Single x-ray absorptiometry method for the quantitative mammographic measure of fibroglandular tissue volume

Sergei Malkov and Jeff Wang

Department of Radiology and Biomedical Imaging, University of California—San Francisco, San Francisco, California 94143

Karla Kerlikowske

Departments of Medicine and Epidemiology and Biostatistics, University of California—San Francisco, San Francisco, California 94143

Steven R. Cummings

Research Institute, California Pacific Medical Center, San Francisco, California 94143

John A. Shepherd^{a)}

Department of Radiology and Biomedical Imaging, University of California—San Francisco, San Francisco, California 94143

(Received 18 March 2009; revised 4 October 2009; accepted for publication 5 October 2009; published 6 November 2009)

Purpose: This study describes the design and characteristics of a highly accurate, precise, and automated single-energy method to quantify percent fibroglandular tissue volume (%FGV) and fibroglandular tissue volume (FGV) using digital screening mammography.

Methods: The method uses a breast tissue-equivalent phantom in the unused portion of the mammogram as a reference to estimate breast composition. The phantom is used to calculate breast thickness and composition for each image regardless of x-ray technique or the presence of paddle tilt. The phantom adheres to the top of the mammographic compression paddle and stays in place for both craniocaudal and mediolateral oblique screening views. We describe the automated method to identify the phantom and paddle orientation with a three-dimensional reconstruction least-squares technique. A series of test phantoms, with a breast thickness range of 0.5–8 cm and a %FGV of 0%–100%, were made to test the accuracy and precision of the technique.

Results: Using test phantoms, the estimated repeatability standard deviation equaled 2%, with a $\pm 2\%$ accuracy for the entire thickness and density ranges. Without correction, paddle tilt was found to create large errors in the measured density values of up to 7%/mm difference from actual breast thickness. This new density measurement is stable over time, with no significant drifts in calibration noted during a four-month period. Comparisons of %FGV to mammographic percent density and left to right breast %FGV were highly correlated ($r=0.83$ and 0.94 , respectively).

Conclusions: An automated method for quantifying fibroglandular tissue volume has been developed. It exhibited good accuracy and precision for a broad range of breast thicknesses, paddle tilt angles, and %FGV values. Clinical testing showed high correlation to mammographic density and between left and right breasts. © 2009 American Association of Physicists in Medicine.
[DOI: [10.1118/1.3253972](https://doi.org/10.1118/1.3253972)]

Key words: FFDM, breast density, risk factors, digital mammography, mammographic density

I. INTRODUCTION

Breast density, which has been studied extensively in large prospective and case-control studies, has been shown to be one of the strongest predictors of breast cancer risk.^{1–5} A high percentage of dense tissue on mammograms is associated with a four to sixfold increased risk of developing breast cancer.¹ Initially, breast density was graded using a semi-quantitative subjective scale known as Wolfe scores.⁶ Wolfe's four classifications took into account the amount (radiopacity) and nature (prominence of ductal structures, hyperplasia) of the density. Other scoring systems have been developed that divide density into categories.^{7,8,1} Mammographic percent density (PD), a more quantitative approach used extensively in clinical research studies, is the ratio of

dense breast area delineated in a mammogram to the total breast area.⁹ These methods are not calibrated measures of fibroglandular tissue volume and thus subject to inter-reader variations. The scoring and threshold methods are subjective and dependent on the training and skill of the reader and thus have limited reproducibility and accuracy. MRI and CT can create accurate 3D images of the dense breast volume and have been used to quantify breast density.^{10–13} However, these modalities are generally used for diagnostic imaging because the exams are more expensive and, in the case of CT, have a higher dose than screening mammography. Thus, there are advantages to measuring breast density in a mammography environment. A full review of methodologies was also recently published.¹⁴

Calibrating image grayscale value to determine volume

and composition of the tissue that created it is difficult. If the x-ray technique, detector characteristics, mammography system geometry, and breast thickness are explicitly known, then image grayscale value can be converted to tissue volume measure. However, in clinical environments, these parameters are rarely controlled. Measurement of tissue thickness can be affected by poorly defined tilt and warp of the compression paddle. Calibration can be complicated for mammography systems since target materials, filters, and x-ray tube voltages are all adjustable by the user. Thus, the clinical measure of breast density should account and adjust for these image-to-image differences.

We describe a method to quantify breast density as either a percent fibroglandular tissue volume (%FGV) or as a fibroglandular tissue volume (FGV) using single-energy x-ray absorptiometry (SXA) for mammography.

SXA is designed to be objective and can be used across different digital mammography systems. The SXA phantom included in the x-ray field during a mammography examination is used to convert pixel grayscale values into the unique volume of adipose and fibroglandular tissue for every given pixel value. A previous work¹⁵ used an in-image biwedge phantom that created a compressible thickness between the compression surfaces to image two reference materials with

the same thickness as the breast. This earlier phantom was successfully used on more than 15 000 film screening exams.¹⁶ However, the biwedge phantom had several limitations including mechanical fatigue of moving parts, removal for mediolateral oblique (MLO)-view images, and limited compression range ability. Our new SXA phantom has the following characteristics: Provides reference attenuations and accurate measurement of breast thickness and paddle tilt; possesses no moving parts; does not obstruct breast tissue; does not require a special imaging technique; and does not interfere with the technologist's positioning of the patient's breast. We present details of the phantom design, algorithms used to quantify fibroglandular tissue volumes, and validation studies for accuracy and precision.

II. MATERIALS AND METHODS

The breast is modeled as a tissue mass of two materials, adipose and fibroglandular tissue. Through breast thickness mapping and examining attenuation of tissue references at various breast thickness areas %FGV can be expressed as a sum of percentage fibroglandular tissue volumes in the pixel at (x,y) location, $\%FGV(x,y)$ weighed by their volumes $V(x,y)$, as shown previously¹⁵

$$\%FGV = \left(\sum_{(x,y) \in \text{Breast}} \%FGV(x,y) \cdot V(x,y) \right) / \left(\sum_{(x,y) \in \text{Breast}} V(x,y) \right). \quad (1)$$

The FGV is calculated by sum of the products $\%FGV(x,y)$ and $V(x,y)$ over the breast

$$FGV = \left(\sum_{(x,y) \in \text{Breast}} \%FGV(x,y) \cdot V(x,y) \right). \quad (2)$$

The SXA phantom provides multiple references for known steps in thickness. This enables derivation of adipose and fibroglandular tissue reference values for the breast thickness obtained from the phantom height and tilt angle. The SXA phantom is measured at the same time and under the same x-ray conditions as the breast and can be used to measure both %FGV and FGV in cubic centimeter.

II.A. SXA phantom design

The current multistep SXA phantom was constructed by the investigators and made of proprietary material from Computerized Imaging Reference Systems, Inc. (CIRS, Inc., Norfolk, VA) equivalent to 80% FGV. The SXA phantom contains nine lead-positioning markers around the base, middle, and top, as illustrated in Fig. 1. Figure 2 illustrates the three-dimensional (3D) geometry and projected mammographic image with the phantom attached to the top of the compression paddle. The SXA phantom is placed on top of the paddle for several reasons. The lead markers can then be

used to calculate the exact compression paddle height and tilt, and unlike our previous phantom design between the paddles,¹⁵ the phantom does not interfere with the technologist's positioning of the patient's breast. The phantom is stepped to provide a compositional range of 1–7 cm in thickness. The height and orientation of the phantom's base is the height and orientation (tilt) of the compression paddle.

FIG. 1. The phantom is imaged in the unused portion of each mammogram.

FIG. 2. Image of the geometry used to find breast thickness and paddle tilt.

Therefore, by utilizing the phantom's position and orientation, we can model the thickness of the breast at all imaged locations.

The phantom is positioned so that it projects into the unused corner of the mammography image. This position was chosen by superimposing left and right craniocaudal (CC) and MLO-view mammograms from 350 patients such that it would reveal areas where breast tissue was unlikely to lie ($<1\%$ probability). Because the phantom is small and has no moving parts, it stays in place for both left and right CC and MLO views without interfering with breast projections. The appearance of the phantom in all four screening views, as would be seen by a radiologist, is shown in Fig. 3. To minimize the phantom's projected footprint on the image, we designed the phantom with a bias angle lean to match the x-ray incident angle for a paddle height of 4.4 cm. The edges of the phantom are flared to broaden the base and increase the valid region of interest (ROI) size.

II.B. 3D phantom locating procedure

To mathematically describe the position of the phantom anywhere in coordinate space of the x-ray gantry (world-space w), the position of the phantom's origin $(x,y,z)_w$ and the orientation of the phantom in terms of rotation angles between the world and local phantom axes $(\alpha,\beta,\theta)_w$ are needed. We first define the detector plane as $z=0$, $(x,y,0)_w$, and the world origin $(0,0,0)_w$ as the point where a line drawn between the detector and the focal spot "normally" intersects the detector plane. A system's x-ray focal spot s cm away from the detector would be at $(0,0,s)_w$, as depicted in Fig. 2. In addition, we define a phantom coordinate system with its origin $(0,0,0)_p$ on the phantom base. The coordinates of the phantom lead markers in the world system of coordinates P_w can be expressed mathematically as

FIG. 3. The phantom appearing in all four screening views, as would be seen from a reading workstation.

$$(P_w) = [R_0]^T(P_p) + [T], \quad (3)$$

where $P_p = (x,y,z)_p$ represents the fixed-point coordinates of the lead markers in the phantom space, $[R_0]$ is the 3D rotational matrix containing the angle difference between world and phantom axes, and $[T]$ is the translation matrix between the origins of the world and phantom space. In the world geometry, the coordinates of a lead marker, such as marker 1, projected onto the detector plane, $P^1_{\text{proj}} = (x^1_{\text{proj}}, y^1_{\text{proj}}, 0)$, can be calculated as

$$x^1_{\text{proj}} = \frac{s * P^1_{w,x}}{(s - P^1_{w,z})} \quad \text{and} \quad (4)$$

$$y^1_{\text{proj}} = \frac{s * P^1_{w,y}}{(s - P^1_{w,z})}, \quad (5)$$

where $P^1_{w,x}$, $P^1_{w,y}$, and $P^1_{w,z}$ are the coordinates of the SXA phantom lead marker 1 in the world system of coordinates. Equations (4) and (5) can be used to determine the true phantom position in space by creating an error function, Eq. (6), of the sum of the squared differences of the projected values to the actual image values for each lead marker

$$\chi^2 = \sum_{i=1}^9 (x^i_{\text{proj}} - x^i_{\text{img}})^2 + (y^i_{\text{proj}} - y^i_{\text{img}})^2, \quad (6)$$

where x^i_{img} and y^i_{img} are image positions of the SXA phantom projected lead markers in relation to the world origin $(0,0,0)_w$.

The position and orientation $(x,y,z,\alpha,\beta,\theta)_w$ that minimizes Eq. (4) are the best estimates. Actually, three parameters are used: tz — z -component or distance from the origin

of phantom system coordinates to the detector plane, rx —tilt angle parallel to chest, and ry —chest-nipple tilt angle.

To find the location of world-space origin coordinates on the detector plane and x-ray source position s , we use a custom phantom consisting of an acrylic rectangular plate 5 cm in thickness with eleven lead sphere pairs attached to each side. The imaged position of world-space origin $O_1 = (0, 0, 0)_w$ was obtained as the interception point of lines defined by the paired projections, as described by Zhang *et al.*¹⁷ We then used the world origin to determine the value of s using the geometry of similar triangles. Once the origin coordinates and source position are determined all calibration procedures are then based on these parameters.

The individual locations of the lead spheres in phantom space were measured. The lead spheres' x-ray image locations $(x_{\text{img}}^i, y_{\text{img}}^i)$ are identified on the image using a local threshold technique in the area of the phantom coupled with a centroid-locating algorithm. We use an unconstrained non-linear optimization with a simplex search method to converge on the exact position and rotation. If the compression paddle were a plane, the breast thickness over the area in contact with the paddle would be explicitly known. However, the paddle always contains some curvature that may not be assumed to be negligible. The breast periphery that is not in contact with the two compression surfaces also has to be addressed separately.

II.C. SXA phantom attenuation values

There are nine reference thicknesses on the phantom, and x-ray attenuation is needed for each thickness in order to generate the image-specific calibration. We create a 3D analysis template that consists of nine square ROIs (ROI_1 – ROI_9), which are the overlapping areas (phantom space coordinates) between the top of the phantom's reference steps and the area of the phantom base (ROI_0). Using the matrices $[R_0]$ and $[T]$, the ten ROIs are projected onto the image. Valid reference attenuations for each step are defined as the region where the x-ray paths travel through the full thickness of the step, that is, passing through the top step ROI and base without passing through another step's ROI. Overlap of the ROIs is caused by parallax in the x-ray projections and is exacerbated by paddle tilt and breast thickness, both being unique for each mammogram. Valid ROIs are shown graphically as black regions in Fig. 4.

II.D. Defining the breast volume and thickness

To determine breast volume, breast thickness was estimated for all pixels either in contact with both compression surfaces and those out of contact in the periphery. To capture paddle tilt and to approximate paddle warpage under compression, “folded book” geometry was used. The top compression paddle was represented as two planes (“book covers”) that intersected approximately along center of the breast image (“book spine”) that corresponded to the line of the breast from the chest wall to the nipple. These two planes are defined in the 3D space relative to the bottom compression surface using the six SXA phantom position and orien-

FIG. 4. Phantom image with valid ROIs shown as the numbered dark areas. The stars represent 3D-reconstructed lead marker locations.

tation $(x, y, z, \alpha, \beta, \theta)$ as described in Sec. II B. Distance between the compression surfaces for the entire image area could then be estimated. The true breast edge in 3D space, where the projected breast thickness in the image goes to zero, was defined to lie at the midpoint between the two compression surfaces. Magnification was taken into consideration to determine how the projected breast edge in the mammogram back projected to this midpoint. Breast shape in the periphery region was approximated as having a semi-circular curvature in cross section. The part of the edge containing the nipple was removed and replaced with a smooth polynomial function fit to the surrounding. Normal bisectors at the midpoints were then defined for all pixels along the true breast edge. The periphery volume was defined by extending these bisectors inward to a point where their lengths were half the breast thickness. The bisector lengths varied along the breast edge because of paddle tilt and warp. The curve connecting these points characterized the inner breast periphery boundary where the breast was modeled as being in full contact with the paddles. The half-circle cross section function was then defined using the periphery region bisector lengths as the radii. The total breast volume was defined as the sum of the individual pixel volumes (pixel area $\times$ thickness). The %FGV in all pixels was found by comparing the breast pixel grayscale values to the SXA phantom for the corresponding thickness. Examples of estimated thickness and %FGV images are shown in Fig. 5.

II.E. Calibration procedure

The %FGV calibration was accomplished by scanning the calibration phantoms of known densities and thicknesses with the SXA phantom in the same mammogram. Then empirical equations from their analysis were derived and two coefficient lookup tables were created.

II.E.1. Description of calibration phantoms and measurements

Two novel phantoms, the seven-step Density Phantom 7 (DSP7) and the smaller three-step density phantom (DSP3), were constructed to calibrate the method and individual SXA phantoms. The DSP7 is shown in Fig. 6 (center) and has

FIG. 5. (a) A mammogram where the gray scale is calibrated to %FGV. (b) The modeled breast thickness from the same mammogram using the SXA method.

constant thickness %FGV steps at 0%, 30%, 45%, 50%, 60%, 70%, and 100%. The DSP7 was designed by the investigators and manufactured by CIRS Inc. by mixing two polymers that are equivalent to adipose and fibroglandular tissue, as defined by Hammerstein *et al.*¹⁸ We made three versions of the DSP7 (of 1, 2, and 4 cm thicknesses) which can be combined to create references for densities ranging from 1 to 8 cm. The DSP3 (Fig. 6, right) with thicknesses ranging from 0.2 to 1 cm in 0.2 cm steps in 0%, 50%, and 100% %FGV was used to calibrate the breast periphery. These two phantoms covering thicknesses from 0.2 to 8 cm were imaged with x-ray techniques from 24 to 35 kVp and 20 to 250 mAs using a dedicated research Hologic Selenia (Hologic Inc., Bedford, MA). The unprocessed raw images (PresentationIntentType=“FOR PROCESSING” in the DICOM header) were used for all calculations. All experimental attenuation images were calculated by log formula normalized to 100 mAs conditions

$$V = -\log((I - I_d) * 100 / (I_o * mAs)), \quad (7)$$

where I_d is the pixel signal of breast position ROI measured by blocking the source with a 5 mm lead plate; I_o is the pixel signal of a blank image at 100 mAs; I is the pixel signal recorded at a particular thickness, density, and x-ray technique; and V is the attenuation.

II.E.2. Factors that influence calibration

The initial calibration measurements demonstrated that although the calibration and SXA phantoms are made of the same materials, there are differences in attenuation values for equal densities and thicknesses. These differences are originated from difference of their sizes and locations and could be caused by flat-field nonuniformity, beam hardening, and

FIG. 6. Three phantoms used to for quality control and validation of the SXA method. (Left) TILT phantom, (center) DSP7, (right) DSP3.

x-ray scattering. Beam hardening effects on the SXA measures were negligible since the SXA phantom was impacted the same way as the DSP7 phantom and thus the breast. To test the influence of flat-field nonuniformity, we imaged a set of uniform-thickness, adipose-equivalent, precision-ground slabs ranging from 1 to 8 cm in thickness, made by CIRS Inc. Each slab was certified by CIRS Inc. to have uniform attenuation and thickness properties across the entire surface. No corrections to the attenuations were necessary for the 4 cm thick measures. This was expected since the Selenia has a weekly flat-field quality control procedure that uses a 4 cm acrylic slab imaged at six different kVp and mAs values. However, small deviations less than 1% were observed for thicker slabs. The largest deviations to the flat uniformity were for thicknesses less than 2 cm where corner pixel values deviated from the center values by approximately 5% for a 1 cm thick slab. We also observed the attenuation profile along the highest step of the SXA phantom is not flat, and there is decreasing attenuation toward the edge direction.

II.E.3. Description of calibration procedure

Differences in attenuation values for equal phantom thicknesses between calibration and SXA phantoms exists, and adipose reference values are needed to complement the SXA phantom values. Thus, lookup tables of two coefficients were created. During mammogram processing, these two coefficients are used to find two references and eventually the density of each breast thickness. These two coefficients happen to be slightly dependent on kVp and mAs, but the dependence on thickness is dominant. The first correction coefficient

$$k_{\text{lean}} = V_{D80} / V_{80} \quad (8)$$

was used to obtain the fibroglandular reference. It was calculated by matching the attenuation measured directly from the SXA phantom (V_{80}) to the DSP7 phantom 80% step attenuation (V_{D80}). The dependence of attenuation signal [Eq. (7)] of V_{D80} and adipose component (V_{D0}) of the DSP7 phantom on thickness (H) in linear approximation can be written as

$$V_{D80}(H) = m_{80} * H + V_0, \quad (9)$$

$$V_{D0}(H) = m_0 * H + V_0. \quad (10)$$

The second coefficient was introduced as $k_m = m_{80} / m_0$, where m_{80} and m_0 are coefficients in Eqs. (9) and (10), representing attenuations of 80% and 0% fibroglandular tissue volumes, respectively. The k_m was used to derive the adipose reference. As the fibroglandular 80% and 0% bars are located at the same image, V_0 in these equations are equal and related to the log of incident x-ray intensity. Thus, by subtracting and regrouping we can derive coefficient k_m

$$k_m = m_{80} * H / (m_{80} * H - V_{D80} + V_{D0}). \quad (11)$$

However, we cannot derive m_{80} from the SXA phantom because its step attenuation's thickness dependence is not lin-

FIG. 7. Calibration image showing measured fibroglandular reference line, derived fat references as dependent on the step height. Quadratic fits plotted as lines and breast region pixel frequencies plotted as contours.

ear, but well described by a second order polynomial equation

$$V_{80}(H) = C_1 * H^2 + C_2 * H + C_3. \quad (12)$$

To take this dependence into account, we used an empirical equation for obtaining k_m by substituting the m_{80} with the rate of attenuation change with thickness $2 * C_1 * H + C_2$

$$k_m = k_{lean} * (2 * C_1 * H^2 + C_2 * H) / (k_{lean} * (2 * C_1 * H^2 + C_2 * H) - V_{D80} + V_{D0}). \quad (13)$$

Thus, using Eqs. (8) and (13), we calculated k_{lean} and k_m coefficients for all screening condition combinations to match the adipose and 80% fibroglandular reference values of the DSP7 and SXA phantoms. We also created the lookup table as the final step of the calibration procedure. For real breast and QC measurements, the fibroglandular reference can be obtained from Eq. (8) and the adipose reference from the following equation:

$$V_{D0} = k_{lean} * (V_{80} + (1/k_m - 1) * (2 * C_1 * H^2 + C_2 * H)), \quad (14)$$

by using known $k_{lean}(H, kVp)$ and $k_m(H, kVp)$ from the lookup table, and C_1 and C_2 coefficients from the SXA phantom quadratic fitting and reconstructed thickness map (H). Figure 7 illustrates the thickness dependence of adipose and fibroglandular reference points and their quadratic fitting curves. To calculate the proper reference values, we first correct the experimental pixel values read out from nine steps by k_{lean} coefficients. Corresponding adipose values are derived from these values based on the equation using k_m calibration coefficients. Next, fibroglandular reference values are calculated by normalizing to 100%. Figure 7 is a graphical representation of the pixel data for a single breast with %FGV=21%. Every pixel in the breast region, including the periphery, is shown as a data point, and 32 level contours that show the number of pixels with a particular thickness and density. The majority of pixels are clustered around the average breast thickness of 6.5 cm. Interpolating between the

two reference lines for a given data point gives the %FGV for any pixel.

II.F. Validation procedure

The purpose of our validation procedure was to (a) estimate the accuracy and precision of the thickness and tilt angle measures, (b) quantify the error in %FGV and FGV due to thickness and tilt angle variations, (c) validate the %FGV measure for thicknesses and tilt angle other than those used in the calibration procedure, (d) quantify the measures' stability over time, and (e) clinically test the SXA method against other methods.

For validation purposes, we developed the "TILT" phantom, which is best described as two DSP7s that have been machined to articulate with each other (Fig. 6, left) to form a partial wedge with angles from 0° to 10° with each of the seven densities. A series of TILT and DSP7 combinations allowed for testing known composition, tilt angle, and thickness distributions. For precision studies, the TILT and DSP7 combinations were scanned multiple times using a simplified compression paddle consisting only of a 4.3 mm thick planar Plexiglas sheet and also using a common mammography compression paddle. The tilt angles were verified to 0.1° using a Smart Tool Angle Sensor (M-D Building Products, Oklahoma City, OK).

To evaluate the %FGV and FGV error due to the thickness deviation in terms of our breast measurement technique, a DSP7 phantom at 2, 4, and 6 cm thicknesses was measured. The error was defined as the difference between %FGV, calculated by using known thickness, and the thicknesses with fixed deviations, thus simulating the wrong thickness estimation during mammogram processing. To estimate error in tilt angle measurement, we measure a set of angles of the TILT phantom in combination with DSP7 phantom and found a difference between two %FGV calculations: (1) The algorithm for processing clinical images using a thickness map, which is obtained with 3D-reconstructed parameters (in our case, tilted planes), and (2) an approximation of a constant thickness equal to that on the edge of a selected ROI.

The calibration stability was evaluated by comparing two sequential calibrations measured on the same mammography machine under all screening parameters and conditions as described in Sec. II E 1.

To test our method under screening conditions, we obtained the CC-view and MLO-view digital mammograms of 14 000 clinical subjects using the SXA phantom with one Hologic Selenia used in clinical practice at the Breast Health Center of the California Pacific Medical Center (CPMC, San Francisco, CA). Although our technique could be used for the MLO-view images, only the CC-view images were analyzed for this study. Mammographic percent density, PD, was compared to %FGV in a subset of randomly selected 300 images. The UCSF PD software was written by the investigators and included as part of the SXA workstation. Pixels containing mostly dense tissue were selected using a histogram threshold. The unpublished validation of the UCSF

TABLE I. Precision and accuracy estimates for phantom thickness at different compression plate angles.

Average of 5 Runs with Plexiglas Plate					One Run with Paddle			
H^a (cm)	Mean tz^b (cm)	Std ^c tz (cm)	Mean ry^b (deg)	Std ry (deg)	Angle ^d (deg)	ry (deg)	H (cm)	Mean tz^e (cm)
1.19	3.176	0.027	0.212	0.019	0.1	-0.35	0	2.318
3.19	5.211	0.014	1.016	0.021	2.3	2.037	1.19	3.571
4.15	6.18	0.015	2.285	0.014	3.4	3.156	3.22	5.523
5.19	7.207	0.012	3.576	0.053	4.8	4.419	5.2	7.527
6.15	8.176	0.021	5.349	0.016	7	6.702	7.23	9.495
7.19	9.249	0.019	6.738	0.01				
8.15	10.15	0.008						
9.19	11.268	0.007						
Precision		0.015		0.022				
Accuracy		0.011		0.022		0.09		0.04

^a H refers to the true thickness.^b tz and ry refer to the phantom system coordinate.^cStd refers to standard deviation.^dAngle refers to the tilt angle measured by the angle sensor.^eMean tz refers to an average of the tz values of seven mammography machines.

PD method is as follows. The PD reader was trained on how to perform PD at the University of Toronto (Toronto, ON, Canada). The interclass correlation coefficient (ICC) agreement between the UCSF setup and a similarly trained reader at the Mayo Clinic (Rochester, MN) using CUMULUS PD software (University of Toronto, Toronto, ON, Canada) was ICC=0.96 for 100 digitized film mammograms. In addition, when comparing 100 women with digitized film mammograms and full-field digital mammograms (FFDM) separated in time by approximately 1 year, the ICC > 0.8. Lastly, the long-term stability of the SXA method was tested by scanning the 4 cm DSP7 calibration phantom weekly using an autocontrast technique.

III. RESULTS

III.A. Thickness and tilt angle accuracy and precision

Table I details the precision and accuracy estimates for phantom thickness at different compression plate angles. We used two regimes: (1) Idealized conditions with a Plexiglas plate instead of a paddle and (2) a standard mammography paddle. In each case, the SXA phantom was adhered to the top surface. For the first regime, we used five scans at eight different thicknesses. As shown in Table I, we found the precision (as standard deviation) in the z direction to be $\sigma_z = 0.15$ mm. Using six compression paddle tilt angles, from 0° to 7° , and five scans per angle, we found the precision to be $\sigma_{\text{angle}} = 0.022^\circ$ for all angles. Moreover, accuracy was calculated using the root mean squared deviation from the linear line fitting. The accuracy was 0.11 mm and 0.022° for thickness and angle, respectively. For the second regime using one run and five angles, we found angle accuracy to be 0.09° , which is four times less than idealized conditions, and a thickness accuracy of 0.04 cm. The accuracy deviation of the two regimes can be explained by the influence of the paddle's flex as well as its deformation during compression in accuracy measurements.

The tz presented in the Table I is equal to sum of the true thickness (H) and distance from the detector surface to the Bucky surface (h). The latter is calculated with the SXA phantom located directly on the Bucky surface. Therefore the measured tz is consistently larger than H presented in the Table I. As we cannot directly measure h , we do the accuracy estimation in the terms of root mean squared deviation from linear fitting line to avoid this systematic permanent offset. It should be noted that every machine has a unique Bucky-detector distance. Therefore we do not have the same values for tz for columns 2 and 9. The angle accuracy is calculated in the same manner as for thickness.

We also compared the reported breast thickness from the Digital Imaging and Communications in Medicine (DICOM) header to the measured values using the phantoms on the seven clinical Hologic Selenia Systems at CPMC's Breast Health Center. Figure 8 shows the Bland-Altman plot of difference between the DICOM thickness (H_d) and our

FIG. 8. Bland-Altman plot of difference between the DICOM thickness (H_d) and our phantom-reported thickness (H_{ph}).

FIG. 9. %FGV error estimation due to thickness variation for three true thicknesses of 6, 4, and 2 cm. The slope of the %FGV error varies as a function of true thickness.

phantom-reported thickness (Hph). Using Bland–Altman analysis,¹⁹ we observed no significant trends in the difference between the Hd and Hph when compare to the average values. However, there was a significant 0.97 cm offset of the average of these differences. Thicknesses reported in the DICOM header can be different than the actual thickness reported by our SXA phantom. We attribute the differences to the fact that the Selenia is calibrated using a rigid spot compression paddle and does not compensate for paddle tilt.

III.B. Calibration stability and error estimation

We imaged the DSP7 phantoms with thicknesses from 0.6 to 8 cm and %FGV steps at 0%, 30%, 45%, 50%, 60%, 70%, and 100% twice over a two week period using the same mammography machine under all screening parameters and conditions described in Sec. II E 1. The %FGV errors, as estimated by standard deviations between two sequential measurements for one machine, were found to be within 1%–2%.

Among many potential factors contributing to deviations in %FGV measurements, the breast thickness error is considered to be the main source of error. We calculated that the error in %FGV is different depending on the true thickness of the breast. Figure 9 demonstrates the density errors of 3%, 4%, and 7% per 1 mm of thickness deviation from true thicknesses of 6, 4, and 2 cm thicknesses, respectively. Using the same set of phantom measurements we estimated the relative error for FGV, due to the phantom thickness deviations, to be 3.5%, 4.5%, and 8.5% per 1 mm of thickness deviation for 6, 4, and 2 cm thicknesses, respectively (see Fig. 10). Another significant factor contributing to %FGV measurement error is paddle tilt error. To characterize the tilt error, we measured combinations of different thicknesses and angles using the abovementioned calibration phantoms, as shown in Fig. 11. Thus, if the tilt is not known, the error in the density measure can be substantial, up to 40%, for a thin (2 cm) breast at 6° tilt.

FIG. 10. FGV error estimation due to thickness variation.

To demonstrate the method's performance and to validate our results, we measured a stack of the TILT and DSP7 (2, 4, or 6 cm) phantoms, thereby allowing us to create a validation phantom with known composition, tilt angle, and thickness distribution. Figure 12 presents the validation results of three compositions, 0%, 50%, and 100% fibroglandular tissue; at 7° tilt angle; and thicknesses ranging from 1.8 to 6.7 cm. This figure looks similar to the breast measurements shown in Fig. 7. There are reference points of %FGV equal to 0% and 100% at different thicknesses with quadratic fit lines and phantom ROI pixel values plotted as dots. It clearly illustrates the validation of the algorithm when using a calibration phantom of known composition and geometrical shape. The phantom ROI pixel values at different thicknesses and compositions lie directly along reference value curves, thereby confirming that values obtained during the reconstruction procedure match those of the TILT phantom. Any deviations of angle and thicknesses from real values place the experimental point outside of the reference curves. The %FGV error (standard deviation) was estimated to be 2% by averaging the differences between the measured and predicted values for TILT and DSP phantom combinations cov-

FIG. 11. %FGV error estimation due to angle variation.

FIG. 12. ROI pixel values of the TILT phantom with %FGV=0, 50/50, and 100% and a tilt angle of 6.8° .

ering a thicknesses range from 1.8 to 6.7 cm and %FGV steps at 0%, 30%, 45%, 50%, 60%, 70%, and 100%.

III.C. Clinical testing

After analyzing volumetric density of clinical screening mammography examinations, we initially found that there were many breasts with %FGV(x,y) values lower than our adipose reference. For example, around 6% of %FGV evaluations were below 0% when the CIRS adipose phantom reference was used. Thus, the majority of pixel values were below the adipose reference value. This difference can be explained by either errors in measuring thickness (the measurement was thinner than in actuality) or by errors in the composition of the adipose reference values. We validated calculation algorithms of the thickness map using our tilting phantoms and precise physical measures and found that our phantom thickness values were very accurate. Thus, negative densities were most likely due to CIRS adipose reference values being radiographically denser than human adipose tissue. Similar results were reported by Johns and Yaffe²⁰ in that the “CIRS Fat” had a higher linear attenuation coefficient than excised breast adipose (0.627 cm⁻¹ versus 0.558 cm⁻¹ for 18 keV, respectively). It should be noted that our previous dual-energy x-ray absorptiometry experiments in measuring breast density using a bone densitometer instead of a mammography system also demonstrated negative densities when referenced to the CIRS adipose phantom.²¹

We followed up by comparing attenuation pixel values of substances more closely related to biological fat, such as stearic acid, canola oil, and Crisco (J. M. Smucker Co., Orville, OH), with the CIRS adipose references of the same thicknesses and x-ray exposure conditions. Using our SXA technique with the CIRS adipose reference, we measured these materials to be -28% , -39% , and -32% , respectively. Thus, to correctly calculate the breast composition, we found it necessary to use the appropriate fat reference closer to our clinical findings. Figure 13 shows the attenuation profile from a low-density breast compared to 100%, 0%, and

FIG. 13. Chest-to-nipple pixel profile values for a single image row of a breast with the appearance of totally fatty. This profile is compared to phantom materials with the same thickness profile but different densities. Note that the breast is more “fatty” than the CIRS fat reference material.

Crisco-referenced %FGV curves with the same thickness profiles. As can be seen, the breast profile has less attenuation than the CIRS adipose reference.

The histograms of two tilt paddle angle distributions for the CC-view digital mammograms of 14 000 clinical subjects are presented in Fig. 14. These distributions are characterized, respectively, by mean= 4.5° and standard deviation= 1.5° for the chest-nipple direction angle and mean= 0.4° and standard deviation= 0.4° for the angle parallel to the chest. These distributions show the importance of considering the paddle tilt angles in SXA %FGV measurements, especially for tilt in the chest-nipple direction. As a consistency test, we compared %FGV, total breast volume and FGV values of left and right breasts for the same women. The agreement was $r^2=0.89$ for %FGV, $r^2=0.94$ for total breast volume, and $r^2=0.77$ for FGV. The PD left and right breast comparison of 102 pair set demonstrated $r^2=0.88$.

The long-term stability of the %FGV measurements is presented in Fig. 15. The standard deviations of the average %FGV values during a four-month period were approximately 1% for the 0%, 50%, and 100% FGV references. There were no significant trends in density values over time.

The results of comparison of %FGV and PD is presented on Fig. 16. Comparison of PD with %FGV values demonstrated good correlation with Pearson coefficient of $r=0.83$.

FIG. 14. Tilt paddle angle distributions for 14 000 subject CC-view mammograms of both right and left breasts.

FIG. 15. Long-term stability of the SXA technique on a Hologic Selenia full-field digital mammography (FFDM).

A linear regression of the dependence of %FGV on mammographic density was characterized by the following equation: $\%FGV = 1.03 * PD + 13.6$.

IV. DISCUSSION AND CONCLUSION

We have presented a method to measure the percent fibroglandular tissue and fibroglandular tissue volumes using digital mammography. Our approach to quantify breast density is unique in that we have designed a reference phantom that allows us to: First, generate an accurate thickness profile of the breast at each image position; and second, calculate an image-specific calibration of the pixel grayscale values to breast composition. The novel aspect of our SXA approach is that the phantom does not interfere with nor modify standard screening procedures.

Our method defines breast composition as containing two compartments, fat and fibroglandular tissue. However, this is not a true molecular or tissue level breast composition model. At the molecular level, the breast is composed of water, proteins, lipids, minerals, and a small residual amount of glycogen and other substances. At the tissue level, the breast is composed of adipose and fibroglandular tissue.

FIG. 16. A comparison of mammographic density and percent fibroglandular tissue volume for 300 CC-view mammograms.

When we used commercial polymers mimicking adipose tissue as our fat reference, fatty breasts were sometimes found to have a negative %FGV. These results are explained by the linear attenuation coefficient for the CIRS fat to be higher than excised breast tissue.²⁰ Using pure fat (Crisco) as the lower reference fits our clinical results better and provides reasonable %FGV values.

We found that unacceptably large errors in quantifying FGV and %FGV occur if the actual thickness and paddle tilt for each image are not known and not included in the calculation of FGV and %FGV. The manufacturer calibration of breast thickness reported on the Selenia image is achieved by compressing a 4 cm reference object with the rigid spot compression paddle at a 10 daN force. This calibration method gives a result that is close to the actual breast thickness, but not close enough for accurate density measures as the technologist has a choice of screening paddle sizes (18 × 24 cm or 24 × 30 cm) and paddle rigidity (spring-loaded, tiltable, or rigidly fixed compression surfaces). There are other methods that could be used to quantify compression thickness and paddle tilt. Yang *et al.*²² have developed a stereo-camera configuration that optically determines the paddle height and shape. This method differs from our own in that it directly measures paddle deformity throughout the entire paddle's complex and also raises concerns about patient privacy.

Other researchers have done similar yet uniquely different work to calculate breast volumetric density. For film mammography, Pawluczyk *et al.*²³ demonstrated that by using a sophisticated model of breast attenuation, breast thickness, x-ray technique, and film response, one could accurately calculate volumetric breast density. They used a small half-value layer aluminum phantom in the image to monitor the variations in exposure, film emulsion, and film processing. A limitation to this model is that it does not monitor actual breast thickness or local thickness variations due to paddle tilt. Another approach, elaborated by Diffey *et al.*,²⁴ uses an aluminum wedge phantom for density and six lead landmarks positioned on the paddle for the thickness and tilt angle. The phantom sits resident on the bottom (Bucky). It is used along with lead spheres distributed around the periphery of the compression paddle positioned not to overlap with the breast image. The markers have the potential to overlay tissue in cases where the breast extends to the edges of the image area.²⁵ Compared to other phantom-based methods our method provides an accurate thickness map that potentially decreases error due to the compression paddle tilt.

Methods that do not use an image reference phantom have also been reported. Highnam *et al.*^{26,27} used a self-referencing method to calibrate the breast tissue fat density of each image and incorporated a full physics model. The work of van Engeland *et al.*²⁸ resembled this approach but used an internal calibration to reduce dependence on breast thickness. However, these approaches require a good knowledge of x-ray properties of materials, scattering conditions and the calibration has a potential source of error for dense breasts. The extensive calibration method of Kaufhold *et al.*,²⁹ rather than a heavy dependence on models, was limited

by the accuracy of the individual mammography systems' breast thickness calibrations. Compared to other volumetric modeling methods, the accuracy of our approach can be verified using simple phantoms of known composition allowing for cross-calibration between machines and the utilization of phantom quality control procedures.

By using a quality assurance phantom, we have shown that our method is stable over time. Quality control (QC) phantoms are required by the Mammography Quality Standards Act³⁰ and are already scanned weekly to confirm system contrast. A single quality assurance phantom could be used for many systems in the same clinic. It may be possible to use the QC phantom to monitor shifts in density calibration over time by using cumulative statistics (CUSUM) approaches^{31,32} that determine the most likely breaking points in the calibration.

A good agreement between mammographic percent density and percent fibroglandular tissue volume provides a strong verification of the feasibility of the designed SXA method for clinical studies. The 13.6% offset for SXA at PD=0 can be explained by our use of a pure fat reference instead of adipose tissue which contains approximately 15% water.³³ A similar offset (20%) and correlation to PD ($r=0.85$) was reported by Boyd *et al.*¹² when comparing PD to breast percent water content using MRI. A good correlation of total breast volume, %FGV and FGV measures between left and right breasts confirms the internal consistency of the method.

We demonstrated our method on one type of digital mammography system. However, there is no technical reason why it could not be used on other makes and models of digital systems including computed radiography methods for mammography using storage phosphors.³⁴

Future research will include validating our technique on other mammography makes and models, studying the method stability and reproducibility in clinical conditions for a longer period of time, machine cross-calibration, as well as further investigating absolute tissue composition accuracy.

ACKNOWLEDGMENTS

The following funding was received in support of this work: Pfizer Independent Research Grant No. G900004, Da-Costa International Breast Cancer Fund, and NIH/NCI Grant No. P01 CA107584. We acknowledge the editorial support of Renée Euchner, R.N., and Erin Reese.

^{a)} Author to whom correspondence should be addressed. Electronic mail: john.shepherd@radiology.ucsf.edu; Telephone: (415) 885-7227; Fax: (415) 353-9425.

¹N. F. Boyd *et al.*, "Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian National Breast Screening Study," *J. Natl. Cancer Inst.* **87**(9), 670–675 (1995).

²C. Byrne *et al.*, "Mammographic features and breast cancer risk: Effects with time, age, and menopause status," *J. Natl. Cancer Inst.* **87**, 1622–1629 (1995).

³A. F. Saftlas and M. Szklo, "Mammographic parenchymal patterns and breast cancer risk," *Epidemiol. Rev.* **9**, 146–174 (1987).

⁴N. F. Boyd *et al.*, "Mammographic features and the risk and detection of breast cancer," *N. Engl. J. Med.* **356**(3), 227–236 (2007).

⁵C. M. Vachon *et al.*, "Mammographic density, breast cancer risk and risk

prediction," *Breast Cancer Res. Treat.* **9**(6), 217 (2007).

⁶J. N. Wolfe, "Breast patterns as an index of risk for developing breast cancer," *AJR, Am. J. Roentgenol.* **126**(6), 1130–1137 (1976).

⁷J. M. Boone *et al.*, "A breast density index for digital mammograms based on radiologists' ranking," *J. Digit. Imaging* **11**(3), 101–115 (1998).

⁸C. Atkinson *et al.*, "Mammographic patterns as a predictive biomarker of breast cancer risk: Effect of tamoxifen," *Cancer Epidemiol. Biomarkers Prev.* **8**, 863–866 (1999).

⁹J. N. Wolfe, A. F. Saftlas, and M. Salane, "Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: A case-control study," *AJR, Am. J. Roentgenol.* **148**(6), 1087–1092 (1987).

¹⁰N. A. Lee *et al.*, "Fatty and fibroglandular tissue volumes in the breasts of women 20–83 years old: Comparison of X-ray mammography and computer-assisted MR imaging [see comments]," *AJR, Am. J. Roentgenol.* **168**(2), 501–506 (1997).

¹¹R. C. Boston *et al.*, "Estimation of the content of fat and parenchyma in breast tissue using MRI T1 histograms and phantoms," *Magn. Reson. Imaging* **23**(4), 591–599 (2005).

¹²N. Boyd *et al.*, "Breast-tissue composition and other risk factors for breast cancer in young women: A cross-sectional study," *Lancet Oncol.* **10**(6), 569–580 (2009).

¹³J. A. Kalef-Ezra *et al.*, "Electron density of tissues and breast cancer radiotherapy: A quantitative CT study," *Int. J. Radiat. Oncol., Biol., Phys.* **41**(5), 1209–1214 (1998).

¹⁴M. J. Yaffe, "Mammographic density. Measurement of mammographic density," *Breast Cancer Res. Treat.* **10**(3), 209 (2008).

¹⁵J. A. Shepherd *et al.*, "Novel use of single x-ray absorptiometry for measuring breast density," *Technol. Cancer Res. Treat.* **4**(2), 173–182 (2005).

¹⁶J. Shepherd *et al.*, Proceedings of the 91st Annual Meeting of the Radiological Society of North America, Chicago, IL, 2005 (unpublished).

¹⁷M. Zhang *et al.*, "On the question of 3D seed reconstruction in prostate brachytherapy: The determination of x-ray source and film locations," *Phys. Med. Biol.* **49**(19), N335–N345 (2004).

¹⁸G. R. Hammerstein *et al.*, "Absorbed radiation dose in mammography," *Radiology* **130**(2), 485–491 (1979).

¹⁹J. M. Bland and D. G. Altman, "Statistical-methods for assessing agreement between 2 methods of clinical measurement," *Lancet* **1**(8746), 307–310 (1986).

²⁰P. C. Johns and M. J. Yaffe, "X-ray characterisation of normal and neoplastic breast tissues," *Phys. Med. Biol.* **32**(6), 675–695 (1987).

²¹J. A. Shepherd *et al.*, "Clinical comparison of a novel breast DXA technique to mammographic density," *Med. Phys.* **33**(5), 1490–1498 (2006).

²²J. Yang, G. Mawdsley, and M. Yaffe, U.S. Patent No. 7,286,632 (23 October 2007).

²³O. Pawluczyk *et al.*, "A volumetric method for estimation of breast density on digitized screen-film mammograms," *Med. Phys.* **30**(3), 352–364 (2003).

²⁴J. Diffey, A. Hufton, and S. Astley, "A new step-wedge for the volumetric measurement of mammographic density," *IWDM 2006, Proceedings of the Eighth International Workshop*, edited by S. Astley, Manchester, UK (Springer, New York, 2006), Paper No. LNCS 4046, pp. 1–9.

²⁵M. Berks *et al.*, "Feasibility and acceptability of stepwedge-based density measurement," *IWDM 2006, Proceedings of the Eighth International Workshop*, edited by S. Astley, Manchester, UK (Springer, New York, 2006), Paper No. LNCS 4046, pp. 355–361.

²⁶R. Highnam, M. Brady, and B. Shephstone, "A representation for mammographic image processing," *Med. Image Anal.* **1**(1), 1–18 (1996).

²⁷R. Highnam *et al.*, "Breast composition measurements using retrospective standard mammogram form (SMF)," *Phys. Med. Biol.* **51**(11), 2695–2713 (2006).

²⁸S. van Engeland *et al.*, "Volumetric breast density estimation from full-field digital mammograms," *IEEE Trans. Med. Imaging* **25**(3), 273–282 (2006).

²⁹J. Kaufhold *et al.*, "A calibration approach to glandular tissue composition estimation in digital mammography," *Med. Phys.* **29**(8), 1867–1880 (2002).

³⁰R. E. Hendrick *et al.*, *American College of Radiology (ACR) Mammography Quality Control Manual for Radiologists, Medical Physicists and Technologists* (American College of Radiology, Reston, 1999).

³¹W. A. Shewhart, *Statistical Method from the Viewpoint of Quality Control* (Graduate School of the Department of Agriculture, Washington, DC, 1939).

- ³²Y. Lu *et al.*, “Dual x-ray absorptiometry quality control: Comparison of visual examination and process-control charts,” *J. Bone Miner. Res.* **11**(5), 626–637 (1996).
- ³³W. S. Snyder *et al.*, *Report of the Task Group on Reference Man* (Pergamon, Oxford, 1975).
- ³⁴S. Rivetti *et al.*, “Comparison of different commercial FFDM units by means of physical characterization and contrast-detail analysis,” *Med. Phys.* **33**(11), 4198–4209 (2006).