

# Comparison of digital mammography and digital breast tomosynthesis in the detection of architectural distortion

Elizabeth H. Dibble<sup>1</sup> · Ana P. Lourenco<sup>1</sup> · Grayson L. Baird<sup>1,2</sup> · Robert C. Ward<sup>1</sup> ·  
A. Stanley Maynard<sup>1</sup> · Martha B. Mainiero<sup>1</sup>

Received: 17 March 2017 / Revised: 7 June 2017 / Accepted: 26 June 2017 / Published online: 14 July 2017  
© European Society of Radiology 2017

## Abstract

**Objectives** To compare interobserver variability (IOV), reader confidence, and sensitivity/specificity in detecting architectural distortion (AD) on digital mammography (DM) versus digital breast tomosynthesis (DBT).

**Methods** This IRB-approved, HIPAA-compliant reader study used a counterbalanced experimental design. We searched radiology reports for AD on screening mammograms from 5 March 2012–27 November 2013. Cases were consensus-reviewed. Controls were selected from demographically matched non-AD examinations. Two radiologists and two fellows blinded to outcomes independently reviewed images from two patient groups in two sessions. Readers recorded presence/absence of AD and confidence level. Agreement and differences in confidence and sensitivity/specificity between DBT versus DM and attendings versus fellows were examined using weighted Kappa and generalised mixed modeling, respectively.

**Results** There were 59 AD patients and 59 controls for 1,888 observations ( $59 \times 2$  (cases and controls)  $\times 2$  breasts  $\times 2$  imaging techniques  $\times 4$  readers). For all readers, agreement improved with DBT versus DM (0.61 vs. 0.37). Confidence was higher with DBT,  $p = .001$ . DBT achieved higher sensitivity (.59 vs. .32),  $p < .001$ ; specificity remained high (>.90).

DBT achieved higher positive likelihood ratio values, smaller negative likelihood ratio values, and larger ROC values.

**Conclusions** DBT decreases IOV, increases confidence, and improves sensitivity while maintaining high specificity in detecting AD.

## Key points

- Digital breast tomosynthesis decreases interobserver variability in the detection of architectural distortion.
- Digital breast tomosynthesis increases reader confidence in the detection of architectural distortion.
- Digital breast tomosynthesis improves sensitivity in the detection of architectural distortion.

**Keywords** Mammography · Breast · Early diagnosis · Digital breast tomosynthesis · Architectural distortion

## Abbreviations

AD	Architectural distortion
DM	Digital mammography
DBT	Digital breast tomosynthesis
IOV	Interobserver variability

## Introduction

Architectural distortion (AD) is a subtle mammographic finding that can be a manifestation of breast cancer. While AD can be due to a variety of malignant and non-malignant causes [1], the positive predictive value for malignancy is approximately 75% [2]. AD may be the earliest manifestation of breast cancer [3] and is the most commonly missed abnormality on false-negative mammograms [4]. Earlier detection of AD may improve patient prognosis more than earlier detection of calcifications [3].

✉ Elizabeth H. Dibble  
edibble@lifespan.org

<sup>1</sup> Department of Diagnostic Imaging, The Warren Alpert Medical School of Brown University/Rhode Island Hospital, Providence, RI, USA

<sup>2</sup> Lifespan Biostatistics Core, Rhode Island Hospital, 593 Eddy Street, Providence, RI 02903, USA

Compared to digital mammography (DM), digital breast tomosynthesis (DBT) is a newer imaging technology that displays thin slices of breast tissue, thus mitigating the effects of tissue overlap. DBT has been shown to increase cancer detection rates and decrease screening call-back rates [5–9]. In addition, use of screening DBT allows more recalled patients to undergo ultrasound alone [10], thus potentially improving diagnostic workflow efficiency for patients recalled from screening. The risk of malignancy in abnormalities detected only by DBT is significant [11], and has been reported at nearly 50% [12]. DBT is particularly helpful in detecting abnormalities in patients with dense breasts [13, 14].

Minimising disagreement in difficult cases is the best way to reduce inconsistencies in screening mammogram interpretation [15]. AD has high interobserver variability (IOV) [16], and while sensitivity for AD is lower than for non-AD manifestations of breast cancer [17], DBT improves detection of AD [11]. The purpose of this study was to compare IOV, reader confidence and sensitivity/specificity in detecting AD on DM versus DBT.

## Materials and methods

### Study design

This reader study, approved by the Institutional Review Board and compliant with the United States Health Insurance Portability and Accountability Act, used a counterbalanced experimental design to estimate the effects of DBT relative to DM regarding sensitivity/specificity, IOV (or reader agreement) and reader confidence in detecting AD. Informed consent was waived.

### Selection of patient images

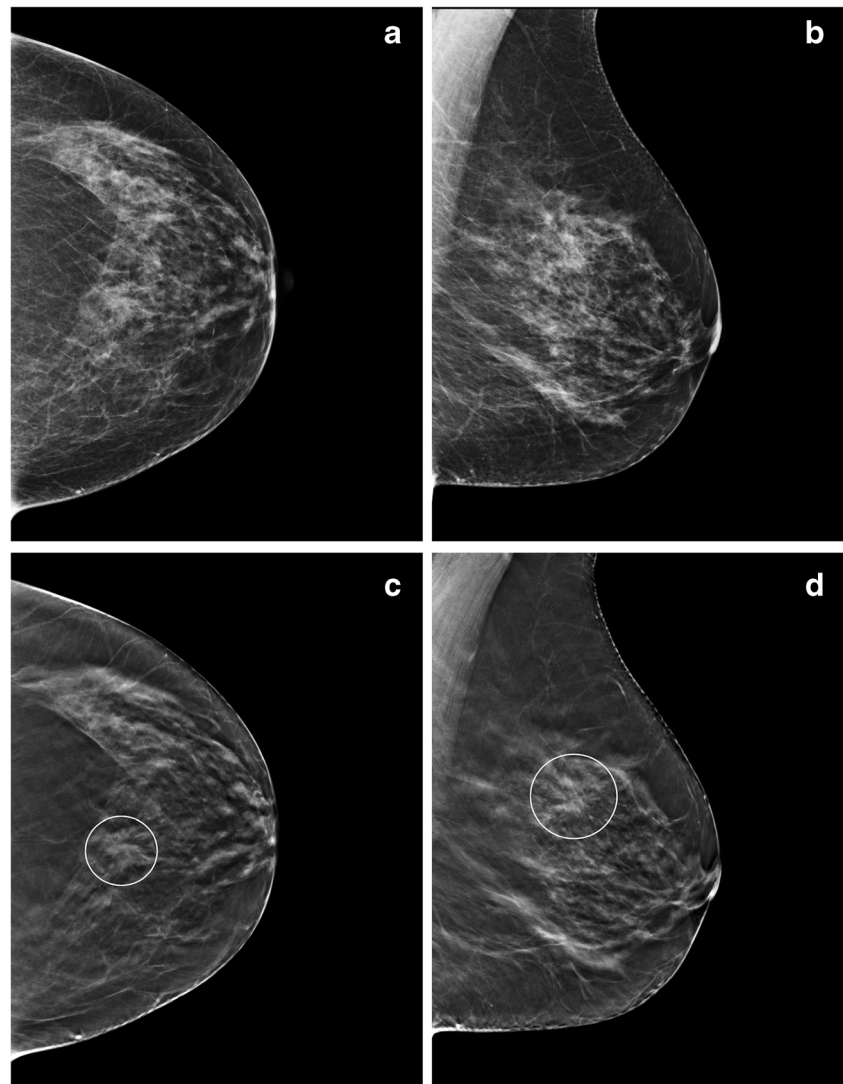
The radiology database at a tertiary breast centre was searched using the PenRad Management Information System (PenRad Technologies Inc., Buffalo, MI, USA) for all reports containing the words ‘architectural distortion’ or ‘possible architectural distortion’ on all screening mammograms performed from 5 March 2012 to 27 November 2013. Unilateral examinations were excluded to decrease a true positive hit by chance and allow each breast to act as a control within a subject. All studies consisted of standard two-view full-field DM images and DBT reconstructions in both the mediolateral oblique and craniocaudal projections. Both DM and DBT images were obtained on 3D units (Selenia Dimensions, Hologic, Marlborough, MA, USA) in a single compression episode for each projection. All patients had both DM and DBT (not synthetic mammography views reconstructed from DBT data). DBT images were obtained through the motion of the x-ray tube over a 15° arc and reconstructed into 1-mm sections.

DM and DBT images from all reports containing the words AD or possible AD were consensus-reviewed to confirm the presence of AD or possible AD (Fig. 1) on screening views on a 5 megapixel liquid crystal display (LCD) diagnostic workstation (SecurView, Hologic, Marlborough, MA, USA) by three radiologists (one breast imaging fellow and two fellowship-trained breast imagers). At the time of the consensus review, the breast imagers had 6 and 16 years of breast imaging experience in practice, respectively. This consensus review took place 3 years prior to the current study. Our institution began using reconstructed two-dimensional images after the consensus review and before this study was performed. To maintain the integrity of the methods for the current experimental study, we did not collect additional AD cases using reconstructed two-dimensional imaging. In addition, the 3-year delay provided the beneficial effects of memory decay and retrograde interference. Control cases were identified through searches on our institution’s PACS (GE Healthcare; Chicago, IL, USA) and through MONTAGE™ Search and Analytics software (Montage Healthcare Solutions, Philadelphia, PA, USA). Controls were matched for age, breast density (as described in the screening mammogram report by the reading radiologist), presence and side of prior malignancy, presence and side of new malignancy on the presented mammogram, presence and side of prior surgery, and date of mammogram when possible. While matching for breast cancer history may have increased the breast cancer risk profile of the study sample in both the AD and non-AD groups compared to the general screening population, it controlled for the possibility that either group would be more complex or at higher risk than its counterpart. The ratio of case/control was one AD patient/one control patient. All cases and controls were bilateral examinations. Patient demographics, imaging findings, pathology findings and follow-up imaging results were obtained from the electronic medical record and recorded. Imaging and clinical follow-up through May 2016 were reviewed via the electronic medical record, thus all cases had at least 2 years of follow-up available.

### Review of images

Two breast radiologists (9 and 19 years experience, respectively) and two breast imaging fellows in the second half of their 1-year breast imaging fellowships who were blinded to patient information, prior imaging, and outcomes independently reviewed the DBT and DM images from screening mammograms done with combined technique in two separate reading sessions. In the first session, for half of the cases (1–59), only the DBT images were reviewed by radiologists and for the other half of the cases (60–118) only the DM images were reviewed. In the second session (at least 1 month later), only the DM images were reviewed for those cases in which DBT images had been previously evaluated (1–59), and only DBT images were reviewed for cases in which DM had been previously

**Fig. 1** A 58-year-old woman with left breast architectural distortion. Left breast (a) craniocaudal and (b) mediolateral oblique digital mammogram images show heterogeneously dense breast tissue. Four out of four study readers did not detect architectural distortion on this patient's digital mammogram images. Left breast (c) craniocaudal and (d) mediolateral oblique digital breast tomosynthesis images show architectural distortion in the 12 o'clock location (white circles). The architectural distortion was detected by four out of four study readers on this patient's digital breast tomosynthesis images. The patient went on to biopsy and subsequent excision revealing invasive ductal carcinoma on pathology



evaluated (60–118). While the order of patient images was held constant across sessions 1 and 2, the order itself was randomly assigned AD versus no AD. This counterbalanced experimental design allowed patient images and radiologists to be held constant while only imaging technique (DBT vs. DM) was manipulated; therefore, observed differences in radiologists' performance can be attributed to the direct effect of imaging technique.

### Measures

For each breast, readers recorded the presence or absence of AD or possible AD (i.e., Yes/No), the location in the breast using clock face position, and their confidence in that interpretation on a scale of 0–4 (i.e., How confident? 0 'Not at all', 1 'Somewhat', 2 'Confident', 3 'Very confident', 4 'Almost sure'). AD or possible AD that was due to post-surgical change and clearly identified as such by the reader was not coded as AD or possible AD. Only unexplained AD or

unexplained possible AD was coded as AD or possible AD for the purposes of this study in order to reproduce the clinical setting in which architectural distortion that is clearly due to prior surgery is not clinically significant. The studies were interpreted on a 5 megapixel LCD diagnostic workstation (SecurView, Hologic).

### Statistical methods

All analyses were conducted using SAS Software 9.4 (SAS Inc., Cary, NC, USA). Agreement was examined using weighted Kappa. Differences in confidence between DBT versus DM and between attending radiologists versus fellows were examined using generalised mixed modeling assuming a binomial distribution. Differences in sensitivity/specificity between DBT versus DM and between attending radiologists versus fellows were examined using generalised mixed modeling assuming a binary distribution. Modeling was accomplished with sandwich estimation using PROC

GLIMMIX, where images were nested within patients and patients nested within radiologists. Positive likelihood ratios were calculated as sensitivity/(1 – specificity) and negative likelihood ratios were calculated as (1 – sensitivity)/specificity. Receiver operating characteristic (ROC) estimates were calculated using PROC LOGISTIC. Statistical significance was established *a priori* at the 0.05 level and all interval estimates are calculated for 95% confidence. Multiple comparisons were examined using Tukey correction.

## Results

Of 25,369 screening DBT examinations, there were 84 reports of AD or possible AD. Thirty-two cases were excluded because AD or possible AD was not confirmed on consensus review. Four cases were excluded because they were unilateral examinations. The 59 remaining patients had 59 AD lesions (all cases were single lesions in one breast of a bilateral screening mammogram) and were matched to 59 controls for a total of 1,888 observations (59 × 2 (cases and controls) × 2 breasts × 2 imaging techniques × 4 readers, or 472 observations per reader). No differences were observed between AD and non-AD patients on matched variables (see Table 1). Results of biopsy or imaging follow-up of AD cases are shown in Table 2. Although the purpose of the present study was to examine consensus-reviewed AD and possible AD, almost half of our identified cases were confirmed AD that persisted on additional imaging and required biopsy (27 of the 59). Thirty-two out of the 59 cases resolved with additional imaging or turned out to be postsurgical change that was not readily apparent to the consensus reviewers.

### Agreement (Tables 3 and 4)

Overall agreement among radiologists was fair to moderate [18] for DM ( $\kappa = 0.37$ ) and moderate to good for DBT ( $\kappa = 0.61$ ). Agreement between attendings was fair to moderate for DM ( $\kappa = 0.40$ ) and good for DBT ( $\kappa = 0.72$ ). Agreement

between fellows was fair for DM ( $\kappa = 0.34$ ) and moderate for DBT ( $\kappa = 0.57$ ). In addition, as seen in Table 4, agreement overall was much higher when using DBT ( $\kappa = .71$ ) relative to DM ( $\kappa = .35$ ); this held true for both attendings ( $\kappa = 0.76$  for DBT versus 0.46 for DM) and fellows ( $\kappa = 0.60$  for DBT vs. 0.34 for DM) for the 27 confirmed AD cases.

### Reader confidence (Tables 3 and 4)

Level of confidence in detecting AD was higher when using DBT compared with DM (3.2 vs. 2.6 on a 0–4 scale),  $p < .001$ . Attendings' level of confidence in detecting AD was higher when using DBT compared with DM (3.7 vs. 3.1 on a 0–4 scale),  $p < .001$ . Fellows' level of confidence in detecting AD was higher when using DBT compared with DM (2.4 vs. 1.8 on a 0–4 scale),  $p < .001$ . As seen in Table 4, confidence also increased when using DBT (2.5) relative to DM (3.1) for confirmed cases of AD,  $p < .0001$ .

### Sensitivity and specificity (Tables 3 and 4)

Overall, DBT achieved higher sensitivity than DM (.59 vs. .32),  $p = .0006$ . Sensitivity for attendings was much higher for DBT than DM (.69 vs. .29),  $p < .0001$ . Sensitivity for fellows was higher for DBT than DM (.49 vs. .35),  $p < .0001$ . Specificity remained high for both DBT and DM for both attendings and fellows (>.90). The small reduction in specificity from .96 to .93 observed for attendings was offset by the increase in specificity from .91 to .94 for fellows, resulting in no significant change overall when combined. In addition, DBT achieved higher positive likelihood ratio values, smaller negative likelihood ratio values, and larger ROC values relative to DM (Table 3). Sensitivity and specificity were also examined for the 27 confirmed AD cases. As seen in Table 4, overall sensitivity increased when using DBT relative to DM (.41 vs. .86). In particular, sensitivity increased dramatically for attendings using DBT relative to DM (.97 vs. .38),  $p < .001$ , for confirmed AD cases; increase in sensitivity was also observed for fellows using DBT relative to DM for confirmed AD (.75 vs. .43),  $p < .0001$ . Specificity remained

**Table 1** Patient demographics

	AD patients	Control patients	<i>p</i> -value
Age (years)	58.9 (range 42–86)	57.5 (range 41–77)	0.621
Heterogeneously or extremely dense breasts	47/59 (79.7%)	46/59 (78.0%)	0.999
History of prior breast surgery	14/59 (23.7%)	15/59 (25.4%)	0.999
History of prior breast malignancy	6/59 (10.2%)	7/59 (11.9%)	0.999
Biopsy-proven cancer on presented mammogram	15/59 (25.4%)	10/59 (16.9%)	0.368
Range of mammogram dates (median days between cases and controls = 23)	5/3/12–27/11/13	5/3/12–27/11/13	N/A

AD architectural distortion, N/A not applicable

**Table 2** Architectural distortion (AD) pathology or imaging follow-up

	No. of AD patients
Pathology based on final excision	26
Cancer	15
High-risk lesion	11
Radial scar	6
ADH	3
Lobular neoplasia	2
Pathology based on biopsy	1
Columnar cell changes	1
Imaging follow-up only (Bi-RADS 1, 2, or 3)	32*

\* AD/possible AD resolved with additional imaging or found to represent post-surgical change after additional history obtained. None of these patients had breast cancer at greater than 2 years of follow-up

ADH atypical ductal hyperplasia

high throughout (.90–.96). Figure 2 shows area under the curve values for attendings and fellows for all AD and possible AD cases versus just for confirmed AD cases.

**Discussion**

AD is a subtle but clinically important mammographic finding that may be the earliest manifestation of breast cancer [3]. The sensitivity for AD is lower than for non-AD manifestations of breast cancer [17], and AD is the most commonly missed abnormality on false-negative mammograms [4]. The results of our study indicate that, compared to DM, DBT decreases IOV, increases reader confidence, increases sensitivity, and maintains high specificity in the detection of AD. Because our study design held patient images constant (i.e., patient images serve as their own controls) and radiologists constant (i.e., radiologists serve as their own controls) while only imaging technique (DBT vs. DM) was manipulated, observed

differences in radiologists’ performance (e.g., sensitivity, specificity) regarding patient images can be attributed to the direct effect of imaging technique.

IOV is an assessment of radiologists’ consensus and, particularly in the setting of expert readers, ambiguity of difficult findings. Because there is no gold standard to confirm mammographic findings other than expert consensus, improving consistency in clinical practice is critical to providing the highest quality care. In a study performed prior to DBT, AD was found to have high IOV compared to other mammographic abnormalities [16]. Our study shows that DBT also decreases IOV compared to DM. Consistency is increased among our four readers together as well as between our two experienced readers and between our two less experienced readers.

In addition, reader confidence was significantly higher when using DBT compared with DM. Confidence in mammogram interpretation is associated with improved accuracy, particularly among low volume readers [19]. Similarly, Tucker et al. recently showed that the increase in sensitivity with DBT is greater for readers with less than 10 years of experience [20]. Compared to DM, DBT allowed for increased detection of AD by all of our readers when examined together as well as by the experienced readers alone (i.e., when examined without the less experienced readers) and by the less experienced readers alone (i.e., when examined without the experienced readers).

The sensitivity for detecting AD in our study was 29–35% with DM and 49–69% with DBT, lower than that found by Suleiman et al. at 87% [17], although all readers in that study were ‘experienced breast screen readers’ and all cases contained a biopsy-proven malignancy. The sensitivity of our attending readers was higher than that of our fellow readers but still lower than that found by Suleiman et al. This is likely due to the differences in case mix. Our cases were consensus reviewed as having AD or possible AD on screening images, although 32/59 cases resolved with

**Table 3** Agreement, reader confidence, sensitivity, specificity, and positive and negative likelihood ratios for architectural distortion (AD) and possible AD cases

	Overall			Attending			Fellow		
	DM	DBT	p-value	DM	DBT	p-value	DM	DBT	p-value
Agreement	0.37	0.61		0.4	0.72		0.34	0.57	
Confidence	2.6 [1.9, 3.1]	3.2 [2.4, 3.7]	0.001	3.1 [2.9, 3.3]	3.7 [3.6, 3.9]	0.0066	1.84 [1.8, 1.9]	2.4 [2.0, 2.8]	0.0086
Sensitivity	0.32 [0.26, 0.39]	0.59 [0.49, 0.69]	0.0006	0.29 [0.22, 0.38]	0.69 [0.69, 0.70]	<.0001	0.35 [0.27, 0.44]	0.49 [0.43, 0.55]	<.0001
Specificity	0.93 [0.91, 0.95]	0.93 [0.93, 0.94]	0.9997	0.96 [0.95, 0.97]	0.93 [0.93, 0.94]	0.0316	0.91 [0.91, 0.91]	0.94 [0.93, 0.95]	<.0001
+LR	4.57	8.43		7.25	9.86		3.89	8.17	
-LR	0.73	0.44		0.74	0.33		0.71	0.54	

DM digital mammography, DBT digital breast tomosynthesis

**Table 4** Agreement, reader confidence, sensitivity, specificity, and positive and negative likelihood ratios for confirmed architectural distortion (AD) cases

	Overall			Attending			Fellow		
	DM	DBT	<i>p</i> -value	DM	DBT	<i>p</i> -value	DM	DBT	<i>p</i> -value
Agreement	0.35	0.71		0.46	0.76		0.34	0.60	
Confidence	2.49 [2.4, 2.6]	3.13 [3.1, 3.2]	<.0001	3.14 [3.0, 3.3]	3.7 [3.7, 3.9]	<.0001	1.84 [1.7, 1.9]	2.5 [2.4, 2.6]	<.0001
Sensitivity	0.41 [.38, .44]	0.86 [.71, .94]	<.0001	0.38 [0.37, 0.39]	0.97 [0.97, 0.97]	<.001	0.43 [0.42, 0.45]	0.75 [0.67, 0.81]	<.0001
Specificity	0.93 [.89, .95]	0.94 [.93, .95]	<.8840	0.96 [0.95, 0.96]	0.93 [0.92, 0.94]	<.0001	0.90 [0.89, 0.90]	0.95 [0.93, 0.95]	<.0001
+LR	5.86	14.33		9.5	13.9		4.3	15	
-LR	0.64	0.15		0.65	0.03		0.63	0.26	

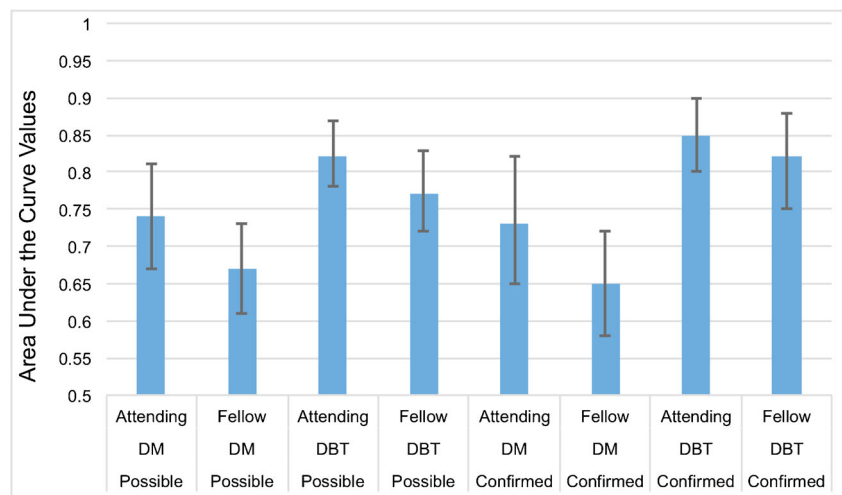
DM digital mammography, DBT digital breast tomosynthesis

additional imaging or turned out to be postsurgical change that was not readily apparent to the consensus reviewers. Scars from prior benign surgeries were not always marked by the technologist, thus some cases of postsurgical distortion were not prospectively identified as such. Many of the cases that were not identified as having AD by our four readers were from this group of 32 cases. When we examined only confirmed AD cases (i.e., only the cases that persisted on additional imaging and required biopsy), sensitivity increased to 97% for attendings and 75% for fellows, which is similar to the sensitivity reported by Suleiman et al. It is possible that additional years of experience with DBT between the time of the consensus review and the time of this study helped the readers to better select only the true cases of AD. Alternatively, it may be that studies selected as having possible AD or AD during the consensus review process were already known to have been read as AD or possible AD by the original radiologist, while in this study the readers also had control cases and were blinded to the original interpretation. Patient information or previous imaging could have contributed to the original designation of AD, but our readers were blinded to patient information and previous

imaging. A study by Partyka et al. showed that some AD is seen better or only on DBT compared to DM [11], and our cases were consensus reviewed using both DM and DBT (as in clinical practice but not in our experiment); this likely accounts for some of the low sensitivity seen with DM alone in our study. Importantly, the benefit of increased sensitivity of DBT compared to DM did not come at a cost to specificity, as specificity remained high for both DBT and DM for both attendings and fellows. Although a statistically significant difference was observed for both groups individually, there was no statistically significant difference when the groups were combined. The 3% change (decrease for attendings and increase for fellows) when specificity remained >90% may not be clinically relevant.

The relative benefit in diagnostic performance of DBT compared with DM can also be seen from the likelihood estimates and area under the curve estimates. Specifically, overall positive likelihood almost doubled when using DBT relative to DM and area under the curve increased by 10%, both indicating a stronger association with AD for DBT relative to DM. Though both groups experienced these gains, fellows in

**Fig. 2** Area under the curve values for all architectural distortion (AD) and possible AD cases ('Possible') versus just for confirmed AD cases ('Confirmed')



particular experienced greater gains in diagnostic performance when using DBT.

Our study design controlled for bias in the following ways: (1) Radiologists were blinded to patient information (e.g., patient identifiers, outcomes), thus reducing the chance of recall due to patient recognition; (2) Reading sessions were separated by 1 month in time and the same patients were viewed using different imaging techniques thus reducing the likelihood that radiologists would recall images from a different imaging technique from the previous month (this time lag of 1 month was selected to ensure the effects of memory decay and retrograde interference of memory given the high volume of breast imaging interpreted by the radiologists in clinical practice); (3) Recall bias, if present, would be the same for DBT and DM because of the counterbalanced design. In addition, although the cases were selected 3 years prior to the study, memory decay and retrograde interference would reduce recall bias after reading thousands of images in clinical practice over the intervening years; (4) To control for order effects, the order of patient images was held constant across sessions 1 and 2, but the order itself was randomly assigned AD versus no AD using a block design.

Our study has some limitations. As this was an efficacy study, a 1:1 matched AD/non-AD design was used to optimise detection in change of sensitivity and specificity equally; positive predictive value and false-positive rate cannot be assessed because of their inherent relationship with prevalence. In addition, as the goal of our study was to examine AD (known to arise from a variety of aetiologies), we included all cases of consensus-reviewed AD, not just those cases due to cancer. A recent study [21] comparing single view DBT and DM showed an increase in recalls for stellate distortions with DBT, most assessed as normal breast tissue after additional imaging; the second most common aetiology was radial scar [21]. Radial scars diagnosed on percutaneous biopsy risk being upgraded to malignancy and thus are often excised, but increased detection of radial scars with DBT compared to DM warrants additional investigation into the need for excision in all cases particularly when there is no evidence of atypia [22, 23]. The inclusion of all AD and possible AD cases in our study, benign and malignant, was to reflect a more real-world collection of cases as seen in clinical practice. While our findings suggest that DBT may improve detection of cancer presenting as AD, our study did not specifically address this. Our study also examined DM alone versus DBT alone, although DBT is used in conjunction with DM or synthesised two-dimensional imaging in clinical practice. When reading DM in conjunction with DBT, the presence of AD on DBT alone (i.e. AD not visible on corresponding DM) would likely prompt further work-up by most radiologists, although we did not specifically address this scenario with our study design. Use of digitally reconstructed two-dimensional images from DBT, as opposed to use of a separate DM acquisition, is

increasing. We believe that our results with DM would translate to reconstructed images because both are two-dimensional techniques as opposed to DBT, although we did not examine this in our study. Finally, our cases of AD were agreed upon consensus review by three radiologists, but different consensus reviewers could disagree on the presence of AD in our case mix.

In conclusion, digital breast tomosynthesis decreases IOV, increases reader confidence, and improves sensitivity while maintaining high specificity in detecting architectural distortion.

#### Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Ana P. Lourenco, MD.

**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.

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

**Statistics and biometry** One of the authors has significant statistical expertise.

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

**Ethical approval** Institutional Review Board approval was obtained.

#### Methodology

- Prospective
- Experimental
- Performed at one institution

#### References

1. Shaheen R, Schimmelpenninck CA, Stoddart L, Raymond H, Slanetz PJ (2011) Spectrum of diseases presenting as architectural distortion on mammography: multimodality radiologic imaging with pathologic correlation. *Semin Ultrasound CT MR* 32:351–362
2. Bahl M, Baker JA, Kinsey EN, Ghate SV (2015) Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy. *AJR Am J Roentgenol* 205:1339–1345
3. Gaur S, Dialani V, Slanetz PJ, Eisenberg RL (2013) Architectural distortion of the breast. *AJR Am J Roentgenol* 201:W662–W670
4. Burrell HC, Sibbering DM, Wilson AR et al (1996) Screening interval breast cancers: mammographic features and prognosis factors. *Radiology* 199:811–817
5. Andersson I, Ikeda DM, Zackrisson S et al (2008) Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings. *Eur Radiol* 18:2817–2825
6. 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
7. Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R Jr (2013) Implementation of breast tomosynthesis in a routine screening practice: an observational study. *AJR Am J Roentgenol* 200:1401–1408
  8. 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
  9. Durand MA, Haas BM, Yao X et al (2015) Early clinical experience with digital breast tomosynthesis for screening mammography. *Radiology* 274:85–92
  10. Lourenco AP, Barry-Brooks M, Baird GL, Tuttle A, Mainiero MB (2015) Changes in recall type and patient treatment following implementation of screening digital breast tomosynthesis. *Radiology* 274:337–342
  11. Partyka L, Lourenco AP, Mainiero MB (2014) Detection of mammographically occult architectural distortion on digital breast tomosynthesis screening: initial clinical experience. *AJR Am J Roentgenol* 203:216–222
  12. Freer PE, Niell B, Rafferty EA (2015) Preoperative Tomosynthesis-guided Needle Localisation of Mammographically and Sonographically Occult Breast Lesions. *Radiology* 275:377–383
  13. Haas BM, Kalra V, Geisel J, Raghu M, Durand M, Philpotts LE (2013) Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology* 269:694–700
  14. Ray KM, Turner E, Sickles EA, Joe BN (2015) Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings. *Breast J* 21:538–542
  15. Beam CA, Conant EF, Sickles EA (2002) Factors affecting radiologist inconsistency in screening mammography. *Acad Radiol* 9: 531–540
  16. Onega T, Smith M, Miglioretti DL et al (2012) Radiologist agreement for mammographic recall by case difficulty and finding type. *J Am Coll Radiol* 9:788–794
  17. Suleiman WI, McEntee MF, Lewis SJ et al (2016) In the digital era, architectural distortion remains a challenging radiological task. *Clin Radiol* 71:e35–e40
  18. Altman DG (1991) *Practical statistics for medical research*, 1st edn. Chapman and Hall, London
  19. Geller BM, Bogart A, Carney PA, Elmore JG, Monsees BS, Miglioretti DL (2012) Is confidence of mammographic assessment a good predictor of accuracy? *AJR Am J Roentgenol* 199:W134–W141
  20. 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
  21. Lang K, Nergarden M, Andersson I, Rosso A, Zackrisson S (2016) False positives in breast cancer screening with one-view breast tomosynthesis: An analysis of findings leading to recall, work-up and biopsy rates in the Malmo Breast Tomosynthesis Screening Trial. *Eur Radiol* 26:3899–3907
  22. Kalife ET, Lourenco AP, Baird GL, Wang Y (2016) Clinical and Radiologic Follow-up Study for Biopsy Diagnosis of Radial Scar/Radial Sclerosing Lesion without Other Atypia. *Breast J* 22:637–644
  23. Ferreira AI, Borges S, Sousa A et al (2017) Radial scar of the breast: Is it possible to avoid surgery? *Eur J Surg Oncol*. doi:10.1016/j.ejso.2017.01.238