# High-resolution $\mu$CT imaging for characterizing microcalcification detection performance in breast CT

**Andrew M. Hernandez,[a,*] Amy E. Becker,[b] Su Hyun Lyu,[b] Craig K. Abbey,[c] and John M. Boone[a,d]**

[a]University of California Davis, Department of Radiology, Sacramento, California, United States

[b]University of California Davis, Biomedical Engineering Graduate Group, Davis, California, United States

[c]University of California Santa Barbara, Psychological and Brain Sciences, Santa Barbara, California, United States

[d]University of California Davis, Biomedical Engineering, Davis, California, United States

## Abstract

**Purpose:** To demonstrate the utility of high-resolution micro-computed tomography ($\mu$CT) for determining ground-truth size and shape properties of calcium grains for evaluation of detection performance in breast CT (bCT).

**Approach:** Calcium carbonate grains (~200 $\mu$m) were suspended in 1% agar solution to emulate microcalcifications ($\mu$Calcs) within a fibroglandular tissue background. Ground-truth imaging was performed on a commercial $\mu$CT scanner and was used for assessing calcium-grain size and shape, and for generating $\mu$Calc signal profiles. Calcium grains were placed within a realistic breast-shaped phantom and imaged on a prototype bCT system at 3- and 6-mGy mean glandular dose (MGD) levels, and the non-prewhitening detectability was assessed. Additionally, the $\mu$CT-derived signal profiles were used in conjunction with the bCT system characterization (MTF and NPS) to obtain predictions of bCT detectability.

**Results:** Estimated detectability of the calcium grains on the bCT system ranged from 2.5 to 10.6 for 3 mGy and from 3.8 to 15.3 for 6 mGy with large fractions of the grains meeting the Rose criterion for visibility. Segmentation of $\mu$CT images based on morphological operations produced accurate results in terms of segmentation boundaries and segmented region size. A regression model linking bCT detectability to $\mu$Calc parameters indicated significant effects of $\mu$Calc size and vertical position within the breast phantom. Detectability using $\mu$CT-derived detection templates and bCT statistical properties (MTF and NPS) were in good correspondence with those measured directly from bCT ($R^2 > 0.88$).

**Conclusions:** Parameters derived from $\mu$CT ground-truth data were shown to produce useful characterizations of detectability when compared to estimates derived directly from bCT. Signal profiles derived from $\mu$CT imaging can be used in conjunction with measured or hypothesized statistical properties to evaluate the performance of a system, or system component, that may not currently be available.

 [DOI: 10.1117/1.JMI.8.5.052107]

**Keywords:** microcalcification; breast computed tomography; breast cancer; detectability; micro-computed tomography.

Paper 21022SSRR received Jan. 26, 2021; accepted for publication Jun. 28, 2021; published online Jul. 20, 2021.

*Address all correspondence to Andrew M. Hernandez, amhern@ucdavis.edu

2329-4302/2021/$28.00 

## 1 Introduction

Breast cancer lesions are generally categorized as either masses or microcalcifications ($\mu$Calcs), and both are important for breast cancer screening. Masses are low-contrast larger soft tissue lesions and $\mu$Calcs are high-contrast small calcium deposits. Digital mammography (DM) is the standard of care for breast cancer screening. However, there has been a sustained effort to develop alternative imaging modalities to overcome some of the limitations of DM. Any new modality will need to perform well at detecting both masses and $\mu$Calcs. Imaging protocols are often developed and optimized using physical phantoms with features that represent disease processes.[1–4] In the case of $\mu$Calcs, this commonly means the use of small calcium grains embedded within some background material.[1,3,4]

One of the advantages of experimental work performed with physical phantom is that ground truth should in principle be available. However, ground truth can be difficult to obtain in the case of small, irregularly shaped grains typically used in $\mu$Calc phantoms. The small size makes them challenging to measure directly (e.g., calipers) and the irregular shapes make it unclear what measurements should even be taken. Another possible approach is the use of a low-power microscope, although we have not pursued this.

There is an additional logistical challenge in measuring individual grains and keeping track of where they are ultimately positioned within a physical $\mu$Calc phantom. In this work, we investigate an alternative strategy for obtaining ground-truth characterization of the calcium grains for use in these phantoms.

High-resolution micro-computed tomography ($\mu$CT) imaging can potentially provide information about calcium grain size and shape that is needed for obtaining ground truth in physical $\mu$Calc phantoms. Commercial $\mu$CT systems provide appreciably higher resolution than most clinical breast imaging modalities and the imaging protocols can be designed to deliver high mA and therefore very low noise levels. If morphological parameters of the $\mu$Calc phantoms can be obtained from these $\mu$CT images, then they can be used to provide ground truth for the $\mu$Calc phantoms.

We are investigating this approach in the context of breast CT (bCT), which is an emerging breast imaging modality that provides 3D reconstructed volume datasets of the breast. Laboratory evaluations[5–7] and preliminary clinical evaluations[8,9] suggest that bCT is better than DM at detecting masses, but that DM is better than an early generation (lower resolution) bCT system at detecting $\mu$Calcs.[8] This is due to several factors that affect the image quality in the reconstructed images. Previous simulation studies show the effect of detector element size,[10] breast diameter,[11] and image reconstruction[12,13] on $\mu$Calc detection in bCT. Previous phantom studies show the effects of tube potential[1] and radiation dose[1,3] on the visibility of $\mu$Calcs by human observers on benchtop systems. These parameters are the focus of ongoing optimization studies to evaluate bCT for clinical breast imaging.

For effective assessment of imaging on small irregularly shaped calcium grains, an accurate ground-truth description is required. For the present study, we developed a physical $\mu$Calc phantom, composed of calcium carbonate grains embedded in agar solution in a cylindrical insert, which is then placed within an adipose-equivalent (polyethylene) realistic breast-shaped background. Ground truth was obtained using a $\mu$CT scanner on the insert (in-air), and segmentation techniques were used to extract morphological parameters. The ultimate purpose in most phantom evaluations is to assess performance of the imaging system. Here, we evaluate detectability using $\mu$CT ground-truth data as a means to accurately estimate system performance and to assess the imaging components that impact performance. Various methods for computing detectability using the $\mu$CT data are described and compared with estimates derived directly from bCT data as a means of testing the accuracy of the results using $\mu$CT as the ground truth. These include (i) fitting a regression model to relate parameters of the calcium grains, derived from $\mu$CT, with bCT detectability, (ii) using $\mu$CT-derived signal profiles convolved with the MTF of the bCT system as a non-prewhitening (NPW) detection template, and (iii) using the $\mu$CT-derived signal profiles and bCT statistical properties (NPS and MTF) to predict NPW detectability. Detectability was evaluated using 23 replicated bCT scans of the $\mu$Calc phantom at 3- and 6-mGy MGD levels.

We show that morphological properties of the calcium grains impact detectability and this suggests that an accurate ground truth is needed for an accurate characterization of calcium-grain

detectability. Additionally, we show that the $\mu$CT data can be used to predict performance characteristics of a breast imaging system using a measured or hypothesized MTF and NPS. This can be of potential benefit for prospectively understanding the performance of a breast imaging system, or system component, in the design phase.

## 2 Methods

### 2.1 *Calcium Grains*

Calcium carbonate ($CaCO_3$) grains (CIRS Inc., Norfolk, Virginia) are used in this study as a surrogate for $\mu$Calcs.[3,14–16] Calcium carbonate is comparable to the chemical composition of breast $\mu$Calcs found clinically which include calcium oxalate in the form of weddellite,[15,17,18] and calcium and phosphorus, mostly in the form of hydroxyapatite[18,19] or tricalcium phosphate.[19] The reconstructed voxel values for the calcium carbonate grains were on average 45% lower than voxel values for clinical $\mu$Calcs as measured in seven example patient scans acquired on the current prototype bCT scanner in our laboratory at the x-ray technique (60 kV/0.2 mm Gd) of interest in this work. While the attenuation coefficient of calcium carbonate is significantly lower than these other forms of calcium, calcium carbonate is a popular choice for $\mu$Calcs because of its availability and its lower contrast makes it more challenging to detect.

Typically, the grains are selected by capturing them between sieves of different pores diameters. This process isolates grains roughly uniform in size, although some variability in the total volume of each calcium grain is to be expected because of oblate and prolate shapes, and other irregularities. The reported size of the grains in this study was 0.212 to 0.224 mm, which represents a size consistent with actionable $\mu$Calcs.[20]

### 2.2 *Microcalcification Phantom Design*

The calcium grains were placed between layers of 1% agar solution in cylindrical polystyrene test tubes (1.5 cm in diameter and 6 cm in length) which in turn could be inserted within a breast-shaped phantom. The reconstructed voxel values for 1% agar solution were 9.1% higher than the corresponding voxel values for fibroglandular tissue measured in seven example patient scans acquired on the current prototype bCT scanner in our laboratory at the x-ray technique (60 kV/0.2 mm Gd) of interest in this work, justifying its use as a reasonable surrogate for fibroglandular tissue in the breast. This phantom design approximates the clinical setting where $\mu$Calcs predominantly reside within areas of fibroglandular tissue or a mass lesion of similar density. A total of three test tubes were filled with the mixture of calcium carbonate grains and agar. Additionally, in each test tube, three Teflon spheres of 2.4 mm in diameter were also embedded (one at the bottom and two at the top) for the purpose of fiducial alignment and estimation of segmentation accuracy as explained in Sec. 2.4. The $\mu$Calc phantom inserts were imaged both individually (in-air), as shown in Figs. 1(a) and 1(b), as well as inserted in a well within a breast-shaped phantom background (in-phantom) as shown in Fig. 1(c).

The breast phantom was fabricated from ultra-high molecular weight polyethylene, which has x-ray attenuation properties similar to adipose tissue. The phantom size and shape parameters were determined from a cohort of 215 patient bCT datasets,[21] and the median (by volume) among six phantoms was used here. As detailed in Table 1 of Hernandez and Boone,[21] the median-sized phantom (V3) has a total volume of 614 $cm^3$, a vertical length ($z$) of 9.8 cm, a chest wall diameter of 12.5 cm, and a diameter of 10.7 cm at the vertical center-of-mass. The arrow in Fig. 1(c) is pointing to a $\mu$Calc phantom insert that is protruding from the top of the V3 phantom background. A pipette was used to displace the small air gap surrounding the test tube with water with care taken to avoid air bubbles.

### 2.3 *Phantom Imaging*

All three $\mu$Calc phantom inserts were imaged separately in-air on both a commercial $\mu$CT system and a prototype bCT system for generation of calcium-grain signal profiles (detailed in Sec. 2.5).

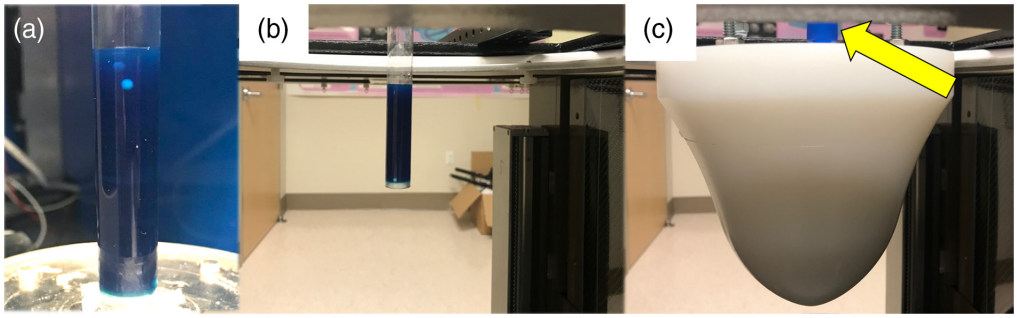


**Fig. 1** Photographs of the microcalcification ($\mu$Calc) phantom insert (a) fixtured on the rotating stage in the $\mu$CT scanner, (b) suspended in-air in the bCT scanner, and (c) placed within a realistic shaped, median-sized breast phantom and suspended in the bCT scanner. The top of the $\mu$Calc phantom insert is indicated by the yellow arrow in (c).

The three $\mu$Calc phantom inserts were also placed within the V3 phantom background (in-phantom) and imaged on the bCT system to demonstrate the utility of these calcium-grain profiles for $\mu$Calc detectability under realistic clinical imaging geometry and dose levels.

### 2.3.1 *In-air microcalcification phantom scans using breast CT and $\mu$CT*

Each $\mu$Calc phantom was first imaged on a commercial $\mu$CT system (MicroXCT-200, Carl Zeiss X-ray Microscopy) within the Center for Molecular and Genomic Imaging at UC Davis. The phantom was fixed to a rotating platform [Fig. 1(a)] and 800 projections were acquired in step-and-shoot mode for a complete 360-deg rotation using a 40-kV x-ray source (80 total mAs). A total of six separate scans were acquired by translating the stage 9.2 mm between consecutive scans in order to cover all agar within the cylindrical test tubes. The projection data were reconstructed, using the vendor-provided filtered backprojection algorithm, into a 3D volume dataset with an isotropic voxel size of 0.034 mm.

The phantom was also imaged on a fourth-generation bCT scanner "Doheny" in our laboratory using 60 kV with 0.2 mm Gd filtration. This choice of kV/filtration combination has been demonstrated to provide a good trade-off for maximizing soft tissue and iodine contrast while still providing enough tube output within the short pulse width (4 ms) of the x-ray tube.[2]

Doheny is equipped with a flat-panel detector (Dexela 2923M, Varex Imaging Corporation, Salt Lake City, Utah) featuring a 0.45-mm-thick thallium-activated structured cesium iodide scintillator coupled to CMOS active pixel sensors. All scans acquired in this work utilized $1 \times 1$ binning mode with a 0.075-mm detector element size and a frame rate of 22.7 fps. Ultra-high mAs scans of the $\mu$Calc phantom inserts suspended in-air were acquired using 800 projections in a 360-deg rotation with the maximum mA at 60 kV (860 total mAs). A total of five repeat scans were acquired for ensemble averaging to minimize noise. Prior to ensemble averaging, pairwise subtraction was used to verify that the phantom did not move and that the x-ray source and detector were stable during scanning.

For all bCT acquisitions, the reconstruction process used a variation of the Feldkamp algorithm with a sharp Shepp–Logan kernel that cuts off at a frequency equal to two times Nyquist. An isotropic voxel size of 0.102 mm was used for all reconstructions, and low-frequency shading-artifact correction was performed using a previously reported correction algorithm.[22]

### 2.3.2 *In-phantom breast CT imaging*

In-phantom imaging was performed at MGD levels of 3 and 6 mGy in $1 \times 1$ binning mode (22.7 fps) with an x-ray pulse width of 4 ms and a total of 250 projections. The decision to use 250 projection images for the in-phantom bCT imaging was based upon a clinically realistic scan time of 10 s given frame rate limitations and the need for breath-hold during acquisition. MGD levels were calculated based upon previously reported normalized glandular dose values[21] and measured air kerma values at the scanner isocenter. The same reconstruction process

explained in Sec. 2.3.1 was used for the in-phantom images. A total of 23 repeat scans were acquired to obtain replicated images of each calcium grain for use in detectability estimation. Pairwise subtraction was used to verify stability of the imaging conditions. These scans were used for assessing $\mu$Calc detectability performance using an NPW-matched filter model observer as explained in Sec. 2.6.

## 2.4 $\mu$CT Size Estimation Accuracy and Segmentation

The $\mu$CT images of the three $\mu$Calc phantoms were used as the gold-standard for estimating the size and shape of each calcium grain. Two segmentation methods were investigated to differentiate the calcium carbonate grain from the agar background. The first approach uses Otsu's method[23] to determine a single threshold value which minimizes the intra-class intensity variance. The second segmentation method is adopted from a previously reported method by Shaheen et al.[24] for segmenting $\mu$Calc clusters from $\mu$CT biopsy specimens. This method uses Sobel edge detection followed by a series of morphological operations to segment the individual grains from the background. In our implementation (Fig. 2), the following steps were employed in order: (i) Sobel edge detection, (ii) 3D morphological dilation using a spherical structuring element with a radius of 1 voxel, (iii) flood fill, (iv) morphological erosion using the same spherical structuring element, (v) morphological opening applied in 2D using a $1 \times 3$ and $3 \times 1$ rectangular structuring element, and (vi) 3D median filtering with a $3 \times 3 \times 3$ voxel filter size.

Once each calcium grain was segmented, several size metrics were determined from the segmented volumes. The equivalent diameter measured in the 3D volume-of-interest (VOI) ($D_{\mathrm{eq,3D}}$) was used as a global comparison of the $\mu$Calc grain sizes to assess consistency with the reported values from the manufacturer (0.212 to 0.224 mm). $D_{\mathrm{eq,3D}}$ was computed as the diameter of a sphere with the same volume as the segmented calcium grain. In addition, two size metrics were measured for the 2D coronal slice through the center of each $\mu$Calc (used in the detectability estimations described in Sec. 2.6) including the area and the eccentricity of an ellipse that has the same normalized second central moments as the segmented calcium grain. If the major axis is defined as $a$ and the minor axis as $b$, then eccentricity is given by $\sqrt{1-(b/a)^2}$.

The Teflon sphere fiducial markers were manufactured to a high tolerance. The Teflon sphere diameter of 2.42 mm was determined by measuring mass using a high-precision digital mass scale and averaging the mass of 10 Teflon spheres; diameter was computed knowing the density of the material. We compare segmentation of these spheres to an ideal sphere with the same diameter to quantify segmentation accuracy. Segmentation accuracy was assessed using the average Hausdorff distance (AHD) for each segmentation method against the ideal sphere.

## 2.5 Generation of Calcium Grain Signal Profiles

The three Teflon spheres embedded in each of the $\mu$Calc phantom inserts were used for an initial, rigid registration between the $\mu$CT and ensemble average bCT reconstructions of the in-air $\mu$Calc phantom scans. After this initial registration, each individual calcium grain was localized in the

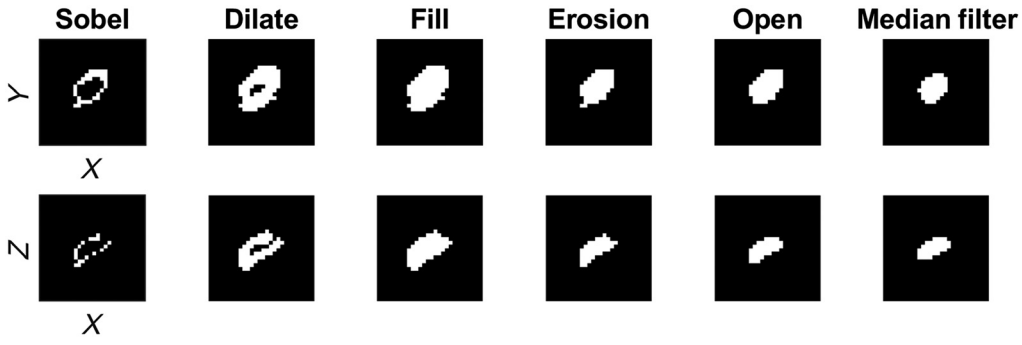


**Fig. 2** A diagram showing each step in the method used for segmentation of the microcalcifications in the $\mu$CT reconstructions.

$\mu$CT reconstructions using a combination of Otsu's method for binary segmentation and connected components analysis for distinguishing between individual grains. Three-dimensional VOIs centered about the binarized image center-of-mass of each individual calcium grain were then extracted from the $\mu$CT and bCT reconstructions. Each VOI was $1.632\ \text{mm} \times 1.632\ \text{mm} \times 1.632\ \text{mm}$ (i.e., $48 \times 48 \times 48$ voxels for $\mu$CT reconstructions and $16 \times 16 \times 16$ voxels for bCT). All VOIs containing either more than one calcium grain or materials other than agar and calcium carbonate were excluded from the analysis.

Fine alignment of the $\mu$CT and bCT VOIs was achieved by minimizing the mean squared error (MSE) of the VOI pairs over rigid translation of the $\mu$CT VOIs as shown in the flow diagram in Fig. 3. Translation of the high-resolution $\mu$CT VOI (as opposed to the bCT VOI) ensures that the interpolation step does not impart unnecessary blurring. The high-resolution $\mu$CT is blurred in order to match the lower resolution of the bCT system. This blurring is accomplished using the bCT MTF since the $\mu$CT resolution is a factor of 4 to 5 times higher than the bCT resolution. The registration procedure involved: (i) blurring each coronal slice of the $\mu$CT VOI using a 2D coronal plane MTF measured on the bCT system, (ii) 3D translation of the $\mu$CT VOI and downsampling to the bCT voxel size, (iii) rescaling the $\mu$CT attenuation values to match the bCT voxels, and (iv) computing the MSE between the VOI pairs. Steps ii to iv were repeated until all possible translations, in steps of 0.034 mm, within a spatial extent of 4 bCT voxels (0.408-mm cube) were compared. Alignment of each calcium grain was determined as the translation that minimized the MSE between the $\mu$CT VOI and bCT VOI. The rescaling process in step iii accounts for differences between the 40-kV spectrum used in the $\mu$CT acquisitions and the 60-kV spectrum used in the bCT acquisitions. This is appropriate for these phantoms in which there are two relevant materials (calcium and an agar background) but may require additional corrections if more materials were involved.

## 2.6 *Estimation of Detectability*

An NPW matched filter model[25] was used for all detectability calculations. In this model, the mean signal profile was used as a detection template. Our goal was to compare detectability estimates derived from the $\mu$CT data with estimates derived directly from the bCT scanner as a test of the accuracy of the results using $\mu$CT as the ground truth.

Detectability calculations required a separate alignment procedure to register the in-air bCT and $\mu$CT scans with the in-phantom bCT data. To address template alignment, a single low-noise bCT image of the in-phantom acquisitions was first generated by taking the ensemble average of all 23 repeat scans. The ensemble average in-phantom bCT image was initially manually registered (using the fiducial markers) with the in-air bCT (or $\mu$CT) reconstruction used for extracting the templates. A $16 \times 16 \times 16$ voxel cube (1.632-mm spatial extent) VOI surrounding each $\mu$Calc

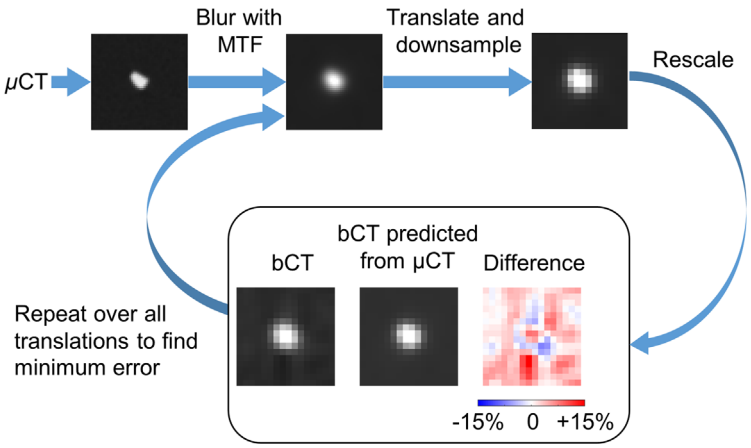


**Fig. 3** Flow diagram for methodology used to register $\mu$CT and bCT VOIs of each individual microcalcification grain. For simplicity, central 2D slices through the $\mu$Calc grain are shown but the methodology was applied in 3D.

grain was then extracted from both the in-phantom and in-air bCT (or $\mu$CT) volumes using a similar process described in Sec. 2.5. Each template VOI was then iteratively translated and compared against the in-phantom VOI in a loop driven by minimizing the MSE computed between the VOI pairs. This allows us to make direct comparisons of detectability using $\mu$CT with estimations from bCT.

### 2.6.1 *Detectability from the bCT scanner*

A central coronal profile through each bCT-derived signal template, $w$ a $16 \times 16$ array, was used as a detection template in the NPW model. The mean background value was computed for each detection template and subtracted from $w$. Template responses were defined by an inner product of the template with a $16 \times 16$ region of interest (ROI) from the in-phantom bCT reconstructions. The signal-present response was computed by taking the inner product of the NPW template with the in-phantom bCT reconstructions at the calcium grain location. The signal-absent response was computed by taking the inner product of the NPW template with the in-phantom bCT reconstruction at 10 different non-signal locations randomly selected within the same slice and contained within the agar region of the $\mu$Calc phantom insert. Let $\lambda_i^+$ represent the signal-present template responses and $\lambda_{i,j}^-$ represent the signal-absent template responses, where $i = 1, \ldots, 23$ indexes the 23 repeat scans and $j = 1, \ldots, 10$ indexes the 10 randomly selected non-signal locations. Let $\bar{\lambda}^+$ and $\bar{\lambda}^-$ represent the average of the signal-present responses and signal-absent responses, respectively. The signal present variance, $\hat{\sigma}^2_{\lambda^+}$, was computed as the sample variance of the signal-present responses. The signal absent variance, $\hat{\sigma}^2_{\lambda^-}$, was computed at each of the 10 locations and then averaged across location. The resulting estimate of detectability is given by

$$d'_T = \frac{\bar{\lambda}^+ - \bar{\lambda}^-}{\sqrt{\frac{1}{2}(\hat{\sigma}^2_{\lambda^+} + \hat{\sigma}^2_{\lambda^-})}}. \tag{1}$$

Detectability was calculated separately for each of the 58 calcium grains. Error in the $d'$ estimation was derived from bootstrapping over the 23 replicated bCT scans with 1000 resamples (with replication). This represents $\mu$Calc phantom detection performance assessed directly from the bCT scanner. Separate estimations were made for the 3- and 6-mGy datasets.

### 2.6.2 *Detectability using $\mu$CT parameters in a regression model*

In the next three sections, different ways to use the $\mu$CT data to estimate detectability were explored. The first method used a regression model to relate parameters of the calcium grains, derived from $\mu$CT images, to $d'_{bCT}$. A linear model was defined with $d'_{bCT}$ as the dependent variable, and with each calcium grain's vertical position ($z$), area (i.e., total area of the segmented grain in the central coronal slice), and eccentricity as the explanatory variables. Vertical position was included because the tapered shape of the V3 breast phantom will likely impact detectability as quantum noise increases with the phantom diameter along $z$. The explanatory variables were all extracted from the $\mu$CT reconstructions of the calcium grains. All three variables were standardized by subtracting off the average and dividing by the standard deviation—producing an average of 0 and a standard deviation of 1 for each variable. The regression model equation for each calcium grain ($m = 1, \ldots, 58$) was of the form

$$d'_m = x_0 + x_1\ z_m + x_2\,\text{area}_m + x_3\,\text{eccentricity}_m + \varepsilon_m, \tag{2}$$

where $x_0 - x_3$ are the regression coefficients, and $\varepsilon_m$ is the residual error. This regression-based approach was used to demonstrate how ground-truth parameters of the calcium grains are associated with detection performance in the bCT images. Separate regressions models were fit for the 3- and 6-mGy estimated detectability measures.

### 2.6.3 *Detectability using $\mu$CT signal profiles as a detection template*

Detectability was estimated in an analogous way to Sec. 2.6.1, except that the $\mu$CT-derived NPW templates were used instead of bCT-derived templates. The $\mu$CT-derived templates were generated by convolving the high-resolution signal profiles with the MTF of the bCT system (as described in Sec. 2.5). The resulting detectability estimate, $d'_{\mu\text{CT}}$, was used to demonstrate that the $\mu$CT-derived signal profiles accurately represent the bCT profiles in a performance calculation. A bootstrapping procedure, identical to Sec. 2.5.1, was used to generate error bars on the estimates. This procedure was applied to both the 3- and 6-mGy data.

### 2.6.4 *Detectability using $\mu$CT data and bCT statistical properties*

In the previous section, the $\mu$CT-derived profiles were used as detection templates and the replicated images were used to obtain the statistical properties of the template responses. However, if the noise power spectrum (NPS) of the bCT images is known, then these statistical properties can be derived directly from theory.[25–27] In this approach, the $\mu$CT-derived templates also represent the mean signal profile ($w = s$), and so the numerator of the detectability formulation is given by the inner product of the template with itself ($\bar{\lambda}^+ - \bar{\lambda}^- = w^T w$). The denominator is calculated by integrating the NPS against the squared magnitude of the Fourier transformed template. For this work, the NPS was estimated from a $90 \times 90$ ROI centered about the $\mu$Calc phantom insert in both the 3- and 6-mGy datasets. The average across the 23 repeat scans was subtracted from each of the replicated images for detrending prior to computing the FFT. For the purpose of computing the variance of the template, the $\mu$CT-derived template was embedded in a $90 \times 90$ zero-padded ROI, $w_{90\times90}$,

$$\sigma_w^2 = \sum_{u=0,v=0}^{89} NPS(u,v)|W_{90\times90}(u,v)|^2, \tag{3}$$

where $W_{90\times90}(u,v)$ is the 2D Fourier transform of the embedded template. The resulting detectability predicted by the $\mu$CT-derived template and the bCT statistical properties is given by

$$d'_{\text{Pred}} = \frac{w^T w}{\sigma_w}. \tag{4}$$

We refer to this quantity as "predicted" because it does not directly utilize the bCT images. However, it does utilize the average of the bCT images to account for contrast differences resulting from the spectral differences between $\mu$CT (40 kV) and bCT (60 kV) x-ray sources.

The measured detectability (described in Sec. 2.6.1) and predicted detectability estimations were compared for all 58 calcium grains and the two dose levels datasets investigated in this work.

### 2.6.5 *Rose criterion*

We used the Rose criterion[28] as a simple measure of how visible the calcium grains are in the bCT images. The original Rose criterion was based on the contrast of the targets and the pixel variance; however, Burgess[29] has shown how the NPW model relates to Rose's original formula for visibility. The Rose criterion is met if $d' \geq 5$ and is a commonly used heuristic indicator of when noise-limited targets should be relatively visible. The Rose criterion is typically used as a performance metric in noise-limited uniform backgrounds, which is consistent with its use in our phantom data. In clinical bCT images, the image background typically varies between adipose and fibroglandular tissues within the breast parenchyma, so the uniform assumption is less clear and may not apply to a $\mu$Calc located near a boundary between adipose and fibroglandular tissues. Thus, our visibility criterion should be interpreted as applying to $\mu$Calcs located within a

uniform fibroglandular region. Clinical $\mu$Calcs located near an adipose/fibroglandular boundary will require other criteria for visibility.

## 3 Results

### 3.1 *Microcalcification Segmentation and Size Estimation from $\mu$CT*

Maximum intensity projections through the $\mu$CT and bCT reconstruction of one $\mu$Calc phantom are shown in Fig. 4 detailing the location of the fiducial markers and the calcium carbonate grains dispersed throughout the phantom volume. The Teflon sphere at the bottom of the test tube shown in Fig. 4 was too close to the bottom of the phantom test tube and was therefore not included in the $\mu$CT segmentation accuracy assessment. Out of the total of nine Teflon spheres in the three $\mu$Calc phantoms, six were used for segmentation accuracy assessment. The AHD [standard deviation] was 0.452 [0.025] using Otsu's method and 0.208 [0.011] using the morphological method across the six Teflon spheres. The morphological method more accurately segmented objects in this imaging setting and was therefore used in this study to segment the calcium grains in the $\mu$Calc phantoms.

A total of 58 distinct $\mu$Calc grains were localized within the $\mu$CT reconstructions of the three $\mu$Calc phantoms after applying the exclusion criteria (explained in Sec. 2.5). An example segmentation of one grain is shown in orthogonal planes in Fig. 5(a). Equivalent diameter results, measured in 3D, for all $\mu$Calc grains are shown in Fig. 5(b) with an average value [range] of 0.249 mm [0.207, 0.316] and 0.207 mm [0.156, 0.287] for the Otsu and morphological segmentation methods, respectively, as indicated by the solid horizontal lines in the figure. The difference for Otsu's method and the morphological method was 12.8% and −5%, respectively, relative to the mean calcium carbonate grain size reported by the manufacturer (0.218 mm) as shown in Fig. 5(b). Together these results provide confidence in the use of the morphological methods for accurate segmentation and estimation of the $\mu$Calc grain sizes in the $\mu$CT reconstructions.

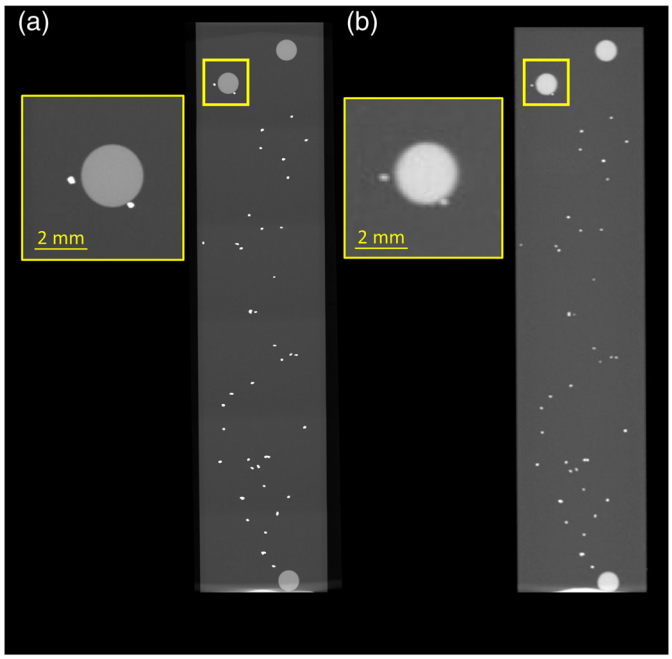


**Fig. 4** Maximum intensity projections of the microcalcification phantom acquired using (a) $\mu$CT and (b) bCT. Images are resized to equivalent scale. Teflon spheres and calcium carbonate grains are shown within the agar background. The inset image shows the high-resolution properties of the $\mu$CT acquisitions.

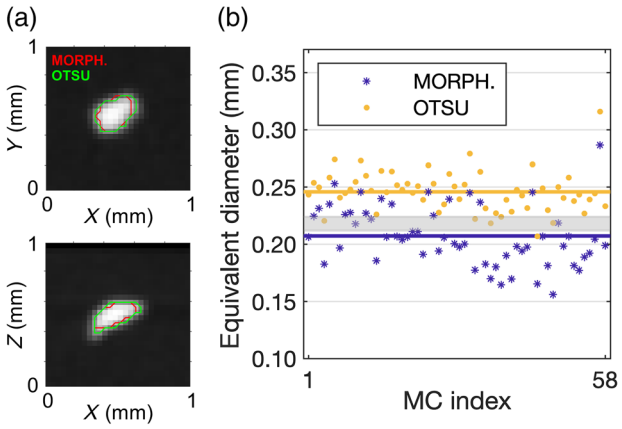


**Fig. 5** $\mu$CT segmentation and size estimations results. Segmentation outlines using the morphological operations[24] "MORPH" and Otsu's thresholding method "OTSU" are shown in orthogonal planes (a), and equivalent diameter estimations (measured in 3D) for all 58 microcalcification grains. The shaded region between the dot-dash lines corresponds to the range of diameter provided by the manufacturer (0.212 to 0.224 mm).

## 3.2 *Microcalcification Signal Profiles*

The predicted $\mu$Calc signal profiles derived from high-resolution $\mu$CT were in strong agreement with those derived directly from the bCT scanner as an average of five high-dose acquisitions with the $\mu$Calc phantom suspended in air. The $\mu$CT profiles compensate for bCT system spatial resolution and differences in the reconstructed voxel values of the two imaging systems. Figure 6 shows an example of the original $\mu$CT image, the original bCT, and the predicted bCT results for 12 example calcium-grain profiles. The ROIs displayed in the figure are cropped to $10 \times 10$ for better visualization, but the full ROI size used for detectability estimation was $16 \times 16$. The absolute difference between the observed bCT profiles and those predicted from the $\mu$CT images ranged from 1.8% to 6.8% for the 58 calcium grains, where the difference was computed as $(\mu\text{CT} - \text{bCT})/\text{bCT}$. The maximum difference for individual voxels within an ROI ranged from $-12.9\%$ to 14.8%.

## 3.3 *Microcalcification Detectability in the Breast CT Scans*

The detectability index [Eq. (1)] of the 58 calcium grains from the 23 repeat scans ranged from 2.5 to 10.6 for 3-mGy MGD and from 3.8 to 15.3 for 6-mGy MGD as shown in Fig. 7. The detectability was uniformly higher for 6 mGy and was well fit by a linear equation with a slope of square root of 2 (i.e., $\sqrt{6\ \text{mGy}/3\ \text{mGy}}$)—suggesting that dose is the limiting factor for this task.

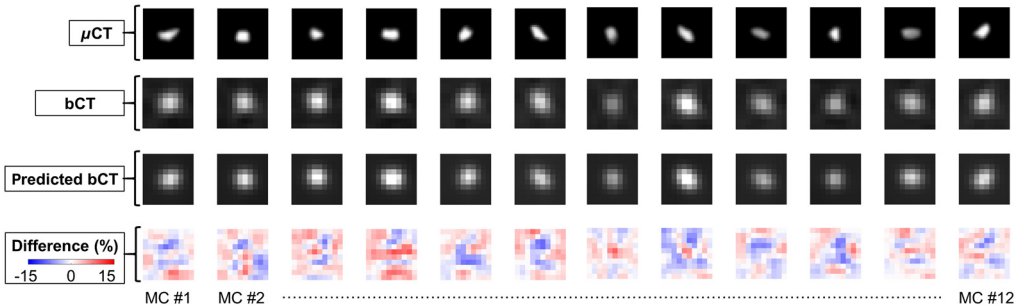


**Fig. 6** Example signal profiles for 12 individual microcalcification grains are shown for a 2D central slice through the high-resolution $\mu$CT images (top row), the central slice through the bCT reconstruction (second row), the predicted bCT derived from the $\mu$CT (third row), and the difference image (bottom row).

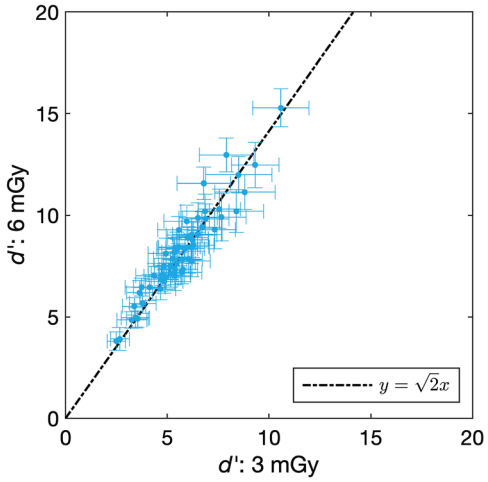


**Fig. 7** Detectability estimations computed using the bCT-derived signal templates plotted for 3 mGy ($x$-axis) and 6 mGy ($y$-axis) MGD. The data are well described by a linear fit with a slope defined by the relative doses. Error bars correspond to the 95% confidence interval derived from bootstrapping across the 23 repeat scans.

### 3.4 *Microcalcification Detection Modeling*

As described in Sec. 2.6.2, linear regression was used to determine a model of $\mu$Calc detectability based on $z$-location, size (area), and shape (eccentricity) at a fixed dose level. The regression showed significant effects ($p < 1.1 \times 10^{-4}$) for $z$-location and area at both dose levels. Shape (i.e., eccentricity) was not a significant factor with $p = 0.4249$ and $p = 0.4362$ for 3 and 6 mGy, respectively. For 3-mGy MGD, the final model was

$$d_i' = 5.7029 + 0.6659 z_i + 0.8890\,\text{area}_i + 0.1220\,\text{eccentricity}_i + \epsilon_i, \tag{5}$$

and for 6 mGy MGD the final model was

$$d_i' = 8.1869 + 0.8489 z_i + 1.2436\,\text{area}_i + 0.1589\,\text{eccentricity}_i + \epsilon_i. \tag{6}$$

The adjusted $R^2$ value for the regression model was 0.505 for 3 mGy and 0.512 for 6 mGy suggesting that the parameters extracted from the $\mu$CT images and $z$-location explain approximately half of the variance in the observed detectability data. The V3 breast phantom diameter tapers from the chest wall to the nipple (i.e., along $z$) and correspondingly the image noise decreases monotonically, resulting in the observed dependence of detectability on $z$-location within the scanner FOV as shown in Fig. 8(a). Additionally, an overall increase in the detectability was observed with increasing calcium-grain size (i.e., area) as shown in Fig. 8(b).

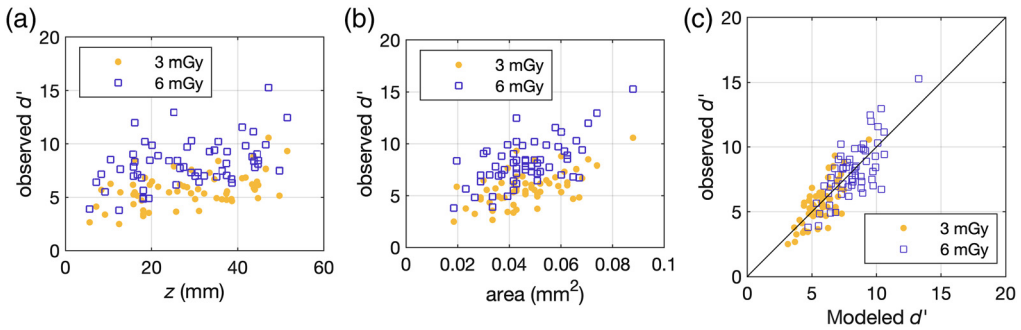


**Fig. 8** Observed detectability plotted as a function of $z$-location (a) and area (b) for 3 mGy (circles) and 6 mGy (open squares) MGD levels. Modeled detectability results are also plotted on the $x$-axis and observed detectability on the $y$-axis for both dose levels (c).

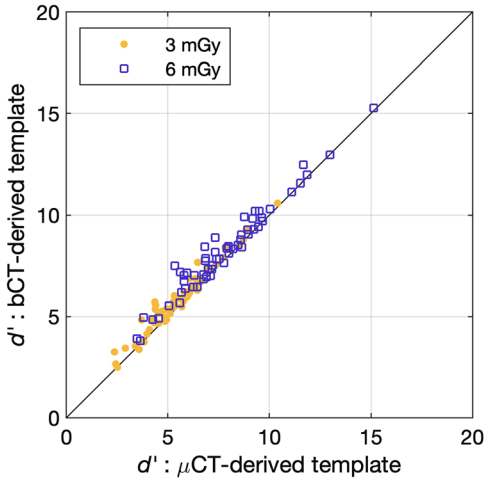


**Fig. 9** Plot of the detectability estimations using bCT-derived signal templates ($y$-axis) and $\mu$CT-derived signal templates ($x$-axis) plotted separately for 3- and 6-mGy MGD.

Figure 8(c) plots the modeled detectability against the observed detectability for both dose levels with clear associations between the modeled and observed detectability values. The variance explained by the regression model was approximately 10 times more than the average variance in detectability due to estimation error, as assessed by the bootstrapping procedure described in Sec. 2.6.1. Thus, information extracted from the $\mu$CT images capture important properties of the calcium grains.

### 3.5 *Detectability using $\mu$CT-derived Templates*

Detectability estimations using $\mu$CT signal profiles blurred by the bCT system MTF (as explained in Sec. 2.5), and applied to the repeated scan data, were in good agreement with detectability generated using $\mu$Calc templates derived directly from the bCT acquisitions as shown in Fig. 9. The coefficients of determination, $R^2$, of the linear regression fit were 0.952 and 0.953 for the dose levels of 3- and 6-mGy MGD, respectively. The slope of the regression fit included unity, and the intercept included zero for both 3- and 6-mGy MGD within a 95% confidence interval. This supports the results in Fig. 6, showing that the good agreement between the signal profiles translates into good agreement in performance when the profiles are used as detection templates in an NPW model.

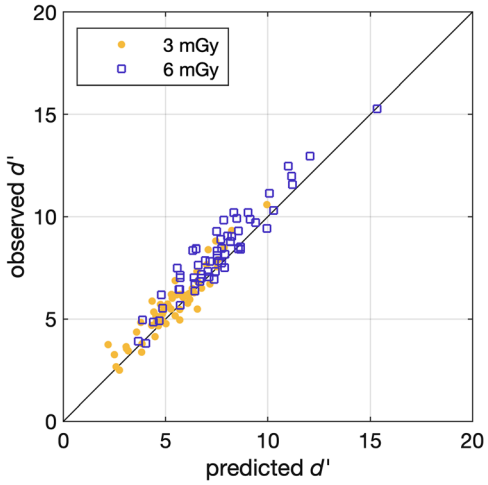


**Fig. 10** Detectability estimations observed directly from the bCT images ($y$-axis) and predicted from $\mu$CT signal profiles and bCT MTF and NPS characterization ($x$-axis) plotted separately for 3- and 6-mGy MGD.

### 3.6 *Detectability using $\mu$CT-derived Templates and Statistical Properties of the Breast CT System*

Detectability predicted from $\mu$CT signal profiles and bCT MTF and NPS are compared to the observed bCT detectability in Fig. 10. The $R^2$ values of the linear regression fit were 0.880 and 0.906 for the dose levels of 3- and 6-mGy MGD, respectively. The slope of the regression fit included unity, and the intercept included zero for both 3- and 6-mGy MGD within a 95% confidence interval. This agreement shows that $\mu$CT-derived templates can be used in conjunction with bCT statistical properties to accurately predict detection performance.

## 4 Discussion

The use of physical phantoms is highly dependent on having ground truth in terms of the objects being imaged. When the objects are small and irregularly shaped, such as the calcium grains, this can be difficult to obtain by direct measurement (e.g., calipers, sieves, etc.). Images from an independent modality with high resolution and low noise, such as $\mu$CT, can be a useful approach to obtain this information. The results shown in Figures 7–10 show that $\mu$CT images of the $\mu$Calc phantom can be used in a variety of ways to understand detectability in the bCT scanner. The regression modeling in Eqs. (5) and (6) shows that the $z$-location and size explain a significant component of variance in the detectability across the 58 calcium grains as shown in Fig. 8. Signal profiles derived from $\mu$CT images appear similar to the profiles directly observed from the bCT scanner (Fig. 6), and when these profiles are used as a detection template, they accurately represent detectability on the bCT scanner (Fig. 9). Furthermore, the signal profiles can be used in conjunction with bCT statistical properties to accurately predict detectability (Fig. 10). Together these results illustrate the importance of characterizing the properties of small calcium grains for use in $\mu$Calc phantoms.

While the purpose of this study was to evaluate $\mu$CT as a method for obtaining ground truth in $\mu$Calc phantoms, the study design does allow for a limited assessment of the bCT scanner for calcification detection. The equivalent diameter (i.e., $D_{\mathrm{eq,3D}}$) of the calcium carbonate grains used in this study ranged in size from 0.156 to 0.287 mm, which are representative of clinically relevant $\mu$Calcs.[20] At a MGD level of 6 mGy, 97% of the calcium grains (56 of 58) met the Rose criterion ($d' > 5.0$) for visibility,[28,29] and 81% of the calcium grains (47 of 58) met the Rose criterion at 3 mGy. These findings suggest that for the specific breast phantom size, calcium grain composition, and position within the scanner FOV investigated in this work, $\sim$200 $\mu$m calcifications are at the threshold where visibility begins to degrade because of increased noise at lower doses. The findings also motivate future studies to give a more complete characterization of calcium-grain detectability.

This work focused on analyzing 2D signal profiles in the coronal plane of the bCT system; however, it is a relatively straightforward process to generate signal profiles, and ultimately detectability estimations, in other planes (e.g., axial or sagittal). Fully 3D assessments are possible as well using the 3D nature of the $\mu$CT scans for ground truth. Additionally, this work focused on the calcium-grain profile as a detection template (i.e., the NPW model). The use of $\mu$CT-derived templates may be of value to assess the $\mu$CT data in the context of other detection templates in future work.

## 5 Conclusion

Imaging microcalcifications remains an ongoing challenge for bCT and other breast imaging modalities, with an important role for physical phantoms in demonstrating and optimizing imaging system performance. We found that $\mu$CT imaging at very high resolution and dose, with appropriate segmentation of the resulting images, is a useful way to characterize ground-truth morphology of calcium grains embedded in these phantoms. There are significant associations between the $\mu$Calc size parameters, measured from the $\mu$CT images, and the detectability of calcifications on the dedicated bCT scanner at clinical dose levels. Thus, the $\mu$CT data allow for a more refined estimate of size and shape for calcium grains that are created using a sieve

process. Additionally, the $\mu$CT data can be used in conjunction with system characterization (MTF and NPS) to obtain accurate measures of detectability. This allows for the possibility of a predictive evaluation of performance for a system, or system component, that may not currently be available. Finally, overall detectability of ~200 $\mu$m calcifications in a breast-sized phantom at clinical dose levels was relatively good, with 81% of the calcification meeting the Rose criterion for visibility at 3 mGy and 97% meeting the criterion at 6 mGy.

## Disclosures

A.M.H. and J.M.B. have patents (pending and issued) pertaining to breast CT. J.M.B. has prior breast CT research support, licensing agreements with, and is a shareholder for Izotropic Imaging Corp. of Canada.

## Acknowledgments

The authors would like to thank Dr. Chris Shaw for providing the calcium carbonate grains used in the phantom fabrication and Dr. Douglas Rowland (Center for Molecular and Genomic Imaging; UC Davis) for the $\mu$CT imaging. The research was supported by a research grant (R01-EB025829) from the National Institute of Biomedical Imaging and Bioengineering and a research grant (R01-CA181081) from the National Cancer Institute.

## References

1. C.J. Lai et al., "Visibility of microcalcification in cone beam breast CT: effects of X-ray tube voltage and radiation dose," *Med. Phys.* **34**(7), 2995–3004 (2007).
2. A. M. Hernandez et al., "Effects of kV, filtration, dose, and object size on soft tissue and iodine contrast in dedicated breast CT," *Med. Phys.* **47**(7), 2869–2880 (2020).
3. H.-M. Cho et al., "Microcalcification detectability using a bench-top prototype photon-counting breast CT based on a Si strip detector," *Med. Phys.* **42**(7), 4401–4410 (2015).
4. L. C. Ikejimba et al., "A four-alternative forced choice (4AFC) methodology for evaluating microcalcification detection in clinical full-field digital mammography (FFDM) and digital breast tomosynthesis (DBT) systems using an inkjet-printed anthropomorphic phantom," *Med. Phys.* **46**(9), 3883–3892 (2019).
5. L. Chen et al., "Simulated lesion, human observer performance comparison between thin-section dedicated breast CT images versus computed thick-section simulated projection images of the breast," *Phys. Med. Biol.* **60**(8), 3347–3358 (2015).
6. N. J. Packard et al., "Effect of slice thickness on detectability in breast CT using a prewhitened matched filter and simulated mass lesions," *Med. Phys.* **39**(4), 1818–1830 (2012).
7. J. Yao et al., "Cone beam CT for determining breast cancer margin: an initial experience and its comparison with mammography and specimen radiograph," *Int. J. Clin. Exp. Med.* **8**(9), 15206–15213 (2015).
8. K. K. Lindfors et al., "Dedicated breast CT: initial clinical experience," *Radiology* **246**(3), 725–733 (2008).
9. A. M. O'Connell, A. Karellas, and S. Vedantham, "The potential role of dedicated 3D breast CT as a diagnostic tool: review and early clinical examples," *Breast J.* **20**(6), 592–605 (2014).
10. X. Gong, A. A. Vedula, and S. J. Glick, "Microcalcification detection using cone-beam CT mammography with a flat-panel imager," *Phys. Med. Biol.* **49**(11), 2183–2195 (2004).
11. A. A. Sanchez, E. Y. Sidky, and X. Pan, "Task-based optimization of dedicated breast CT via Hotelling observer metrics," *Med. Phys.* **41**(10), 101917 (2014).
12. M. Han, B. Kim, and J. Baek, "Human and model observer performance for lesion detection in breast cone beam CT images with the FDK reconstruction," *PLoS One* **13**(3), e0194408 (2018).
13. J. Bian et al., "Investigation of iterative image reconstruction in low-dose breast CT," *Phys. Med. Biol.* **59**(11), 2659–2685 (2014).

14. Y. Shen et al., "Cone beam breast CT with a high pitch (75 mum), thick (500 mum) scintillator CMOS flat panel detector: visibility of simulated microcalcifications," *Med. Phys.* **40**(10), 101915 (2013).
15. P. Henrot et al., "Breast microcalcifications: the lesions in anatomical pathology," *Diagn. Interv. Imaging* **95**(2), 141–152 (2014).
16. S. Vedantham and A. Karellas, "X-ray phase contrast imaging of the breast: analysis of tissue simulating materials," *Med. Phys.* **40**(4), 041906 (2013).
17. V. Barth, E. D. Franz, and A. Scholl, "Microcalcifications in mammary glands," *Naturwissenschaften* **64**(5), 278–279 (1977).
18. A. S. Haka et al., "Identifying microcalcifications in benign and malignant breast lesions by probing differences in their chemical composition using Raman spectroscopy," *Cancer Res.* **62**(18), 5375–5380 (2002).
19. O. Hassler, "Microradiographic investigations of calcifications of the female breast," *Cancer* **23**(5), 1103–1109 (1969).
20. I. Willekens et al., "High-resolution 3D micro-CT imaging of breast microcalcifications: a preliminary analysis," *BMC Cancer* **14**, 9 (2014).
21. A. M. Hernandez and J. M. Boone, "Average glandular dose coefficients for pendant-geometry breast CT using realistic breast phantoms," *Med. Phys.* **44**(10), 5096–5105 (2017).
22. P. Ghazi et al., "Shading artifact correction in breast CT using an interleaved deep learning segmentation and maximum-likelihood polynomial fitting approach," *Med. Phys.* **46**(8), 3414–3430 (2019).
23. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
24. E. Shaheen et al., "The simulation of 3D microcalcification clusters in 2D digital mammography and breast tomosynthesis," *Med. Phys.* **38**(12), 6659–6671 (2011).
25. K. J. Myers et al., "Effect of noise correlation on detectability of disk signals in medical imaging," *J. Opt. Soc. Am. A* **2**(10), 1752–1759 (1985).
26. J. G. Ott et al., "Update on the non-prewhitening model observer in computed tomography for the assessment of the adaptive statistical and model-based iterative reconstruction algorithms," *Phys. Med. Biol.* **59**(15), 4047–4064 (2014).
27. C. Abbey and M. Eckstein, "Observer models as a surrogate to perception experiments," in *The Handbook of Medical Image Perception and Techniques*, pp. 240–250 (2010).
28. A. Rose, "The sensitivity performance of the human eye on an absolute scale," *J. Opt. Soc. Am.* **38**(2), 196–208 (1948).
29. A. E. Burgess, "The rose model, revisited," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **16**(3), 633–646 (1999).

**Andrew M. Hernandez** is an assistant adjunct professor of radiology at the University of California Davis. He received his BS degree in nuclear engineering from UC Berkeley in 2011 and his PhD in biomedical engineering from UC Davis in 2017. His research interests include dosimetry and x-ray imaging for applications in radiology, nuclear medicine, and radiation oncology.

**Amy E. Becker** is a PhD candidate at UC Davis. She received her BS degree in imaging science from Rochester Institute of Technology in 2016 and is currently pursuing a PhD in biomedical engineering at UC Davis. Her research focuses on multisource x-ray imaging systems, including digital breast tomosynthesis and breast CT.

**Su Hyun Lyu** is a graduate student researcher at UC Davis. She received her BS degree in biomedical engineering from UC Davis in 2018 and is currently pursuing a PhD in biomedical engineering. Her research interests include contrast-enhancement and deep learning for medical imaging applications.

**Craig K. Abbey** is a researcher in the Department of Psychological and Brain Sciences at UC Santa Barbara. His training is in the field of applied mathematics, and his research focuses on assessment of medical imaging devices and image processing in terms of performance in diagnostic and quantitative tasks.

**John M. Boone** is a professor of radiology at UC Davis and holds an appointment in the Department of Biomedical Engineering. He received his training in medical physics and is board-certified by the American Board of Radiology in diagnostic radiological physics. His research interests focus on breast dosimetry and the development of breast computed tomography for breast cancer screening and diagnostic evaluation.