



# Impact of super-resolution and image acquisition on the detection of calcifications in digital breast tomosynthesis

Bruno Barufaldi<sup>1</sup> · Raymond J. Acciavatti<sup>1</sup> · Emily F. Conant<sup>1</sup> · Andrew D. A. Maidment<sup>1</sup>

Received: 23 February 2023 / Revised: 9 June 2023 / Accepted: 8 July 2023 / Published online: 12 August 2023  
© The Author(s), under exclusive licence to European Society of Radiology 2023

## Abstract

**Objectives** A virtual clinical trial (VCT) method is proposed to determine the limit of calcification detection in tomosynthesis.

**Methods** Breast anatomy, focal findings, image acquisition, and interpretation ( $n = 14$  readers) were simulated using screening data ( $n = 660$  patients). Calcifications ( $0.2\text{--}0.4\text{ mm}^3$ ) were inserted into virtual breast phantoms. Digital breast tomosynthesis (DBT) acquisitions were simulated assuming various acquisition geometries: source motion (continuous and step-and-shoot), detector element size (140 and  $70\text{ }\mu\text{m}$ ), and reconstructed voxel size ( $35\text{--}140\text{ }\mu\text{m}$ ). VCT results were estimated using multiple-reader multiple-case analyses and  $d'$  statistics. Signal-to-noise (SNR) analyses were also performed using BR3D phantoms.

**Results** Source motion and reconstructed voxel size demonstrated significant changes in the performance of imaging systems. Acquisition geometries that use  $70\text{ }\mu\text{m}$  reconstruction voxel size and step-and-shoot motion significantly improved calcification detection. Comparing  $70$  with  $100\text{ }\mu\text{m}$  reconstructed voxel size for step-and-shoot, the  $\Delta\text{AUC}$  was  $0.0558$  ( $0.0647$ ) and  $d'$  ratio was  $1.27$  ( $1.29$ ) for  $140\text{ }\mu\text{m}$  ( $70\text{ }\mu\text{m}$ ) detector element size. Comparing step-and-shoot with a continuous motion for a  $70\text{ }\mu\text{m}$  reconstructed voxel size, the  $\Delta\text{AUC}$  was  $0.0863$  ( $0.0434$ ) and the  $d'$  ratio was  $1.40$  ( $1.19$ ) for  $140\text{ }\mu\text{m}$  ( $70\text{ }\mu\text{m}$ ) detector element. Small detector element sizes (e.g.,  $70\text{ }\mu\text{m}$ ) did not significantly improve detection. The SNR results with the BR3D phantom show that calcification detection is dependent upon reconstructed voxel size and detector element size, supporting VCT results with comparable agreement (ratios:  $d' = 1.16 \pm 0.11$ ,  $\text{SNR} = 1.34 \pm 0.13$ ).

**Conclusion** DBT acquisition geometries that use super-resolution (smaller reconstructed voxels than the detector element size) combined with step-and-shoot motion have the potential to improve the detection of calcifications.

**Clinical relevance** Calcifications may not always be discernable in tomosynthesis because of differences in acquisition and reconstruction methods. VCTs can identify strategies to optimize acquisition and reconstruction parameters for calcification detection in tomosynthesis, most notably through super-resolution in the reconstruction.

## Key Points

- Super-resolution improves calcification detection and SNR in tomosynthesis; specifically, with the use of smaller reconstruction voxels.
- Calcification detection using step-and-shoot motion is superior to that using continuous tube motion.
- A detector element size of  $70\text{ }\mu\text{m}$  does not provide better detection than  $140\text{ }\mu\text{m}$  for small calcifications at the threshold of detectability.

**Keywords** Digital breast tomosynthesis · Imaging phantoms · Breast cancer · Computer simulations

## Abbreviations

AUC	Area under the curve
CC	Cranial-caudal
CHO	Channelized Hotelling observer
DBT	Digital breast tomosynthesis
ML	Medio-lateral
MRMC	Multiple-reader multiple-case
PMMA	Polymethyl methacrylate
ROC	Receiver operating characteristic

✉ Bruno Barufaldi  
Bruno.Barufaldi@pennteam.upenn.edu

<sup>1</sup> Department of Radiology, Hospital of University of Pennsylvania, 3400 Spruce Street 1 Silverstein, Philadelphia, PA 19103, USA

ROI	Region of interest
SNR	Signal-to-noise ratio
VCT	Virtual clinical trial

## Introduction

Digital breast tomosynthesis (DBT) has been shown to increase cancer detection and reduce false positive recalls in screening [1], but there is still a concern that small, and/or low-density calcifications may not be depicted well given all the changes that have occurred in breast screening [2, 3]. Commercially available DBT systems use a limited-angle acquisition and a small number of low-dose projections that produce an under-sampled dataset, compromising the detection of clinically significant calcifications [3]. Calcification detection at screening is important because they are one of the earliest signs of breast cancer [4].

Calcification detection is degraded by noise and blurring in DBT [5]. DBT systems with continuous tube motion have shorter scan times compared to step-and-shoot systems; however, the trade-off is increased blurring due to focal spot motion [6]. A previous study has shown that detector element size affects the spatial resolution of flat panel detectors [7], influencing the effective signal–noise ratio and the overall system modulation transfer function [8]. Preliminary data has been previously reported to demonstrate that the reconstructed voxel size also affects spatial resolution and image blur in DBT reconstructions [9]. However, there is a limit to how much the image quality can be improved by varying the reconstructed voxel size [10].

Virtual clinical trials (VCTs) have been widely used to evaluate the performance of virtual readers to detect calcifications [11–15]. VCTs are beneficial for the development and optimization of novel imaging systems prior to clinical trials [11, 13], making the design process more cost-effective and rapid [16]. VCTs are used to simulate human anatomy, image acquisition, and image interpretation [16]. VCTs have the potential to support regulatory and pre-clinical evaluations [15] and to evaluate and optimize the performance of imaging systems using virtual phantoms [11]. Simulations of the breast anatomy require high resolution to depict accurately the fine details associated with linear structures (e.g., ductal network, blood vessels, and Cooper’s ligaments) and more focal findings (e.g., calcifications, masses, and architectural distortion) [16].

This work presents a VCT method to determine the limit of calcification detection in DBT. Single calcifications of various sizes and densities were inserted into *in-silico* breast phantoms [11]. DBT acquisitions were simulated assuming various acquisition geometries. Variations of x-ray tube motion, detector element size, and reconstructed voxel size (allowing for super-resolution) were simulated to evaluate

detection performance. Channelized Hotelling observers [17] (CHOs) were used to evaluate calcification detection with multiple-reader multiple-case (MRMC) receiver operating characteristic (ROC) analyses [18].

## Material and methods

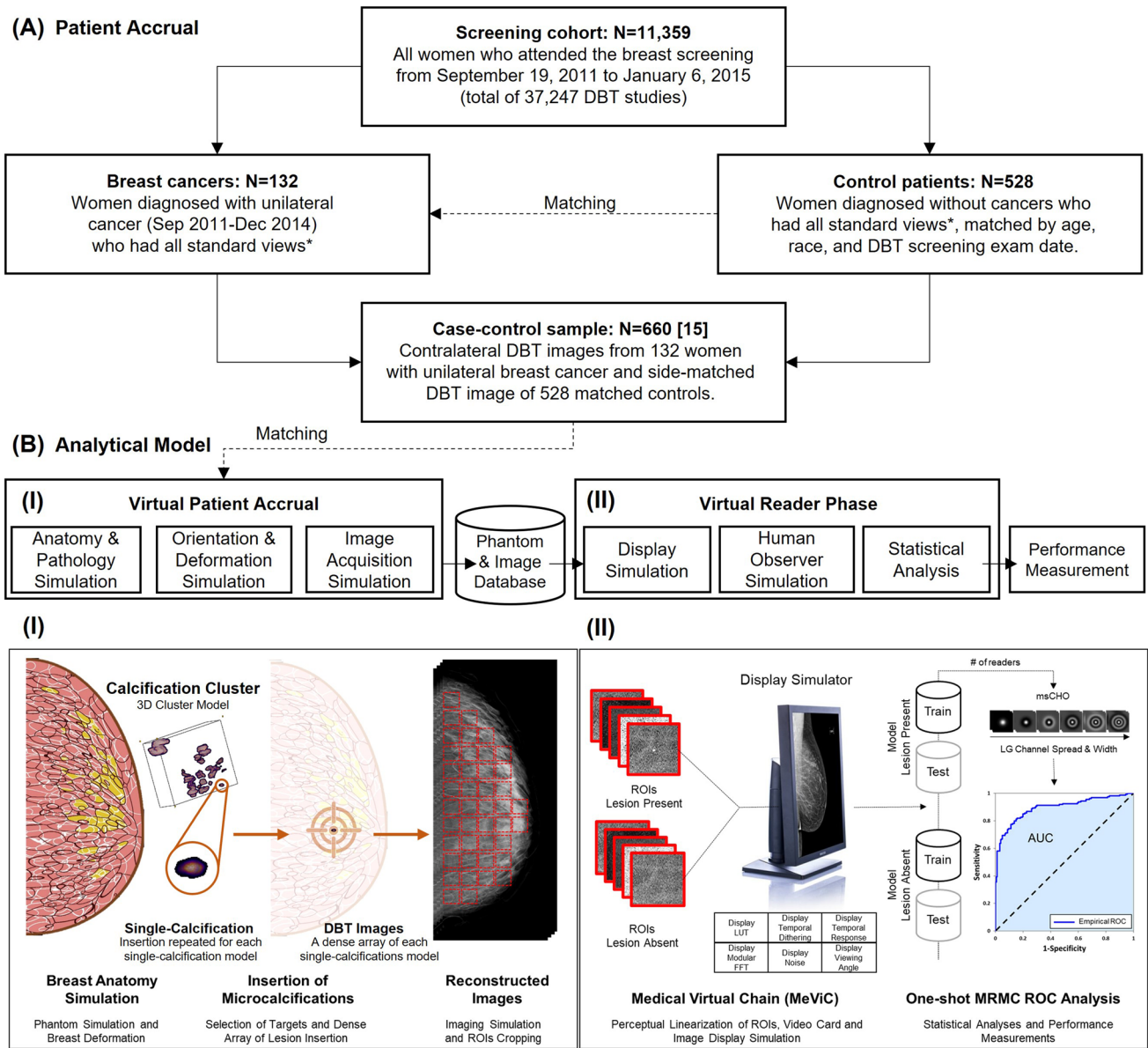
Screening data was collected retrospectively (Fig. 1A) [19]; institutional review board approved and Health Insurance Portability and Protection Act compliant. Screening exams were performed on four Selenia Dimensions systems (Hologic) from Sep 2011 to Jan 2015. A case–control cohort consisting of 132 cancer patients and 528 matched controls was selected for modeling [19]. Control patients were matched using compressed breast thickness (CBT), body mass index (BMI), and BI-RADS density [20].

A VCT pipeline was developed to model and conduct clinical trials of imaging systems (Fig. 1B). The steps used to conduct the VCTs include image pre-processing, image post-processing, and data analysis. The VCT pipeline consists of two major phases: virtual patient accrual and virtual reader simulation. We used this pipeline to calibrate calcifications below, at, and (slightly) above the threshold of lesion detectability (i.e., various levels of case difficulty) [21].

### Determining representative calcifications at the threshold of detectability

In total, 600 anthropomorphic phantoms (voxel size of  $100 \mu\text{m}^3$ ) were simulated using a principal component analysis (PCA) method to create compressed breast outlines [22] and a GPU-based recursive partitioning algorithm [23]. The phantoms were simulated using a comparable distribution of compressed breast thickness (CBT = 57.75 [49.93, 66.68] mm,  $p < 0.001$ ) and volumetric breast density (VBD = 13.33 [8.55, 20.76]%,  $p = 0.24$ ) data of a large screening population (Table 1) [22]. In addition, 24 breast phantoms (calibration set) were simulated to calibrate the composition of single-calcification models [21]. The lesion calibration is needed to match clinically-known AUCs reported in previous reader study thereby duplicating the detection difficulty [24]. The CBT and VBD of the calibration set varied from [30, 80] mm (steps of 10 mm) and [5, 35]% (steps of 10%), respectively.

Single-calcifications acquired from microcalcification clusters were inserted into each breast phantom. The models of calcification clusters were extracted from diagnostic clinical images using a prone stereotactic breast biopsy system (Fischer MammoVision™) [25]. The single-calcification models varied in size ( $2^3$ ,  $3^3$ , and  $4^3$  voxels). The composition of the calcifications was simulated using partial volumes of hydroxyapatite [11, 21]. The



\*Only complete DBT studies (raw and processed images) were considered for VCT modeling.

**Fig. 1** (A) Patient accrual for developing analytical model of screening population. (B) VCT Pipeline and flowchart used to insert calcifications and simulate synthetic DBT images, and to model virtual

observers to collect and compare performance measurements. The method of calcification insertion and reading trials with mathematical models (CHOs) are illustrated in I and II, respectively

**Table 1** Summary of patient demographics. *CBT* compressed breast thickness, *BMI* body mass index, *VBD* volumetric breast density

Patients, (#)	660
Age, mean ± SD (years)	59.3 ± 10.7
BMI, mean ± SD (kg/m <sup>2</sup> )	28.7 ± 7.7
Density BI-RADS® {1,2,3,4}, # (%)	{79 (12), 367 (56), 201 (30), 13 (2)}
Race {Asian, Black, White, Other}, # (%)	{25 (4), 309 (47), 320 (48), 6 (1)}
CBT, median [1 <sup>st</sup> , 3 <sup>rd</sup> ] Quantile (mm)	60.50 [51.50, 69.5]
VBD, median [1 <sup>st</sup> , 3 <sup>rd</sup> ] Quantile (%)	12.75 [8.40, 20.00]

fraction of hydroxyapatite in each model was varied using a previously validated method (Figure S1) [21]. In total, 40 single-calcifications of each model were inserted in the central plane (cranial-caudal or CC view) corresponding to the mid-thickness of 24 breast phantoms selected for calibration. The calcifications were randomly distributed in the central plane of the phantoms; calcifications were distributed throughout the plane allowing some background breast tissue to surround each calcification within its bounding box (size of 6.4 mm with air exclusion, Fig. 1BI). This insertion method was repeated for each single-calcification model; results were collected and stratified as a function of calcification size (Figure S1). We simulated central projections of a digital breast tomosynthesis (DBT) system (Table 2) using clinical exposure settings reported in the literature [26]; the dose was varied as a function of CBT and VBD for each phantom. The projections were simulated using the X-ray mass attenuation reported by the ICRU [27]. A ray-tracing algorithm produced images that were processed using commercial software (Briona, version 9.02, real-time tomography LLC, Villanova, PA). Regions of interest (ROIs) of size  $4.48 \times 4.48 \text{ mm}^2$  ( $64 \times 64$  pixels) were presented with and without calcifications to virtual readers modeled as channelized Hotelling observers (CHOs, Fig. 1BII) [17]. For calibration, we used 15 channels and 9 spread as parameters of each CHO. The threshold of detectability of calcifications was matched within the 95% confidence interval (CI) reported in a previous reader study (Figure S1) [24]. The VCT method and the parameters used to calibrate lesions are described in previous publication [21].

**Table 2** Summary of parameters used to simulate the combination of various DBT acquisition geometries. NA not applicable

Simulation parameters	(A) Predicate	(B) Studied modality	
	modality (DM, calibration)	(DBT, evaluation)	
Number of acquisitions (n)	1	15	
Angular range (degrees)	0	−7.5 to 7.5	
Source image distance (mm)	700 mm		
Detector type	a-Se		
Tube motion	Stationary	Continuous	Step-and-shoot
Detector element size ( $\mu\text{m}$ )	70	140	70
Reconstructed voxel size ( $\mu\text{m}$ )	NA	100	70

## Calculating changes in performance between virtual readers

Similarly, 20 calcifications of each model were selected and randomly distributed in the central slice (CC view) of the 600 breast phantoms selected for evaluating the limit of calcification detection using various DBT acquisitions. DBT projections of phantoms with and without calcifications were simulated. The acquisition geometries were varied in terms of detector element size, tube motion, and reconstructed voxel size (Table 2B) to evaluate changes in the performance of virtual readers trained for detecting calcifications. The DBT projection and reconstruction methods described in the previous section were used.

ROIs were selected and centered on each calcification in the central slice of the reconstructed DBT image. The ROIs also measured  $4.48 \times 4.48 \text{ mm}^2$ . We varied the spread parameter of each CHO for ROIs reconstructed using 35, 50, 70, 100, and 140  $\mu\text{m}$  voxel sizes. For each CHO, training and testing image sets for each acquisition condition included 600 ROIs with simulated lesions and 600 lesion-free ROIs.

The ROIs were separated into two main sets, with and without calcifications. Each main set was then separated into independent training and testing ROI sets. Each reader was trained with an independent training set but tested with a common test set, simulating a scenario where independent readers read the same cases. The ROIs used for training and testing were selected randomly for each reader. The scores of each reader were pooled to calculate a cumulative ROC [24].

## Statistical analyses: ROC with AUCs and signal detection with $d'$ index

One-shot MRM ROC analysis was used to estimate the AUC and its variance for lesion detectability with multiple lesions per case [18]. We simulated different blurring conditions to evaluate the virtual readers' performance. Blurring conditions were analyzed as a function of calcification size: two tube motions (continuous and step-and-shoot), two detector element sizes (70  $\mu\text{m}$  and 140  $\mu\text{m}$ ), and various reconstructed voxel sizes ([35, 100]  $\mu\text{m}$ ).

The AUC and  $d'$  were calculated based on the pooled ROC curves of the 14 virtual readers. AUC differences between two paired ROCs were reported using 2000 stratified bootstrap replicates computed using the Delong method [28]. The 95% CI AUC was also reported using 2000 stratified bootstrap replicates computed based on the Delong method. R packages “pROC” (version 1.18.0) and “psycho” (version 0.6.1) were used to collect the ROC and  $d'$  (sensitivity index) statistics, respectively [29]. Differences ( $\Delta$ ) in AUC and ratios ( $\lambda$ ) of  $d'$  were calculated to evaluate changes in performance between acquisition geometries. AUC and

$d'$  were categorized by tube motion, detector element size, reconstructed voxel size, and calcification size.

## Experimental validation

To validate our VCTs, we acquired DBT images of a BR3D breast phantom (Mod. 020, CIRS Inc.). The BR3D phantom consists of six individual semi-elliptical slabs (semi-major axis = 180 mm, semi-minor axis = 100 mm). Each slab is 10 mm thick and composed of a heterogeneous 50–50% adipose-glandular mixture of tissue-equivalent material. The slab that contains an assortment of target lesions was positioned at the central slice of the BR3D phantom (slab #3).

Images of the BR3D phantom were acquired on two different Hologic DBT systems: Selenia Dimensions and 3Dimensions (Hologic Inc.). The Hologic DBT systems acquire 15 equal dose projections of the breast over a 15° angle using continuous X-ray tube motion. The Selenia Dimensions and 3Dimensions systems have detector element sizes of 140  $\mu\text{m}$  (binned) and 70  $\mu\text{m}$ , respectively. Images of the BR3D phantom were first acquired on the Selenia Dimensions system using the AEC mode. The AEC mode corresponded to an acquisition technique of 33 kV, 75 mAs, and W/AI target-filter combination for a thickness of 60 mm. These same technique settings were then used to image the BR3D phantom on the 3Dimensions system. The projections were reconstructed using voxel sizes of 35, 50, 70, 100, and 140  $\mu\text{m}$  (Piccolo, version 5.02, real-time tomography).

The signal–noise-ratio ( $SNR = \mu/\sigma$ , where  $\mu$  is the mean signal of calcifications and  $\sigma$  is the standard deviation of the surrounding background) was calculated for each calcification in the reconstructed slice where the calcifications are most in-focus. Single-calcifications were segmented using the Otsu method [30] within ROIs equal to the nominal size of each speck group [30]. The mean  $SNR$  of each calcification size was calculated by averaging the  $SNR$  of each individual calcification (95% CI is also shown). The  $SNR$  ratio between the two imaging systems was then calculated.

## Results

### VCT analysis

The pooled ROC results indicate that reconstruction with 70  $\mu\text{m}$  voxels is superior to 100  $\mu\text{m}$  voxels (Fig. 2A). Comparing 70 to 100  $\mu\text{m}$  using continuous motion, the difference in AUC  $\Delta = 0.0236$  (0.0651) and  $d'$  ratio  $\lambda = 1.12$  (1.33) show an improvement in calcification detection ( $p < 0.001$ ) for  $\text{del} = 140 \mu\text{m}$  (70  $\mu\text{m}$ ). For the similar comparison with step-and-shoot motion,  $\Delta = 0.0558$  (0.0647) and  $\lambda = 1.27$

(1.29) also show that an improvement in calcification detection ( $p < 0.001$ ) for  $\text{del} = 140 \mu\text{m}$  (70  $\mu\text{m}$ ).

The pooled ROC results indicate that step-and-shoot is superior to continuous tube motion (Fig. 2B). Comparing step-and-shoot with continuous using a 100  $\mu\text{m}$  reconstructed voxel,  $\Delta = 0.0541$  (0.0438) and  $\lambda = 1.24$  (1.23) show an improvement in calcification detection ( $p < 0.001$ ) for  $\text{del} = 140 \mu\text{m}$  (70  $\mu\text{m}$ ). For similar comparison using a 70  $\mu\text{m}$  reconstructed voxel size,  $\Delta = 0.0863$  (0.0434) and  $\lambda = 1.40$  (1.19) also show that there is an improvement in calcification detection ( $p < 0.001$ ) for  $\text{del} = 140 \mu\text{m}$  (70  $\mu\text{m}$ ).

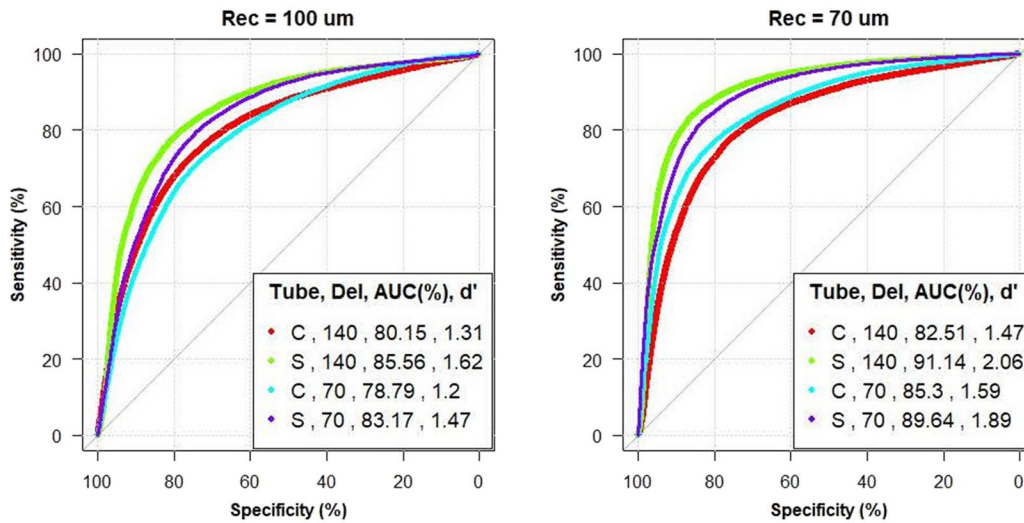
Finally, the pooled ROC results did *not* indicate that a detector element size of 70  $\mu\text{m}$  is superior to 140  $\mu\text{m}$  (Fig. 2C) for most test conditions. The only exception was for continuous tube motion in combination with a reconstructed voxel size of 70  $\mu\text{m}$  ( $\Delta = 0.0259$  and  $\lambda = 1.08$ ,  $p < 0.001$ ).

We observed that calcification size affected lesion detectability. The performance is summarized by calcification size and blurring source conditions in Table 3. Note that a reconstructed voxel size of 70  $\mu\text{m}$  and step-and-shoot motion show the greatest improvement in calcification detection for most acquisition conditions (Table 3A, B), especially for the smallest calcifications (8 and 27 voxels). Since we simulate small calcifications at the threshold of detectability, noise is dominant for a detector element size of 70  $\mu\text{m}$ , substantially affecting the performance in calcification detection (Table 3C).

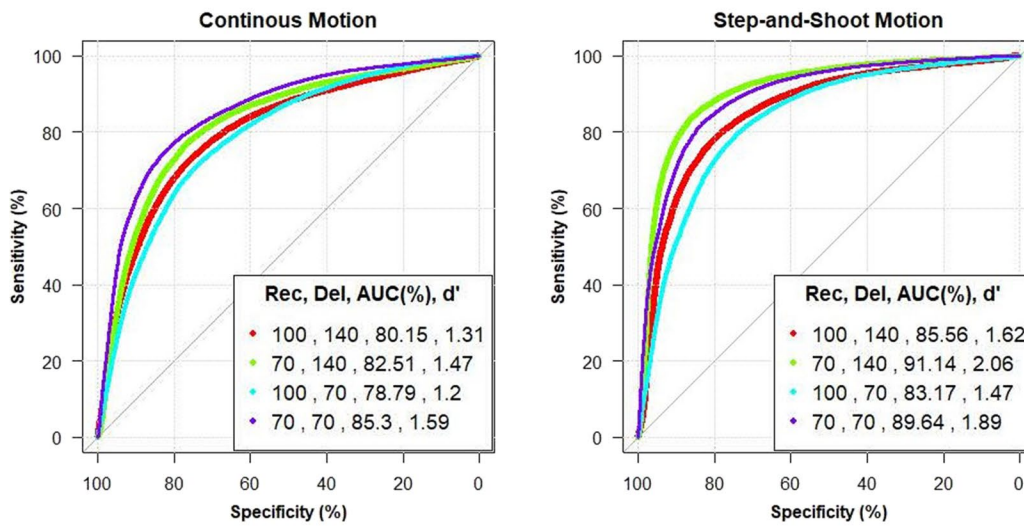
### Experimental analysis

The  $SNR$  calculated using DBT images of a BR3D phantom produced results that closely matched the computer simulations (Table 4). Images reconstructed with super-resolution can increase the  $SNR$  of calcifications. For 140  $\mu\text{m}$  detector element size, images reconstructed with 35  $\mu\text{m}$ –100  $\mu\text{m}$  voxel size (super-resolution) resulted in higher  $SNR$  for small specks. Large calcifications benefitted from larger reconstructed voxel size; this is consistent with signal detection theory in which objects larger than the blurring kernel do not lose contrast and benefit from noise blurring.

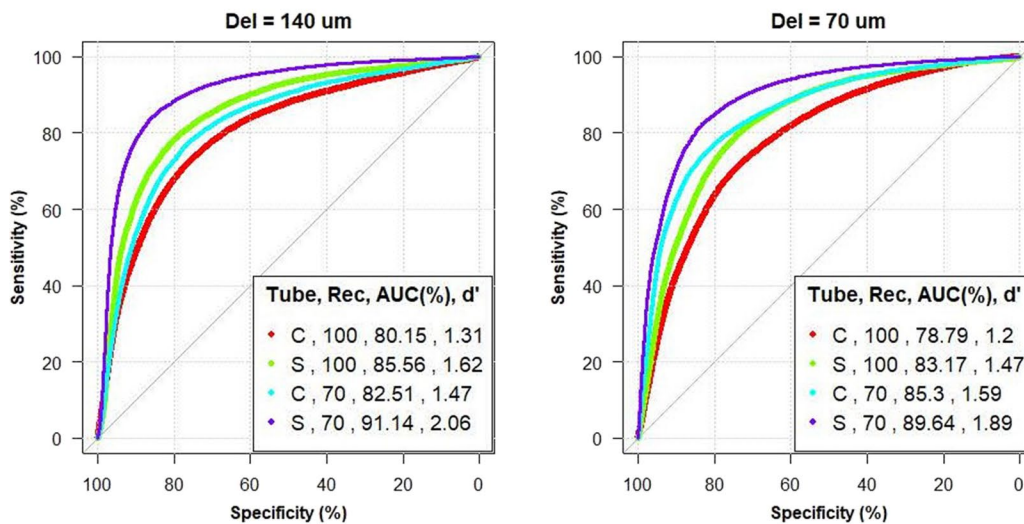
Smaller detector element sizes did not improve calcification detection. In fact, smaller detector elements resulted in noisier images with high  $\sigma$  values (Fig. 3B), reducing the  $SNR$  (Table 4B). For 70  $\mu\text{m}$  detector element size, images reconstructed at 100  $\mu\text{m}$  resulted in the highest  $SNR$ . The combination of the 70  $\mu\text{m}$  detector element with the smallest reconstructed sizes does not improve the  $SNR$  of the calcifications, likely because the detection of small calcifications is dose limited. Computer simulations also demonstrated that a smaller detector element size does not improve calcification detection (Table 4D, E). The benefit of larger detector



(A)



(B)



(C)

◀**Fig. 2** Pooled ROC curves from MRMC study. Results are categorized by (A) reconstructed voxel size (Rec), (B) tube motion (Tube), and (C) detector element size (Del). Abbreviations ‘C’ is used in plot legends for continuous and ‘S’ for step-and-shoot tube motion

elements is consistent between experiment and simulation (Table 4C, E;  $1.34 \pm 0.13$  vs.  $1.16 \pm 0.11$ ).

## Discussion and conclusions

This study shows that tomosynthesis acquisition geometry plays an important role in calcification detection, affecting the performance of virtual observers. Both source motion and reconstructed voxel size show significant changes in the performance of imaging systems. The overall impact of small detector element size does not affect substantially the detection of calcifications, except when tomosynthesis is used in combination with an additional blurring source (e.g., continuous tube motion). The greatest impact of these parameters is observed in the detection of small calcifications at the threshold of detectability. The best geometries made use of step-and-shoot motion and small reconstructed voxel sizes, taking advantage of the super-resolution abilities of tomosynthesis. Experimental results showed similar effects of the acquisition geometry.

The limited angle and the limited number of projections in DBT produce an under-sampled dataset that may compromise the detection of clinically significant calcifications. Early DBT studies showed that some calcium-only lesions were less conspicuous when compared with conventional DM [31]. In a survey, 61% of radiologists responded that DBT did not show an improvement in calcification characterization due to the poor image quality of synthetic 2D mammograms in comparison to DM [32]. Calcifications without an associated mass or area of distortion may be the earliest sign of breast cancer, specifically ductal carcinoma in situ (DCIS). In addition, in denser breasts, invasive cancers may often be detected by the identification of suspicious calcifications and these cancers may often have a worse prognosis than some non-calcified cancers [33]. A study conducted with over a million women shows that DBT should be considered the new standard of care for breast cancer screening [34]; therefore, improving calcification detection with the optimization of imaging systems is necessary since DBT has already been shown to improve substantially the detection and characterization of lesions, especially for small lesions composed of soft tissue.

VCTs and other task-based methods have become increasingly complex over the last few years. In this study, we have shown that VCTs can predict with high confidence the physical properties of imaging systems to provide guidance on how to improve calcification detection. Validations of

imaging systems currently rely on experimental evaluation of physical phantoms and small clinical datasets [24, 35]. Physical phantoms provide preliminary data to indicate changes in performance between systems but may under-sample the full range of patients and disease complexity. Our VCT results can predict clinical trials that are similar to the performance of clinical systems [11]. Importantly, the US Food and Drug Administration (FDA) supports the use of VCTs to decrease human trial size and duration [15].

Reconstructed DBT images of a BR3D phantom were used to compare and validate our VCT results. The DBT projections were acquired using two commercially available systems and reconstructed using third-party software at 35, 50, 70, 100, and 140  $\mu\text{m}$  voxel sizes. We observed that images reconstructed with super-resolution [36] can improve the *SNR* of calcification specks for projections acquired using 140  $\mu\text{m}$  detector elements. We observed that a smaller detector element size (70  $\mu\text{m}$ ) did not substantially improve calcification detection; small detector elements will increase the influence of quantum noise at the detector, resulting in sharper but noisier images. It is important to mention that the DBT projections of the BR3D phantom were acquired using two clinical systems designed with continuous tube motion. The benefits of step-and-shoot should also be investigated experimentally.

Experimental validation was performed with technique factors that matched using the automatic exposure settings of the Selenia Dimensions system. Radiation dose can affect the threshold of detectability of calcifications; Hadjipanteli et al have demonstrated with a human reader experiment that the limit in calcification detection varied for DBT images acquired at different doses, namely  $0.243 \pm 0.014$ ,  $0.211 \pm 0.011$ , and  $0.204 \pm 0.011$  mm for 1.25, 2.5, and 5.0 mGy, respectively [5]. Because of this, in our study, the kV and exposure (mAs) were kept the same for both tomosynthesis systems to eliminate the effect of dose from our comparison, and to analyze only the effect of detector element size in calcification detection. We acknowledge that the systems have small differences and that detection may be different if both systems were operated by a phototimer.

The effect of angular range in DBT was not investigated in this work. Hadjipanteli et al have shown that narrow-angle DBT improves the limit in calcification detection ( $0.211 \pm 0.011$  mm) as compared to wide-angle DBT ( $0.257 \pm 0.015$  mm) [5]. In this study, we wanted to evaluate the impact of super-resolution and the effect of detector element size, tube motion, and reconstructed voxel size in calcification detection; the effect of angular range in DBT could be explored in a separate work. Importantly, the single-calcification models—defined at the threshold of detectability in our study—are in the same range of size reported by Hadjipanteli et al ( $\approx 0.2$  mm for narrow-angle DBT) [5]. However, some differences in  $d'$  are observed;

**Table 3** Summary of MRMC analyses (%AUC with 14 virtual readers) and  $d'$  calculations of each imaging design. Results are categorized by calcification size, (A) reconstructed voxel size, (B) tube motion, and (C) detector element size.  $p$  values ( $p$ ) calculated using pair-wise comparison between pooled ROCs (DeLong method with bootstrap  $n=2000$ )

...			(B) Tube Motion							
Calcification size, # voxels (Volume, mm) <sup>3</sup>	(A) reconstructed voxel size, $\mu\text{m}$	Metric	Continuous			Step-and-Shoot			$p$	
			(C) Detector element size ( $\mu\text{m}$ )			(C) Detector element size ( $\mu\text{m}$ )			Detector element size ( $\mu\text{m}$ )	
			70	140	$P$	70	140	$P$	70	140
<b>8</b> (0.2×0.2×0.2)	<b>70</b>	<b>AUC</b>	81.13	79.87	<b>&lt;.001</b>	86.26	87.52	<b>0.06</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[80, 82]	[79, 80]		[86, 87]	[87, 88]			
		<b><math>d'</math></b>	1.40	1.37		1.65	1.76			
	<b>100</b>	<b>AUC</b>	75.55	75.03	<b>0.83</b>	80.09	82.30	<b>&lt;.001</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[75, 76]	[74, 76]		[79, 81]	[82, 83]			
		<b><math>d'</math></b>	1.03	1.08		1.30	1.48			
<b><math>P</math></b>	...	<b>&lt;.001</b>	<b>&lt;.001</b>	...	<b>&lt;.001</b>	<b>&lt;.001</b>	...	...		
<b>27</b> (0.3×0.3×0.3)	<b>70</b>	<b>AUC</b>	83.89	79.34	<b>&lt;.005</b>	88.73	91.80	<b>&lt;.001</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[83, 84]	[79, 80]		[88, 89]	[91, 92]			
		<b><math>d'</math></b>	1.50	1.31		1.85	2.28			
	<b>100</b>	<b>AUC</b>	74.52	77.32	<b>&lt;.001</b>	80.66	83.45	<b>&lt;.001</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[74, 75]	[77, 78]		[80, 81]	[83, 84]			
		<b><math>d'</math></b>	1.09	1.20		1.35	1.53			
<b><math>P</math></b>	...	<b>&lt;.001</b>	<b>&lt;.001</b>	...	<b>&lt;.001</b>	<b>&lt;.001</b>	...	...		
<b>64</b> (0.4×0.4×0.4)	<b>70</b>	<b>AUC</b>	90.28	86.17	<b>&lt;.001</b>	93.53	91.94	<b>&lt;.001</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[90, 91]	[86, 87]		[93, 94]	[91, 92]			
		<b><math>d'</math></b>	1.91	1.76		2.25	2.21			
	<b>100</b>	<b>AUC</b>	85.79	85.42	<b>0.78</b>	88.26	88.89	<b>0.06</b>	<b>&lt;.001</b>	<b>&lt;.001</b>
		<b>[95% CI]</b>	[85, 86]	[85, 86]		[88, 89]	[88, 89]			
		<b><math>d'</math></b>	1.60	1.68		1.85	1.91			
<b><math>P</math></b>	...	<b>&lt;.001</b>	<b>0.35</b>	...	<b>&lt;.001</b>	<b>&lt;.001</b>	...	...		

we attributed these differences to the composition of the simulated calcifications.

While there was general agreement, the  $d'$  the ratio in the computer simulations did not exactly match with the change in SNR observed in the physical experiments as a function of reconstructed voxel size. There are several possible reasons for the observed differences, including differences in the shape and size of the calcifications, limits to the x-ray physics and detector models, and constraints in the computational resources available. Importantly, SNR differs from  $d' - d'$  includes the effects of the observer's eye filter and internal noise. That said, the metrics are intrinsically correlated, and the average results comparing the two detector element sizes are remarkably similar.

The breast phantom population in the VCT was constrained by the simulation parameters. The breast phantoms and calcifications used a voxel size of 100  $\mu\text{m}$ . The computational time and GPU memory required to simulate breast phantoms, DBT projections, and reconstructions increase with smaller voxel sizes and larger breast volumes. Thus, we did not use smaller voxels based on our limits on simulation

time and computer memory. In addition, the calcifications were only inserted at the central plane (CC view) of phantoms to improve accuracy in estimating the precise location of the calcifications in the reconstructed images, avoiding rounding errors.

Although patient motion is also considered a blurring source, it was not simulated in this study. Patient motion affects lesion detectability and is more apparent in images acquired using step-and-shoot motion due to the longer scan time. Future work should investigate whether the benefits of step-and-shoot motion in calcification detection could potentially be outweighed by the blurring effects of patient motion; preliminary data has shown that patient motion affects image quality more than tube motion [37]. To conduct VCTs would require an accurate model of patient motion.

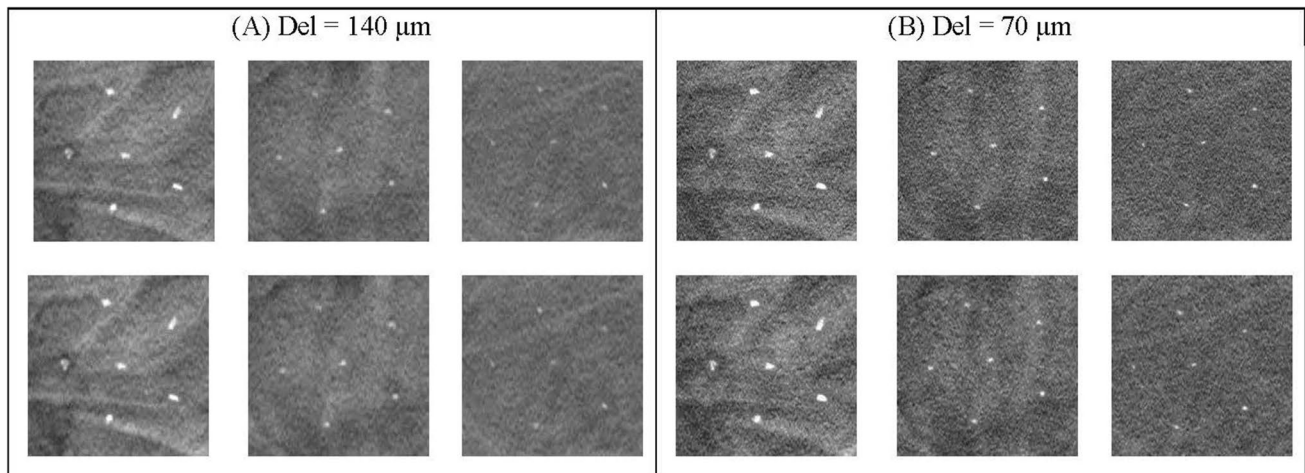
The field will benefit from the adoption of standards in VCTs, since clinical trials are limited by cost and duration. Clinical trials of DBT involve ionizing radiation that requires repeated irradiation of volunteers, which may be practically unfeasible today. In particular, screening



**Table 4** Summary SNR of each calcification speck and ratio of SNR between (A) Selenia Dimensions and (B) 3Dimensions systems. Calcification size of 0.136 mm (signal) was not detectable by human observers for signal-to-noise ratio (SNR) calculations

Speck, size (mm)	35 (μm)	50 (μm)	70 (μm)	100 (μm)	140 (μm)	Average
<b>(A) Experimental—Del= 140 μm</b>						
Reconstructed Voxel Size (SNR, 95% CI)						
0.400	18.2 [15.8, 20.6]	19.0 [16.8, 21.3]	19.2 [16.8, 21.7]	20.4 [17.6, 23.1]	20.6 [17.7, 23.5]	<b>19.4 [18.6, 20.4]</b>
0.290	16.1 [14.8, 17.4]	17.0 [15.4, 18.5]	17.3 [15.7, 18.9]	17.3 [15.2, 19.5]	17.6 [15.2, 19.9]	<b>17.1 [16.4, 17.7]</b>
0.230	14.5 [13.0, 16.1]	15.6 [14.4, 16.8]	15.5 [14.3, 16.7]	16.4 [14.7, 18.0]	15.4 [14.5, 16.3]	<b>15.5 [15.0, 16.0]</b>
0.196	13.7 [12.6, 14.8]	14.0 [12.5, 15.4]	14.2 [13.2, 15.2]	13.6 [11.6, 15.6]	12.3 [9.9, 14.8]	<b>13.6 [12.9, 14.2]</b>
0.165	13.1 [11.7, 14.5]	13.5 [11.9, 15.1]	13.6 [12.2, 15.0]	13.7 [12.4, 15.0]	12.9 [11.5, 14.4]	<b>13.4 [12.9, 13.9]</b>
<b>Mean, 95% CI</b>	<b>15.1 [14.3, 16.0]</b>	<b>15.8 [14.9, 16.7]</b>	<b>16.0 [15.0, 16.9]</b>	<b>16.3 [15.1, 17.4]</b>	<b>15.8 [14.4, 17.1]</b>	<b>15.8 [15.3, 16.2]</b>
<b>(B) Experimental—Del= 70 μm</b>						
Reconstructed Voxel Size (SNR, 95% CI)						
0.400	11.3 [10.0, 12.6]	12.7 [11.3, 14.1]	12.9 [11.5, 14.2]	15.1 [13.4, 16.8]	15.9 [14.2, 17.6]	<b>13.6 [12.7, 14.4]</b>
0.290	10.7 [10.2, 11.2]	11.9 [11.2, 12.6]	12.0 [11.6, 12.5]	14.2 [13.2, 15.2]	14.7 [13.1, 16.3]	<b>12.7 [12.0, 13.3]</b>
0.230	9.7 [8.8, 10.6]	11.1 [10.1, 12.1]	11.5 [10.4, 12.5]	13.3 [12.2, 14.4]	13.3 [11.6, 15.0]	<b>11.8 [11.1, 12.4]</b>
0.196	9.3 [8.4, 10.1]	10.2 [9.4, 11.1]	10.7 [10.0, 11.5]	12.1 [10.5, 13.8]	10.6 [8.5, 12.7]	<b>10.6 [10.0, 11.2]</b>
0.165	9.4 [8.7, 10.2]	9.7 [8.5, 10.9]	10.2 [8.4, 11.9]	11.4 [10.0, 12.8]	10.5 [9.7, 11.2]	<b>10.2 [9.8, 11.2]</b>
<b>Mean, 95% CI</b>	<b>10.1 [9.7, 10.5]</b>	<b>11.1 [10.6, 11.7]</b>	<b>11.5 [10.9, 12.0]</b>	<b>13.2 [12.5, 13.9]</b>	<b>13.0 [12.0, 14.0]</b>	<b>11.8 [11.4, 12.1]</b>
<b>(C) Ratio SNR, reconstructed voxel size (A/B)</b>						
0.400	1.61	1.50	1.49	1.35	1.29	<b>1.45 ± 0.13</b>
0.290	1.51	1.43	1.44	1.22	1.20	<b>1.36 ± 0.14</b>
0.230	1.50	1.40	1.35	1.23	1.16	<b>1.33 ± 0.13</b>
0.196	1.48	1.37	1.32	1.12	1.16	<b>1.29 ± 0.15</b>
0.165	1.39	1.39	1.34	1.20	1.24	<b>1.31 ± 0.09</b>
<b>Mean, SD</b>	<b>1.50 ± 0.08</b>	<b>1.42 ± 0.05</b>	<b>1.39 ± 0.07</b>	<b>1.23 ± 0.08</b>	<b>1.21 ± 0.05</b>	<b>1.34 ± 0.13</b>
<b>(D) Simulated—Del= 140 μm</b>						
AUC [95% CI]	92.30 [92, 93]	92.04 [92, 93]	92.26 [91, 93]	84.09 [83, 85]	91.41 [91, 92]	<b>90.42 [90, 91]</b>
<i>d'</i>	2.26	2.24	2.28	1.53	2.12	<b>2.09 ± 0.31</b>
<b>(E) Simulated—Del= 70 μm</b>						
AUC [95% CI]	93.25 [93, 94]	91.10 [91, 92]	88.80 [88, 89]	80.94 [80, 82]	85.51 [85, 86]	<b>87.92 [87, 88]</b>
<i>d'</i>	2.20	2.05	1.85	1.35	1.62	<b>1.81 ± 0.34</b>
$\lambda$ (D/E)	<b>1.03</b>	<b>1.09</b>	<b>1.23</b>	<b>1.13</b>	<b>1.31</b>	<b>1.16 ± 0.11</b>

Values in bold represent mean [95% CI], mean ±SD, or ratios of data reconstructed with voxel size 35–140 μm



**Fig. 3** Images of speck of calcifications acquired using a (A) Selenia Dimensions (140 μm detector size, binned) and (B) 3Dimension system (70 μm detector size) and reconstructed with (top) 35 μm and

(bottom) 70 μm voxel size. Examples of speck sizes of 0.400, 0.230, and 0.165 mm are shown, respectively

populations do not have a high incidence of cancer; therefore, many cancer-free women are irradiated. In DBT, the incorporation of personalized acquisition geometries and reconstruction methods will allow better detection and characterization of breast findings.

In summary, calcification detection in DBT is degraded due to blurring from reconstructed voxel size, detector element size, and tube motion. The blurring is more evident in small calcifications. The impact on calcification detection ranked from most to least significant is (i) reconstruction voxel size, (ii) tube motion during acquisition, and (iii) detector element size. Acquisition geometries that use smaller reconstructed voxels and step-and-shoot tube motion have the potential to further improve the detection of calcifications.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1007/s00330-023-10103-6>.

**Funding** Funding was supported by the Department of Defense Breast Cancer Research Program (W81XWH-18-1-0082), Burroughs Wellcome Fund (IRSA 1016451), Susan G. Komen® (IRR13264610 and PDF14302589), National Institute of Health (R37CA273959, NCI P30 CA016520, and R01CA154444), American Association of Physicists in Medicine (2020 AAPM Research Seed Funding Grant), and Breast Cancer Alliance, Inc. (2022 Young Investigator Grant). Equipment support was provided by Analogic Inc., Barco NV, and Real Time Tomography LLC. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

## Declarations

**Guarantor** The scientific guarantor of this publication is the senior author Dr. Andrew D.A. Maidment.

**Conflict of interest** Andrew D. A. Maidment and Raymond J. Acciavatti are inventors of patents and patent applications related to next-generation tomosynthesis technology. Andrew D. A. Maidment has received research support from Hologic Inc., Barco NV, and Analogic Corporation; is the spouse to an employee and stockholder of Real Time Tomography (RTT), LLC; is a member of the scientific advisory board of RTT; and is an owner of Daimroc Imaging, LLC.

Emily F. Conant: Grants or contracts from iCAD, Hologic, and OM1; payment or honoraria from MedScape and Aunt Minnie; lecture travel support from the RSNA for European Congress of Radiology 2022; advisory board, iCAD and Hologic; current board member, Society of Breast Imaging.

**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.

**Study subjects or cohorts overlap** Study subjects or cohorts have been modelled based on clinical data previously reported in Radiology in 2021 ( $n=660$ ) [24].

## Methodology

- Retrospective
- observational
- performed at one institution

## References

1. Gur D, Abrams GS, Chough DM et al (2009) Digital breast tomosynthesis: observer performance study. *AJR Am J Roentgenol* 193:586–591. <https://doi.org/10.2214/AJR.08.2031>
2. Mackenzie A, Thomson EL, Mitchell M et al (2021) Virtual clinical trial to compare cancer detection using combinations of 2D mammography, digital breast tomosynthesis and synthetic 2D imaging. *Eur Radiol*. <https://doi.org/10.1007/s00330-021-08197-x>
3. Renaldo A, Miller M, Caley M et al (2022) 2D or synthetic 2D? A reader study of visualization of amorphous calcifications. *J Breast Imaging* 4(1):9–24. <https://doi.org/10.1093/jbi/wbab094>
4. Mordang JJ, Gubern-Mérida A, Bria A et al (2018) The importance of early detection of calcifications associated with breast cancer in screening. *Breast Cancer Res Treat*. <https://doi.org/10.1007/s10549-017-4527-7>
5. Hadjipanteli A, Elangovan P, Mackenzie A et al (2017) The effect of system geometry and dose on the threshold detectable calcification diameter in 2D-mammography and digital breast tomosynthesis. *Phys Med Biol* 62(3):858–877. <https://doi.org/10.1088/1361-6560/aa4f6e>
6. Shaheen E, Marshall N, Bosmans H (2011) Investigation of the effect of tube motion in breast tomosynthesis: continuous or step and shoot? In: *Proceedings Volume 7961, Medical Imaging 2011: Physics of Medical Imaging* 79611E
7. Yaffe MJ, Mainprize JG (2004) Detectors for digital mammography. *Technol Cancer Res Treat* 3:309–324
8. Hardie RC, Droege DR, Dapore AJ, Greiner ME (2015) Impact of detector-element active-area shape and fill factor on super-resolution. *Front Phys* 3:1–17. <https://doi.org/10.3389/fphy.2015.00031>
9. Barufaldi B, Bakic PR, Maidment ADA (2019) Multiple-reader, multiple-case ROC analysis for determining the limit of calcification detection in tomosynthesis. In: *Proc. SPIE 10948, Medical Imaging 2019: Physics of Medical Imaging*, 109480N (1 March 2019). <https://doi.org/10.1117/12.2512884>
10. Neves FS, Vasconcelos TV, Vaz SLA et al (2012) Evaluation of reconstructed images with different voxel sizes of acquisition in the diagnosis of simulated external root resorption using cone beam computed tomography. *Int Endod J* 45:234–239. <https://doi.org/10.1111/j.1365-2591.2011.01966.x>
11. Barufaldi B, Maidment ADA, Dustler M et al (2021) Virtual clinical trials in medical imaging system evaluation and optimisation. *Radiat Prot Dosimetry* 195:363–371. <https://doi.org/10.1093/rpd/ncab080>
12. Borges LR, Barufaldi B, Caron RF et al (2019) Technical note: noise models for virtual clinical trials of digital breast tomosynthesis. *Med Phys* 46:2683–2689. <https://doi.org/10.1002/mp.13534>
13. Zeng R, Samuelson FW, Sharma D et al (2020) Computational reader design and statistical performance evaluation of an in-silico imaging clinical trial comparing digital breast tomosynthesis with full-field digital mammography. *J Med Imaging (Bellingham)* (4):042802. Epub 2020 Feb 26. <https://doi.org/10.1117/1.jmi.7.4.042802>
14. Chase JG, Preiser JC, Dickson JL et al (2018) Next-generation, personalised, model-based critical care medicine: a state-of-the-art review of in silico virtual patient models, methods, and cohorts,

- and how to validation them. *Biomed Eng Online* 17(1):24. <https://doi.org/10.1186/s12938-018-0455-y>
15. Badano A, Graff CG, Badal A et al (2018) Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. *JAMA Netw Open* 1:1–12. <https://doi.org/10.1001/jamanetworkopen.2018.5474>
  16. Abadi E, Segars WP, Tsui BMW et al (2020) Virtual clinical trials in medical imaging: a review. *J Med Imaging* 7:042805-1–42840. <https://doi.org/10.1117/1.JMI.7.4.042805>
  17. Platiša L, Goossens B, Vansteenkiste E et al (2011) Channelized hotelling observers for the assessment of volumetric imaging data sets. *J Opt Soc Am A Opt Image Sci Vis* 28(6):1145–63. <https://doi.org/10.1364/josaa.28.001145>
  18. Gallas BD (2006) One-shot estimate of MRMC variance: AUC. *Acad Radiol* 13:353–362. <https://doi.org/10.1016/j.acra.2005.11.030>
  19. Gastounioti A, Pantalone L, Scott CG et al (2021) Fully automated volumetric breast density estimation from digital breast tomosynthesis. *Radiology* (3):561–568. Epub 2021 Sep 14. <https://doi.org/10.1148/radiol.2021210190>
  20. Zuckerman SP, Conant EF, Keller BM et al (2016) Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening. *Radiol J* 000:1–7. <https://doi.org/10.1148/radiol.2016160366>
  21. Barufaldi B, Vent TL, Bakic PR, Maidment ADA (2022) Computer simulations of case difficulty in digital breast tomosynthesis using virtual clinical trials. *Med Phys* 49:2220–2232. <https://doi.org/10.1002/mp.15553>
  22. Rodriguez-Ruiz A, Feng SSJ, Van Zelst J et al (2017) Improvements of an objective model of compressed breasts undergoing mammography: Generation and characterization of breast shapes. *Med Phys* 44:2161–2172. <https://doi.org/10.1002/mp.12186>
  23. Petkovic MD, Bakic PR, Maidment ADA, Pokrajac D (2015) Asymptotic number of  $Z^3\Delta$  cells covering  $C(1)$  surface on uniform grid and complexity of recursive-partitioning simulation of septal tissue regions. *Appl Math Comput* 252:263–272. <https://doi.org/10.1016/j.amc.2014.11.111>
  24. Rafferty EA, Park JM, Philpotts LE et al (2013) Digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology* 266:104–113. <https://doi.org/10.1148/radiol.12120674/-DC1>
  25. Maidment ADA, Albert M, Conant EF, Feig SA (1996) Three-dimensional calcification reconstruction from a limited number of views. *Proc. SPIE 2708, Medical Imaging 1996: Physics of Medical Imaging*, (11 April 1996); 378–389. <https://doi.org/10.1117/12.237800>
  26. Feng SSJ, Sechopoulos I (2012) Clinical digital breast tomosynthesis system: dosimetric characterization. *Radiology* 263:35–42
  27. International Commission on Radiation Units and Measurements (ICRU) Report 44. Tissue substitutes in radiation dosimetry. Washington: International Commission on Radiation Units and Measurements; 1989
  28. Robin X, Turck N, Hainard A et al (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:1–8. <https://doi.org/10.1186/1471-2105-12-77>
  29. Makowski D (2018) The psycho package: an efficient and publishing-oriented workflow for psychological science. *J Open Source Softw* 3(22):470. <https://doi.org/10.21105/joss.00470>
  30. Otsu N (1979) Threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybernet* 9(1):62–66. <https://doi.org/10.1109/tsmc.1979.4310076>
  31. Gilbert FJ, Tucker L, Gillan MGC et al (2015) Accuracy of digital breast tomosynthesis for depicting breast cancer subgroups in a UK retrospective reading study (TOMMY Trial). *Radiology*. <https://doi.org/10.1148/radiol.2015142566>
  32. Zuckerman S, Sprague BL, Weaver DL et al (2019) Survey results regarding uptake and impact of synthetic digital mammography with tomosynthesis in the screening setting. *J Am Coll Radiol*. <https://doi.org/10.1016/j.jacr.2019.07.020>
  33. Nyante SJ, Lee SS, Benefield TS et al (2017) The association between mammographic calcifications and breast cancer prognostic factors in a population-based registry cohort. *Cancer*. <https://doi.org/10.1002/cncr.30281>
  34. Conant EF, Talley MM, Parghi CR et al (2023) Mammographic screening in routine practice: multisite study of digital breast tomosynthesis and digital mammography screenings. *Radiology*. <https://doi.org/10.1148/radiol.221571>
  35. U.S. Food and Drug Administration (FDA) (2015) Mammomat inspiration with tomosynthesis option premarket approval: P140011/S003. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=P140011>
  36. Acciavatti RJ, Maidment ADA (2012) Observation of super-resolution in digital breast tomosynthesis. *Med Phys* 39:7518–7539. <https://doi.org/10.1118/1.4757583>
  37. Acciavatti RJ, Maidment ADA (2012) Optimization of continuous tube motion and step-and-shoot motion in digital breast tomosynthesis systems with patient motion. *Proc. SPIE 8313, Medical Imaging 2012. Phys Med Imag* 831306 (2 March2012). <https://doi.org/10.1117/12.911016>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.