04 ± 1.85 s. Notably, we also found that reduction in interpretation time was consistently observed in novice and experienced readers. Our findings are similar to those reported by Benedikt et al [13]. In that study, 240 DBT cases with 68 malignancies were reviewed using a CAD system. The authors reported a 29.2% reduction in reading time in the presence of CAD. They also reported that the radiologist's performance with CAD was non-inferior to that without CAD.

The interpretation times in our study were longer than those in previous studies [13, 14], and in these studies, the reading times were reported as 48–65 s in the absence of CAD and 39–46 s in the presence of CAD. In comparing our methods with those of previous studies, a few possible explanations could be considered. First, our DBT review workstation is different from theirs (IDI MammoWorkstation, GE Healthcare). Second, the proportion of breast cancer cases in our study was higher than that in the previous studies (61/240, 25% [13]; 23/80, 29% [14]). Although the differences may be noted for specific parameters, the overall conclusions remain the same.

Over the last few years, mammography has been moving to the DBT as they had evolved from film-screen to full-field digital mammography (FFDM). DBT has been shown to overcome the limitations of digital mammography, mainly by reducing the effects of tissue overlap [21–23]. Previous investigations of DBT have reported that DBT in combination with FFDM offers the dual benefit of significantly increased diagnostic accuracy and reduced recall rates for non-cancer cases, compared with FFDM alone [9–11, 24–26]. To take advantage of the time-reducing benefits of DBT-CAD, it is important that radiologist's performance in breast cancer detection does not deteriorate. In our current study, the AUC values of DBT interpretation studies for the detection of malignant lesions using the CAD system are not significantly different from the AUC values obtained without using the CAD system. Differences

Table 4 Performance for novice and experienced readers with and without CAD using modified BI-RADS scale

Reader		Sensitivity	Specificity	Positive predictive value	Negative predictive value
Novice	Without CAD	0.743 (0.664–0.808)	0.667 (0.539–0.774)	0.825 (0.753–0.880)	0.524 (0.418–0.628)
	With CAD	0.771 (0.695–0.834)	0.667 (0.539–0.774)	0.825 (0.752–0.880)	0.524 (0.418–0.628)
	p value	0.246	1	0.753	0.352
Experienced	Without CAD	0.807 (0.733–0.864)	0.667 (0.539–0.774)	0.827 (0.757–0.881)	0.584 (0.472–0.689)
	With CAD	0.800 (0.726–0.858)	0.667 (0.539–0.774)	0.827 (0.756–0.881)	0.584 (0.472–0.689)
	p value	0.739	1	0.942	0.788
All	Without CAD	0.775 (0.722–0.820)	0.667 (0.578–0.745)	0.826 (0.777–0.866)	0.553 (0.475–0.628)
	With CAD	0.786 (0.734–0.830)	0.667 (0.578–0.745)	0.826 (0.778–0.866)	0.553 (0.475–0.628)
	p value	0.512	1	0.873	0.598

Note: Data in parentheses indicate the 95% confidence intervals in per-case sensitivity and specificity were not statistically significant in the presence or absence of CAD system. One noteworthy finding is that these results were consistent across three different rating scales. We also found that results obtained from using CAD were similar regardless of the level of experience of the reader. Although the impact of adding DBT to FFDM on the diagnostic performance may vary according to the experience level of readers [27], the effect of CAD in DBT interpretation could be consistent regardless of their experience. In this study, the novice readers performed as well as the experienced readers in all aspects with and without CAD. This could be explained as the number of proven cancer in the data set is high (70 out of 100), and the tumor size is quite large (average 23.2 mm in size).

This study had several limitations. First, this was a reader study with a small sample size. Second, the data set for this study included only a limited portion (14%) of patients for screening and a large number of cancer patients (70%). The proportion of patients in our study was different from those in the general population, and this was probably because the study was conducted at a single tertiary referral center. Because the incidence of breast cancer is low in the general screening population, a prospective study design will require much larger sample sizes. Next, reader variability may be an issue. However, the results of this study were similar for experienced and novice readers and were consistent across three different scales. Lastly, because we assessed the presence or absence of cancer in the DBT interpretation, real life would require more time to examine for additional findings other than the most obvious cancer. In conclusion, this reader performance study enriched with cancer cases showed that use of concurrent DBT-CAD system shortens interpretation time without the loss of diagnostic performance to novice and experienced readers, and therefore improves the efficiency of DBT interpretation. Further studies are necessary to validate and generalize the findings of the current study to the overall population. And it would need to validate these findings using other CAD systems as well.

Funding This study has received funding by the R&D Convergence Program (Grant Number: CAP-13-3-KERI) of the National Research Council of Science & Technology (NST) of the Republic of Korea.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Hak Hee Kim.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry Mingyu Han kindly provided statistical advice for this manuscript.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
• retrospective
• experimental
• performed at one institution

References

1. Tabár L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA (2001) Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 91:1724–1731
2. Independent UK Panel on Breast Cancer Screening (2012) The benefits and harms of breast cancer screening: an independent review. Lancet 380:1778–1786
3. Boyd NF, Guo H, Martin LJ et al (2007) Mammographic density and the risk and detection of breast cancer. N Engl J Med 356:227–236
4. Harvey JA, Bovbjerg VE (2004) Quantitative assessment of mammographic breast density: relationship with breast cancer risk. Radiology 230:29–41
5. Kolb TM, Lichy J, Newhouse JH (2002) Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology 225:165–175
6. Chiu SY, Duffy S, Yen AM, Tabar L, Smith RA, Chen HH (2010) Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. Cancer Epidemiol Biomarkers Prev 19:1219–1228
7. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V (1996) Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 276:33–38
8. Skaane P, Bandos AI, Gullien R et al (2013) Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. Eur Radiol 23:2061–2071
9. Skaane P, Bandos AI, Gullien R et al (2013) Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology 267:47–56
10. Ciatto S, Houssami N, Bernardi D et al (2013) Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 14:583–589
11. Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S (2016) Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. Eur Radiol 26:184–190
12. Dang PA, Freer PE, Humphrey KL, Halpern EF, Rafferty EA (2014) Addition of tomosynthesis to conventional digital mammography: effect on image interpretation time of screening examinations. Radiology 270:49–56
13. Benedikt RA, Boatsman JE, Swann CA, Kirkpatrick AD, Toledano AY (2018) Concurrent computer-aided detection improves reading time of digital breast tomosynthesis and maintains interpretation performance in a multireader multicase study. AJR Am J Roentgenol 210:685–694
14. Balleyguier C, Arfi-Rouche J, Levy L et al (2017) Improving digital breast tomosynthesis reading time: a pilot multi-reader, multi-case

study using concurrent computer-aided detection (CAD). Eur J Radiol 97:83–89
15. Duda RO, Hart PE (1972) Use of the Hough transformation to detect lines and curves in pictures. Commun ACM 15:11–15
16. Ballard DH (1981) Generalizing the Hough transform to detect arbitrary shapes. Pattern Recogn 13:111–122
17. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, p 770–778
18. Jeong JW, Chae SH, Cho YL et al (2017) A deep convolutional neural network based false positive reduction in mass detection algorithm on digital breast tomosynthesis images. Int J CARS 12(Suppl 1):S273
19. Choi WJ, Kim HH, Lee SY et al (2016) A comparison between digital breast tomosynthesis and full-field digital mammography for the detection of breast cancers. Breast Cancer 23:886–892
20. Chae EY, Kim HH, Cha JH, Shin HJ, Choi WJ (2016) Detection and characterization of breast lesions in a selective diagnostic population: diagnostic accuracy study for comparison between one-view digital breast tomosynthesis and two-view full-field digital mammography. Br J Radiol 89:20150743
21. Gur D, Abrams GS, Chough DM et al (2009) Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol 193:586–591
22. Poplack SP, Tosteson TD, Kogel CA, Nagy HM (2007) Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography. AJR Am J Roentgenol 189: 616–623
23. Vedantham S, Karellas A, Vijayaraghavan GR, Kopans DB (2015) Digital breast tomosynthesis: state of the art. Radiology 277:663–684
24. Rafferty EA, Park JM, Philpotts LE et al (2013) Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. Radiology 266:104–113
25. Hooley RJ, Durand MA, Philpotts LE (2017) Advances in digital breast tomosynthesis. AJR Am J Roentgenol 208:256–266
26. McCarthy AM, Kontos D, Synnestvedt M et al (2014) Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. J Natl Cancer Inst 106(11). https://doi.org/10.1093/jnci/dju316
27. Tucker L, Gilbert FJ, Astley SM et al (2017) Does reader performance with digital breast tomosynthesis vary according to experience with two-dimensional mammography? Radiology 283:371–380