European Radiology (2019) 29:319–329
https://doi.org/10.1007/s00330-018-5585-x

BREAST

Comparison of synthetic and digital mammography with digital breast tomosynthesis or alone for the detection and classification of microcalcifications

Ji Soo Choi[1] · Boo-Kyung Han[1] · Eun Young Ko[1] · Ga Ram Kim[2] · Eun Sook Ko[1] · Ko Woon Park[1]

Received: 9 January 2018 / Revised: 4 May 2018 / Accepted: 1 June 2018 / Published online: 21 June 2018

Abstract

Objective To compare the performance of synthetic mammography (SM) and digital mammography (DM) with digital breast tomosynthesis (DBT) or alone for the evaluation of microcalcifications.

Methods This retrospective study includes 198 mammography cases, all with DM, SM, and DBT images, from January to October 2013. Three radiologists interpreted images and recorded the presence of microcalcifications and their conspicuity scores and final BI-RADS categories (1, 2, 3, 4a, 4b, 4c, 5). Readers' area under the ROC curves (AUCs) were analyzed for SM plus DBT vs. DM plus DBT and SM alone vs. DM alone using the BI-RADS categories for the overall group and dense breast subgroup.

Results Conspicuity scores of detected microcalcifications were neither significantly different between SM and DM with DBT nor alone (p>0.05). In predicting malignancy of detected microcalcifications, no significant difference was found between readers' AUCs for SM and DM with DBT or alone in the overall group or dense breast subgroup (p>0.05).

Conclusions Diagnostic performances of SM and DM for the evaluation of microcalcifications are not significantly different, whether performed with DBT or alone.

Key Points

- *In DBT-imaging, SM and DM show comparable performances when evaluating microcalcifications.*
- *For BI-RADS classification of microcalcifications, SM and DM show similar AUCs.*
- *DBT with SM may be sufficient for diagnosing microcalcifications, without DM.*

Keywords Microcalcification · Digital breast tomosynthesis · Digital mammography · Synthetic image · Diagnosis

Abbreviations

AUC	Area under the receiver operating characteristic curve
BI-RADS	Breast Imaging Reporting and Data System
DBT	Digital breast tomosynthesis
DM	Digital mammography
ICC	Intraclass correlation coefficient
SM	Synthetic mammography

✉ Boo-Kyung Han
bkhan@skku.edu

[1] Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-Ro Gangnam-gu, Seoul 135-710, Korea

[2] Department of Radiology, Inha University Hospital, Inha University School of Medicine, Incheon, Korea

Introduction

Calcifications are frequently found on mammography and can be seen in both benign and malignant cases. Calcifications associated with malignancy usually present as microcalcifications with specific morphologic features [1, 2]. Mammography is considered the most important diagnostic tool in the detection and characterization of microcalcifications, because microcalcification morphology cannot be fully interpreted with other breast imaging modalities (e.g., ultrasound, magnetic resonance imaging) alone. Two-view digital mammography (DM) can depict microcalcifications with extreme accuracy, because of its inherent high contrast resolution, tailored to reveal calcifications [3, 4]. Since DM was adopted as the primary imaging modality for breast cancer screening, it has played an important role in differentiating malignant and benign microcalcifications [5, 6].

Digital breast tomosynthesis (DBT) provides a three-dimensional view of the breast through a series of thin-slice images. This technique can facilitate discrimination between normal breast tissue and lesions by reducing tissue overlap which may mask or mimic breast abnormalities [7]. Early clinical studies showed that adding two-view DBT to standard DM has at least equal or higher accuracy than DM alone, but these results were mainly due to increased diagnostic accuracy for non-calcified lesions [3, 8–12]. However, there is little evidence that DBT improves the diagnostic accuracy for microcalcifications [4, 13, 14]. Rather, some studies reported that DBT alone has a lower sensitivity in detecting and characterizing calcifications compared to DM [4, 14]. Because of concerns with difficulty in detecting and interpreting microcalcification groups with DBT alone, DBT has been used with DM for breast cancer screening [4]. In multiple population-based studies, this combination of DBT and DM has shown higher accuracy than DM alone in both screening and diagnostic settings for breast cancer [9, 15–19]. However, combining DBT and DM increases radiation dose compared to DM alone, and this is a major drawback when screening the general population [20, 21]. To address this issue, the DBT industry developed two-dimensional (2-D) synthetic mammography (SM) from data obtained during DBT acquisition as an alternative to DM. Several recent clinical studies showed that SM plus DBT or SM alone has comparable performance to DM plus DBT or DM alone [22–24], but, to our knowledge, very few reports have compared SM and DM for microcalcifications. Hence, the purpose of this study was to compare the performance of SM and DM with DBT or alone for the evaluation of microcalcifications.

Materials and methods

Case selection

This study was conducted retrospectively and was approved by the institutional review board. Informed consent was waived. From January to October 2013, 1421 women underwent combined DM and DBT at a single tertiary referral centre, for diagnostic or screening purposes (Fig. 1). Inclusion criteria for the microcalcification group were as follows: women who had (1) microcalcifications (diameter < 1 mm) that presented solely without associated mammographic findings, (2) no clinical symptoms, and microcalcifications detected with screening DM performed at our hospital or other hospitals which referred the women for proper management, and (3) final diagnoses for microcalcifications based on histopathologic results or based on typical imaging findings only if the microcalcifications were stable on follow-up imaging. Of 1421 patients, 88 patients had microcalcifications that presented without associated mammographic findings. Among them, 11 patients were excluded (Fig. 1). Finally, 154 mammography cases were obtained from 77 women who underwent mammography for both breasts (Fig. 1). Among the 154 cases, there were 90 microcalcification cases and 64 negative mammography cases.

Sample size was calculated to provide 80% power to detect a difference of 10 % in area under the receiver operating characteristic curve (AUC) for microcalcification detection between SM and DM with DBT or alone, with a two-sided significance level of 5%. Based on previous studies on microcalcification detection, we assumed an AUC of 0.88 for DM [12, 23], a ratio of 1.2 between no detection and detection cases and a rating correlation coefficient of 0.4 among cases within the same subjects. Therefore, with 90 microcalcification cases, 108 normal cases were required to have 80% power [25]. Forty-four additional negative mammography cases were thus required and selected for the final study cohort (SPSS statistical software). We referenced previous studies published from our institution and the Asian population breast density distribution [26–28], and considered the dense breast (heterogeneously or extremely dense) to non-dense breast (almost entirely fatty or scattered areas of fibroglandular density) ratio to be 3:1 [26–28]. We therefore collected 32 dense breast cases from 16 women and 12 non-dense breast cases from six women who had a negative mammography during the study period.

The final study cohort included 90 mammography cases with microcalcifications and 108 negative mammography cases from 99 women (mean age, 50.9 years), with DBT, SM and DM images (Fig. 1).

Imaging technique

All patients underwent DBT and DM (two-view: craniocaudal and mediolateral oblique) for both breasts on the same mammography unit (Selenia Dimensions; Hologic, Bedford, MA, USA). DBT and DM images of each breast were sequentially obtained during a single breast compression per view. Automated exposure control was employed. The imaging system automatically estimated the average glandular dose (AGD). SM images were generated from the DBT dataset for each case with image processing software (C-View for Hologic; version 1.0.0.1) [29, 30]. Finally, three image sets (DM, SM, and DBT) were available for each case.

Reading study

Three breast radiologists (E.Y.K., C.J.S. and G.R.K. with 15, 7, and 4 years of experience in breast imaging, respectively) participated in this reading study. The radiologists were asked to interpret data sets composed of a combination of 2-D images (SM or DM) and associated DBT images, or a combination of 2-D images (SM or DM) alone. The cases within each

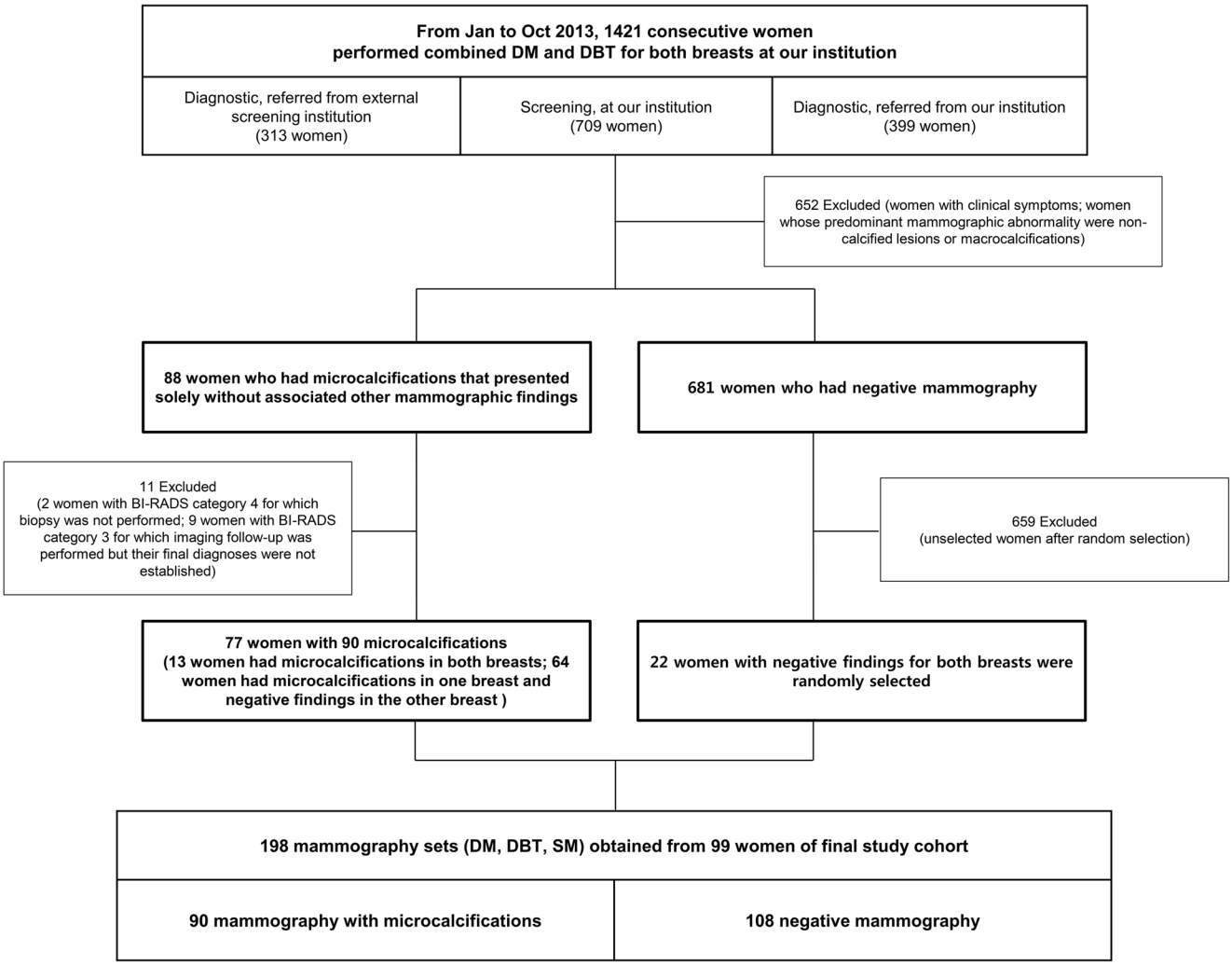

Fig. 1 Flow chart of the study population

data set were arranged randomly. Data sets were reviewed with an interval of at least three weeks. To minimize learning bias, we concealed the names, ages, identification numbers, and image parameters of the patients during the review. Any information on patient history, findings detected with other imaging modalities, or histopathologic diagnosis was also concealed and the radiologists interpreted images at the same workstation (SecurView; Hologic) independently.

The readers recorded the presence of microcalcifications in each breast. If the readers detected microcalcifications on SM or DM images, they recorded the group size, anatomic location (left or right, clock position, and distance from nipple) and conspicuity scores [1 if microcalcifications were possibly present (low conspicuity), 2 if probably present (medium conspicuity), and 3 if definitely present (high conspicuity)]. When the readers detected multiple lesions in the same breast, they documented the size and location of each microcalcification group to avoid confusion during data analysis. Then, they assessed the morphology and distribution of the detected microcalcifications with the BI-RADS lexicon by using the magnifier window offered by the dedicated workstation. Finally, the readers assessed the detected lesions by BI-RADS category (1, 2, 3, 4a, 4b, 4c, or 5) to predict malignant calcifications [1], and they did not consider findings of other images including additional magnification view obtained for proper management.

Data and statistical analysis

Final diagnoses of the 90 microcalcification cases were based on histopathologic results of surgical excision (n=57), stereotactic vacuum-assisted biopsy (n=10), or based on typical benign findings (BI-RADS category 2) if they were stable on follow-up mammography for over two years prior to the commencement of this study (n=23). Surgical excision or stereotactic biopsies were performed on microcalcifications assessed as BI-RADS 4a or higher (n=63) and category 3 microcalcifications at patient request (n=4). There were 42 malignant and 48 benign microcalcifications in 77 patients (Table 1) (Fig. 1.).

Table 1 Final diagnoses of 90 microcalcification cases in 77 women

Type	Histopathologic diagnosis	Number of cases (%)
Malignant (*n*=42)	Invasive ductal carcinoma	26 (61.9)
	Ductal carcinoma in situ	12 (28.5)
	Invasive lobular carcinoma	2 (4.8)
	Mucinous carcinoma	2 (4.8)
Benign (*n*=48)	Fibrocystic change	5 (10.4)
	Columnar cell change	5 (10.4)
	Atypical ductal hyperplasia	3 (6.3)
	Lobular carcinoma in situ	3 (6.3)
	Mucocele-like lesion	3 (6.3)
	Stromal fibrosis	3 (6.3)
	Intraductal papilloma	2 (4.2)
	Sclerosing adenosis	1 (2.0)
	N/A*	23 (47.8)

Thirteen women had calcifications in both breasts, and 64 women had calcifications in one breast

*These were assessed as BI-RADS 2, and all were stable or decreased on follow-up mammography for over two years

N/A, not available

To establish reference mammographic findings of the study cohort, two radiologists (B.K.H. and E.S.K.) determined the presence of microcalcifications and the morphology of the microcalcifications on SM and DM images in consensus. These radiologists did not participate in the reading study, and were informed of final histological diagnoses and findings with other imaging modalities. They also rated mammographic breast density according to BI-RADS classes (almost entirely fatty, scattered areas of fibroglandular density, heterogeneously dense, or extremely dense) with DM images [1].

To compare microcalcification conspicuity, conspicuity scores of the calcifications detected with SM and DM with DBT and alone were compared by Wilcoxon's signed rank test for each reader. Readers' decisions on morphology, distribution, and BI-RADS category for detected microcalcifications were compared between SM and DM with DBT and alone using the chi-square test.

We evaluated the AUC, sensitivity and specificity of SM and DM with DBT and alone in the assessment of 198 mammography cases in 99 patients. The sensitivity and specificity as well as the AUC of each reader were also assessed based on final diagnosis for SM and DM with DBT or alone.

AUC was calculated for the prediction of malignant microcalcifications using the BI-RADS categories (1, 2, 3, 4a, 4b, 4c, 5). Readers' AUCs for SM and DM with DBT or alone were analysed and compared using a nonparametric approach. For the detection of microcalcifications, the AUCs could not be calculated because there were no false positive cases for any of the readers [31].

Sensitivities and specificities for SM and DM were assessed. Positive and negative results were defined differently for detecting overall calcifications and for predicting malignant calcifications. For detecting overall calcifications, BI-RADS categories 2-5 defined positive results and BI-RADS category 1 defined negative results. Sensitivity was calculated as case number of test positives per case number of overall calcifications (both benign and malignant), and specificity as case number of test negatives per case number of overall negative mammograms. For predicting malignant calcifications, BI-RADS categories 3-5 defined positive results and BI-RADS categories 1-2 defined negative results. Sensitivity was calculated as case number of test positives per case number of malignant calcifications, and specificity as case number of test negatives per case number of negative mammograms and typical benign calcifications. Statistical differences in sensitivities and specificities for SM and DM were analysed by the generalized estimating equation (GEE) to adjust for correlation of mammographic outcomes within radiologists [32].

Interobserver agreements between the three readers for the final BI-RADS category and conspicuity score were analysed with the intraclass correlation coefficient (ICC) and 95% Bland-Altman limits of agreement. The ICC defined interobserver agreements as follows: poor, 0.00-0.20; fair, 0.21-0.40; moderate, 0.41-0.60; good, 0.61-0.80; and excellent, 0.81-1.00.

Statistical analyses were performed using SAS statistical software (SAS system for Windows, version 9.1.3; SAS institute, Cary, NC, USA) for the overall study cohort and dense breast subgroup (heterogeneously or extremely dense on mammographic density), respectively. $P < 0.05$ was considered statistically significant.

Results

AGD values for DBT and DM were 1.91±0.48 (mean±standard deviation) and 1.94±0.70 mGy, respectively, for the craniocaudal view, and 1.90±0.52 mGy and 1.87±0.81, respectively, for the mediolateral oblique view. The breast

Table 2 Conspicuity scores of detected calcifications with SM and DM for each reader

	With DBT			Without DBT		
	SM	DM	*p*	SM	DM	*p*
Reader 1	2 (2-3)	2 (2-3)	0.14	3 (2-3)	2 (2-3)	0.07
Reader 2	3 (2-3)	3 (2-3)	0.67	3 (3-3)	3 (3-3)	0.57
Reader 3	2 (1-2)	2 (1-2)	0.24	2 (1-3)	2 (1-3)	0.90

Data are presented as median values (interquartile range)

DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, digital mammography

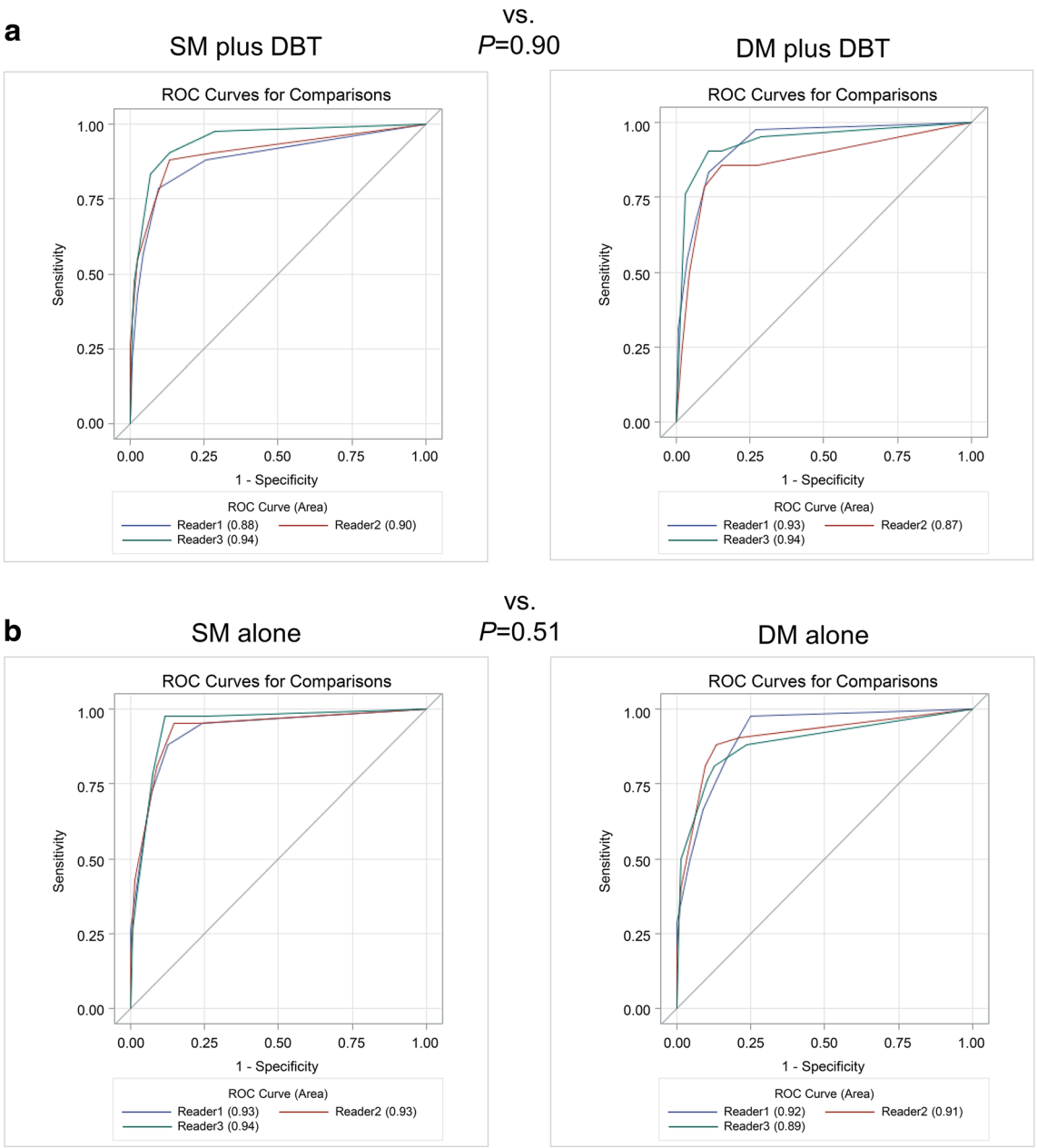

Fig. 2 ROC curves based on the probability of malignancy ratings for all breasts; (**a**) Readers' AUCs for SM plus DBT (left; range 0.88-0.94) and DM plus DBT (right; range 0.87-0.94) were not significantly different (p=0.90). **b** Readers' AUCs for SM (left; range 0.93-0.94) and DM alone (right; range 0.89-0.92) were not significantly different (p=0.51). ROC, receiver operating characteristic; AUC, area under the ROC curve. DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, full-field digital mammography

density distribution of the 198 mammographies were as follows: four almost entirely fatty (2.0%), 22 scattered areas of fibroglandular density (11.1%), 108 heterogeneously dense (54.6%), and 64 extremely dense (32.3%). Of these, heterogeneously or extremely dense (n=172) cases were included in the dense breast group.

Conspicuity scores of the microcalcifications were not significantly different between SM and DM with DBT or alone for all readers (p>0.05) (Table 2). There was also no statistical difference in the radiologists' decisions on morphology, description, and BI-RADS category of detected microcalcifications between SM and DM with DBT or alone.

There was no statistically significant difference in readers' AUCs between SM and DM with DBT or alone for predicting the probability of malignancy with BI-RADS ratings in all breasts (p>0.05) (Fig. 2). For the dense breast group, AUC differences between SM and DM with DBT or alone were not significantly different (p>0.05) (Fig. 3).

Sensitivities and specificities for detecting overall microcalcifications are presented in Table 3. When

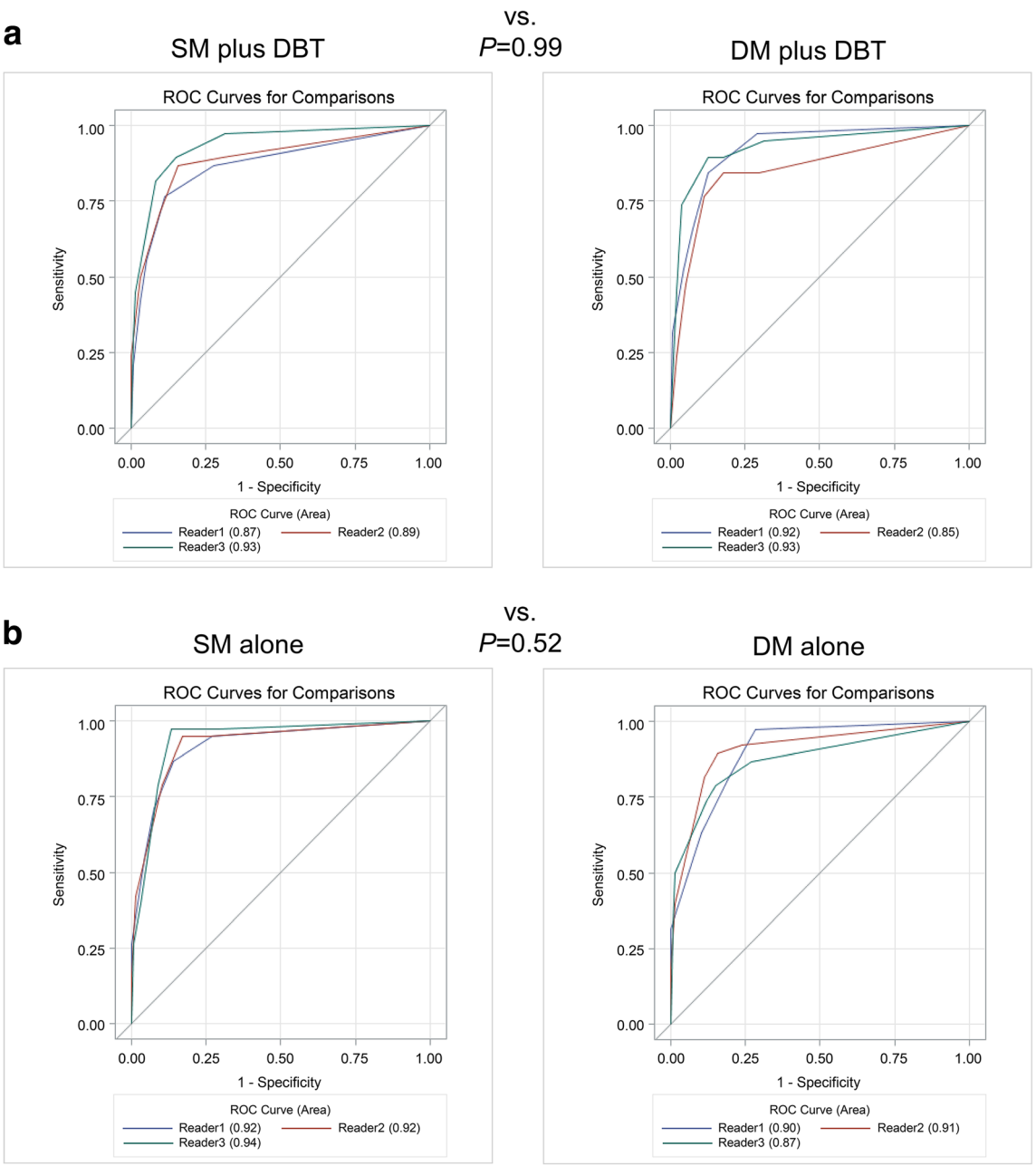

Fig. 3 ROC curves based on the probability of malignancy ratings for the dense breast group; (**a**) Readers' AUCs for SM plus DBT (left; range 0.87-0.93) and DM plus DBT (right; range 0.85-0.93) were not significantly different (p=0.99). **b** Readers' AUCs for SM (left; range 0.92-0.94) and DM alone (right; range 0.87-0.91) were not significantly different (p=0.52). ROC, receiver operating characteristic; AUC, area under the ROC curve. DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, full-field digital mammography

detecting calcifications with SM plus DBT and DM plus DBT, the readers' sensitivities were not significantly different for all breasts (SM plus DBT 85.6-95.6% vs, DM plus DBT 87.8-94.4%), while specificities were 100%. For the dense breast group, there was no statistically significant difference in sensitivities and specificities between SM and DM with DBT or alone (p >0.05). For SM and DM alone for all breasts, reader 2 and 3 showed higher sensitivities with SM than DM, while the sensitivity of reader 1 was higher with DM than SM. Overall, the difference between SM and DM alone was borderline statistically significant (p=0.05). For the dense breast group, there were no significant differences either in the readers' sensitivities or specificities with SM and DM alone (p >0.05).

Sensitivities and specificities for predicting malignant microcalcifications are presented in Table 4. For predicting malignant calcifications with SM plus DBT and DM plus DBT for all breasts, the readers' sensitivities and specificities were not statistically significantly different between the two combined modes (p>0.05). For predicting malignant

Table 3 Sensitivities and specificities of SM and DM for each reader in detecting overall microcalcifications

	With DBT			Without DBT		
	SM	DM	*p*	SM	DM	*p*
All breasts (*n*=198)*						
Sensitivity			0.62			0.05
Reader 1	85.6 (77/90; 78.3-92.8)	92.2 (83/90; 86.7-97.8)		86.7 (78/90; 79.6-93.7)	88.9 (80/90; 82.4-95.4)	
Reader 2	91.1 (82/90; 85.2-97.0)	87.8 (79/90; 81.0-94.5)		84.4 (76/90; 77.0-91.9)	78.9 (71/90; 70.5-87.3)	
Reader 3	95.6 (86/90; 91.3-99.8)	94.4 (85/90; 89.7-99.2)		90.0 (81/90; 83.8-96.2)	82.2 (74/90; 74.3-90.1)	
Specificity			N/A			N/A
All Readers‡	100.0 (108/108; 100.0-100.0)	100.0 (108/108; 100.0-100.0)		100.0 (108/108; 100.0-100.0)	100.0 (108/108; 100.0-100.0)	
Dense breasts (*n*=172)†						
Sensitivity			0.62			0.10
Reader 1	84.3 (70/83; 76.5-92.2)	91.6 (76/83; 85.6-97.5)		86.7 (72/83; 79.5-94.0)	90.4 (75/83; 84.0-96.7)	
Reader 2	90.4 (75/83; 84.0-96.7)	86.7 (72/83; 79.5-94.0)		85.5 (71/83; 78.0-93.1)	80.7 (67/83; 72.2-89.2)	
Reader 3	95.2 (79/83; 90.6-99.8)	94.0 (78/83; 88.9-99.1)		91.6 (76/83; 85.6-97.5)	83.1 (69/83; 75.1-91.2)	
Specificity			N/A			N/A
All Readers‡	100.0 (89/89; 100.0-100.0)	100.0 (89/89; 100.0-100.0)		100.0 (89/89; 100.0-100.0)	100.0 (89/89; 100.0-100.0)	

*Breast-based assessment of 198 breasts in 99 patients

†Breast-based assessment of 172 dense breasts in 86 patients

Positive test results for the presence of microcalcifications were defined with BI-RADS category 2 or higher

Sensitivities are presented as percentages (case number of test positives/ case number of overall calcifications; 95% confidence interval of sensitivities)

Specificities are presented as percentages (case number of test negatives/ case number of overall negative mammograms; 95% confidence interval of specificities).

‡All readers showed 100% specificities.

DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, digital mammography

N/A: not available

calcifications with SM and DM alone, the readers' sensitivities were statistically significantly higher with SM compared to DM.(p=0.01) for all breasts and the dense breast group (p=0.02), but the readers' specificities were not significantly different for all breasts and the dense breast group (p>0.05) (Fig. 4).

Interobserver agreements of the three readers are presented in Table 5. For BI-RADS final assessment category, SM and DM showed good to moderate interobserver agreement both with DBT and alone. For conspicuity scoring, the combined modes of SM and DM with DBT showed moderate agreement, while the SM and DM alone modes showed fair agreement.

Discussion

In this study, we compared the diagnostic performances of SM and DM with DBT and alone for the evaluation of microcalcifications. For SM and DM with DBT, there were no significant differences in the readers' sensitivities and specificities either in detecting overall microcalcifications or in predicting malignant microcalcifications. Additionally, the readers' AUCs for SM plus DBT were not significantly different compared to DM plus DBT for BIRADS category ratings. For SM and DM alone, all readers showed statistically significantly higher sensitivities for detecting overall microcalcifications with SM compared to DM alone for the dense breast group, although the difference in sensitivities between SM and DM only showed borderline statistical significance for all breasts. In predicting malignant microcalcifications; however, all readers showed significantly higher sensitivities with SM alone than DM alone for both all breasts and the dense breast group. This may reflect that SM may be slightly more sensitive than DM for detecting and characterizing microcalcifications, which, in our experience, may be due to the more dense appearance of microcalcifications on SM than DM. Consequently, we believe that SM plus DBT may replace DM plus DBT for the diagnosis of microcalcifications as an alternative approach which has already been substantiated in the diagnosis of non-calcified breast lesions [11, 12].

In a previous study comparing DM and DBT to characterize microcalcifications, lower BI-RADS categories were assigned with DBT compared to DM in a few cases (11/107), although most microcalcifications were classified

Table 4 Sensitivities and specificities of SM and DM for each reader in predicting malignant microcalcifications

	With DBT			Without DBT		
	SM	DM	*p*	SM	DM	*p*
All breasts (*n*=198)*						
Sensitivity			0.83			0.01
Reader 1	78.6 (33/42; 66.2-91.0)	83.3 (35/42; 72.1-94.6)		88.1 (37/42; 78.3-97.9)	83.3 (35/42; 72.1-94.6)	
Reader 2	88.1 (37/42; 78.3-97.9)	85.7 (36/42; 75.1-96.3)		95.2 (40/42; 88.8-100.0)	88.1 (37/42; 78.3-97.9)	
Reader 3	90.5 (38/42; 81.6-99.4)	90.5 (38/42; 81.6-99.4)		97.6 (41/42; 93.0-100.0)	81.0 (34/42; 69.1-92.8)	
Specificity			0.19			0.35
Reader 1	90.4 (141/156; 85.8-95.0)	89.1 (139/156; 84.2-94.0)		87.2 (136/156; 81.9-92.4)	83.3 (130/156; 77.5-89.2)	
Reader 2	86.5 (135/156; 81.2-91.9)	84.6 (132/156; 79.0-90.3)		85.3 (133/156; 79.7-90.8)	86.5 (135/156; 81.2-91.9)	
Reader 3	86.5 (135/156; 81.2-91.9)	84.6 (132/156; 79.0-90.3)		88.5 (138/156; 83.4-93.5)	87.2 (136/156; 81.9-92.4)	
Dense breasts (*n*=172)†						
Sensitivity			0.67			0.02
Reader 1	76.3 (29/38; 62.8-89.8)	84.2 (32/38; 72.6-95.8)		86.8 (33/38; 76.1-97.6)	81.6 (31/38; 69.3-93.9)	
Reader 2	86.8 (33/38; 76.1-97.6)	84.2 (32/38; 72.6-95.8)		94.7 (36/38; 87.6-100.0)	89.5 (34/38; 79.7-99.2)	
Reader 3	89.5 (34/38; 79.7-99.2)	89.5 (34/38; 79.7-99.2)		97.4 (37/38; 92.3-100.0)	78.9 (30/38; 66.0-91.9)	
Specificity			0.13			0.27
Reader 1	88.8 (119/134; 83.5-94.1)	87.3 (117/134; 81.7-92.9)		85.8 (115/134; 79.9-91.7)	80.6 (108/134; 73.9-87.3)	
Reader 2	84.3 (113/134; 78.2-90.5)	82.1 (110/134; 75.6-88.6)		82.8 (111/134; 76.5-89.2)	84.3 (113/134; 78.2-90.5)	
Reader 3	85.1 (114/134; 79.0-91.1)	82.1 (110/134; 75.6-88.6)		86.6 (116/134; 80.8-92.3)	85.1 (114/134; 79.0-91.1)	

*Breast-based assessment of 198 breasts in 99 patients

†Breast-based assessment of 172 dense breasts in 86 patients

Positive test results for malignant microcalcifications were defined with BI-RADS category 3 or higher

Sensitivities are presented as percentages (case number of test positives/ case number of malignant calcifications; 95% confidence interval of sensitivities)

Specificities are presented as percentages (case number of test negatives/ case number of overall negative mammograms and typical benign calcifications; 95% confidence interval of specificities).

DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, digital mammography

similarly [14]. The authors suggested that some dust-like microcalcifications were less visible on DBT images compared to DM to explain this result. In our study, we found that for all readers, the sensitivities of SM plus DBT (78.6-90.5%) tended to be lower than SM alone (88.1-97.6%) for predicting malignant calcifications, although this tendency was not clear for DM with DBT or alone. In fact, all malignant cases which were assessed as positive with SM alone but as negative with SM plus DBT showed lower conspicuity scores with SM plus DBT than SM alone, and most presented as amorphous or punctate microcalcifications on SM (Reader 1, 5/6; Reader 2, 3/4; Reader 3, 3/3). Considering the previous study [14] and our results together, SM plus DBT might have shown a tendency for lower sensitivity compared to SM alone because some microcalcifications may be less conspicuous on DBT compared to 2-D mammography and this may cause an underclassification of the BI-RADS categories with the combined DBT plus 2-D mammography mode. Therefore, we believe that if radiologists determine the final BI-RADS category of microcalcifications with more focus on SM findings than DBT findings, the sensitivities of SM plus DBT may be further improved.

For the dense breast subgroup, SM and DM showed similar diagnostic performances to the overall group both with DBT and alone. This may be because a majority of our patients had dense breasts (86.9%, 172/198), which was concordant with proportions (66-88%) previously reported for Asian women [26–28]. SM and DM also showed similar ICC values with DBT and alone. SM and DM with DBT and alone showed moderate to good interobserver agreements (0.60-0.61) for the final assessments of BI-RADS categories, and fair to moderate agreements (0.31-0.57) in conspicuity scoring. Low ICC values of SM and DM in conspicuity scoring is probably due to the fairly subjective aspect of its method. ICC values of SM and DM for the BI-RADS classification may have been relatively high compared to ICC values for conspicuity scoring, because the radiologists of our reading studies were highly trained in interpreting microcalcifications with BI-RADS [1]. Our study was not a non-inferiority trial, but our overall results suggest that SM may replace dose-requiring DM for detecting and

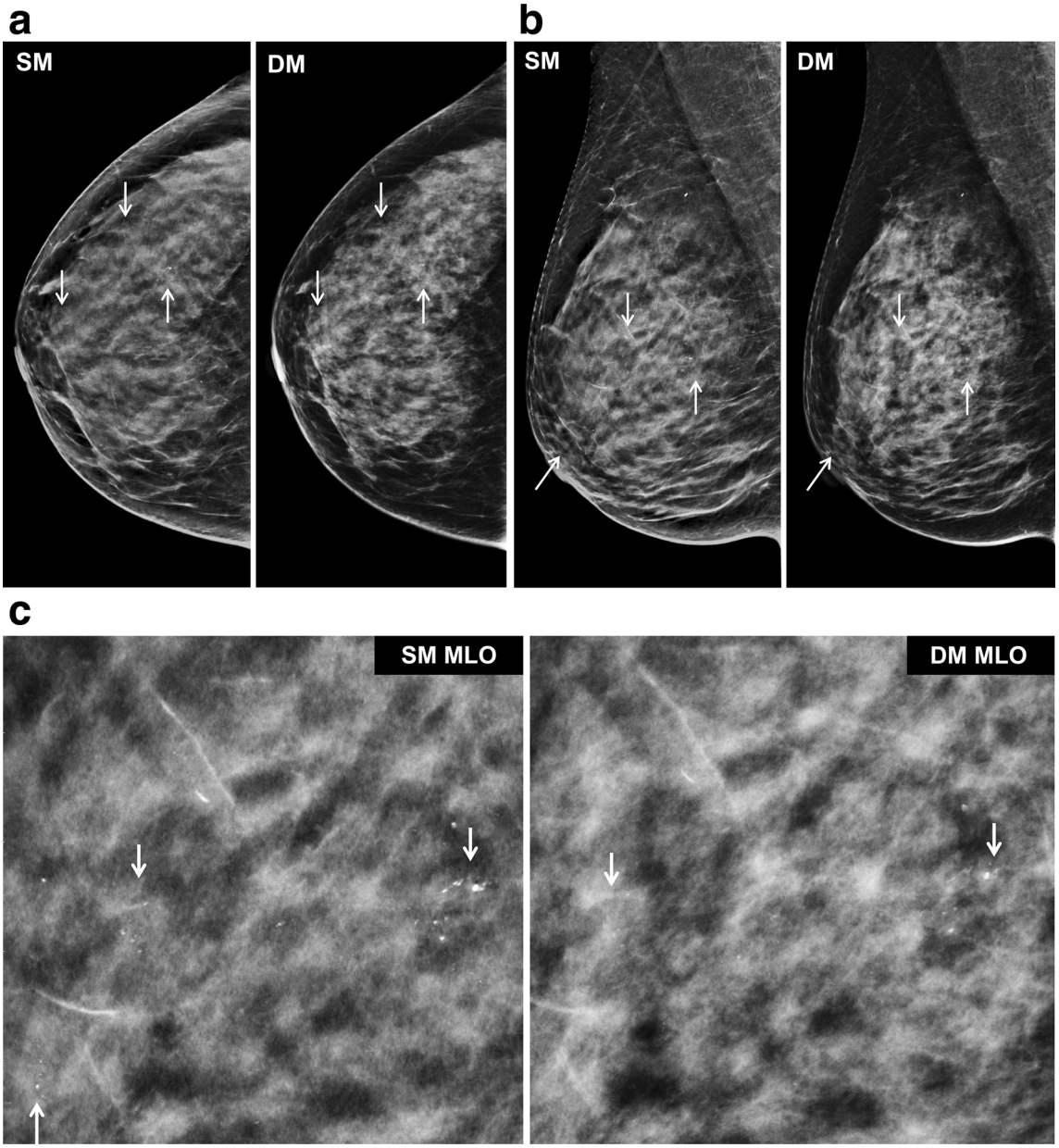

Fig. 4 A 51-year-old woman with microinvasive ductal carcinoma and extremely dense breast tissue: (**a**) Craniocaudal SM (left) and DM (right) images and (**b**) mediolateral oblique SM (left) and DM (right) images demonstrate multiple grouped microcalcifications (arrows) in the right upper outer beast. (**c**) Magnified region of interest for mediolatereal oblique SM (left) and DM (right) images demonstrate suspicious microcalcifications (arrows). All readers detected microcalcifications with SM and DM alone. They assessed the lesions as punctate or amorphous and categorized them as BI-RADS 3 or 4a. These microcalcifications were diagnosed as microinvasive ductal carcinoma by stereotactic vacuum-assisted biopsy. SM, synthetic mammography; DM, digital mammography

Table 5 Interobserver agreements of diagnostic performances of SM and DM

	With DBT		Without DBT	
	SM	DM	SM	DM
BI-RADS category	0.60 (0.55-0.64)	0.62 (0.58-0.66)	0.61 (0.56-0.65)	0.61 (0.57-0.66)
Conspicuity score	0.57 (0.36-0.72)	0.53 (0.30-0.69)	0.31 (0.02-0.55)	0.39 (0.08-0.61)

Data represented by the intraclass correlation coefficient value (95% confidence interval)

DBT, digital breast tomosynthesis; SM, synthetic mammography; DM, digital mammography

characterizing microcalcifications in patients undergoing DBT-based imaging without any loss in performance.

A recent study which compared microcalcification detectability in an anthropomorphic phantom between DM, SM, and DBT images, found that the rates of correct assessments for the number, size, and shape of simulated microcalcifications were significantly higher in DM than in SM [33]. By comparing this study's results with ours, we inferred the reason for different results on in vitro and in vivo settings. The phantom study evaluated diagnostic accuracy for size, number, and shape of microcalcifications, separately. However, in our study, readers diagnosed microcalcifications by evaluating individual characteristics (size, number, shape) comprehensively, as if in clinical practice. Our readers diagnosed microcalcifications as BI-RADS category 4 or higher if any of the individual characteristics showed suspicious features [1]. Therefore, we think that the differences in assessing individual characteristics of microcalcifications between SM and DM did not directly lead to the differences in diagnostic performances with BI-RADS ratings on microcalcifications between the two modes.

This study had several limitations. First, this was a single-institution, retrospective study, and we included a limited number of cases with microcalcifications (n=90) because we only included cases that presented solely as microcalcifications without associated mammographic findings. Further studies with more microcalcifications are needed to validate SM as a replacement for DM when evaluating microcalcifications for DBT-based imaging. However, the results found here are reasonably encouraging and thought to be of value because SM showed similar or higher sensitivities in diagnosing microcalcifications compared to DM. Second, the proportion of women with mammographically dense breasts in our study cohort (89%, 172/198) was higher than those of the western population. So, the non-dense breast subgroup could not be statistically analysed due to the small number of cases. However, our results may have clinical value when considering that the dense breast group is a more challenging cohort than the non-dense breast group in mammography performance. Third, our institution is a referral centre which in itself might affect the study population as we have a higher proportion of patients with malignant microcalcifications. Therefore, the proportion of malignancy in our cohort is higher than that of the general population. Fourth, we evaluated the morphology of detected microcalcifications with the magnifier window offered by the dedicated workstation. Therefore, we did not assess the findings of additional magnification views, which were obtained to decide whether or not to perform a biopsy at the time of diagnosis, because we were focused on comparing the diagnostic performances of SM and DM for evaluating microcalcifications in patients undergoing DBT-based imaging. Although we assessed the BI-RADS categories of microcalcifications without the magnification view findings, readers' AUCs of SM and DM were high (0.88-0.94). However, some microcalcifications can be less conspicuous on DBT and enhanced ligaments may mimic calcifications on the SM reconstruction algorithm [34]. Therefore, in patients with microcalcifications, for which it is hard to make final assessments only by SM plus DBT, additional magnification views may help to fully characterize the morphology and distribution of microcalcifications.

In conclusion, diagnostic performance of SM and DM for the evaluation of microcalcifications are not significantly different, whether performed with DBT or alone. Our results may reflect that SM plus DBT is sufficient in the evaluation of microcalcifications during DBT-based imaging, without the addition of dose-requiring DM.

Acknowledgements The authors are grateful to Seonwoo Kim, Ph.D and Min-Ji Kim, MS from Samsung Biomedical Research Institute, Samsung Medical Centre, for help in the statistical analyses.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Boo-Kyung Han.

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 Seonwoo Kim, Ph.D, Samsung Biomedical Research Institute, Samsung Medical Center helped the statistical analyses of the study.

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

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- diagnostic or prognostic study
- performed at one institution

References

1. D'Orsi C, Sickles E, Mendelson E, Morris E (2013) ACR BI-RADS® Atlas, breast imaging reporting and data system. American College of Radiology, Reston
2. D'Orsi CJ, Mendelson EB, Ikeda DM (2003) ACR breast imaging and reporting data system: breast imaging atlas. American College of Radiology, Reston
3. Conant EF (2014) Clinical implementation of digital breast tomosynthesis. Radiol Clin North Am 52:499–518
4. Spangler ML, Zuley ML, Sumkin JH et al (2011) Detection and classification of calcifications on digital breast tomosynthesis and

2D digital mammography: a comparison. AJR Am J Roentgenol 196:320–324
5. Skaane P, Hofvind S, Skjennald A (2007) Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo II study. Radiology 244:708–717
6. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P (2008) Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. Eur Radiol 18:183–191
7. Gilbert FJ, Tucker L, Young KC (2016) Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. Clin Radiol 71:141–150
8. Gur D, Abrams GS, Chough DM et al (2009) Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol 193:586–591
9. Skaane P, Gullien R, Bjorndal H et al (2012) Digital breast tomosynthesis (DBT): initial experience in a clinical setting. Acta Radiol 53:524–529
10. Wallis MG, Moa E, Zanca F, Leifland K, Danielsson M (2012) Two-view and single-view tomosynthesis versus full-field digital mammography: high-resolution X-ray imaging observer study. Radiology 262:788–796
11. Zuley ML, Bandos AI, Ganott MA et al (2013) Digital breast tomosynthesis versus supplemental diagnostic mammographic views for evaluation of noncalcified breast lesions. Radiology 266:89–95
12. Gilbert FJ, Tucker L, Gillan MG et al (2015) Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). Radiology 277: 697–706
13. Kopans D, Gavenonis S, Halpern E, Moore R (2011) Calcifications in the breast and digital breast tomosynthesis. Breast J 17:638–644
14. Tagliafico A, Mariscotti G, Durando M et al (2015) Characterisation of microcalcification clusters on 2D digital mammography (FFDM) and digital breast tomosynthesis (DBT): does DBT underestimate microcalcification clusters? Results of a multicentre study. Eur Radiol 25:9–14
15. 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
16. 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
17. Gur D, Bandos AI, Rockette HE et al (2011) Localized detection and classification of abnormalities on FFDM and tomosynthesis examinations rated under an FROC paradigm. AJR Am J Roentgenol 196:737–741
18. Gennaro G, Toledano A, di Maggio C et al (2010) Digital breast tomosynthesis versus digital mammography: a clinical performance study. Eur Radiol 20:1545–1553
19. 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
20. Olgar T, Kahn T, Gosch D (2012) Average glandular dose in digital mammography and breast tomosynthesis. Rofo 184:911–918
21. Houssami N, Skaane P (2013) Overview of the evidence on digital breast tomosynthesis in breast cancer detection. Breast 22:101–108
22. Garayoa J, Chevalier M, Castillo M et al (2018) Diagnostic value of the stand-alone synthetic image in digital breast tomosynthesis examinations. Eur Radiol 28:565–572
23. Zuley ML, Guo B, Catullo VJ et al (2014) Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. Radiology 271:664–671
24. Skaane P, Bandos AI, Eben EB et al (2014) Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. Radiology 271:655–663
25. Hajian-Tilaki K (2014) Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform 48:193–204
26. Choi JS, Han BK, Ko EY et al (2016) Comparison between two-dimensional synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for the detection of T1 breast cancer. Eur Radiol 26:2538–2546
27. Park IH, Ko K, Joo J et al (2014) High volumetric breast density predicts risk for breast cancer in postmenopausal, but not premenopausal, Korean women. Ann Surg Oncol 21:4124–4132
28. Choi JS, Lee CW, Seo HJ et al. (2012) Mammographic density assessment: comparison of Volpara™ software and visual BI-RADS classification. J Korean Soc Breast Screen 9: 127–132.
29. Gur D, Zuley ML, Anello MI et al (2012) Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. Acad Radiol 19:166–171
30. Kromke M, Palomino-Schatzlein M, Mayer H et al (2016) Profiling human blood serum metabolites by nuclear magnetic resonance spectroscopy: a comprehensive tool for the evaluation of hemodialysis efficiency. Transl Res 171(71-82):e71–e79
31. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
32. Barlow WE, Chi C, Carney PA et al (2004) Accuracy of screening mammography interpretation by characteristics of radiologists. J Natl Cancer Inst 96:1840–1850
33. Peters S, Hellmich M, Stork A et al (2017) Comparison of the Detection Rate of Simulated Microcalcifications in Full-Field Digital Mammography, Digital Breast Tomosynthesis, and Synthetically Reconstructed 2-Dimensional Images Performed With 2 Different Digital X-ray Mammography Systems. Invest Radiol 52:206–215
34. Zuckerman SP CE, Weinstein S, Korhonen K, Synnestvedt M, McDonald E (2015) Early implementation of synthesized 2D in screening with digital breast tomosynthesis: a pictorial essay of early outcomes [abstr]. In: Radiological Society of North America Scientific Assembly and Annual Meeting Program. Oak Brook, Ill: Radiological Society of North America 220