IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine and Biology

Phys. Med. Biol. **59** (2014) 4391–4409 doi:10.1088/0031-9155/59/15/4391

Comparison of volumetric breast density estimations from mammography and thorax CT

N Geeraert[1,2,3], R Klausz[2], L Cockmartin[1], S Muller[2], H Bosmans[1] and I Bloch[3]

[1] Department of Radiology—LUCMFR, KU Leuven, Herestraat 49, Leuven, Belgium
[2] GE Healthcare, 283 Rue de la Miniere, 78533 Buc, France
[3] Institut Mines-Télécom, Telecom ParisTech, CNRS LTCI, 46 Rue Barrault, Paris, France

E-mail: nausikaa.geeraert@student.kuleuven.be

Received 20 December 2013, revised 11 June 2014
Accepted for publication 12 June 2014
Published 22 July 2014

Abstract

Breast density has become an important issue in current breast cancer screening, both as a recognized risk factor for breast cancer and by decreasing screening efficiency by the masking effect. Different qualitative and quantitative methods have been proposed to evaluate area-based breast density and volumetric breast density (*VBD*). We propose a validation method comparing the computation of *VBD* obtained from digital mammographic images (VBD_{MX}) with the computation of *VBD* from thorax CT images (VBD_{CT}). We computed VBD_{MX} by applying a conversion function to the pixel values in the mammographic images, based on models determined from images of breast equivalent material. VBD_{CT} is computed from the average Hounsfield Unit (*HU*) over the manually delineated breast volume in the CT images. This average *HU* is then compared to the *HU* of adipose and fibroglandular tissues from patient images. The VBD_{MX} method was applied to 663 mammographic patient images taken on two Siemens Inspiration ($hosp_L$) and one GE Senographe Essential ($hosp_J$). For the comparison study, we collected images from patients who had a thorax CT and a mammography screening exam within the same year. In total, thorax CT images corresponding to 40 breasts ($hosp_L$) and 47 breasts ($hosp_J$) were retrieved. Averaged over the 663 mammographic images the median VBD_{MX} was 14.7% . The density distribution and the inverse correlation between VBD_{MX} and breast thickness were found as expected. The average difference between VBD_{MX} and VBD_{CT} is smaller for $hosp_J$ (4%) than for $hosp_L$ (10%). This study shows the possibility to compare VBD_{MX} with the *VBD* from thorax CT exams, without additional examinations. In spite of the limitations caused by poorly defined breast limits, the calibration of mammographic images to local VBD provides opportunities for further quantitative evaluations.

0031-9155/14/154391+19$33.00

Keywords: volumetric breast density, mammography, thorax CT, HU

(Some figures may appear in colour only in the online journal)

1. Introduction

The classification of breasts in groups based on the characteristics of parenchymal patterns in the mammographic images was originally developed by Wolfe in 1976 (Wolfe 1976). Wolfe's classification was later converted to the BIRADS classification (Dórsi 2003), where breast density is estimated from the ratio of fibroglandular tissue to total breast tissue as observed in mammograms. The ratio of the *area* of fibroglandular tissue to the *area* of the breast is referred to as the *areal breast density (ABD)*. Later, Kopans (Kopans 2008) put forward the *volumetric breast density (VBD)* as the ratio of the *volume* of the fibroglandular tissue to the *volume* of the breast, on which Highnam, Brady and Shepstone (Highnam *et al* 1996) and others had been working.

A systematic review and meta-analysis of publications on breast density and parenchymal patterns in relation to breast cancer risk was published in 2006 by McCormack and dos Santos (2006). The study demonstrated that the ABD, and to a lesser extent Wolfe grades and BIRADS classification, are strong predictors of the risk of developing breast cancers. Van Gils *et al* (1999) and Ting *et al* (2012) also found that the evolution of the density as a function of time was indicative for the risk of developing breast cancers: the faster the changes, the higher the risk. For women who are being treated for breast cancer with Tamoxifen, the effect on the fibroglandular tissue was found indicative for survival rate (Li *et al* 2013), with a higher survival rate for women whose fibroglandular tissue decreases. Reproducible and consistent methods of quantification of breast density have therefore become important tools in breast cancer epidemiology.

Several methods of local (Beckett and Kotre 2000, Desponds and Klausz 1994) and global (Highnam *et al* 1996, 2006, 2010Kaufhold *et al* 2002, Hartman *et al* 2008, Malkov *et al* 2008, Shepherd *et al* 2005b, van Engeland *et al* 2006) breast density computation have been published, all relying on measurements of the transmission of x-rays through the breast. Almost all methods for digitized and digital mammographic images compare breast image pixel values to image pixel values of a breast tissue equivalent material (Heidsieck 1989) as system calibration. The most important problem encountered by all methods was the lack of an accurate measure of breast thickness.

Two commercially available products are based on the work of Highnam, Brady *et al* (Highnam *et al* 2006): R2 QUANTRA™ (Hologic, Bedford, MA) and Volpara® (Matakina® Inc., Wellington, New Zealand). They handle the thickness estimation in a different way. Whereas the thickness estimation from the method of Highnam, Brady *et al* (Highnam *et al* 2006) was based only on the hypothesis of a purely adipose uncompressed breast region and some image smoothness criteria, the R2 QUANTRA™ software uses an improved image-based thickness correction as described in Hartmann *et al* (2008), with the recorded paddle height as the initial value for the compressed breast thickness. According to Highnam, Brady *et al* (Highnam *et al* 2010) the Volpara® software is searching for a purely adipose region in the image, based on phase congruency, to compute the *VBD* with a high precision compressed breast thickness. Shepherd dealt with the compression paddle orientation by placing a phantom with lead markers on the paddle (Malkov *et al* 2008).

All these methods make some hypotheses. It is therefore important that they can be validated against known values. The common approach is to quantify the respective amounts of fibroglandular and adipose tissues in 3D images of the same breast. Unfortunately the limits of

the breast are anatomically not well defined and determining the volume of the breast is inherently difficult. This is also the case for 2D mammography. As a result, while good correlations can be obtained between breast volume estimates from different modalities, it is more difficult to obtain a good match between these volumes (van Engeland *et al* 2006, Wang *et al* 2013). Despite this difficulty, the correlation between *VBD* as obtained from different images and from different modalities is possible. For the previously described methods (Van Engeland *et al* 2006, R2 QUANTRA™ 2008, Volpara® 2010) the authors compared the *VBD* from mammographic images to the *VBD* obtained by delineating manually the breast and the fibroglandular tissue in breast magnetic resonance (MR) images of the same patient. The comparison between the *VBD* computed by single x-ray absorptiometry (Shepherd *et al* 2005a), QUANTRA™ and Volpara® on the one hand and MR on the other hand was described recently by Wang *et al* (2013). Using clinical MR images has however some drawbacks. Due to in-plane inhomogeneity of the pixel values it is not possible to set a threshold for automatic segmentation of the fibroglandular tissue and therefore more sophisticated methods are needed. Segmenting manually the fibroglandular tissue is difficult due to its irregular borders and the presence of small fibrous structures. Next, MR images represent MR characteristics of the tissues, and the subsequent tissue classification may be different from that based on x-ray imaging. Finally, breast MR is not performed for the screening population but mainly for patients being suspicious for a malignant disease. The presence of a disease may disturb the breast density measurement, in case the hypothesis of the two-compartment model, with only adipose and fibroglandular tissues, is not valid anymore. Alonzo-Proulx *et al* (2010a) proposed dedicated breast CT images as a basis for comparison. They computed the volumes of fibroglandular and adipose tissues in breast CT by thresholding the images. Then they simulated mammographic images from the breast CT acquisitions for which they then computed the *VBD* with their 2D method. Also Vedantham *et al* (2012) showed the possibility to measure the fibroglandular tissue and the volumetric glandular fraction based on breast CT images. Breast CT seems to be a good solution to establish reliable *VBD* values but is only available in a small number of research centers. We propose a validation method based on regular thorax CT images. We exploit the large number of thorax CT procedures performed on the breast screening population, which include most of the time the complete breast. The method can easily be performed by users of breast density applications.

In January 2014 Salvatore *et al* (2014) published a study where the authors show a promising agreement between BI-RADS density classifications in mammography and thorax CT of the same patients. Their work indicates a possibility to compute also the *VBD* from thorax CT images. Salvatore *et al* (2012) segmented the breast with a semi-automatic computer algorithm prototype. They computed the breast density by partitioning the segmented breast region. We propose to use the characteristic of the Hounsfield Units (HU) provided by CT reconstruction algorithms being a linear function of the average attenuation of the material over the volume corresponding to the spatial resolution of the CT scanner in the *X*-*Y*-*Z* directions. As a consequence, this value is a direct function of the linear attenuation coefficients of the local tissue components, and the *VBD* can be simply computed from the average *HU* in the breast volume and the *HU* of adipose and fibroglandular tissues, without further in-breast tissue segmentation. Another advantage of the use of CT images instead of MR images is that the common use of x-rays in CT and mammography ensures the same classification of tissues, based on x-ray attenuation properties of the materials. Thorax CT images with lesion free breasts can be collected in a relatively short time period. For example, at the university hospital *Universitair Ziekenhuis Gasthuisberg Leuven (BE)* over 30 women per month undergo a screening exam within 12 months of a CT thorax exam. Therefore, the collection of these data does not require additional medical examinations, so there is neither extra radiation given to the patients nor extra cost and time are asked from the radiologists.

In this work, we show the possibility to use thorax CT images obtained from a routine CT scanner to verify the accuracy of algorithms for *VBD* estimations. We have tested this proposal for an otherwise verified *VBD* computation algorithm that we derived from the method by Kaufhold *et al* (2002). The additional verification consisted of left/right and CC/MLO correlations and comparison of population distribution of VBD with previous results from literature.

2. Material and methods

2.1. VBD for phantoms from digital mammographic images

Our model is based on the work of Kaufhold *et al* (2002) and in particular Formula 8 therein:

$$\%G = 100\frac{g_{\text{patientID}} - F_{\text{patientID}}}{G_{\text{patientID}} - F_{\text{patientID}}}, \tag{1}$$

with $\%\ G$ the percentage density, and $g_{\text{patientID}}$, $F_{\text{patientID}}$ and $G_{\text{patientID}}$ the 'mAs-normalized negative log' intensity values for respectively the pixel where the density is evaluated, and pure adipose and fibroglandular tissues with the same thickness. This expression was derived from the mono-energetic attenuation laws of Beer-Lambert by replacing mono-energetic photon fluence with detector signal, integrated over the complete spectrum. This includes scatter and beam hardening and thus a dependence on the object thickness and tube potential. By denoting m the current-time product and p the offset-corrected pixel values (Perry *et al* 2006), VBD_{MX} the *VBD* based on mammographic images and x, A and G indicating the tissue of unknown density, purely adipose tissue and purely fibroglandular tissue respectively, our equation can be written as:

$$VBD_{\text{MX}} = \frac{\ln(p/m)_x - \ln(p/m)_A}{\ln(p/m)_G - \ln(p/m)_A} \tag{2}$$

Kaufhold *et al* (2002) previously calibrated the values for $F_{\text{patientID}}$ and $G_{\text{patientID}}$ as a function of thickness separately for eleven anode-filter-tube potential combinations by imaging phantoms of breast tissue equivalent material with thickness ranging from 2 to 7 cm. We have performed the calibration for a broader range of spectra (see table 1) and we modeled $\ln(p/m)_A$ as a second order polynomial function of both the tube potential and the thickness. In our computations we preferred to model $\ln(p/m)_{G-A} = \ln(p/m)_G - \ln(p/m)_A$ as a single term because this gave smaller errors for VBD than modeling $\ln(p/m)_G$ and $\ln(p/m)_A$ separately and subtracting afterwards, i.e. 0.6% instead of 1.7% on average over intermediate points. This can be understood knowing that both terms, $\ln(p/m)_G$ and $\ln(p/m)_A$, depend on the incoming spectrum in the same way, which cancels out in the combined term.

For the implementation of equation (2) we acquired images of breast tissue simulating phantoms (CIRS Inc, Norfolk, VA). We used $24 \times 18 \times 1$ cm^3 plates with attenuation equivalent to 0% and 100% fibroglandular tissue. Phantoms with different *VBD* and thicknesses as presented in figure 1 were realized by combining the different plates.

The acquisitions were done for two Siemens Inspiration systems in *Universitair Ziekenhuis Gasthuisberg, Leuven, Belgium*, hereafter called hosp_L, and a GE Senographe Essential in *Centre Hospitalier Jolimont-Lobbes, Entité Jolimontoise, La Louvière, Belgium*, hereafter called hosp_J. Experimental conditions are listed in table 1.

For all acquisitions the compression paddle and the anti-scatter grid were in place. Acquisitions were performed in manual exposure mode with current-time product as close as possible to the one used in automatic exposure mode for the corresponding thickness. With the manual exposure mode we extended the thickness-tube potential range of the automatic

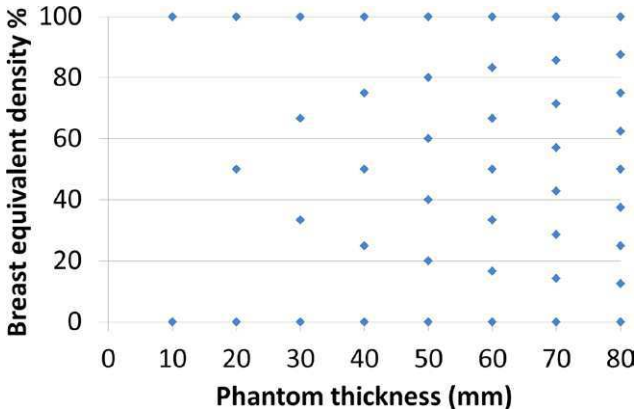

Figure 1. Overview of combinations of phantom thickness and *VBD* used for the implementation of equation (2). All combinations are obtained from plates of 0% and 100% *VBD* of 1 cm thick.

Table 1. Summary of experimental conditions for the images of the phantoms. These data were used to implement equation (2).

System	GE essential	Siemens inspiration
Exposure mode	AUTOMATIC, MANUAL	AUTOMATIC, MANUAL
Anode	Molybdenum (Mo), Rhodium (Rh)	Tungsten (W)
Filter	Molybdenum (Mo), Rhodium (Rh)	Rhodium (Rh)
Tube potential (kV)	24–27 (MoMo) 25–29 (MoRh) 27–32(RhRh)	24–32 (W/Rh)
Thickness (mm)	10–50 (MoMo) 20–70 (MoRh) 20–80 (RhRh)	20–80 (W/Rh)

exposure mode. The values of p of the images were determined as the average of a square region of interest of 1 cm^2 in the middle of the plate at 6 cm from the chest wall.

First we show the validity of equation (2) for our implementation. Therefore we computed the *VBD* for the phantoms in figure 1 and compared the results to their known *VBD*. To do so, we used equation (2) with $\ln(p/m)_A$ and $\ln(p/m)_G$ computed with the values of p measured in the images of 0% and 100% phantoms. Secondly we evaluated the *VBD* model by comparing the results of equation (2) for the phantoms, computed with $\ln(p/m)_A$ and $\ln(p/m)_{G-A}$ from the second order polynomial, to the known *VBD*.

2.2. Application to mammographic images

To compute the VBD_{MX} of a breast, the local VBD as obtained with equation (2) is first determined in all pixels of the mammographic image. The VBD_{MX} is then obtained by multiplication of the local VBD_{MX} and the local thickness, giving the local glandular content, and further integration of these values over the breast area and normalization by the total volume. The anode, filter, tube potential and current-time product are retrieved from the DICOM header of the image. However, since the models for $\ln(p/m)_A$ and $\ln(p/m)_{G-A}$ are

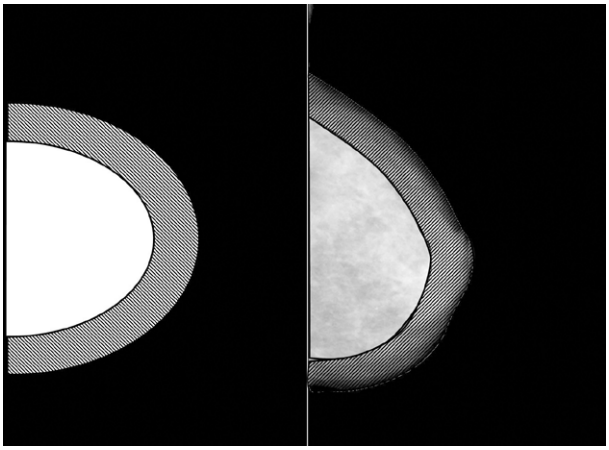

Figure 2. Segmentation of the mammographic image according to three thickness regions: the background (dark region), the region where the breast under compression is in contact with the compression paddle (white region), and the peripheral region (shaded region). Left: schematic view, right: patient example.

highly dependent on the total thickness, the breast thickness in every pixel must be known accurately.

Therefore the image is segmented into three zones (see figure 2): the background (outside the breast), the region where the breast under compression is in contact with the compression paddle, and the peripheral region of the breast in between these zones.

The background region is segmented by thresholding the image using a threshold value determined automatically from the histogram of the image. The thickness of the area in contact with the compression paddle has been considered to be constant and equal to the compressed breast thickness value, stored in the DICOM header. The peripheral region has been set as a band parallel to the border of the breast and with a width equal to half the thickness of the breast. In this band the thickness profiles are semi-circular (van Engeland *et al* 2006).

The skin is excluded, from both the fibroglandular and the breast volume, by subtracting a constant gland thickness of 2 × 1.5 mm (Yaffe *et al* 2009) from the thickness maps. The region of the pectoral muscle is excluded too. The segmentation was kept simple by segmenting a triangle formed by the line from the middle of the breast at the long side of the image to the start of the breast at the upper short side of the image.

The knowledge of the acquisition parameters, and, for each point of the image, the pixel value and the breast thickness, allows computing $\ln(p/m)_A$, $\ln(p/m)_{G-A}$ and $\ln(p/m)_x$. This results in a density map of the image and the VBD_{MX} of the breast. Highly attenuating objects are automatically excluded if their VBD_{MX} is over 100% and also pixels with values lower than 0%, mainly in the peripheral region, are excluded from integration for the VBD_{MX}.

2.3. Computation of volumetric breast density from CT images

The *VBD* was obtained from CT images ($\mathrm{VBD_{CT}}$) for comparison with the VBD_{MX} from mammographic images. The voxel values in CT images are defined by:

$$HU_x = \frac{\mu_x - \mu_{\mathrm{water}}}{\mu_{\mathrm{water}}} \times 1000 \tag{3}$$

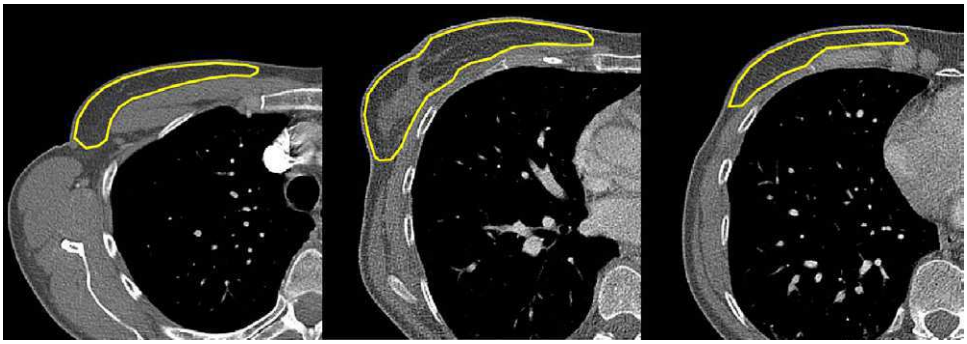

Figure 3. Delineation of the CT images. An upper slice (left), middle slice (middle) and lower slice (right) are presented with the delineation of the breast, determining the volume of the breast. The average *HU* is computed in the delineated part. In each case the skin is excluded.

We can write VBD_{CT} as (Deslattes 1969):

$$VBD_{\mathrm{CT}} = \frac{\mu_x - \mu_A}{\mu_G - \mu_A} = \frac{HU_X - HU_A}{HU_G - HU_A} \tag{4}$$

and

$$HU_X = (HU_G - HU_A) \cdot VBD_{\mathrm{CT}} + HU_A \tag{5}$$

with HU_X, HU_A and HU_G the average *HU*, the *HU* of the adipose tissue and the *HU* of the fibroglandular tissue, respectively (Geeraert *et al* 2013). equation (4) can be applied per voxel, but can as well be applied to a larger volume such as the entire breast, where HU_X is the average *HU* over the entire volume, so only the breast and not the detailed glandular structure has to be segmented in the CT images.

We applied the computation of VBD_{CT} to the database of breast images described in section 2.4. The breast was delineated manually slice per slice for the CT acquisitions (see figure 3). The skin was excluded from the volume. The pectoral muscle was taken as the chest wall border of the breast.

As for mammography, the CT measurement method relies on the knowledge of the reference materials, adipose and fibroglandular tissues, for which the *HU* values must be determined in the same conditions as for the clinical images. Therefore we identified regions in the breast of the patient images that showed up as purely adipose or purely fibroglandular tissue. Ten adipose and ten fibroglandular regions were manually identified in ten patients (see figure 4). HU_A and HU_G were fixed to the average *HU* of these regions.

2.4. Database of mammographic and CT images

We have used three databases: one containing only mammographic images and two containing both mammographic images and CT images (see table 2). All patient images were acquired for medical reasons, and no extra exams were acquired only for this study (Trial ID NTR3357 at *Nederlands Trial Register* according to the declaration of Helsinki, 2008). All images were collected in the two hospitals where the VBD_{MX} model was calibrated: hosp_L and hosp_J. Database 1 consists of mammographic images for testing the VBD_{MX} computation algorithm. The VBD_{MX} distribution and the mean and standard deviation over the population were computed. The VBD_{MX} distribution was also plotted versus breast

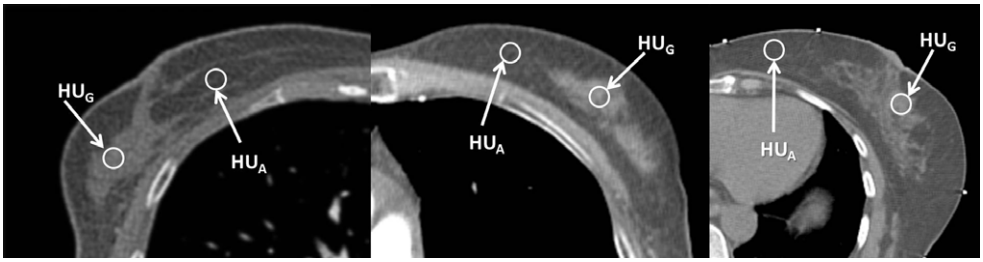

Figure 4. Regions of purely adipose and purely fibroglandular tissues are selected in patients images. HU_A and HU_G were set to the average HU of the regions.

Table 2. Overview of the databases used for the computation of VBD_{MX} and VBD_{CT}.

	Database 1	Database 2	Database 3
Type of images	MX	MX + CT	MX + CT
# of MX images	663	50	103
# of breasts	235	40	47
# of CT images	—	25	27
Hospital	$hosp_L$, $hosp_J$	$hosp_L$	$hosp_J$
MX system	Siemens Inspiration, GE Senographe Essential	Siemens Inspiration	GE Senographe Essential
CT system	—	Siemens Sensation 64	Siemens SOMATOM Definition Flash and Definition AS
CT tube potential	—	120 kV	120 kV
CT slice thickness	—	5 mm	3 mm (22 cases), 5 mm (5 cases)
CT pixel size range	—	0.63×0.63 mm^2 to 0.89×0.89 mm^2	0.43×0.43 mm^2 to 0.98×0.98 mm^2

thickness. Databases 2 and 3 consist of thorax CT and standard screening mammography images of patients who underwent both exams within the same year. We generally collected one CT series and four mammographic images per patient: one CC and one MLO mammographic image per breast. However for some CT series one breast was not fully present in the FOV or some patients had a mastectomy, and not all patients had two mammographic images per breast. Therefore we do not always have four mammographic images per CT exam. These databases were used to study the correlation between VBD_{MX} and VB_{CT} values.

3. Results

3.1. VBD_{MX} for phantoms

The VBD_{MX} of the phantoms were computed using equation (2) and the values of $\ln(p/m)_A$ and $\ln(p/m)_{G-A}$ directly derived from the measurements, and then compared to their nominal values figure 5(*a*). The resulting maximum deviation from the nominal values figure 5(*b*) was 3.8% for the Siemens system and 1.5% for the GE system. The average errors were 0.6% and 0.3% with standard deviations 1.2% and 0.7% respectively. When the VBD_{MX} was computed from the $\ln(p/m)_A$ and $\ln(p/m)_{G-A}$ models, the average errors and standard deviations are 2% (2%) and 0.1% (0.5%) respectively figure 5(*c*).

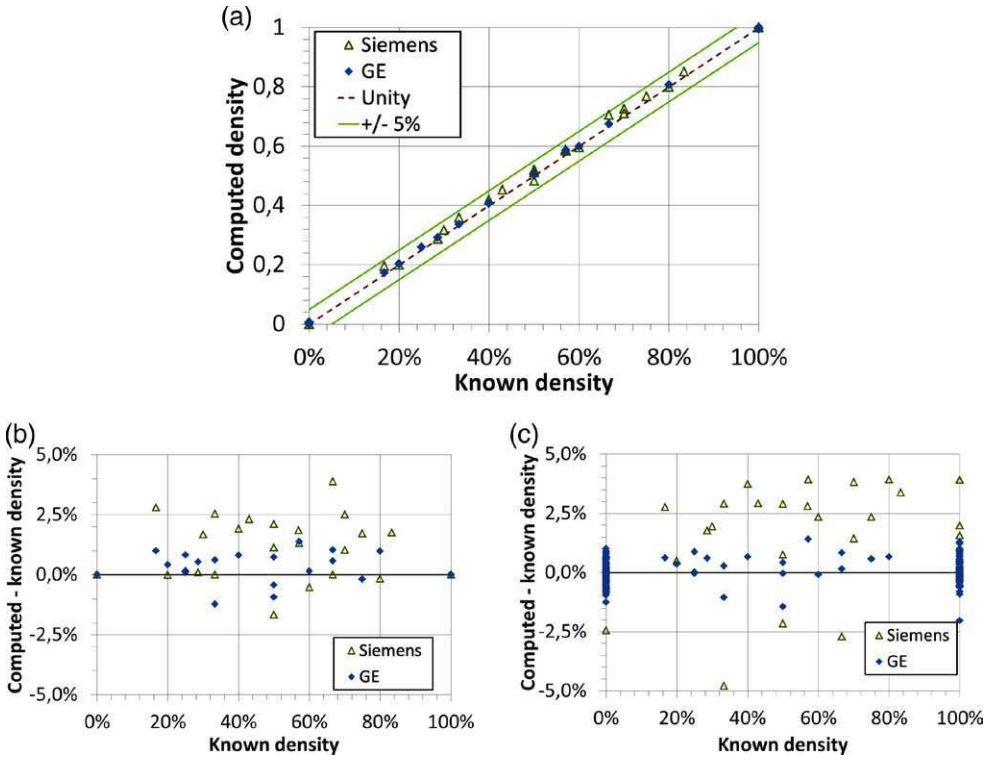

Figure 5. Verification of equation (2). (*a*) The VBD_{MX} computed with the measured $\ln(p/m)_A$ and $\ln(p/m)_G$ as a function of the *VBD* of the phantom. (*b*) The absolute difference between the VBD_{MX} computed with the measured $\ln(p/m)_A$ and $\ln(p/m)_G$ and the known *VBD* as a function of known *VBD*. (*c*) The absolute difference between the VBD_{MX} computed with the modeled $\ln(p/m)_A$ and $\ln(p/m)_{G-A}$ and the known *VBD* as a function of known *VBD*.

The impact of variations of the input parameter values (breast thickness, tube potential, current-time product and detector gain) separately was analyzed. The error on the *VBD* can be considered as a linear function of the errors on the input parameters. The deviations generating a 5% *VBD* error are given in table 3.

3.2. Calibration of the CT method

Sufficiently large homogeneous regions could be found in the thorax CT images except for one patient for adipose tissue and two patients for fibroglandular tissue. The HU_A and HU_G values averaged over all patients were found to be −109 and +13, with standard deviations 9.1 and 11.9 respectively.

3.3. VBD_{MX} and VBD_{CT} for the databases

The VBD_{MX} computation method was applied to Database 1 (663 images). Figure 6 represents the distribution of the VBD_{MX} for all images. The maximum of the distribution is found at 10% and the median density of the population is 14.7% . The skewness is 1.55. Our results

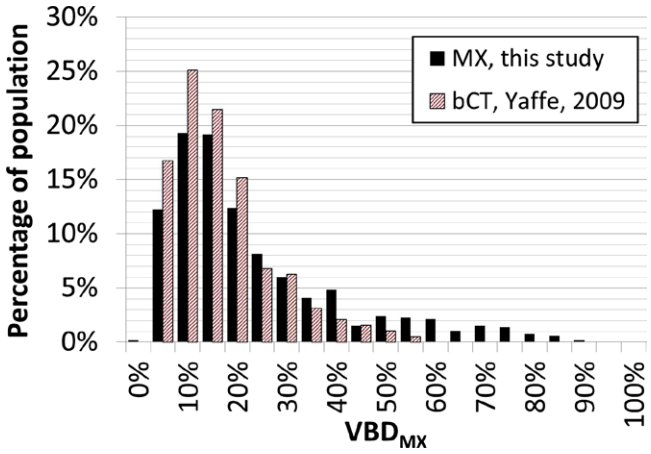

Figure 6. The distribution of the *VBD* computed for a large population (Siemens (560 images) and GE (103 images)) compared to the distribution published by Yaffe *et al* (2009). Both are skewed distributions with their maximum at 10%.

Table 3. The errors on the input parameters generating a 5% error on VBD.

Input parameter	Deviation generating a 5% *VBD* error
Breast thickness	2.1%
Tube potential	0.8%
Current-time product	4.6%
Detector gain	4.8%

were compared with the results obtained by Yaffe *et al* (2009), who segmented automatically fibroglandular tissues in dedicated breast CT images to compute the *VBD*. They also found a skewed distribution with a maximum at 10% and a skewness of 1.68.

Figure 7 shows the average *VBD* per thickness category as a function of compressed breast thickness. It is decreasing with increasing compressed breast thickness. Dance *et al* 2000 published this decreasing trend for the local breast density based on computations from the exposure parameters of film-screen mammographic images acquired under automatic exposure control (AEC). The AEC sensor was supposed to be manually placed over a dense region of the breast, so the resulting *VBD* is local and more representative of higher densities in the breast. We therefore computed the maximum local breast density manually for a subset of 129 images of Database 1 to compare to Dances results (see figure 7).

We have then compared our VBD_{CT} data, based on the patient calibration, to the results of volumetric glandular fraction (VGF) published by Vedantham *et al* Vedantham *et al* (2012). The characteristics of the distributions (see table 4) are in good agreement, despite the small number of cases (25 for $hosp_L$ and 27 for $hosp_J$ versus 150 for Vedantham *et al*). The distribution of our measurements is compared to the distributions published by Yaffe *et al* (2009) and Vedantham *et al* (2012) in figure 8, all excluding the skin.

The correlations between CC and MLO images for VBD_{MX} and between left and right breasts for VBD_{MX} and VBD_{CT} are shown in figures 9–11. The good correlation coefficients and the slope of the graphs close to 1 confirm the confidence in the method.

In figure 12 some examples of breast density maps for eight different women are shown together with their exclusion map indicating in black the pixels with density below 0% and

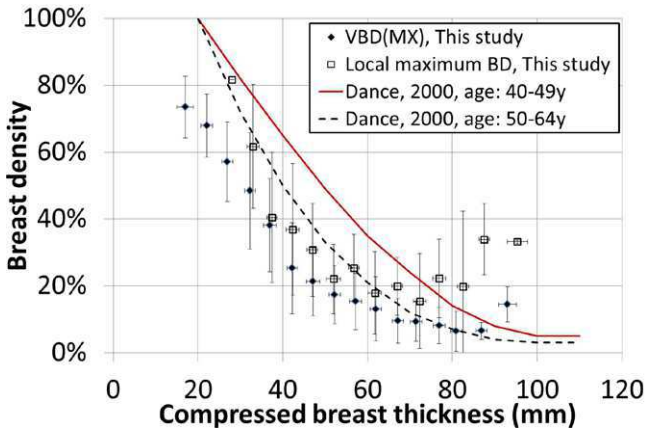

Figure 7. The average VBD_{MX} per thickness category computed for Database 1 (full dots), as well as the maximum local breast density for the subset of 129 GE images (empty dots), compared to the local breast density described by Dance *et al* (2000).

Table 4. Distribution characteristics of the VBD_{CT} in comparison with the volumetric glandular fraction (VGF), published by Vedantham *et al* (2012).

	Our dataset		Vendantham
	$hosp_L$	$hosp_J$	*et al* (2012)
Mean	0.18	0.10	0.17
Minimum	0	–0.02	0.01
Median	0.14	0.07	0.14
Maximum	0.51	0.63	0.72
First quartile	0.02	0.03	0.07
Third quartile	0.33	0.12	0.24

over 100%, excluded for the VBD_{MX} computation. In figure 12(*b*) the white lines in the density maps indicate the exclusion from the pectoral muscle estimation. Comparable findings were reported by Alonzo-Proulx *et al* (2010b) and Zoetelief (*et al* (2006).

3.4. Correlation between VBD_{CT} and VBD_{MX}

We checked first how good the correlation is of the volume obtained from the mammographic images compared to the volume obtained from the delineated thorax CT images figure 13. The correlation could be improved, but we have good reasons to believe that a volume match is not critical to find a *VBD* correlation (Bakic *et al* 2009). We thus plotted the correlation of the VBD_{MX} with *HU* in figure 14, with the characteristics summarized in table 5. Each dot of the graph represents a mammographic image, with on the *x*-axis the VBD_{MX} and on the *y*-axis the average *HU* of the corresponding breast delineated in the CT images. A linear regression is applied to the points of both Databases 2 and 3 separately. The solid line represents the expected *HU* on the *y*-axis based on equation (5) and the values of HU_A and HU_G measured in the thorax CT images for the given *VBD* on the *x*-axis. Figures 14(*b*) and (*c*) show separately the MLO images, respectively the CC images. It can be seen that the linear regression in Database 3 is in closer agreement with the expected curve than the linear regression in Database 2, and that the linear regression on the MLO images of Database 3 is in closer

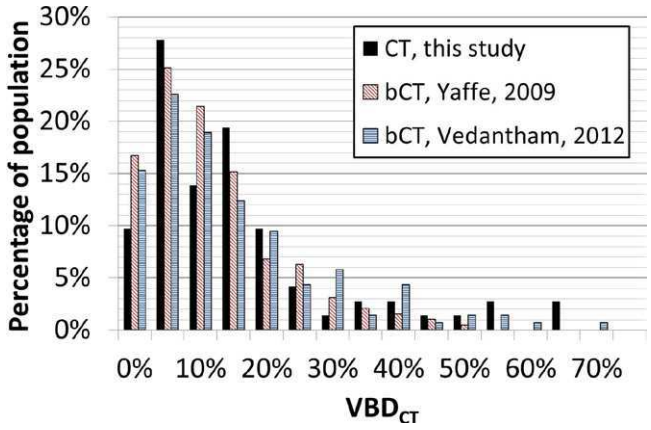

Figure 8. Distribution of VBD from CT measurements (our data) and breast CT (Yaffe *et al* (2009) and Vedantham *et al* (2012), all excluding the skin.

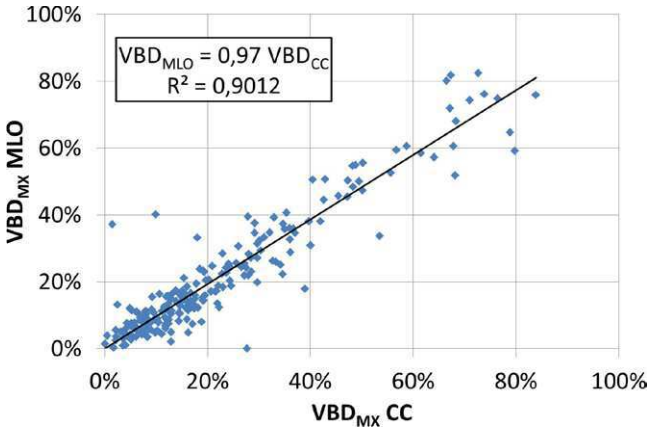

Figure 9. Correlation between CC and MLO VBD_{MX} for the images of Database 1.

agreement with the expected curve than the linear regression on the CC images. The error for a point is computed as the shortest distance of the point to the expected curve. The value of Δ_1 is the average of the errors for all points per database. Table 5 gives an overview of the slopes, intercepts, correlation coefficients R^2 and Δ_1 for the calibration and the least square fits.

4. Discussion and conclusion

The original motivation for the present study was the need to validate an algorithm for the *VBD* computation from mammographic images. We first verified the algorithm on the images of the phantoms used for calibration, then applied the algorithm to patient images. Finally we compared these results to previously published statistics. To extend the field of validations, we have proposed a method using routine thorax CT images.

When applied to phantoms of known density, the accuracy of the presented VBD_{MX} method can be compared to results of Highnam *et al* in (2010): the authors obtained an average error

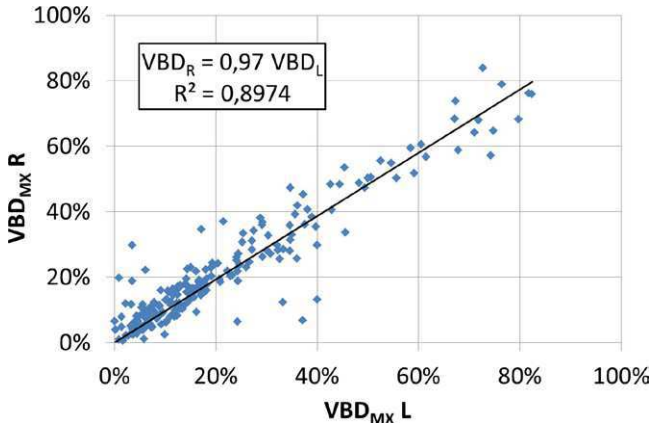

Figure 10. Correlation between left and right VBD_{MX} for the images of Database 1.

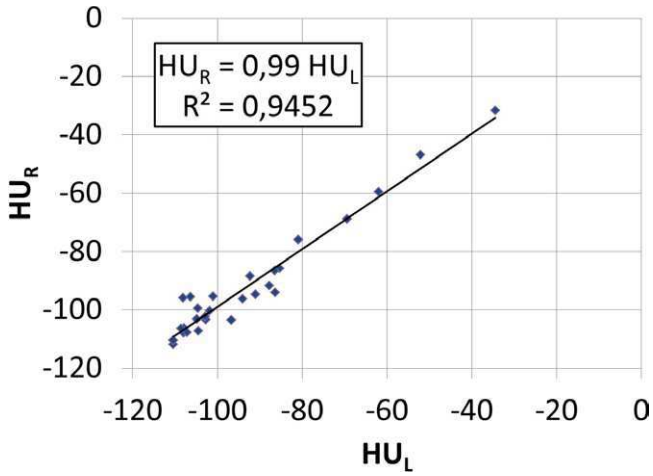

Figure 11. Correlation between left and right VBD_{CT} for the images of Database 2 and Database 3.

on a GE Essential of 1.11% for 25 measurements over a density range from 0% to 37.5% for Molybdenum–Rhodium 28 kV, Molybdenum–Molybdenum 26 kV and 28 kV spectra. If we restrict our results to that range of spectra we obtained an average error of 0.1%.

When applied to patient images, our VBD_{MX} computation method gave population distributions comparable to those obtained from other methods based on breast CT's (Yaffe *et al* 2009, Vedantham *et al* 2012) as shown in figure 6. Using a different method Beckett and Kotre (2000) studied breast densities of large sets 1258 women of patient images. They computed the *VBD* at the position of the AEC-sensor, manually placed by the radiographer in an area supposed to be representative of a dense part of the breast. Dance *et al* (2000) used those results to obtain the local *VBD* as a function of compressed breast thickness for the screening populations which has been proposed in the European Guidelines for Quality Control (Perry *et al* 2006). Since these methods computed the density in a dense part of the breast, an overestimation of the average *VBD* was expected, and demonstrated in figure 7.

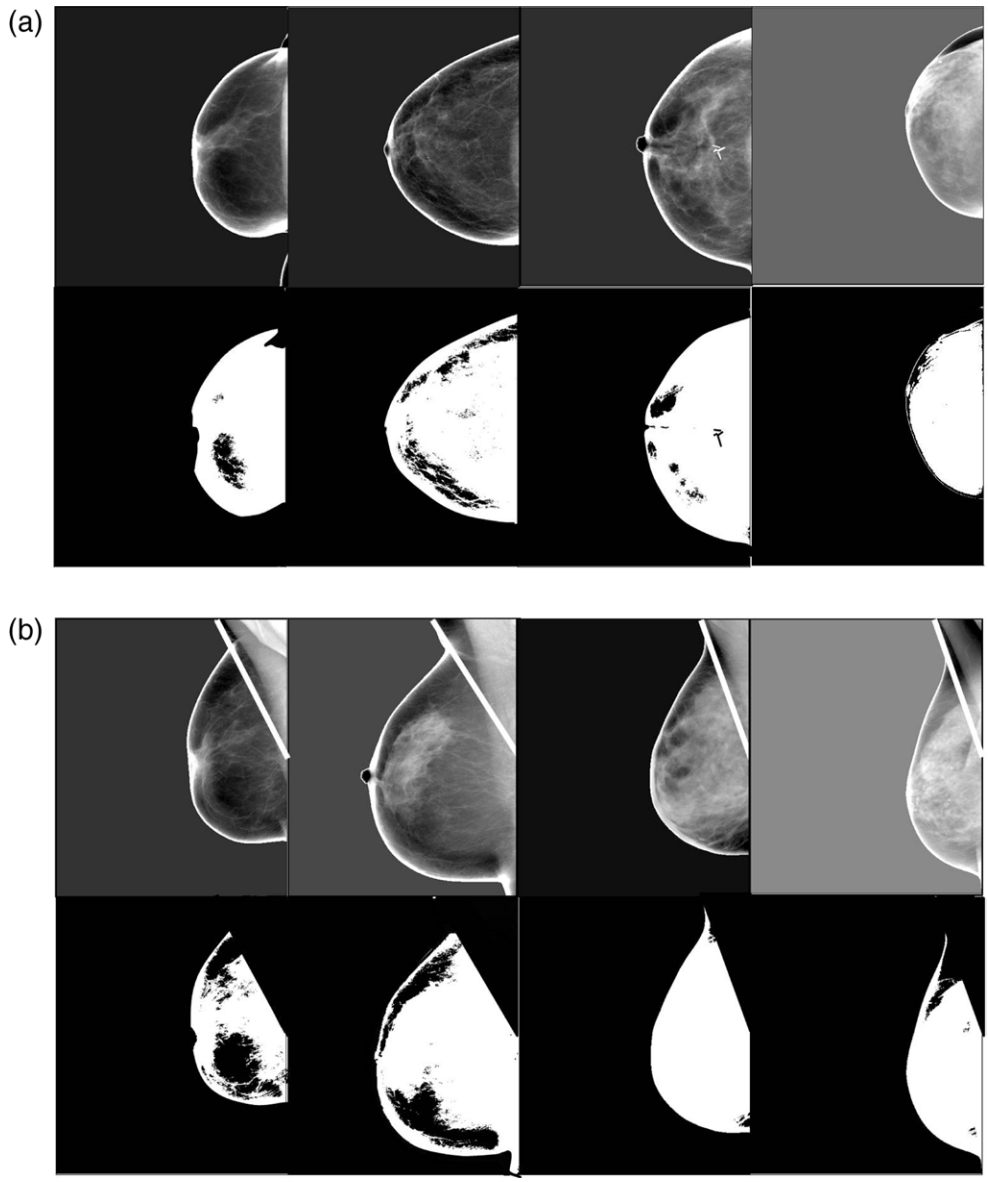

Figure 12. Examples of breast density maps (up) and their corresponding exclusion maps (down). (*a*) CC (*b*) MLO. The black areas of the exclusion maps are excluded from the VBD_{MX} computation. On MLO images the diagonal bars indicate the pectoral muscle delineation.

The assessment of *VBD* from mammographic images, however, is penalized by several factors. First of all, the thickness of the breast is known with a limited accuracy. This is even a concern for the compressed part (Hauge *et al* 2012). Secondly, we supposed a two-compartment breast with adipose and fibroglandular tissues. In order to take the skin into account, a general correction was made, with the thickness of the skin assumed to be 1.5 mm at both sides of the breast. This is however based on a standard value instead of a value derived from

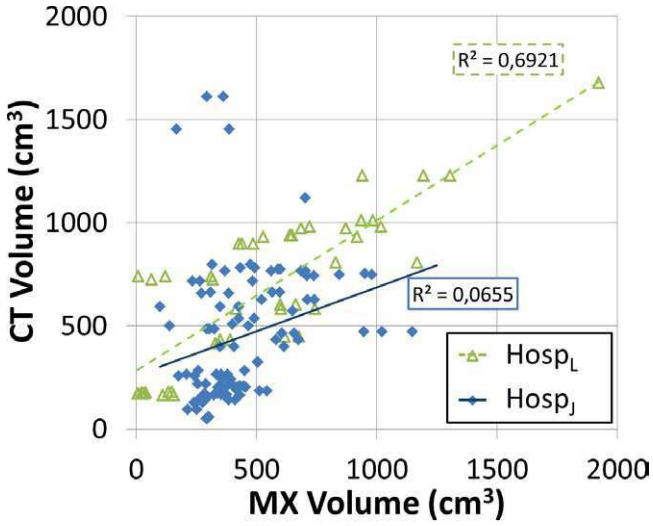

Figure 13. Volumes of the breast in mammographic images compared to the volumes of the same breast measured in CT images.

the individual patient data. Finally, the VBD_{MX} method relies on the correspondence in attenuation coefficients between the phantoms and real breast tissue.

We built a database of patients who underwent both a thorax CT and a mammographic exam within one year. In large hospitals, over 5000 CT exams per month are performed. The chance for a woman at screening age undergoing a thorax CT exam also passing a screening mammographic exam within one year is approximately equal to the participation rate to the breast screening. We found over 30 patients per month, resulting in over 100 mammographic images for which 3D *VBD* computation is available. The biggest limitation of collecting VBD_{CT} is the manual delineation of the breast in the CT images, which is time-consuming. The good correlation of the VBD_{MX} with the VBD_{CT} (table 5) compared to the weaker correlation of the mammographic and CT volumes figure 13 demonstrated that the thorax CT method is robust against errors in the determination of the breast volume. However the total breast volume in CT segmentation can be made easier using semi-automatic methods as described by Salvatore *et al* (2014).

The slope of the trend line for the $hosp_L$ images in figure 14 is different from the one for the $hosp_J$ images. The trend line for the MLO images of $hosp_J$ figure 14(*b*) is in closer agreement with the expected curve than for the CC images figure 14(*c*). For $hosp_L$ there is no difference between the MLO and CC images. There are several limitations that can cause these deviations. First of all, the CT images were delineated manually by two different physicists, which can cause a systematic bias due to a different interpretation of breast limits. Secondly, the mammographic images were acquired by two different teams of radiographers on different equipment. Due to different positioning techniques, this can cause a systematic difference in the imaged volumes between CC and MLO images. Systematic differences in positioning do not affect the correlation, but only the slope and intercept of the least square fits. Missing adipose tissue (as in CC images) in a mammographic image would increase the measured VBD_{MX} and decrease the slope of the least square curve in figure 14. In CT, it was difficult to accurately determine the limits of the breast in each slice as well as the first and last slices of series containing breast tissue.

We compared the results in figure 14 to the results of the correlation between the Volpara® software and MR (figure 2 in Highnam *et al* (2010)). The range of values in their data cloud at 25% VBD_{MX} was between 17% and 40% VBD_{MR}. For our method, we found a range between

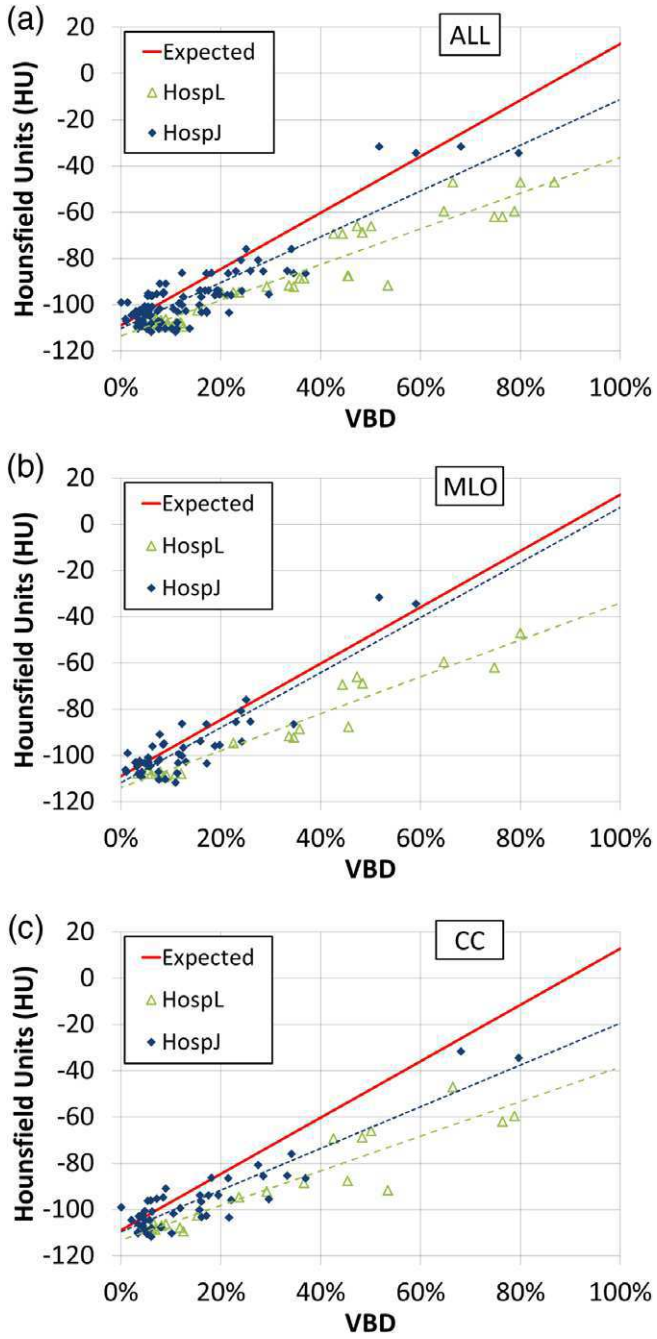

Figure 14. The average *HU* over the breast volume in thorax CT plotted versus VBD_{MX}. Each dot represents a breast for which the average *HU* is available (Databases 2 and 3) together with a VBD_{MX} estimation. They are represented separately for the two hospitals, with the corresponding linear regressions. The solid line represents the expected *HU* based on equation (5) and calibration values of HU_A and HU_G. (*a*) Grouped CC and MLO mammographic images, (*b*) points relative to MLO mammographic images only, (*c*) points relative to CC mammographic images only.

Table 5. Characteristics of the graphs presented in figure 14.

	Slope	Intercept	R^2	Δ_1 (%)
Expected curve	122	−109	—	—
$hosp_L$ (ALL)	77	−113	0.91	10
$hosp_J$ (ALL)	99	−110	0.78	4
$hosp_L$ (MLO)	80	−113	0.93	10
$hosp_J$ (MLO)	119	−112	0.80	3
$hosp_L$ (CC)	75	−113	0.86	10
$hosp_J$ (CC)	90	−110	0.83	5

4% and 32% VBD_{CT}. This suggests that the comparison between VBD_{MX} and a 3D imaging method as VBD_{CT} or VBD_{MR} is limited by comparable problems.

This opens the question of the determination of the 'true' breast volume and as a consequence the overall volumetric breast density. Anatomically however the limits of the breast as an organ are not clear. Radiologists tend to consider the content of the mammographic images as setting the breast limits, even if different in CC and MLO views. This problem emphasizes the limitations of the volumetric breast density as a quantitative risk factor. We must keep in mind that the concept of breast density was originally introduced based on the appearance of film images, including the texture of the tissue (Wolfe 1976). The same concept was used in most studies on associated risk, and formalized in the standardized reporting method (McCormack and dos Santos 2006). It might therefore be completed with for example the absolute volume of the glandular tissue, or the texture in the unprocessed image as was explored by several groups (Byng *et al* 1996, Huo *et al* 2000, Reiser *et al* 2011). These analyses can benefit from quantitative images calibrated to local *VBD* values (Highnam *et al* 2006, Fowler *et al* 2013).

In this study we demonstrated the possibility of validating VBD_{MX} computation methods by correlating them with *VBD* values from routine thorax CT exams. Therefore we have built a database of mammographic images of breasts for which corresponding CT series were available and showed a good correlation of the VBD_{MX}, computed with a state of the art method, to the *HU* measured in CT images. The correlation between the *VBD* computed from mammography and the *VBD* computed from 3D-imaging as CT, as shown in this study, shows the possibility to compute a volumetric quantity representing the dense tissue in the breast and its 2D distribution, which can be used for texture analysis and for possible dose applications.

Acknowledgments

We would like to thank the group of medical physics at UZ-KU Leuven for its valuable inputs and experiments and in particular Marieke Durnez for testing the method on the Siemens Inspiration in UZ Leuven. We would like to thank the physicists at Jolimont hospital for their flexibility in accessing the system and for their help with the data, and the GE Healthcare Field Engineer for his support.

This work was done within the frame of the French ANRT grant CIFRE 2011/0416.

References

Alonzo-Proulx O, Packard N, Boone J M, Al-Mayah A, Brock K K, Shen S Z and Yaffe M J 2010a Validation of a method for measuring the volumetric breast density from digital mammograms *Phys. Med. Biol.* **55** 3027–44

Alonzo-Proulx O, Mainprize J G, Packard N, Boone J M, Al-Mayah A, Brock K K and Yaffe M J 2010b Development of a peripheral thickness estimation method for volumetric breast density

measurements in mammography using a 3d finite element breast model *Int. Workshop on Digital Mammography* vol 6136, ed J Martí *et al* (Berlin: Springer) pp 467–73
Bakic P R, Carton A K, Kontos D, Zhang C, Troxel A B and Maidment A D A 2009 Breast percent density: estimation on digital mammograms and central tomosynthesis projections *Radiology* **252** 40–9
Beckett J R and Kotre C J 2000a Dosimetric implications of age related glandular changes in screening mammography *Phys. Med. Biol.* **45** 801–13
Byng J W, Boyd N F, Fishell E, Jong R A and Yaffe M J 1996 Automated analysis of mammographic densities *Phys. Med. Biol.* **41** 909–23
Dance D R, Skinner C L, Young K C, Beckett J R and Kotre C J 2000 Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol *Phys. Med. Biol.* **45** 3225–40
Desponds L and Klausz R 1994 Automatic estimation of breast composition with mammographic x-ray systems *Annual Meeting of the Radiological Society of North America* (Oak Brook, IL: RSNA)
Deslattes R D 1969 Estimates of x-ray attenuation coefficients for the elements and their compounds *Acta Crystallogr.* A **25** 89–93
Dórsi C J 2003 *BI-RADS: Mammography* 4th edn (Reston, VA: American College of Radiology) pp 179–80
Fowler E E, Sellers T A, Lu B and Heine J J 2013 Breast imaging reporting and data system (BI-RADS) breast composition descriptors: automated measurement development for full field digital mammography *Med. Phys.* **40** 113502
Geeraert N, Klausz R, Giudici P, Cockmartin L and Bosmans H 2012 Dual-energy CT characterization of x-ray attenuation properties of breast equivalent material plates *SPIE Medical Imaging 2012* (*Physics of Medical Imaging* vol 8313) ed N J Pelc *et al* (Bellingham, WA: SPIE)
Hauge I H R, Hogg P, Szczepura K, Connolly P, McGill G and Mercer C 2012 The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units *Med. Phys.* **39** 263–71
Hartman K, Highnam R, Warren R and Jackson V 2008 Volumetric assessment of breast tissue composition from FFDM images *Int. Workshop on Digital Mammography* vol 5116, ed E A Krupinski (New York: Springer) pp 33–9
Heidsieck R 1989 FR 2648229 (A1): Procédé d'étalonnage d'un système radiologique et de mesure de l'épaisseur équivalente d'un objet Patent EP0402244 A2
Highnam R, Brady M and Shepstone B 1996 A representation for mammographic image processing *Med. Imag. Anal.* **1** 1–8
Highnam R, Brady S M, Yaffe M J, Karssemeijer N and Harvey J 2010 Robust breast composition measurement: Volpara *Int. Workshop on Digital Mammography* vol 6136, ed J Mart *et al* (Berlin: Springer) p 342
Highnam R, Pan X, Warren R, Jeffreys M, Smith G D and Brady M 2006 Breast composition measurements using retrospective standard mammogram form (SMF) *Phys. Med. Biol.* **51** 2695–713
Huo Z, Giger M L, Wolverton D E, Zhong W, Cumming S and Olopade O I 2000 Computerized analysis of mammographic parenchymal patterns for breast cancer risk assessment: feature selection *Med. Phys.* **27** 4–12
Kaufhold J, Thomas J A, Eberhard J W, Galbo C E and Gonzalez D E 2002 A calibration approach to glandular tissue composition estimation in digital mammography *Med. Phys.* **29** 1867–80
Kopans D B 2008 Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk *Radiology* **246** 348–53
Li J, Humphreys K, Eriksson L, Edgren G, Czene K and Hall P 2013 Mammographic density reduction is a prognostic marker of response to adjuvant tamoxifen therapy in postmenopausal patients with breast cancer *J. Clin. Oncol.* **31** 1–0
Malkov S, Wang J and Shepherd J 2008 Improvements to single energy absorptiometry. Method for digital mammography to quantify breast tissue density *International Workshop on Digital Mammography* vol 5116, ed E A Krupinski (New York: Springer) pp 1–8
McCormack V A and dos Santos I 2006 Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis *Cancer Epidemiol. Biomarkers Prev.* **15** 1159–69
Perry N, Broeders M, de Wolf C, Törnberg S, Holland R and von Karsa L 2006 *European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis* 4th edn (Luxembourg: Office for Official Publications of the European Communities)
Reiser I, Lee S and Nishikawa R M 2011 On the orientation of mammographic structure *Med. Phys.* **38** 5303–6

Salvatore M, Margolies L, Kale M, Wisnivesky J, Kotkin S, Henschke C I and Yankelevitz D F 2014 Breast density: comparison of chest CT with mammography *Radiology* **270** 67–73

Shepherd J A, Herve L, Landau J, Fan B, Kerlikowske K and Cummings S R 2005a Novel use of single x-ray absorptiometry for measuring breast density *Technol. Cancer Res. Treat.* **4** 173–82

Shepherd J A, Herve L, Landau J, Fan B, Kerlikowske K and Cummings S R 2005b Novel use of single x-ray absorptiometry for measuring breast density *Technol. Cancer Res. Treat.* **4** 173–82

Ting C, Astley S M, Morris J, Stavrinos P, Wilson M, Barr N, Boggies C and Sergeant J C 2012 Longitudinal change in mammographic density and association with breast cancer risk: a case-control study *Int. Workshop on Digital Mammography* vol 7361, ed A D A Maidment *et al* (New York: Springer) pp 205–11

van Gils C H, Hendriks J H C L, Holland R, Karssemeijer N, Otten J D M, Straatman H and Verbeek A L M 1999 Changes in mammographic breast density and concomitant changes in breast cancer risk *Eur. J. Cancer Prev.* **8** 509–15

van Engeland S, Snoeren P R, Huisman H, Boetes C and Karssemeijer N 2006 Volumetric breast density estimation from full-field digital mammograms *Inst. Electr. Electron. Eng.* **25** 273–82

Vedantham S, Shi L, Karellas A and O'Connell A M 2012 Dedicated breast CT: fibroglandular volume measurements in a diagnostic population *Med. Phys.* **39** 7317–28

Wang J, Azziz A, Fan B, Malkov S, Klifa C, Newitt D, Yitta S, Hylton N, Kerlowski K and Shepherd J A 2013 Agreement of mammographic measures of volumetric breast density to MRI *PLOS One* **8** e81653

Wolfe N J 1976 Breast patterns as an index of risk for developing breast cancer *Am. J. Radiol.* **126** 1130–9

Yaffe M J, Boone J M, Packard N, Alonzo-Proulx O, Huang S-Y, Peressotti C L, Al-Mayah A and Brock K 2009 The myth of the 50–50 breast *Med. Phys.* **36** 5437–43

Zoetelief J, Veldkamp W J H, Thijssen M A O and Jansen J T M 2006 Glandularity and mean glandular dose determined for individual women at four regional breast cancer screening units in The Netherlands *Phys. Med. Biol.* **51** 1807–17