Population of 224 realistic human subject-based computational breast phantoms

David W. Erickson[a]
Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705 and Medical Physics Graduate Program, Duke University,
Durham, North Carolina 27705

Jered R. Wells[a],[b]
Clinical Imaging Physics Group and Carl E. Ravin Advanced Imaging Laboratories, Duke University
Medical Center, Durham, North Carolina 27705 and Medical Physics Graduate Program, Duke University,
Durham, North Carolina 27705

Gregory M. Sturgeon
Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705

Ehsan Samei
Department of Radiology and Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705 and Departments of Physics, Electrical and Computer Engineering,
and Biomedical Engineering, and Medical Physics Graduate Program, Duke University,
Durham, North Carolina 27705

James T. Dobbins III
Department of Radiology and Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705 and Departments of Physics and Biomedical Engineering and Medical Physics
Graduate Program, Duke University, Durham, North Carolina 27705

W. Paul Segars
Department of Radiology and Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705 and Medical Physics Graduate Program, Duke University,
Durham, North Carolina 27705

Joseph Y. Lo
Department of Radiology and Carl E. Ravin Advanced Imaging Laboratories, Duke University Medical Center,
Durham, North Carolina 27705 and Departments of Electrical and Computer Engineering and Biomedical
Engineering and Medical Physics Graduate Program, Duke University, Durham, North Carolina 27705

Purpose: To create a database of highly realistic and anatomically variable 3D virtual breast phantoms based on dedicated breast computed tomography (bCT) data.

Methods: A tissue classification and segmentation algorithm was used to create realistic and detailed 3D computational breast phantoms based on 230+ dedicated bCT datasets from normal human subjects. The breast volume was identified using a coarse three-class fuzzy C-means segmentation algorithm which accounted for and removed motion blur at the breast periphery. Noise in the bCT data was reduced through application of a postreconstruction 3D bilateral filter. A 3D adipose nonuniformity (bias field) correction was then applied followed by glandular segmentation using a 3D bias-corrected fuzzy C-means algorithm. Multiple tissue classes were defined including skin, adipose, and several fractional glandular densities. Following segmentation, a skin mask was produced which preserved the interdigitated skin, adipose, and glandular boundaries of the skin interior. Finally, surface modeling was used to produce digital phantoms with methods complementary to the XCAT suite of digital human phantoms.

Results: After rejecting some datasets due to artifacts, 224 virtual breast phantoms were created which emulate the complex breast parenchyma of actual human subjects. The volume breast density (with skin) ranged from 5.5% to 66.3% with a mean value of $25.3\% \pm 13.2\%$. Breast volumes ranged from 25.0 to 2099.6 ml with a mean value of 716.3 ± 386.5 ml. Three breast phantoms were selected for imaging with digital compression (using finite element modeling) and simple ray-tracing, and the results show promise in their potential to produce realistic simulated mammograms.

Conclusions: This work provides a new population of 224 breast phantoms based on *in vivo* bCT data for imaging research. Compared to previous studies based on only a few prototype cases, this dataset provides a rich source of new cases spanning a wide range of breast types, volumes, densities, and parenchymal patterns. © *2016 American Association of Physicists in Medicine.*
[http://dx.doi.org/10.1118/1.4937597]

1. INTRODUCTION

More than 1 in 8 women who live for 80 years are diagnosed with breast cancer in their lifetime making it the second leading cause of cancer death for women in the United States. Early detection of breast cancer has been found to significantly improve survival rates. This underscores the need for early detection to enable diagnosis and effective treatment. Among other clinically available techniques, screening x-ray mammography is presently the most effective for early detection of breast cancer. Despite the success of mammography, there is still substantial room for improvement.[1,2] A major limitation to mammography is that volumetric information is inferred from 2D projection images: the tissues are superimposed on one another making the detection of breast cancer challenging.[1] Overall, the breast lesion detection sensitivity of screening mammography falls between 63% and 88% and is dependent upon patient age, family medical history,[3] breast density,[4] and technical factors (e.g., image acquisition technique factors, dose, and quality of the imaging system) among other factors. To address some of the limitations of mammography, several promising new breast imaging technologies have emerged: digital breast tomosynthesis (DBT),[5,6] contrast-enhanced mammography,[7,8] and dedicated breast computed tomography (bCT).[9,10] However, it can be challenging to fully evaluate each technology, its merits, and its shortcomings, especially in the preclinical phases. While clinical trials still serve as the gold standard for making these evaluations, they are costly, time consuming, and sometimes involve risk to the patient (e.g., radiation dose). In response to these challenges, one proposal is the use of virtual clinical trials involving the use of computerized phantoms in realistic imaging scenarios.

Two main approaches have typically been utilized to create computerized phantoms for breast imaging research: (a) rule-based models where the anatomy is defined by mathematical equations or relationships and (b) voxelized models based upon human imaging data. Rule-based models[11–16] utilize a combination of voxel matrices and geometric primitives to create a phantom that may include the breast surface, the duct system, terminal ductal lobular units, Cooper's ligaments, the pectoral muscle, and breast abnormalities. Because of their mathematical basis, phantoms such as the ones made by Bakic *et al.* can produce a multitude of variable models for an imaging study.[17–19] In spite of their complexity, however, they are limited in their level of realism when compared to actual patient data. Alternatively, voxelized phantoms have been investigated as a means to produce more realistic models. These types of phantoms have been largely based on dedicated bCT data from mastectomy specimens.[20,21] However, these specimens do not adequately represent intact breasts. Currently, only a small number of voxelized phantoms using this method exist due to limited access to specimens.

We have taken a different approach to address the limitations of mathematical and voxelized phantoms through the segmentation of *in vivo* bCT data. Our current work is based on the original work of Li *et al.*[22] and several key improvements to that work contributed by Segars *et al.*[23] Further modification, improvements, and more bCT segmentations have been performed in order to create a larger cohort of over 200 realistic and detailed 3D computational breast phantoms based on *in vivo* human subject bCT data with the goal of incorporating them into the 4D XCAT suite of phantoms. The phantoms are hybrid models: they are derived from imaging data like voxelized phantoms but like mathematical phantoms, they use geometric surfaces to define the anatomy.[24] The anatomical basis provides a foundation for building digital breast phantoms which emulate actual breast anatomy, while the mathematical basis provides flexibility to manipulate and deform the models—a benefit which, for example, allows for realistic breast compression and manipulation for multimodality imaging simulations. Furthermore, previous work has shown that a nearly limitless number of realistic phantoms can be derived by merging and manipulating a small number of 3D computational breast phantoms.[25] These methods can be used in conjunction with the larger cohort of phantoms presented in the current work to produce an even larger and more diverse population of digital breast models. Ultimately, the breast phantoms developed in this work will be packaged into a user-friendly software application to distribute for breast imaging research.

2. METHODS

2.A. bCT segmentation

The phantoms produced in this work are based on 231 dedicated bCT datasets obtained through a cooperative agreement with John Boone at the University of California, Davis (UC Davis).[26–28] The majority of the women imaged were at a high degree of suspicion for breast cancer (BIRADS 4 or 5), and many were diagnosed with malignant breast cancers; however, all bCT data provided in this study were derived from unaffected breasts. A more detailed description of the bCT dataset demographics is provided by Huang *et al.*[26] Seven of the original 231 datasets received were removed from analysis due to image artifacts, and each of the remaining datasets was segmented using a semiautomated method. Several other authors have developed bCT segmentation algorithms. Notable among them are Chen,[29] Nelson *et al.*,[30] and Yang *et al.*[31] Chen used histogram partitioning into five subintervals followed by interval thresholding to decompose the breast into five segmentation classes based on voxel intensity. The methods of Nelson *et al.* employed a two compartment histogram model to classify adipose, glandular, and skin tissues. This was done using a median filter (to reduce noise) followed by region growing. The methods of Yang *et al.* made several novel contributions to the (nontrivial) task of bCT segmentation. They corrected adipose nonuniformities with a bias field correction method, denoised the volume using

TABLE I. Overview of steps in the segmentation routine.

BCFCM3D masking of bCT volume
↓
Noise reduction of reconstructed bCT data
↓
Voxel intensity normalization
↓
3D nonuniformity (bias field) correction
↓
Conversion to cubic voxels
↓
BCFCM3D bCT segmentation
↓
Skin segmentation
↓
Surface model generation

a bilateral filter, and segmented the bCT volume into three classes using a fuzzy C-means (FCM) classification approach. The methods of Yang *et al.*[31] inspired several improvements to the original work of Li *et al.*[22] which were necessary to develop a sufficiently generalized approach capable of adequately segmenting a larger and highly variable data set. Table I provides an overview of the updated segmentation algorithm which can be compared and contrasted with the segmentation steps of Li *et al.*[22] and Yang *et al.*[31] Sections 2.B–2.H explain in detail the major components of the new segmentation routine.

2.B. Breast volume masking

Masking of the breast volume to separate it from the background is the first step in the segmentation process. The original volume segmentation method by Li *et al.* implemented a thresholding scheme that required manual tuning of the threshold value for each breast volume.[22] This process was automated through the use of a 3D bias-corrected fuzzy C-means (BCFCM3D) algorithm.[32] The standard FCM algorithm objective function was modified to involve the influence of immediate neighbors,[33]

$$J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^p \| y_k - \beta_k - v_i \|^2$$
$$+ \frac{\alpha}{N_R} \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^p \left(\sum_{y_r \in \mathcal{N}_k} \| y_r - \beta_r - v_i \|^2 \right). \quad (1)$$

Equation (1) is comprised of two parts—a voxel intensity component and a spatial component. The first intensity-based component involves y_k which are the original image voxel values corrected by bias field β_k and compared to cluster means v_i. u_{ik}^p is the probability of FCM class membership, and p is a weighting component that determines the "fuzziness" of the resultant classification. The second spatial component of the algorithm serves as a regulizer which biases voxel classification toward that of its neighbors $\mathcal{N}_k$, the set of voxels neighboring y_k. α is another weighting term controlling the influence of neighbor classification, and N_R is the cardinality of $\mathcal{N}_k$. Bias field estimation and iterative objective function minimization schema are well-described by other authors.[32,33] Three voxel intensity values were used to define three cluster means: background, breast, and background/breast average intensities. The third class was used to exclude breast boundary blurring and artifacts due to the partial volume effect and patient motion during bCT data acquisition as depicted in Fig. 1. The breast volume mask was finalized through morphological smoothing with a 3D spherical structuring element with 1 mm radius to smooth the breast mask external boundary.[34]

It should be noted that patient motion may affect both the rendition of the breast boundary and structures internal to the breast. While the methods described in this section are designed to correct motion artifacts affecting the breast external boundary, no attempt has been made to correct artifacts associated with structures within the breast volume. This topic is discussed at length in Sec. 4.

2.C. Denoising

Of the 224 bCT datasets, projection space denoising was previously applied to 53 using partial diffusion equation-based techniques as part of work by Xia *et al.*[35] The remaining 171 datasets were reconstructed without projection space
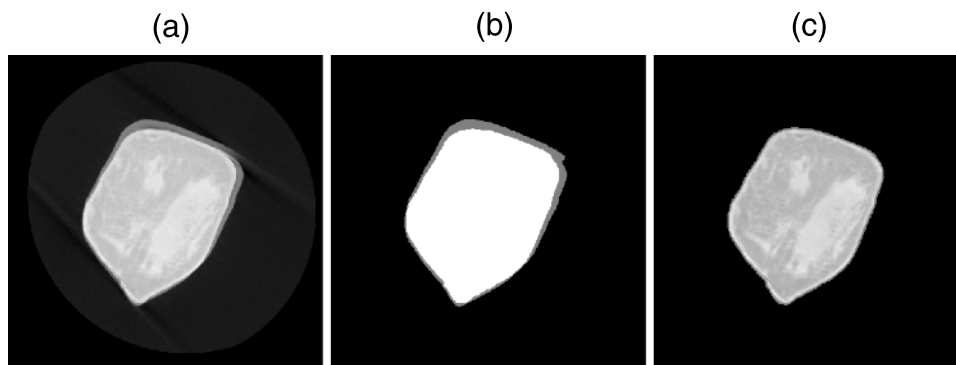
(a)   (b)   (c)

FIG. 1. Breast volume masking. (a) Original bCT coronal slice showing severe motion artifact. (b) BCFCM3D segmentation of (a) depicting three segmentation classes: background (black), breast, and background/breast average (gray). (c) Masked version of (a) with intermediary (blurred) boundary removed.

denoising and contained considerable noise due to quantum mottle. This required a postreconstruction denoising method to improve segmentation quality. A 3D bilateral filter was applied to the reconstructed bCT volumes to preferentially smooth the noise while attempting to preserve the sharpness of tissue boundaries and other anatomical structures,[36]

$$I_{\text{bilateral}}(x,y,z) = \sum_{(x_i,y_j,z_k) \in \Omega} I(x_i,y_j,z_k)$$
$$\times e^{-\left[(x_i-x)^2+(y_j-y)^2+(z_k-z)^2\right]/2\sigma_D^2}$$
$$\times e^{-\left[I(x,y,z)-I(x_i,y_j,z_k)\right]^2/2\sigma_R^2}. \qquad (2)$$

The bilateral filter works by weighting the influence of both Euclidian distance (governed by the σ_D term) and radiometric differences (governed by the σ_R term) between a pixel and its neighbors (within local neighborhood Ω). Doing so increases the amount of denoising (smoothing) in regions with small fluctuations in voxel intensity and decreases the amount of denoising (smoothing) across large intensity gradients (edges). In order to be properly implemented, bCT volume-specific spatial and intensity filtering parameter values must be determined. The intensity parameter is set to the full width at half maximum (FWHM) of the adipose peak of the breast volume histogram. The spatial parameter is set to the average voxel dimension (measured in mm). These parameter values can be multiplied by a scalar before they are implemented to increase or decrease the overall effect of the bilateral filter. Parameter values were manually optimized by one observer on a case-by-case basis in order to maximize noise reduction without compromising the visibility of finer glandular structures.

2.D. Normalization and 3D nonuniformity correction

The cupping artifact in bCT data attributed to scattered radiation is a well-known phenomenon commonly corrected through the use of a second-order polynomial fit and subtraction of the low-frequency artifact from the original data.[37] This correction method was used by Li *et al.* in a slice-by-slice application.[22] However, such a method is not necessarily applicable to nonuniformities which are not well-suited to second-order modeling. For example, breasts with ovular or noncircular coronal cross section will not have circularly symmetric cupping artifact due to scatter. Furthermore, breasts with nonconvex surfaces (e.g., teardrop-shaped breasts) can have even more complex nonuniformity throughout the breast volume. Because of this, a simple second-order polynomial fit is not suitable in general for addressing retrospective (i.e., postreconstruction) adipose inhomogeneity correction.

The current segmentation routine considers adipose inhomogeneity as a 3D artifact with neither specific origin nor form:[38] Unlike other techniques, it does not explicitly depend on breast symmetry to perform the correction.[37] Adipose nonuniformity was corrected in the new protocol using a nonparametric method initially developed for application to the slowly varying, positive, multiplicative bias field in MRI

data with additive noise[38] and later applied by Yang *et al.*[31] to bCT data,

$$H(X) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \frac{\sum_{l=1}^{L} \delta(X_l-i)\delta(M_l-j)}{\sum_{i=1}^{N_1} \sum_{j=2}^{N_2} \sum_{l=1}^{L} \delta(X_l-i)\delta(M_l-j)}$$
$$\times \log\left[\frac{\sum_{l=1}^{L} \delta(X_l-i)\delta(M_l-j)}{\sum_{i=1}^{N_1} \sum_{j=2}^{N_2} \sum_{l=1}^{L} \delta(X_l-i)\delta(M_l-j)}\right]. \qquad (3)$$

The above equation describes the Shannon entropy $H(X)$ of the normalized intensity gradient joint histogram associated with image X containing L pixels and N_1 grey levels. Its gradient magnitude of X is defined by M with N_1 distinct grey levels. δ is the delta function. The entropy function is minimized by modeling multiplicative bias in X using B-spline basis functions $B(\omega)$ of iteratively increasing frequency until no noticeable differences were observed between iterations,

$$E(\omega) = \arg\min_\omega H\{\log[X/B(\omega)]\}. \qquad (4)$$

Because this method expands on the widely used Shannon entropy homogeneity measurement, it was first necessary to convert the intensity values of the bCT data onto a zero to unity range. Flattening the adipose intensity profile using this uniformity correction prepared the bCT volume for fibroglandular segmentation.

2.E. Resampling and BCFCM3D segmentation

The original bCT data were collected and reconstructed into image matrices with pixel size determined by the diameter of the coronal breast cross section, and slice thickness was determined by the length of the breast. Therefore, prior to segmentation, the bCT volumes were resampled with bicubic interpolation to produce isotropic voxels (with minimum original voxel dimension) to ensure the integrity of the segmentation. Once resampled, the BCFCM3D method of Ahmed *et al.* was applied a second time to the bCT volume to segment the breast tissue into background, adipose, fibroglandular, and skin classes.[32] A total of six tissue segmentation classes were defined to account for the partial volume effect of finite pixel resolution yielding multiple classes with different fractional fibroglandular composition. Six classes were chosen due to a trade-off between fine-structure preservation and computation time. With more classes, computation time and the number low-intensity misclassifications increased; with too few classes, fine glandular structures were improperly classified as fat voxels. The adipose and glandular class means were chosen based on histogram analysis of the breast volume data. The adipose and glandular peaks were automatically detected, and five linearly spaced class means between the peaks were chosen as segmentation seeds corresponding to 0%, 25%, 50%, 75%, and 100% glandular density. However, due to the fuzzy nature
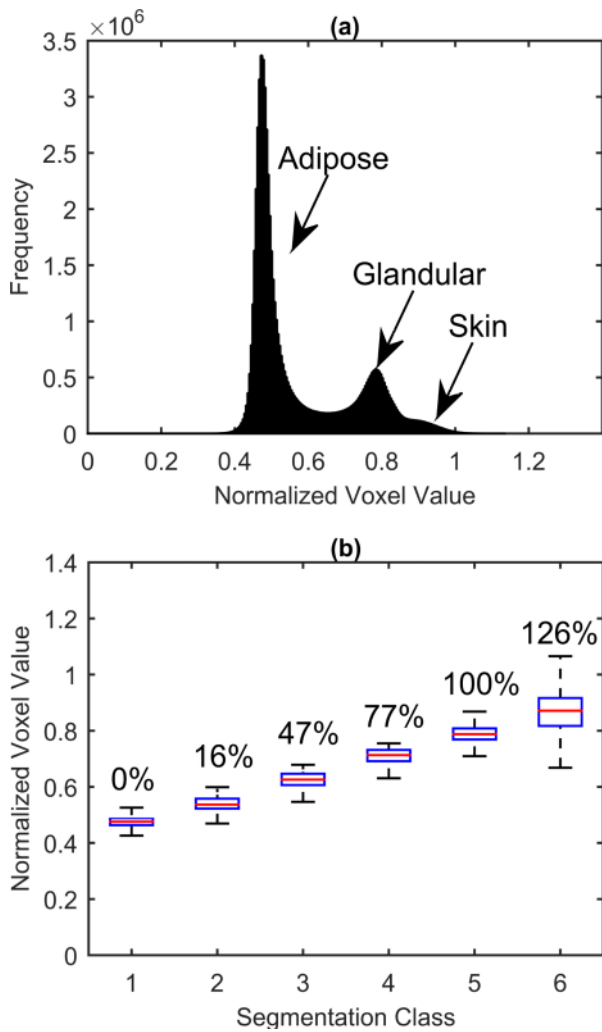
Fig. 2. Normalized voxel intensity distribution and class assignments from sample bCT dataset. (a) Histogram of voxel intensities clearly shows adipose, glandular, and skin peaks. (b) Box-whisker plot of BCFCM3D segmentation of data into six classes ranging from adipose (0% glandular density—class 1) to glandular (100% glandular density—class 5). Note that skin (class 6) tends to be more dense than glandular tissue making its equivalent glandular density >100%.

of the segmentation routine, the actual resultant class means were subject to change as shown in Fig. 2. The skin intensity peak was used to determine the mean of the final (sixth) tissue segmentation class (if and only if the skin peak was determined to be significantly higher than the glandular peak).

2.F. Skin segmentation

The final step in the segmentation algorithm was to accurately locate the internal skin boundary. To maintain the realism of the breast models in this work, it was important to correctly identify the depths at which the glandular and adipose tissues interdigitated with the hypodermal, dermal, and epidermal skin layers. To do this, the external breast boundary was first identified from the perimeter of the breast mask segmentation. Normal vectors were then determined at 50 000–100 000 surface voxels. An iterative approach was

then used to make unit-voxel steps toward the breast interior until an adipose voxel was found. A realistic minimum skin thickness limit was set to 0.5 mm, while the maximum skin thickness limit was set to 5 mm.[39] Skin thickness estimates were smoothed, and a mask or "skin shell" was formed which contained the external skin surface and the internal breast-skin transition region. The skin was then defined as all nonadipose voxels within this skin shell which contained voxels belonging to the skin class (class 6) in addition to a variety of fractional glandular-equivalent densities (classes 2–5) present at the skin-adipose interface. These glandular-equivalent voxels are attributed primarily to the partial volume effect. The use of a skin shell to define the skin preserved the rough interior surface of the skin while still accurately differentiating the skin from glandular structures and blood vessels internal to the breast. Any voxels classified by the BCFCM3D algorithm as skin-class voxels within the breast parenchyma (e.g., due to microcalcifications or very dense glandular tissue) were forced to assume the highest density glandular class.

2.G. Statistical analysis

The output breast segmentations (8-bit unsigned integer type) along with statistics such as breast volume and volume breast density (VBD) were saved to MATLAB data (*.mat) files. Breast volume was computed as the total volume of all nonbackground voxels in coronal slices between the end of the pectoral muscle (if visualized) and the nipple. The reported VBD was computed from these same slices using methods similar to those of Yaffe *et al.*[27]

$$\text{VBD}_{\text{NSk}} = 100\% \times \frac{V_{\text{fg}}}{V_{\text{fg}} + V_{\text{ad}} + V_{\text{sk}}}, \tag{5}$$

$$\text{VBD}_{\text{Sk}} = 100\% \times \frac{V_{\text{fg}} + V_{\text{sk}}}{V_{\text{fg}} + V_{\text{ad}} + V_{\text{sk}}}, \tag{6}$$

where V_{fg} is the volume of fibroglandular tissue, V_{ad} is the volume of adipose tissue, and V_{sk} is the volume of skin. VBD without skin (VBD_{NSk}) is computed as the percentage of total breast volume ($V_{\text{fg}} + V_{\text{ad}} + V_{\text{sk}}$) comprised of fibroglandular tissue, whereas VBD with skin (VBD_{Sk}) is the percentage of total breast volume comprised of fibroglandular tissue or skin. In a departure from the methods of Yaffe *et al.* who estimated V_{fg}, V_{ad}, and V_{sk} from the total number of voxels identified in fibroglandular, adipose, and skin segmentations, respectively,[27] the segmentations in this work contained estimates of the fractional glandular composition of every voxel. Therefore, V_{fg}, V_{ad}, and V_{sk} required the following formulae to estimate the respective volumes:

$$V_{\text{fg}} = V_{\text{voxel}} \sum_i F_{\text{fg},i} N_i, \tag{7}$$

$$V_{\text{ad}} = V_{\text{voxel}} \sum_i F_{\text{ad},i} N_i, \tag{8}$$

$$V_{\text{sk}} = V_{\text{voxel}} N_{\text{sk}} \tag{9}$$

where V_{voxel} is the volume of a single voxel, $F_{\text{fg},i}$ is the fractional glandular composition of segmentation class i, $F_{\text{ad},i}$ is the fractional adipose composition of segmentation class i,
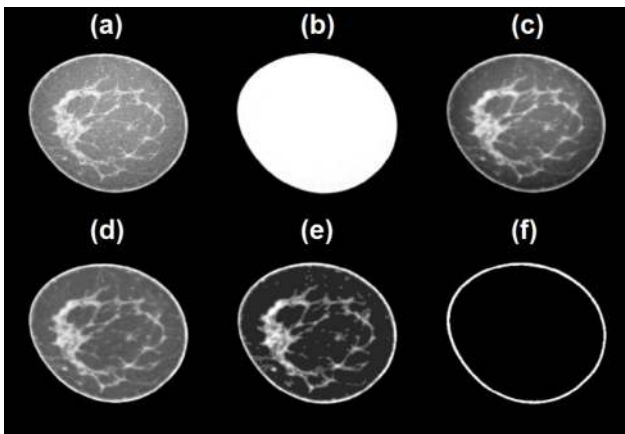
FIG. 3. Visual depiction of bCT segmentation processing. (a) Original coronal slice from breast volume. (b) Breast volume mask. (c) Image (a) after masking and denoising with bilateral filter. (d) Image (c) after uniformity correction. (e) BCFCM3D segmentation of (d). (f) Skin mask.

N_i is the number of voxels belonging to segmentation class i, and N_{sk} is the number of voxels.

2.H. XCAT phantom creation

Segmented volumes were used to create a mathematically defined 3D breast phantom in a format congruent with the 4D XCAT suite of phantoms.[40] Surface meshes were used to represent the boundaries between fibroglandular classes. Since the intensities of the segmentation correspond to the glandular fraction, increasing from adipose (n_1) to skin (n_{skin}), a mesh at the $n_{k+0.5}$ isosurface creates a boundary

between classes n_k and n_{k+1}. The resulting meshes created with the Marching Cubes algorithm[41] had a nested structure, such that the n_{k-1} mesh lies completely within mesh n_k. Prior to creating the meshes, the background was set to an integer value of $n_{skin} + 1$. These meshes are capable of representing complicated geometry of arbitrary topology, which are capable of being deformed to model differing patient positions, different states of mechanical compression, and physics of the imaging process.[42–45]

These types of deformations can be solved for using the finite element method, which discretizes a complex domain into a system of smaller simpler elements. Use of the finite element method to model breast deformations and compression has been applied in several applications including multimodality registration,[46–48] image-guided surgery,[49] and breast augmentation.[50,51] Our finite element model uses approximately 60 000 hexahedral elements with six materials of varying stiffness, which are assigned based on the fibroglandular classes within each element. The compression [performed using FEBio (http://mrl.sci.utah.edu/software/febio)] was defined as a displacement constrained problem with a nonlinear contact analysis between the skin surface and the plates (friction neglected), and gravity applied as a body force. The deformation from the finite element model was then applied to the mesh representation of the phantom. A ray-triangle intersection algorithm can be used in conjunction with the resulting deformed meshes to determine the thickness of each material that was intersected by each ray. For the proof-of-concept mammographic images in Fig. 7, a parallel beam projection was performed with the integer of the fibroglandular classes used as a correlate for attenuation.
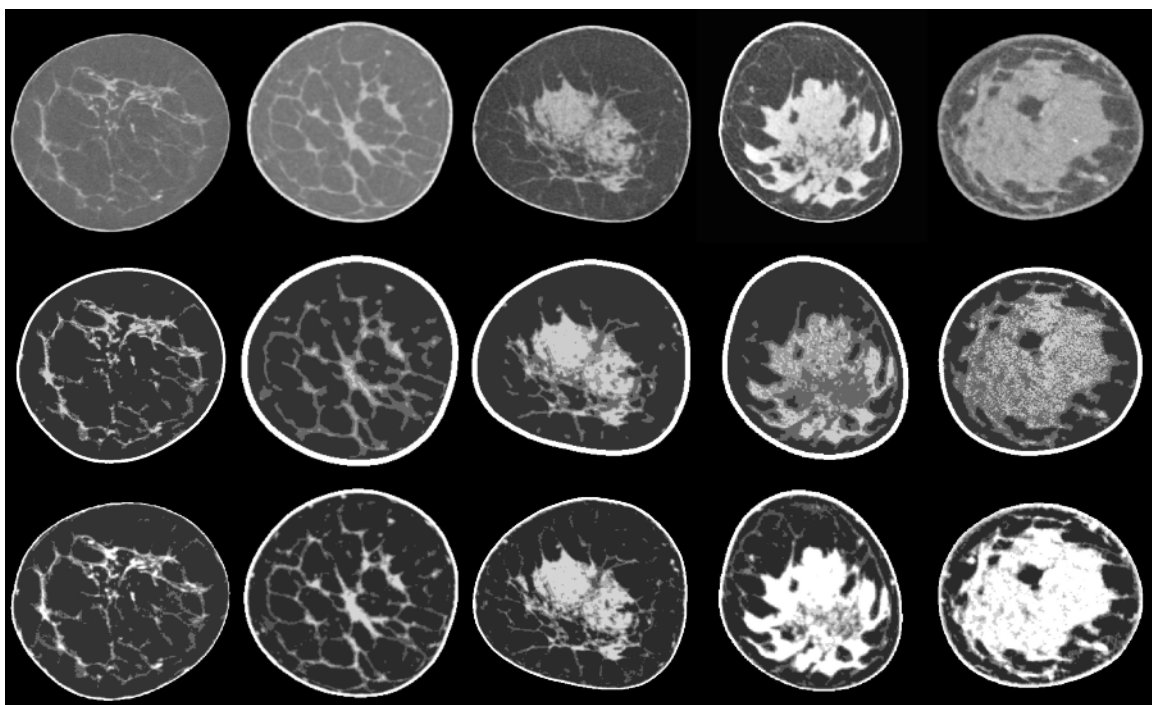

FIG. 4. (Top) Denoised and nonuniformity-corrected coronal bCT images. (Middle) Phantoms produced with the methods of Li *et al.* with five integer density classes. (Bottom) Corresponding phantoms from the current work with six segmentation classes shown as integer grayscale levels.

3. RESULTS

The segmentation algorithm required about 30 min of runtime per case using a 3.4 GHz Intel i7 Processor with 16 GB of RAM running a Windows 7 (64-bit) operating system and MATLAB R2014a. Refer to Fig. 3 for an illustration of each step in the segmentation routine. Additionally, images from five of the 224 segmented bCT volumes are shown in Fig. 4 to demonstrate how closely our segmented volumes resemble the actual patient data sets with a variety of breast sizes, geometries, glandular distributions, and densities. Our results are also shown alongside segmentations made using the methods of Li *et al.*[22] Based on subjective comparisons, the methods proposed in this paper provide better retention of fine fibroglandular detail and more reasonable skin boundary segmentations.

Figure 5 shows the distributions of VBD_{NSk} and VBD_{Sk} measured using our methods. The VBD_{NSk} estimates ranged from 1.4% to 65.3% with a mean value of $18.0\% \pm 13.4\%$ while VBD_{Sk} ranged from 5.5% to 66.3% with a mean value of $25.3\% \pm 13.2\%$. Segmented breast volumes ranged from 25.0 to 2099.6 ml with a mean value of 716.3 ± 386.5 ml. Figure 5 and the mean and range of breast densities can be compared with the results of Yaffe *et al.* who found distributions of VBD_{NSk} to be $14.3\% \pm 10.3\%$ and VBD_{Sk} to be $25.6\% \pm 12.6\%$. Figure 6 demonstrates the distribution of breast model densities versus breast model volume in the newly available library of phantoms.

Figure 7 shows three examples of simulated mammograms after a modeled compression and simple ray-tracing from three different breast phantoms with VBD_{Sk} of 11.8%, 24.7%, and 34.6%. The breasts were compressed to 45% of their original diameter resulting in compressed breast thicknesses of 5.5, 4.9, and 5.4 cm, respectively.
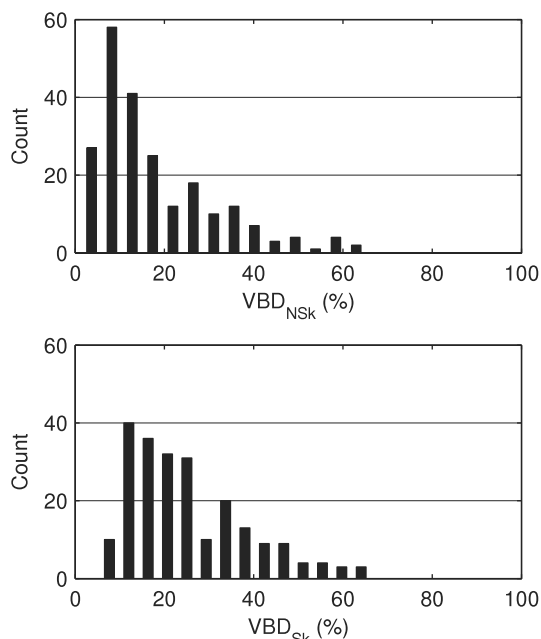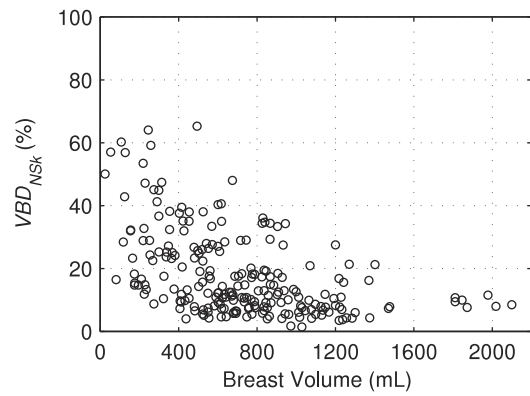
FIG. 6. VBD_{NSk} versus breast volume for 224 bCT subjects.

4. DISCUSSION

This work has produced a large library of anatomically diverse virtual breast phantoms which, because of their anatomic origins, mirror the natural appearance of breast anatomy and its natural variation in a patient population. This contribution compliments other recent work in the growing field of breast modeling and simulation research. While modifications to prior bCT segmentation methods have resulted in the satisfactory production of over 200 virtual breast phantoms, this approach to breast modeling has several benefits and challenges which merit discussion.

Segmentation of the breast external boundary involved methods which accounted for blur and artifacts attributed to the partial volume effect and patient motion during data acquisition. In Sec. 2.B, it was noted that patient motion may affect both the rendition of the external breast boundary and structures internal to the breast, and that no correction of motion artifacts associated with these internal structures would be attempted. As a result, bCT segmentations derived from cases containing substantial patient motion may be subject to a number of issues depending on the nature of the patient motion. For patients whose breast shifts position during acquisition, a "doubling" of the internal structures may occur due to improper alignment and registration of the internal structures during reconstruction. This doubling would affect the appearance of the glandular architecture and the apparent breast density. In almost all cases involving patient motion, some amount of blurring is involved. Blurring of the fibroglandular structures in the bCT data can cause degradation in the delineation between fibroglandular and adipose tissue in addition to potentially altering assessment of breast attributes such as VGF. Therefore, caution should be exercised when dealing with data potentially contaminated by patient motion, and those data subject to severe patient motion should be screened prior to segmentation and model generation.

The segmentation of 224 bCT volumes presented here has expanded on the earlier work of Li *et al.*[22] In the previous work, the segmentation process was optimized and demonstrated on a single, high-quality breast case. In this work, we have made many modifications to the segmentation algorithm allowing for more reliable and efficient application
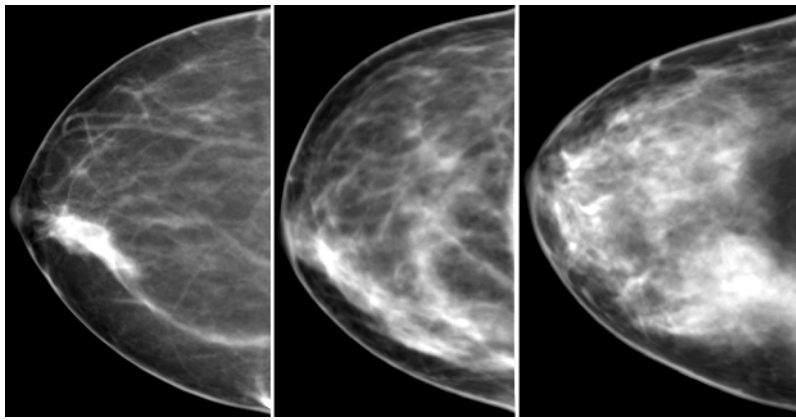
FIG. 5. Volume breast density distributions of the virtual breast phantom library.

Fig. 7. Simulated projection images (with breast compression and ray-tracing). From left to right, VBD$_{Sk}$ are 11.8%, 24.7%, and 34.6%, respectively.

to hundreds of datasets. One advance was the inclusion of a 3D bilateral filter step to accommodate imaging data that had not been denoised prior to reconstruction. The 3D bilateral filter was used to preferentially smooth the noise while attempting to preserve the sharpness of tissue boundaries and anatomical structures. Another major update to the segmentation routine was the use of a 3D bias-field correction to flatten gross adipose nonuniformities within the breast. This is done to provide more consistent tissue classification across the volume. Compared to the second-order polynomial fit method of Altunbas *et al.*,[37] the new method provided a more generalized approach to nonuniformity correction by not relying on geometric assumptions about the general shape of the nonuniformity. Therefore, nonuniformities due to scatter, beam-hardening, breast positioning, and any incomplete sampling in the bCT data could all be corrected in a single step. A final update to the segmentation routine was the complete revision of the tissue classification algorithm. Previously, iterative histogram classification was applied independently to each axial slice of the bCT data.[22] Although this method provided reasonable results in the case of high-quality bCT volumes, it lacked utilization of the 3D voxel neighbor classification information, a characteristic which made the BCFCM3D algorithm more robust, especially when segmenting higher-noise volumes.

Comparing our work to that of Yaffe *et al.* yielded interesting results. Although the two studies draw on bCT datasets from similar populations, there are key differences between VBD assessments. There are several possible explanations as to why our VBD distributions differ. Yaffe *et al.* used only the left breast of 191 individuals within their cohort (to maintain independence of the results); our cohort contained bilateral studies from twenty subjects (40 bCT volumes) and 184 unilateral (left or right) breast data from 184 different subjects. Both left and right breast data were used in our study to increase the total number of phantoms produced. It is also possible that our definition of classes with fractional glandularity has positively biased our assessments of VBD. In comparison of our VBD$_{NSk}$ measurements to those of Yaffe, use of classes with fractional granularity may partially explain why our assessment of VBD$_{NSk}$ is 3.7% higher. On

the other hand, the average of our VBD$_{Sk}$ estimates is only 0.3% lower. This suggests that our methods for segmenting the skin result in lower overall skin density. Although we allowed the skin to assume density which was greater than that of 100% glandular tissue (as Yaffe's methods also did), our BCFCM3D segmentation algorithm routinely classified voxels at the skin-adipose boundary as mixed glandular + adipose bringing down the overall skin density.

There are some notable limitations to our method of creating phantoms from bCT data. The inherent resolution of the current bCT data is not sufficient to resolve some small anatomical structures (e.g., Cooper's ligaments and microcalcifications) that are often of interest to researchers and clinicians. The lack of very fine detail structures in turn can negatively impact the realism of the model. To restore these missing data, the work of Lau *et al.*[52] could be used to add detail to our existing models at the high frequencies absent in the bCT data. Their work was specifically developed to enhance existing anthropomorphic breast models. Other authors have taken different approaches to the challenge of simulating realistic breast textures. Elangovan *et al.*[53] have recently demonstrated how breast structure extracted from reconstructed DBT planes can be combined to produce realistic breast tissue background for observer studies in mammography and DBT. Such an approach is attractive since it takes advantage of the high resolution of a digital flat panel detector combined with the quasi-3D resolution of tomosynthesis. Both are interesting works that could be used to improve the realism of the (high-frequency) detail in our population of breast models.

In the same way that microcalcifications and other fine details might be added to enhance the current bCT phantoms, synthetic lesions may also be added to support the utility of these phantoms in virtual clinical trials. This was one motivation for using lesion-free data to produce the initial population of bCT models. The inclusion of lesions in a "healthy" breast model population could affect the outcomes of future virtual clinical trials. Instead, since we elected to start with lesion-free cases, synthetic lesions with controlled size, contrast, shape, etc., can be inserted at will. There is already considerable research on breast lesion modeling,[13,54–57] and the feasibility of lesion insertion in our breast model has

already been demonstrated in other work suggesting the feasibility of such an approach for virtual clinical trials.[58] There is however a caveat to performing virtual clinical trials with a limited (finite) number of available data sets. Previous studies have shown that thousands of models may be required for sufficient statistical power in virtual imaging studies.[59,60] While mathematical phantoms have the freedom to create an essentially limitless number of phantoms, our phantoms (based on human subject data) are limited by the number of available bCT datasets. While we recognize this as a potential limitation, bCT data are becoming more readily available to researchers and more common to the clinic. Not only has clinical research involving bCT grown[28,61] but the recent FDA approval of the Koning Breast CT system may further increase the availability of clinical cases.[62] Additionally, our group has demonstrated the ability to morph between models in order to create a much larger number of phantoms with user-defined anatomical characteristics.[25]

A second limitation of this segmentation routine is the sensitivity of the algorithm to noise in the bCT data set. As a consequence of using *in vivo* data, acquisitions are dose-limited and thus limited by the resulting noise which increases the risk of voxel misclassification during segmentation. Our response to this limitation was 2-fold—through the use of a 3D bilateral filter to reduce the noise and through the incorporation of a BCFCM3D segmentation method. With this combined approach to handling image noise, the new segmentation paradigm was successfully applied to produce over 200 digital breast models.

A final limitation and consequence of using the reconstructed bCT images for segmentation is the innate dependence on the reconstruction algorithm. Reconstruction algorithm choice can influence CT image resolution, noise, and artifacts; changes to any of these image quality characteristics may affect the quality of the final segmentation. However, it is important to note that the goal of this work is not necessarily to recreate the exact anatomy of each subject, but rather to generate breast phantoms that portray each bCT dataset with a high degree of realism. Therefore, modest deviations from truth may be tolerable without unduly affecting the utility of the phantom.

5. CONCLUSION

This study provides a new population of 224 breast phantoms for imaging research. As these models are based upon actual human subject bCT imaging data, they offer a degree of realism that may improve the clinical relevance of virtual studies. Compared to previous studies based on only a few prototype cases, this new dataset provides a rich source of cases spanning a wide range of breast types, volumes, densities, and parenchymal patterns. As these phantoms are distributed to the research community, they may be used for a variety of purposes, including virtual clinical trials. Ongoing work seeks to improve the realism of these models as the topics of breast imaging and simulation continue to be subjects of major research and clinical interest.

a)D. W. Erickson and J. R. Wells contributed equally to this work.
b)Electronic mail: jered.wells@duke.edu

[1]V. P. Jackson, R. E. Hendrick, S. A. Feig, and D. B. Kopans, "Imaging of the radiographically dense breast," Radiology **188**, 297–301 (1993).

[2]M. J. Yaffe, "What should the burden of proof be for acceptance of a new breast-cancer screening technique?," Lancet **364**, 1111–1112 (2004).

[3]K. Kerlikowske, P. A. Carney, B. Geller, M. T. Mandelson, S. H. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg, and R. Ballard-Barbash, "Performance of screening mammography among women with and without a first-degree relative with breast cancer," Ann. Intern. Med. **133**, 855–863 (2000).

[4]T. M. Kolb, J. Lichy, and J. H. Newhouse, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations," Radiology **225**, 165–175 (2002).

[5]J. A. Baker and J. Y. Lo, "Breast tomosynthesis: State-of-the-art and review of the literature," Acad. Radiol. **18**, 1298–1310 (2011).

[6]J. M. Park, E. A. Franken, Jr., M. Garg, L. L. Fajardo, and L. T. Niklason, "Breast tomosynthesis: Present considerations and future applications," Radiographics **27**, S231–S240 (2007).

[7]R. A. Jong, M. J. Yaffe, M. Skarpathiotakis, R. S. Shumak, N. M. Danjoux, A. Gunesekara, and D. B. Plewes, "Contrast-enhanced digital mammography: Initial clinical experience," Radiology **228**, 842–850 (2003).

[8]J. M. Lewin, P. K. Isaacs, V. Vance, and F. J. Larke, "Dual-energy contrast-enhanced digital subtraction mammography: Feasibility," Radiology **229**, 261–268 (2003).

[9]J. Boone, A. C. Kwan, K. Yang, G. Burkett, K. Lindfors, and T. Nelson, "Computed tomography for imaging the breast," J. Mammary Gland Biol. Neoplasia **11**, 103–111 (2006).

[10]M. P. Tornai, R. L. McKinley, C. N. Bryzmialkiewicz, P. Madhav, S. J. Cutler, D. J. Crotty, J. E. Bowsher, E. Samei, and J. C. E. Floyd, "Design and development of a fully 3D dedicated x-ray computed mammotomography system," Proc. SPIE **5745**, 189–197 (2005).

[11]P. R. Bakic, C. Zhang, and A. D. A. Maidment, "Development and characterization of an anthropomorphic breast software phantom based upon region-growing algorithm," Med. Phys. **38**, 3165–3176 (2011).

[12]K. Bliznakova, S. Suryanarayanan, A. Karellas, and N. Pallikarakis, "Evaluation of an improved algorithm for producing realistic 3D breast software phantoms: Application for mammography," Med. Phys. **37**, 5604–5617 (2010).

[13]B. Chen, J. Shorey, R. S. Saunders, Jr., S. Richard, J. Thompson, L. W. Nolte, and E. Samei, "An anthropomorphic breast model for breast imaging simulation and optimization," Acad. Radiol. **18**, 536–546 (2011).

[14]J. Näppi, P. B. Dean, O. Nevalainen, and S. Toikkanen, "Algorithmic 3D simulation of breast calcifications for digital mammography," Comput. Methods Programs Biomed. **66**, 115–124 (2001).

[15]D. D. Pokrajac, A. D. A. Maidment, and P. R. Bakic, "Optimized generation of high resolution breast anthropomorphic software phantoms," Med. Phys. **39**, 2290–2302 (2012).

[16]P. Taylor, R. Owens, and D. Ingram, "Simulated mammography using synthetic 3D breasts," in *Digital Mammography*, edited by N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning (Springer, Netherlands, 1998), Vol. 13, pp. 283–290.

[17]P. R. Bakic, M. Albert, D. Brzakovic, and A. D. A. Maidment, "Mammogram synthesis using a 3D simulation. I. Breast tissue model and image acquisition simulation," Med. Phys. **29**, 2131–2139 (2002).

[18]P. R. Bakic, M. Albert, D. Brzakovic, and A. D. A. Maidment, "Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture," Med. Phys. **29**, 2140–2151 (2002).

[19]P. R. Bakic, M. Albert, D. Brzakovic, and A. D. A. Maidment, "Mammogram synthesis using a three-dimensional simulation. III. Modeling and evaluation of the breast ductal network," Med. Phys. **30**, 1914–1925 (2003).

[20]C. Hoeschen, U. Fill, M. Zankl, W. Panzer, D. Regulla, and W. Döhring, "A high-resolution voxel phantom of the breast for dose calculations in mammography," Radiat. Prot. Dosim. **114**, 406–409 (2005).

[21]J. M. O'Connor, M. Das, C. S. Dider, M. Mahd, and S. J. Glick, "Generation of voxelized breast phantoms from surgical mastectomy specimens," Med. Phys. **40**, 041915 (12pp.) (2013).

[22]C. M. Li, W. P. Segars, G. D. Tourassi, J. M. Boone, and J. T. Dobbins III, "Methodology for generating a 3D computerized breast phantom from empirical data," Med. Phys. **36**, 3122–3131 (2009).

[23]W. P. Segars, A. I. Veress, J. R. Wells, G. M. Sturgeon, N. Kiarashi, J. Y. Lo, E. Samei, and J. T. Dobbins, "Population of 100 realistic, patient-based computerized breast phantoms for multi-modality imaging research," Proc. SPIE **9033**, 90331X–90336X (2014).

[24]W. P. Segars, Ph.D. dissertation, University of North Carolina, 2001.

[25]C. M. L. Hsu, M. L. Palmeri, W. P. Segars, A. I. Veress, and J. T. Dobbins III, "Generation of a suite of 3D computer-generated breast phantoms from a limited set of human subject data," Med. Phys. **40**, 043703 (11pp.) (2013).

[26]S.-Y. Huang, J. M. Boone, K. Yang, N. J. Packard, S. E. McKenney, N. D. Prionas, K. K. Lindfors, and M. J. Yaffe, "The characterization of breast anatomical metrics using dedicated breast CT," Med. Phys. **38**, 2180–2191 (2011).

[27]M. J. Yaffe, J. M. Boone, N. Packard, O. Alonzo-Proulx, S. Y. Huang, C. L. Peressotti, A. Al-Mayah, and K. Brock, "The myth of the 50-50 breast," Med. Phys. **36**, 5437–5443 (2009).

[28]K. K. Lindfors, J. M. Boone, T. R. Nelson, K. Yang, A. L. C. Kwan, and D. F. Miller, "Dedicated breast CT: Initial clinical experience," Radiology **246**, 725–733 (2008).

[29]Z. Chen, "Histogram partition and interval thresholding for volumetric breast tissue segmentation," Comput. Med. Imaging Graphics **32**, 1–10 (2008).

[30]T. R. Nelson, L. I. Cervino, J. M. Boone, and K. K. Lindfors, "Classification of breast computed tomography data," Med. Phys. **35**, 1078–1086 (2008).

[31]X. Yang, S. Wu, I. Sechopoulos, and B. Fei, "Cupping artifact correction and automated classification for high-resolution dedicated breast CT images," Med. Phys. **39**, 6397–6406 (2012).

[32]M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," IEEE Trans. Med. Imaging **21**, 193–199 (2002).

[33]J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition* (IEEE, New York, NY, 1992).

[34]D. R. Nayak, "Image smoothing using fuzzy morphology," Int. J. Comput. Appl. **1**, 43–48 (2012).

[35]J. Q. Xia, J. Y. Lo, K. Yang, C. E. Floyd, Jr., and J. M. Boone, "Dedicated breast computed tomography: Volume image denoising via a partial-diffusion equation based technique," Med. Phys. **35**, 1950–1958 (2008).

[36]C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *International Conference on Computer Vision* (IEEE, Bombay, 1998), Vol. 6, pp. 839–846.

[37]M. C. Altunbas, C. C. Shaw, L. Chen, C. Lai, X. Liu, T. Han, and T. Wang, "A post-reconstruction method to correct cupping artifacts in cone beam breast computed tomography," Med. Phys. **34**, 3109–3118 (2007).

[38]J. V. Manjón, J. J. Lull, J. Carbonell-Caballero, G. García-Martí, L. Martí-Bonmatí, and M. Robles, "A nonparametric MRI inhomogeneity correction method," Med. Image Anal. **11**, 336–345 (2007).

[39]L. Shi, S. Vedantham, A. Karellas, and A. M. O'Connell, "Technical note: Skin thickness measurements using high-resolution flat-panel cone-beam dedicated breast CT," Med. Phys. **40**, 031913 (6pp.) (2013).

[40]W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," Med. Phys. **37**, 4902–4915 (2010).

[41]W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," ACM SIGGRAPH Comput. Graphics **21**, 163–169 (1987).

[42]X. Li, E. Samei, T. Yoshizumi, J. G. Colsher, R. P. Jones, and D. P. Frush, "Experimental benchmarking of a Monte Carlo dose simulation code for pediatric CT," Proc. SPIE **6510**, 65102A–65110A (2007).

[43]R. S. Saunders, Jr. and E. Samei, "A Monte Carlo investigation on the impact of scattered radiation on mammographic resolution and noise," Proc. SPIE **6142**, 61423A–61427A (2006).

[44]W. P. Segars, M. Mahesh, T. Beck, E. C. Frey, and B. M. W. Tsui, "Validation of the 4D NCAT simulation tools for use in high-resolution x-ray CT research," Proc. SPIE **5745**, 828–834 (2005).

[45]W. P. Segars, B. M. W. Tsui, E. C. Frey, and E. K. Fishman, "Extension of the 4D NCAT phantom to dynamic x-ray CT simulation," in *IEEE Nuclear Science Symposium, Conference Record* (IEEE, Portland, 2003), Vol. 5, pp. 3195–3199.

[46]L. Han, J. Hipwell, Z. Taylor, C. Tanner, S. Ourselin, and D. Hawkes, "Fast deformation simulation of breasts using GPU-based dynamic explicit finite element method," in *Digital Mammography*, edited by J. Martí, A. Oliver, J. Freixenet, and R. Martí (Springer, Berlin, Heidelberg, 2010), Vol. 6136, pp. 728–735.

[47]T. Hopp, M. Dietzel, P. A. Baltzer, P. Kreisel, W. A. Kaiser, H. Gemmeke, and N. V. Ruiter, "Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization," Med. Image Anal. **17**, 209–218 (2013).

[48]C. Tanner, J. A. Schnabel, D. L. G. Hill, D. J. Hawkes, M. O. Leach, and D. R. Hose, "Factors influencing the accuracy of biomechanical breast models," Med. Phys. **33**, 1758–1769 (2006).

[49]L. Han, J. Hipwell, T. Mertzanidou, T. Carter, M. Modat, S. Ourselin, and D. Hawkes, "A hybrid fem-based method for aligning prone and supine images for image guided breast surgery," in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, Chicago, IL, 2011), pp. 1239–1242.

[50]A. P. del palomar, B. Calvo, J. Herrero, J. López, and M. Doblaré, "A finite element model to accurately predict real deformations of the breast," Med. Eng. Phys. **30**, 1089–1097 (2008).

[51]L. Roose, W. De Maerteleire, W. Mollemans, and P. Suetens, "Validation of different soft tissue simulation methods for breast augmentation," Int. Congr. Ser. **1281**, 485–490 (2005).

[52]B. A. Lau, I. Reiser, R. M. Nishikawa, and P. R. Bakic, "A statistically defined anthropomorphic software breast phantom," Med. Phys. **39**, 3375–3385 (2012).

[53]P. Elangovan, A. Rashidnasab, A. Mackenzie, D. R. Dance, K. C. Young, H. Bosmans, W. P. Segars, and K. Wells, "Performance comparison of breast imaging modalities using a 4AFC human observer study," Proc. SPIE **9412**, 94121T–94127T (2015).

[54]K. Bliznakova, Z. Bliznakov, V. Bravou, Z. Kolitsi, and N. Pallikarakis, "A three-dimensional breast software phantom for mammography simulation," Phys. Med. Biol. **48**, 3699–3719 (2003).

[55]L. de Sisternes, J. G. Brankov, A. M. Zysk, R. A. Schmidt, R. M. Nishikawa, and M. N. Wernick, "A computational model to generate simulated three-dimensional breast masses," Med. Phys. **42**, 1098–1118 (2015).

[56]A. Rashidnasab, P. Elangovan, M. Yip, O. Diaz, D. R. Dance, K. C. Young, and K. Wells, "Simulation and assessment of realistic breast lesions using fractal growth models," Phys. Med. Biol. **58**, 5613–5627 (2013).

[57]E. Shaheen, F. De Keyzer, H. Bosmans, D. R. Dance, K. C. Young, and C. V. Ongeval, "The simulation of 3D mass models in 2D digital mammography and breast tomosynthesis," Med. Phys. **41**, 081913 (17pp.) (2014).

[58]A. Rashidnasab, P. Elangovan, A. Mackenzie, D. R. Dance, K. C. Young, H. Bosmans, and K. Wells, "Virtual clinical trials using inserted pathology in clinical images: Investigation of assumptions for local glandularity and noise," Proc. SPIE **9412**, 94122D–94127D (2015).

[59]K. Fukunaga and R. R. Hayes, "Effects of sample size in classifier design," IEEE Trans. Pattern Anal. Mach. Intell. **11**, 873–885 (1989).

[60]S. Young, P. R. Bakic, K. J. Myers, R. J. Jennings, and S. Park, "A virtual trial framework for quantifying the detectability of masses in breast tomosynthesis projection data," Med. Phys. **40**, 051914 (15pp.) (2013).

[61]N. D. Prionas, K. K. Lindfors, S. Ray, S.-Y. Huang, L. A. Beckett, W. L. Monsky, and J. M. Boone, "Contrast-enhanced dedicated breast CT: Initial clinical experience," Radiology **256**, 714–723 (2010).

[62]A. M. O'Connell and D. Kawakyu-O'Connor, "Dedicated cone-beam breast computed tomography and diagnostic mammography: Comparison of radiation dose, patient comfort, and qualitative review of imaging findings in BI-RADS 4 and 5 lesions," J. Clin. Imaging Sci. **2**, 1–8 (2012).