Estimation of Volumetric Breast Density from Digital Mammograms

by

Olivier Alonzo-Proulx

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Medical Biophysics
University of Toronto

© Copyright by Olivier Alonzo-Proulx 2014
Abstract

Estimation of volumetric breast density from digital mammograms

Olivier Alonzo-Proulx

Doctor of Philosophy

Department of Medical Biophysics, University of Toronto

2014

Mammographic breast density (MBD) is a strong risk factor for developing breast cancer. MBD is typically estimated by manually selecting the area occupied by the dense tissue on a mammogram. There is interest in measuring the volume of dense tissue, or volumetric breast density (VBD), as it could potentially be a stronger risk factor. This dissertation presents and validates an algorithm to measure the VBD from digital mammograms. The algorithm is based on an empirical calibration of the mammography system, supplemented by physical modeling of x-ray imaging that includes the effects of beam polychromaticity, scattered radiation, anti-scatter grid and detector glare. It also includes a method to estimate the compressed breast thickness as a function of the compression force, and a method to estimate the thickness of the breast outside of the compressed region. The algorithm was tested on 26 simulated mammograms obtained from computed tomography images, themselves deformed to mimic the effects of compression. This allowed the determination of the baseline accuracy of the algorithm. The algorithm was also used on 55,087 clinical digital mammograms, which allowed for the determination of the general characteristics of VBD and breast volume, as well as their variation as a function of age and time. The algorithm was also validated against a set of 80 magnetic resonance images, and compared against the area method on 2688 images. A preliminary study comparing association of breast cancer risk with VBD and MBD was also performed, indicating that VBD is a stronger
risk factor. The algorithm was found to be accurate, generating quantitative density measurements rapidly and automatically. It can be extended to any digital mammography system, provided that the compression thickness of the breast can be determined accurately.
Acknowledgments

I would like to thank the following people who helped, encouraged and supported my throughout my Ph.D. degree:

- My supervisor Dr. Martin Yaffe and members of my supervisory committee Dr. Stuart Foster and Dr. Jean-Philippe Pignol;
- James Mainprize for his invaluable help in Matlab coding and in debugging my brain;
- Gordon Mawdsley who imparted to me a no-nonsense, practical approach to physics;
- My fellow graduate students Mellisa Hill and Gang Wu for their help and support;
- All members of the Yaffe lab, past and present, who contributed to a pleasant and dynamic work environment;
- My research collaborators from Sunnybrook Research Institute, University Health Network, University of California at Davis and University of Virginia;
- Samuel Richard for his friendship and advice as a senior graduate student;
- My partner Maude Parent for loving and supporting me all those years;
- My friends and family for their continued love and support.
Table of Contents

Acknowledgments........................................................................................................ iv
Table of Contents........................................................................................................ v
List of Tables ................................................................................................................. viii
List of Figures ............................................................................................................... xi
List of Abbreviations .................................................................................................... xxi
List of Publications Included as Part of the Thesis ......................................................... xxiii
Chapter 1 Introduction .................................................................................................. 1

1.1 Motivation ............................................................................................................... 1
  1.1.1 Breast Cancer .................................................................................................... 1
  1.1.2 Breast Tissue ................................................................................................... 2
1.2 Mammography ....................................................................................................... 3
1.3 Mammographic breast density ............................................................................... 5
1.4 Mammography Systems ....................................................................................... 8
  1.4.1 Attenuation and Scatter .................................................................................. 9
  1.4.2 X-ray Production .......................................................................................... 10
  1.4.3 Breast Compression ....................................................................................... 12
  1.4.4 Anti-Scatter Grid .......................................................................................... 12
  1.4.5 Screen-Film Mammography ......................................................................... 14
  1.4.6 Digital Mammography ................................................................................ 15
  1.4.7 Image Quality ............................................................................................... 18
1.5 Volumetric Breast Density .................................................................................... 21
  1.5.1 Limitations of Mammographic Breast Density ............................................. 21
  1.5.2 Volumetric Breast density in Mammography .............................................. 24
1.6 Overview of Thesis ............................................................................................... 26
1.6.1 Hypothesis and Aims ................................................................. 26
1.6.2 Outline of Thesis ................................................................. 27

Chapter 2 .................................................................................. 28

2.1 Introduction ........................................................................... 28
2.2 Calibration Approach .......................................................... 29
   2.2.1 Phantom Calibration ....................................................... 29
   2.2.2 Limitations of the Calibration Method ............................. 34
2.3 Modeling Approach .............................................................. 39
   2.3.1 X-Ray Propagation ......................................................... 39
   2.3.2 Scatter Simulation ......................................................... 42
   2.3.3 Glare Characterization ............................................... 43
   2.3.4 Anti-Scatter Grid Characteristics ................................. 48
2.4 Simulation Results ............................................................... 53
2.5 Thickness Sensitivity ........................................................... 57
2.6 Summary ............................................................................. 60

Chapter 3 .................................................................................. 62

3.1 Introduction ........................................................................... 62
3.2 Material and Methods .......................................................... 64
   3.2.1 Breast CT ................................................................. 64
   3.2.2 Deformation .............................................................. 67
   3.2.3 Mammogram Simulation ............................................ 68
3.3 Results ............................................................................... 72
   3.3.1 CT Analysis and Deformation ...................................... 72
   3.3.2 Mammogram Simulation ............................................ 74
   3.3.3 Comparison of the VBD Algorithm with the CT Data .... 76
3.4 Discussion and Conclusion ................................................. 79
Chapter 4 ............................................................................................................................................. 83

4.1 Introduction ................................................................................................................................. 83

4.2 Thickness Estimation .................................................................................................................. 84

  4.2.1 Compression Thickness Correction ...................................................................................... 84

  4.2.2 Peripheral Thickness Correction ......................................................................................... 90

4.3 Volumetric Breast Density Characteristics of a Large Clinical Image Set ......................... 98

  4.3.1 Methods ................................................................................................................................. 98

  4.3.2 Results ................................................................................................................................. 101

  4.3.3 Discussion ............................................................................................................................. 109

Chapter 5 ........................................................................................................................................... 115

5 ..................................................................................................................................................... 115

5.1 Introduction ................................................................................................................................. 115

5.2 Comparison of the Density Algorithm with Magnetic Resonance Imaging ....................... 115

  5.2.1 Segmentation of Dense Tissue on MR Images ..................................................................... 116

  5.2.2 Results ................................................................................................................................. 119

  5.2.3 Discussion ............................................................................................................................. 122

5.3 Comparison between the Risk Factors Associated with VBD and with the Manual Area Density Method .................................................................................................................. 127

  5.3.1 Area Density Measurement ................................................................................................. 127

  5.3.2 Comparison between the Area and Volume Density .......................................................... 133

  5.3.3 Risk Calculation ................................................................................................................... 135

5.4 Summary and Final Words ......................................................................................................... 140

References ........................................................................................................................................... 142
List of Tables

Table 1-1: Risk factors for breast cancer. For small probabilities, the relative risk equals the odds ratio. .................................................................................................................................................. 8

Table 1-2: Values of the parameter $\mu/\Delta\mu$ as a function of density $m$ and x-ray energy. ............... 25

Table 1-3: List of published work relevant to each aim and chapter..................................................... 27

Table 2-1: glare parameters for three similar CsI detectors. The values in parenthesis correspond to the errors in the fitted parameters. ..................................................................................... 47

Table 2-2: Primary transmission factor $T_P$ for a Senograph DS as a function of thickness, kV and anode/filter for a plastic breast phantom of 30% density. The experimental error was on average ± 0.01........................................................................................................................................ 53

Table 2-3: sensitivity in percent density per mm as a function of thickness, for a uniformly thick breast imaged at 29 kV with a Rh/Rh beam. The sensitivity is computed as the slope of the relation between $\Delta m$ and $\Delta T$, where $\Delta T$ ranges from – 5 to 5 mm..................................................... 60

Table 3-1: Elastic properties of breast tissue. From ref. [105]. .............................................................. 67

Table 3-2: Correspondence between the spectrum and breast thickness............................................. 71

Table 3-3: Sensitivity $S_s$ in %VBD/ADU/mAs. ..................................................................................... 76

Table 3-4: Linear regression and rms analysis between VBD$_{CT}$ and VBD$_{mammo}$ (VBD with and without the skin) and between $V_{CT}$ and $V_{mammo}$ (total breast volume $V$ and skin volume $V_{sk}$). 79

Table 3-5: Observed and predicted difference as a function of kV for the Rh/Rh image set ...... 81

Table 4-1: Coefficients of the force response for the hinged paddles. .................................................... 86

Table 4-2: linear least-square fit coefficients for the thickness correction parameter. .................. 87

Table 4-3: Distribution in the paddle and machine for the clinical cohort. ......................................... 88

Table 4-4: thickness correction and compression force data.............................................................. 88
Table 4-5: Mean VBD, breast volume and thickness as a function of mammography machine.. 90

Table 4-6: Comparison between the true volumes and densities and those determined using the peripheral detection algorithm. ................................................................................................................................. 94

Table 4-7: Differences between the true VBD and volumes and those calculated using the peripheral correction algorithm................................................................................................................................. 94

Table 4-8: density measurement in the breast and in the periphery. ........................................... 95

Table 4-9: Summary of the age, VBD, VBD with skin excluded, breast volume, dense volume, skin volume and normalized skin volume (skin volume/breast volume × 100) data from the image set. ......................................................................................................................................................... 103

Table 4-10: Values of the mean, median, rms difference and linear fir coefficients describing the difference between the left and right VBD and breast volume................................................................................................................................. 104

Table 4-11: Mean VBD and number of images in the volume groups........................................... 105

Table 4-12: Mean and median VBD, dense volume and breast volume, and number of images for each age groups. The BMI was obtained from ref. 115................................................................. 107

Table 4-13: Number of cases, mean VBD difference and mean relative breast volume difference for each time bin. ......................................................................................................................................................... 108

Table 4-14: Imaging techniques used for the 55 087 images of the study. ................................. 109

Table 4-15: Comparison of the percent breast density and breast volume between this work and those measured with 3D imaging techniques. The values are the mean ± standard deviation. .. 109

Table 4-16: Mean density and volume measurement done on mammograms found in the literature. Some values are estimated from the published data................................................................. 113

Table 5-1: Summary of the patients’ age and time difference between the MR and mammography exams. ......................................................................................................................................................... 120
Table 5-2: Comparison of the measured density and volume between MR and mammography. .................................................................................................................. 121

Table 5-3: RMS difference, correlation and the linear least square fit coefficients between the MR and mammography measurements. .................................................................................................................. 122

Table 5-4: RMS difference, correlation and linear least-square fit coefficients between the MR and mammography VBD values and the ABD values........................................................................................................ 122

Table 5-5: RMS difference and Pearson correlation of breast density and breast volume, between the left and right breast for the MR and mammography results. ............................................. 123

Table 5-6: Results of the comparison of the ABD measurement between the repeats and with an experienced reader. ICC stands for intraclass correlation. *For the last row, we removed 4 outliers in the data. .................................................................................................................. 128

Table 5-7: Results of the comparison between the repeat images. The first three rows are for the entire repeat set of 350 images. The last two rows are for the intra-read (N=300) and inter-read (N=50) repeat images.................................................................................................................. 129

Table 5-8: Linear fit coefficients, correlation value and RMS difference between the ABD measurement from the different images. .................................................................................................................. 132

Table 5-9: Mean and standard deviation of the total area, dense area and area density for the three image types. ........................................................................................................................................ 132

Table 5-10: Age (in years) distribution of the case and control groups ........................................ 136

Table 5-11: Lower and upper quintiles for the VBD and ABD for the control population (N = 2188)........................................................................................................................................ 137

Table 5-12: Linear parameters and correlations between readers of ABD .................................................................................................................. 139
List of Figures

Figure 1.1: Breast anatomy. 1. chest wall; 2. pectoralis muscle; 3. lobules; 4 nipple; 5. aerola; 6. ducts; 7. fatty tissue; 8. skin. Image credits: Patrick J. Lynch; illustrator; C. Carl Jaffe; MD; radiologist; Yale University Center for Advanced Instructional Media................................. 2

Figure 1.2: X-ray linear attenuation coefficients of fibroglandular, adipose, and ductal carcinoma as a function of x-ray energy. From [28] .................................................................................. 5

Figure 1.3: BIRADS density categories. From left to right: almost entirely fat (<25% density by area); scattered fibroglandular densities (25-50%); heterogeneously dense (50-75%), hard to see small masses; and extremely dense (>75%), can lead to missing cancers................................. 7

Figure 1.4: Schematic of a mammography system .................................................... 9

Figure 1.5: Example of spectra in mammography. The mean energy for the Mo/Mo and Rh/Rh spectra are 17 and 19 keV, respectively................................................................. 11

Figure 1.6: Schematic of an anti-scatter grid................................................................. 13

Figure 1.7: Characteristic response to exposure of radiographic film ....................... 15

Figure 1.8: Indirect detector. The light is emitted isotropically. The closer the x-ray absorption is from the detector array, the less distance the light will travel laterally............................. 16

Figure 1.9: Direct detector .......................................................................................... 17

Figure 1.10: Typical readout array in a digital detector. The detector element can be an amorphous silicon photosensor or an electrode................................................................. 17

Figure 1.11: Thresholding method to estimate mammographic density. The left image is the mammogram (negative logarithm of the raw image). The middle and right images represent two thresholds levels leading to an area density of 28% and 39%, respectively. As we can see, both thresholds lead to a reasonable delineation of the dense tissue, yet the area density measure differs by an absolute value of 11 %. The thresholds levels differed by 5 units on a scale of 1 to 256 (2%).................................................................................................................. 22
Figure 2.1: Image of the calibration phantom showing the fat and fibroglandular cylindrical inserts. The phantom in this image is 3 cm thick.

Figure 2.2: Effective attenuation for $f_0$ (fat, dashed line) and $f_1$ (fibroglandular, solid line) as a function of thickness, for the Rh/Rh anode/filter combination on a GE Essential unit. The dashed and solid lines represent the polynomial fits for $f_0$ and $f_1$, respectively, and the data markers represent the experimental calibration data. For each of the two groups of curves, the kVs (increasing from bottom to top) are 28, 30, 32 and 34.

Figure 2.3: Effective attenuation for fat (left) and fibroglandular (right) as a function of kV, for the Rh/Rh anode/target combination on a GE essential unit. The dashed and solid lines represent the polynomial fits and the data markers represent the experimental calibration data. For the two groups of curves, the thickness increases from top to bottom in 1 cm steps up to 8 cm.

Figure 2.4: $S_0$ for a GE essential unit, as a function of kV for the Rh/Rh and Mo/Rh anode filter combination. The lines represent the polynomial fit and the data markers represent the experimental data points.

Figure 2.5: Histogram of the true phantom density minus the calculated density, for the entire calibration image set. The mean difference is -0.9 %VBD, with a standard deviation of 3.1 %VBD.

Figure 2.6: Mean density difference (calculation minus truth), as a function of anode/filter combination (top left), kV (top right), thickness (bottom left) and density (bottom right). For each data point, the absolute density error is averaged for all the parameters (kV, target/filter, thickness and density) except for the parameter on the x-axis. The errors bars correspond to ± the standard deviation in the density difference.

Figure 2.7: Histogram of the true phantom density minus the calculated density (using the location-specific $S_0$), for the entire calibration image set. The mean difference is -0.2 %VBD, with a standard deviation of 2.2 %VBD.

Figure 2.8: Attenuation of breast tissue and of breast plastic phantoms. From Byng et al. [91].
Figure 2.9: MTF of a GE 2000D system. Note the drop in the MTF, of approximately 20%, in the low frequencies, up to approximately 0.5 mm⁻¹. ................................................................. 44

Figure 2.10: DTF for a GE essential system. The solid line represents a least-square fit to the data, explained in the text below. .................................................................................. 46

Figure 2.11: Frequency response of the glare on a 2000D system (left); effect of glare deconvolution on a lead disk image on a Essential system (right). .................................................. 47

Figure 2.12: Parameters of the DTF (see equation (2-25) and Table 2-1) as a function of kV for the Senograph DS. The circles, squares and dots are for the Mo/Mo, Mo/Rh and Rh/Rh beams, respectively. A single error bar is shown for clarity; all the data points had similar errors. .......... 48

Figure 2.13: Signal (normalized to the signal at the 0.74 mm aperture) as a function of aperture diameter. The raw signal and the signal after glare deconvolution are shown. Note the break in the x-axis. The x-ray beam is 28kV with a Rh/Rh anode/filter, and the machine is a Senograph DS. ............................................................................................................................ 50

Figure 2.14: Ratio of the off-focus radiation to the total incident radiation as a function of kV and for different anode/filter combinations. The circles, squares and dots are for the Mo/Mo, Mo/Rh and Rh/Rh beams, respectively. The upper data points correspond to the raw data while the lower data points are computed after removing the glare. The dotted lines correspond to the average off-focus ratio in each set. .................................................................................. 51

Figure 2.15: Primary transmission factor (top row and bottom left) and scatter transmission factor (bottom right) for a Senograph DS system, as a function of thickness kV and anode/filter. A single set is shown with error bars in each figure for clarity; the errors were similar for the other points. ................................................................................................. 52

Figure 2.16: projection geometry of the simulation........................................................................................................ 54

Figure 2.17: Comparison of the calibrated and simulated effective attenuation on the Essential units, for the Rh/Rh 29 kV beam (left), and the Mo/Mo 25 kV beam (right). ............................... 55

Figure 2.18: Density errors resulting from the differences in the simulated and calibrated effective attenuation as a function of thickness and kV. Top figures show the density difference
for the Rh/Rh beam, for the fat attenuation (left) and fibroglandular attenuation (right). The bottom figures show the differences for the Mo/Mo beam, for the fat attenuation (left) and fibroglandular attenuation (right).

Figure 2.19: Comparison of the plastic effective attenuation (measured) and breast effective attenuation (calculated), for an Essential machine, Rh/Rh 29 kV. FG stands for fibroglandular tissue.

Figure 2.20: Density error $\Delta VBD$ as a function of thickness and density. From bottom to top, the solid line, dotted line, dashed line, circles and squares represent a density of 0, 25, 50, 75 and 100 %, respectively. The left figure is for a 29 kV Rh/Rh beam and the right figure is for a 25 kV Mo/Mo beam.

Figure 2.21: Density error $\Delta VBD$ as a function of the thickness error $\Delta T$ for different density values, for a 5 cm thick breast (left) and for a 2 cm breast (right), both imaged at 29 kV with a Rh/Rh anode/filter. A thin attenuator has a non-linear sensitivity response.

Figure 2.22: Average sensitivity of the volumetric breast density (VBD) measurement per mm of thickness error, for an error range $\Delta T = [-5,5]$ mm. The computations were done assuming a beam of 29 kV with a Rh/Rh anode/filter.

Figure 3.1: Flow chart summary of the validation process. From top to bottom: sagittal slice from a segmented CT volume, sagittal slice of the deformed volume, negative logarithm of the simulated digital mammogram and calculated density map of the mammogram. Used with permission from IOP Science [79].

Figure 3.2: Histogram of the CT attenuation for one case in the data set. We can easily identify the threshold at low CT numbers that separates the air from the breast. We can see that the peaks in attenuation corresponding to the fat and fibroglandular tissue are poorly separated. An initial two-compartment Gaussian fit gave a good estimate of the threshold between the two tissues, but it was adjusted manually to a lower value to obtain a better separation between the two tissue types.
Figure 3.3: Illustration of segmentation of coronal CT slices from two cases (top and bottom). Left: original slice in units of effective attenuation displaced with a narrow window; right: segmented slice. Used with permission from IOP Science [79] ................................................................. 66

Figure 3.4: Illustration of the finite-element model (FEM) deformation process. U represents the magnitude of the displacement in cm. Used with permission from IOP Science [79] .................. 69

Figure 3.5: Illustration of the finite-element deformation. Left: original segmented CT image; right: displaced fractional density image. From top to bottom: sagittal, transverse and coronal slices. The white vertical line represents the cropping to remove the non-compressed part of the breast. Used with permission from IOP Science [79] ................................................................. 73

Figure 3.6: Distribution of ADU/mAs values on simulated and experimental images of a plastic breast phantom of thickness 5 cm (a), and 7 cm (b), imaged at 28 and 30 kV respectively, with a Rh/Rh anode/filter. The distributions are computed over an area of approximately 107 cm$^2$ on the projected area of the phantom. Used with permission from IOP Science [79] .................. 74

Figure 3.7: (a) Comparison between the means of the experimental (solid line) and simulated (dashed line) phantom images as a function of thickness and kV. The marker size is representative of the error in the experimental value. (b) Corresponding difference $\Delta$VBD between experiment and simulation. Used with permission from IOP Science [79] .................. 75

Figure 3.8: Image signal in ADU/mAs as a function of density and thickness, for a 29kV Rh/Rh beam. The sensitivity $S_d$ for each thickness is the inverse of the slopes of the dashed lines. ..... 75

Figure 3.9: Frequency distribution of the VBD difference between the true density of the phantom and the calculated density using the algorithm. The results show are for the Rh/Rh beam. Used with permission from IOP Science [79] ................................................................. 76

Figure 3.10: Comparison between the VBD from CT and from the mammographic breast density algorithm, for the set of images simulated with a Rh/Rh beam. The dashed line is the identity function. Used with permission from IOP Science [79] ................................................................. 77
Figure 3.11: Comparison between the VBD from CT and from the mammographic breast density algorithm, for the set of images simulated with a Mo/Rh beam (a) and a Mo/Mo beam (b). The dashed line is the identity function. Used with permission from IOP Science [79].

Figure 3.12: Comparison between the total breast volume (left) and skin volume (right) from the CT image and the mammographic density algorithm. The dashed line is the identity function.

Figure 4.1: Left: picture of the thickness measurement apparatus and phantom, showing, from top to bottom the height gauge, the perforated plate, the compression paddle, the phantom and the breast support table. Right: schematic of the measured points.

Figure 4.2: Thickness measurements of the phantom for the 19×23 cm semi-rigid paddle (left) and for the 24×30 cm hinged paddle (right). The compression force is 10 dN. The measurements and the average thickness are shown by the grey and white planes, respectively.

Figure 4.3: Thickness (left) and percent slope (right) of the phantom as a function of compression force. The left figure is for the 19×23 cm semi-rigid paddle; the right figure is for the 24×30 cm hinged paddle. The dashed lines represent a linear least-square fit to the data. The error bars originate from the standard deviation in the thickness measurement.

Figure 4.4: Measured thickness of the phantom minus the readout thickness of the mammography machine versus compression force for the semi-rigid 19×23 cm paddle (left) and for the hinged 24×30 cm paddle (right). The dashed lines represent the linear least square fit.

Figure 4.5: VBD as a function of paddle type and mammography unit. Paddles 1, 2, 3 and 4 represent the 19×23 cm semi-rigid paddle, the hinged 19×23 cm paddle, the semi-rigid 24×30 cm paddle and the hinged 24×30 cm paddle, respectively.

Figure 4.6: Breast volume (left) and corrected thickness (right) as a function of paddle type and mammography unit. Refer to Figure 4.6 for the paddle type legend.

Figure 4.7: Illustration of the peripheral detection algorithm. On the left we see the breast image, the outer breast edge, the points $r_0$ (dots), the smoothed approximate periphery $C'$ (solid line), the eroded periphery $C$ (dotted line), the central point $M$ and the radial lines in white. On the
right we see an example of a radial line, along with the baseline linear fit, the detected periphery point $r_0$, the fitting limit $R_c$ and the intensity range $\Delta$.

Figure 4.8: Picture of the plastic phantom with a machined semi-circular peripheral region (left). On the right we see the extracted thickness profile (solid line), as well as the point $(\delta, \beta)$. For comparison, the dashed curve represents the thickness profile for a parallel beam geometry,

$$T = \sqrt{1 - r^2}$$

Figure 4.9: Comparison between the true VBD (left) and volumes (right) with those obtained using the peripheral detection algorithm (left). The dashed line is the identity function.

Figure 4.10: Histogram of the rms difference in thickness within the peripheral region between the true and the calculated thickness.

Figure 4.11: Example of a density map calculated on a mammogram (left). The window and level was adjusted so that density values above 70% are saturated. The right image shows where the points above 100% and below 0% are located, in white and light gray, respectively. The periphery contour at 93% compression thickness is shown by the line in a darker gray.

Figure 4.12: Distribution of the thickness were extrema errors occurred. The left figure shows the under-0% errors and the right figure shows the over-100% errors.

Figure 4.13: Age distribution for the image set. With permission from IOP science [112].

Figure 4.14: VBD (left) and VBD with no skin (right) distributions for the image set. With permission from IOP science [112].

Figure 4.15: Total breast volume (left) and dense volume (right) distributions for the image set. With permission from IOP science [112].

Figure 4.16: Scatter density plot of the relation between VBDsk and VBD (including the skin). The dashed line represents the identity function, and the solid line the linear least-square fit. The gray scale (light to dark) represents a decrease in the density of data points. With permission from IOP science [112].
Figure 4.17: Scatter density plot for the VBD (left) and breast volume (right) of the right breast versus the left breast for 22,619 patients. The dashed line represents the identity function, and the solid line the linear least-square fit. The gray scale (light to dark) represents a decrease in the density of data points. See Table 4-10 for the values of the linear regression and other descriptive parameters. The outliers described above are indicated with a circle. The gray scale (light to dark) represents a decrease in the density of data points. With permission from IOP science [112].

Figure 4.18: Boxplot of the VBD vs. breast volume. The volume groups include the upper bound but not the lower. The mean VBD for each volume group is shown by the dots joined with a line. With permission from IOP science [112].

Figure 4.19: Age dependence of VBD (left) and breast volume (right). The age groups include the upper bound but not the lower. The mean VBD or breast volume of each age group is shown by the dots joined with a line. See Table 4-12. With permission from IOP science [112].

Figure 4.20: Boxplot of the VBD difference (left) and relative volume difference (right) versus time between examinations. The mean difference for each time interval group is shown by the dots joined with a line. See Table 4-12. With permission from IOP science [112].

Figure 5.1: The left image shows original breast in black and the oversampled and smoothed breast edge in as the line. The right image shows the calculated partial volume fraction in greyscale.

Figure 5.2: Illustration of the soft-threshold transformation function. The $I_f$ and $I_{fg}$ values are illustrated by the left and right dashed lines, respectively.

Figure 5.3: Original MR slice (left) and the corresponding segmented image (right) where fibroglandular tissue appears dark and bright, respectively.

Figure 5.4: Left: age distribution of the patients at the time of the mammography and MR exams. Right: distribution of the time difference between the MR and mammography exams.

Figure 5.5: Comparison between the VBD as measured by MR and by mammography. The dashed line represents the identity function.
Figure 5.6: Comparison between the breast volume (left) and dense volume (right) as measured with MR and mammography. The dashed line represents the identity function. .............................. 120

Figure 5.7: Comparison of the VBD from the mammograms (left) and from the MR (right) with the ABD measurement. .................................................................................................................. 121

Figure 5.8: Illustration of the original and segmented MR images containing spurious low intensity regions. The signal decreases towards the top, bottom and chest wall side of the breast, leading to high density in the segmentation. Including as much breast tissue as possible leads to estimates of VBD and volume of 51 % and 1030 cm$^3$, respectively, while excluding the problematic regions lead to VBD and volume estimates of 41 % and 796 cm$^3$ respectively. We favoured the latter method. ............................................................................................................................ 124

Figure 5.9: Density versus MR slice (left), and corresponding digital mammograms (right). The mammographic VBD is 90 % and 79 % for the left and right breast, respectively. As we can see, some slices in the MR volume had density comparable to the mammographic VBD, however some slices near the side had lower density, which brought the average down. ...................... 125

Figure 5.10: Plot of the dense area (left) and area density (right) for the repeat images. The identity function and the linear fit are represented by the dashed and solid lines, respectively. 129

Figure 5.11: Film-like conversion algorithm. The left image is the negative logarithm of the raw mammogram, which is converted into a film-like mammogram (right) using a reference sensitometric curve (bottom). The speed point is indicated by the dot. A low exposure leads to a high digitized value, i.e. a low OD. .......................................................................................................................... 130

Figure 5.12: Film-like ABD vs. the processed ABD (left) and screen-film ABD (right). The dashed line is the identity function and the solid line a linear fit. See table Table 5-8. ............... 131

Figure 5.13: Comparison between the different types of processing on mammograms used to measure area density. In clockwise order beginning at lower right: screen-film image; processed image, negative log. of the raw image; film-like converted image ........................................ 133

Figure 5.14: Comparison between the volumetric and area breast density. The outliers are shown with the circles. .......................................................................................................................... 134
Figure 5.15: Example of an outlier in the ABD vs. VBD relation. From left to right: film-like converted digital image; thresholded image (ABD = 22.4 %, the pectoralis muscle was excluded); volumetric density map (VBD = 72.8 %). In the density map, the background density ranged between 50 and 60 %. ................................................................. 134

Figure 5.16: Left: age distribution of the controls and cancer cases. Right: time of the mammography imaging minus the time of the diagnosis the cancer cases. The mammogram closest in time to the diagnostic date was always chosen. We removed the cases with exactly zero time difference for clarity. ........................................................................................................... 137

Figure 5.17: Example of the age distribution of removed cases in the matched analysis. In this iteration the ORs for VBD and ABD were respectively 2.1 and 1.1.................................................. 138

Figure 5.18: Histograms of the OR for VBD (left) and ABD (right) resulting from the 10 000 iterations in the matched case-control analysis................................................................. 138

Figure 5.19: Comparison between the inexperienced (ABD₁) and experienced (ABD₂, ABD₃) readers in measuring area density on a subset of 105 images from the case-control dataset. The dashed and solid lines represent the identity function and the linear fit, respectively.............. 139
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABD</td>
<td>Area Breast Density</td>
</tr>
<tr>
<td>ADU</td>
<td>Analog to Digital Units</td>
</tr>
<tr>
<td>BIRADS</td>
<td>Breast Imaging-Reporting and Data System</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CC</td>
<td>Cranio-Caudal</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>DBCT</td>
<td>Dedicated Breast Computed Tomography</td>
</tr>
<tr>
<td>DCIS</td>
<td>Ductal Carcinoma In Situ</td>
</tr>
<tr>
<td>DICOM</td>
<td>Digital Imaging and Communication in Medicine</td>
</tr>
<tr>
<td>DQE</td>
<td>Detective Quantum Efficiency</td>
</tr>
<tr>
<td>DTF</td>
<td>Disk Transfer Function</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite Element Analysis</td>
</tr>
<tr>
<td>FFDM</td>
<td>Full-Field Digital Mammography</td>
</tr>
<tr>
<td>ICC</td>
<td>Intraclass Correlation</td>
</tr>
<tr>
<td>MBD</td>
<td>Mammographic Breast Density</td>
</tr>
<tr>
<td>MLO</td>
<td>Medio Lateral Oblique</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MTF</td>
<td>Modulation Transfer Function</td>
</tr>
<tr>
<td>NEQ</td>
<td>Noise Equivalent Quanta</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>RMS</td>
<td>Root Mean Square</td>
</tr>
<tr>
<td>ROI</td>
<td>Region Of Interest</td>
</tr>
<tr>
<td>RR</td>
<td>Relative Risk</td>
</tr>
<tr>
<td>SHSC</td>
<td>Sunnybrook Health Science Centre</td>
</tr>
<tr>
<td>SID</td>
<td>Source to Image Distance</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>TDLU</td>
<td>Terminal Ductal Lobular Unit</td>
</tr>
<tr>
<td>TFT</td>
<td>Thin Film Transistor</td>
</tr>
<tr>
<td>V</td>
<td>Breast volume</td>
</tr>
<tr>
<td>$V_D$</td>
<td>Dense (fibroglanular) volume</td>
</tr>
<tr>
<td>VBD</td>
<td>Volumetric Breast Density</td>
</tr>
</tbody>
</table>
List of Publications Included as Part of the Thesis

Chapters 2 and 3


Chapter 2 and 4

1.1 Motivation

1.1.1 Breast Cancer

Breast cancer is the most common cancer among women in Canada. Approximately 11% of women will be diagnosed with breast cancer in their lifetime, and approximately 3.5% of women will die from breast cancer. In 2012, this corresponded to an estimated 22,700 new breast cancer cases and 5100 breast cancer deaths in Canada; breast cancer deaths are second only to lung cancer deaths in women. Breast cancer occurs in men as well, with approximately 200 new cases and 55 deaths in 2012 in Canada [1]. Thanks to breast cancer screening, and improvements in detection and treatments, the mortality rate has decreased by approximately 35% since 1986 [1]. Moreover, the 5-year survival rate for women is 88% (79% for men), up from 72% in 1975-1986 [2]. Breast cancer can unfortunately recur; the survival rate after initial diagnosis declines steadily to 70% at 20 years post-diagnosis [2].

Breast cancer originates as an unrestrained proliferation of cells in the epithelium of the breast’s glandular tissue (the ducts and lobules). There are multiple subtypes of breast cancer. The most common types are ductal carcinoma in situ (DCIS), infiltrating ductal carcinoma, infiltrating lobular carcinoma and medullary carcinoma. The causes of breast cancer are still largely unknown: only a small fraction (5-10%) of breast cancers can be linked to hereditary factors. In particular, mutations of the BRCA 1 and BRCA 2 genes have been linked to breast and ovarian cancers.

Early detection can significantly improve prognosis; for DCIS, or an invasive lesion less than 2 cm confined to a localized area of the breast, the 5-year survival rate is 96%. If the tumour is 2-5 cm in size and has not spread to the auxiliary lymph nodes, the 5-year survival rate is 86%. For later stage breast cancers, where the tumour is larger than 5 cm with metastases in the lymph nodes or with metastases in the other organs, the 5-year survival is much worse 59% or 26%, respectively [2].
There are several techniques available for the detection and diagnosis of breast cancer: physical breast exam, ultrasound, magnetic resonance imaging (MRI) and mammography. In addition, genetic testing of the BRCA 1 or 2 genes can be done. However, screening and diagnostic x-ray mammography remains the most common imaging method. X-ray mammography is a dedicated breast imaging procedure that produces a 2-D image of the breast with good contrast and high resolution at a low radiological dose to the breast tissue.

1.1.2 Breast Tissue

The breast tissue extends from the clavicle and the axilla to the sternum, covering most of the chest. The primary function of the breast is to secrete milk in order to feed an infant. The breast contains 15-20 lobes, a collection of smaller lobules whose basic unit is the terminal duct lobular unit (TDLU), the gland that produces the milk. A network of branching terminal ducts drain the milk from the TDLUs and converge to larger lactiferous ducts which then drain to the nipple. The glandular tissue is closely supported by fibrous connective tissue, collagen and elastin. In addition, the fibrous Cooper’s ligaments provide the suspensory function in the breast. Cooper’s ligaments run from the clavicle and form a network of connective tissue that radiates from the superficial fascia (the layer of tissue under the subcutaneous fat) to the dermis. The breast also contains adipose tissue, which is present in a layer below the skin and in front of the pectoralis muscle, as well as being distributed more generally within the breast. See Figure 1.1.
Most of the lymph in the breast is drained to the axillary lymph nodes, which are also connected to nodes that drain the pectoral muscle and scapula. The remainder of the lymph is drained to nodes under the sternum to the other breast and to abdominal nodes. Breast cancer can metastasize to other parts of the body by means of the lymphatic system.

Breast tissue changes with hormonal changes and age. At menarche, the breast tissue develops and proliferates. During menstruation, estrogen and progesterone induce proliferation of the glandular tissue. Cell hypertrophy and water retention cause an increase in volume. At the end of the menstrual cycle, deprivation of sexual hormones induces cell death, and the breast tissue returns to its original state. During pregnancy, lactation and breastfeeding, similar hormonal changes occur and the glandular tissue increases in volume, and the breasts are enlarged and firmer. At menopause, the levels of estrogen and progesterone decrease, and the lobular tissue regresses, resulting in a proportionally fattier breast. Moreover, the size and composition of the breast can change with changes in body weight.

1.2 Mammography

Mammography is a dedicated radiographic system for breast imaging. The goal of mammography is in the detection of breast cancer, on the basis of the following types of signs:

- Mass lesions. Tumours can appear as masses of increased density on the image. A spiculated mass with ill-defined margins is typically malignant, while a rounded mass with a regular border is typically benign.
- Microcalcifications. Small deposits of calcium found along the ducts or lobules. They occur naturally with age and are generally benign, but some configurations are indicative of disease, especially DCIS.
- Architectural distortion occurs when a lesion, visible or not, causes the neighbouring normal tissue to contract and distort.
- Asymmetry in the patterns of breast tissue between the left and right mammograms.

Mammography is used for screening (on asymptomatic patients), as well as for investigating symptomatic patients (diagnostic mammography). A diagnostic examination can involve specialized mammography procedures, such as a magnification view, to further investigate a
patient. A radiological diagnosis is confirmed with a surgical or needle biopsy, often guided using mammography.

Generally, a mammography exam consists of 4 views, the cranial-caudal (CC) and medio-lateral oblique (MLO) views, for the left and right breast. The CC is acquired vertically along the head-toe axis, while the MLO is acquired at a 45 degree angle from the vertical direction, in order to include in the image as much of the axilla and pectoralis muscle as possible. Having two views also gives the radiologist some insight in the 3D configuration of the breast tissue.

Screening mammography is generally performed from the age of 40 or 50 years up to 70 years, every 1-2 years. The efficacy of screening mammography is quantified by its sensitivity and specificity. The sensitivity is the probability of the test giving a correct positive cancer diagnosis when the disease is truly present, i.e. the ratio of true positives to the total number of cancer cases, including the false negatives. The specificity is the probability of the test giving a correct negative cancer diagnosis when the disease is truly absent, i.e. the ratio of true negatives to the total number of non-cancer cases, including the false positives. The sensitivity of mammography depends on age, mammography technology, and cancer type. Values generally range from 70% to 85% [3-6]. The specificity of mammography is generally between 90% and 98% [3-6]. The high specificity is partly explained by the low incidence of breast cancer, of approximately 0.5% in a screening population. The recall rate of mammography is relatively high at 8-10%. In Ontario in 2008, for every 1000 women screened, 915 are normal and 85 are abnormal. Seventy-one women (84% of the abnormal screens) are found benign or normal after a non-invasive work-up, while the remaining 14 (16 %) undergo an invasive work-up (needle/core or surgical biopsy), 5 of which (6% of the abnormals) are found to have cancer [7]. A study has shown that the diagnostic performance of digital mammography is superior to that of film screen mammography in breast cancer screening [6]. While the performance of the two in the general screening population was similar, digital mammography performed better for women under 50 years, for women with high breast density (see Section 1.3), and for pre- and perimenopausal women.

There is ample evidence of the benefits of breast cancer screening from a number of randomized controlled trials, showing a reduction in breast cancer mortality due to screening [8-15]. Large trials are difficult to implement and may suffer from contamination in the control group (when
women randomized not to be screened do get screened elsewhere). Moreover, they may differ in the technology used, in the imaging protocol and in the experience of the image readers. Nevertheless, meta-analyses and reviews of the combined trials have shown significant long-term mortality reduction due to mammography screening, from to 20% to 35%, for women aged 40-69 [16-18]. Some studies have claimed that breast cancer screening is ineffective in reducing mortality [19-22], but those claims have generally been refuted [23-27].

### 1.3 Mammographic breast density

Mammographic breast density refers to the appearance and amount of the glandular and fibrous tissue on the mammogram. Because the radiographic properties of the glandular and the fibrous or connective tissue are very similar, they are generally referred to collectively as fibroglandular tissue. Mammography systems are designed to offer optimal contrast between the fibroglandular tissue structures and the fatty tissue, in order to allow the detection of lesions that originate in the glandular tissue. That contrast is due to the difference in the attenuation coefficient between adipose and fibroglandular tissue [28], as illustrated in Figure 1.2. The fibroglandular tissue is more opaque to x-rays than adipose tissue, thus the fibroglandular structures appear as “densities” on the image, in contrast to the translucent appearance of the fatty tissue. The fibroglandular tissue is physically denser than fatty tissue; the x-ray attenuation coefficient increases with the electron density in the tissue, which is in turn related to the physical density.

![Figure 1.2: X-ray linear attenuation coefficients of fibroglandular, adipose, and ductal carcinoma as a function of x-ray energy. From [28].](image-url)
Breast density is strongly related to the quality of the mammogram. As seen on Figure 1.2, the attenuation of cancerous breast tissue is very similar to that of normal fibroglandular tissue, and consequently there is little to no contrast between the two types of tissue. Therefore, normal fibroglandular tissue can mask a lesion by tissue superposition, and that effect is exacerbated in a dense breast. The sensitivity of mammography decreases to 60% due to high density [29], and as a result it may be appropriate to direct women with high density toward alternative imaging methods such as MRI or ultrasound.

In 1976, Wolfe [30,31] suggested a classification scheme in which mammograms were categorized in four classes described by different patterns of dense tissue, with emphasis on ductal patterns. It was found that women with mammograms with extensive ductal patterns had a higher risk of developing breast cancer compared to women with fatty breasts. Since that work was done, multiple density classification schemes have been developed. The classification is typically subjective and involves 4-6 discrete categories that describe features of dense tissue (the Wolfe [30,31] and Tabăr [32] scales) or based on a semi-quantitative estimation of the proportion of dense tissue on the image (the Boyd scale [33], planimetry [34]). The Breast Imaging Reporting and Data System (BIRADS) density scale is commonly used in north America, alternatively with descriptive or semi-quantitative categories [35,36] (see Figure 1.3). The BIRADS scale is intended to allow the radiologist to communicate concern regarding the possibility of a lesion being obscured in the mammogram and therefore not detected due to density. A computerized thresholding method, Cumulus [37], has also been developed to allow users to estimate the proportion of dense tissue on a continuous quantitative scale of percent density. See Figure 1.11 in Section 1.5.1.

Initially, Wolfe’s results regarding breast density and risk were not reproduced due to the subjectivity of the categories. But in the past two decades, multiple case-control studies [38-46], all using the quantitative percent density measure, have confirmed that increased mammographic density is a risk factor for developing breast cancer. Reviews of those studies [47,] showed a 4 to 5 fold increase in risk between the highest and lowest breast density categories, with a follow-up period of 5-10 years. McCormak and Dos Santos Silva [49] conducted a meta-analysis of 42 studies and found a consistent association of breast density to breast cancer risk.
Moreover, the work of Boyd [39] determined that the cancer risk associated with breast density was independent of other risk factors such as age, ethnicity and menopausal status. In addition, the risk association persisted in the long term (6-8 years) and thus could not be explained by the masking effect of high density [39]. Table 1-1 shows the association of breast cancer with various other factors [50]. Breast density is among the highest risk factors for developing breast cancer, after sex and age, and it has been suggested that breast density may be a factor in 30% of cancers [38,47]. Certain factors can change breast density. 20-30% of the variation in mammographic density is attributable to changes in age, menopausal status and parity, while studies done with twins indicate that part of the remaining variation in density (50-65%) is due to heritability [47]. Tamoxifen treatment, as well as increases in body weight and the number of live births are associated with a reduction of density. A family history of breast cancer and hormone replacement therapy (estrogen and progestin) is associated with a higher density [47].

It may seem paradoxical that breast cancer incidence increases with age, while breast density typically decreases with age. Pike suggested that the cumulative exposure of hormones and growth factors to breast tissue (tissue aging) can describe the incidence of cancer [51]. Thus, the age of menarche, the number of pregnancies and the age of menopause would have greater influence on future breast cancer incidence than the chronological age. The same factors that influence the exposure of breast tissue in the Pike model also influence the breast composition and thus mammographic density: breast density may be a correlate of the rate of tissue aging [52].
Table 1-1: Risk factors for breast cancer. For small probabilities, the relative risk equals the odds ratio.

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Relative risk or odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: female vs. male</td>
<td>100</td>
</tr>
<tr>
<td>Age: &gt;65 vs. &lt;65</td>
<td>5.8</td>
</tr>
<tr>
<td>Breast density: high vs. low</td>
<td>4-5</td>
</tr>
<tr>
<td>BRCA1 gene mutation: present vs. absent (age 60-69)</td>
<td>15</td>
</tr>
<tr>
<td>Personal history of breast cancer: yes vs. no</td>
<td>6.8</td>
</tr>
<tr>
<td>Number of first degree relatives with breast cancer: 1 vs. 0</td>
<td>1.8</td>
</tr>
<tr>
<td>Age at first live birth: &lt;20 years vs. &gt;30 years or nulliparous</td>
<td>1.7-1.9</td>
</tr>
<tr>
<td>Age at menarche: &lt;12 years vs. &gt;15 years</td>
<td>1.3</td>
</tr>
<tr>
<td>Hormone replacement therapy: yes vs. no</td>
<td>1.3</td>
</tr>
<tr>
<td>Body mass index: 80th percentile vs. 20th percentile, &gt;55 years</td>
<td>1.2</td>
</tr>
</tbody>
</table>

A causal link between breast density and breast cancer has not been established. Since breast cancer originates in epithelial cells, the number of epithelial cells should be related to the probability of genetic damage that may lead to cancer. The number of epithelial cells is related to the breast composition and thus to breast density. Moreover, the connective tissue also contributes to breast density, and may facilitate tumour invasion: it has been shown that the interactions between the connective and glandular tissues influences the development and changes in the breast during pregnancy, lactation, menopause and tumourigenesis [53-58].

1.4 Mammography Systems

Mammography was introduced in the 1960’s and mammography technology has steadily improved up to the present time. Full-field digital mammography (FFDM) has been in use for over a decade, and has largely supplanted the screen-film systems in North America and Europe.

A modern conventional FFDM system uses an x-ray tube that produces an x-ray spectrum with a mean energy between 17 and 20 keV, low enough to enhance the tissue contrast between adipose tissue and fibroglandular tissue (Figure 1.2). Breast tissue is strongly attenuating at those energies, and thus to obtain an adequate signal on the image receptor, a relatively high dose is received by the breast (1-3 mGy, vs. 0.05 mGy for a chest x-ray). The breast is compressed in order to spread out superimposed tissue, and the resulting smaller thickness reduces the
necessary dose, as well as the ratio of the amounts of scattered and directly transmitted (primary) radiation. To further reduce the scatter, an anti scatter grid, which preferentially removes scattered radiation versus primary radiation, is located in the “bucky” assembly between the breast and image receptor. FFDM uses an x-ray-sensitive flat-panel electronic detector. The signals arising from the x-rays transmitted through the breast are digitized in the detector and transferred to a computer for storage, processing and display. See Figure 1.4.

Figure 1.4: Schematic of a mammography system

1.4.1 Attenuation and Scatter

The basic law of x-ray interaction with matter is described by the Lambert-Beer law. It states that for a narrow beam of $N$ photons impinging on a thin layer of $\delta x$ thickness of some material, the change in the number of photons in the beam, $\delta N$, is proportional to $N$ and $\delta x$.

$$\delta N = -\mu N \delta x,$$  \hspace{1cm} (1-1)

where $\mu$ is the constant of proportionality called the linear attenuation coefficient, in units of reciprocal length. In other words, the number of photons absorbed is a fraction of the incident number, and scales with the thickness of material; more material causes more absorption. By integrating both sides of equation (1-1), we obtain that

$$N = N_0 e^{-\mu x},$$  \hspace{1cm} (1-2)
that is, the beam, initially containing $N_0$ quanta is attenuated exponentially in a material of thickness $x$. The value of $\mu$ represents the probability per unit length of material that a photon will interact with the material and be removed from the beam by absorption or scattering. The attenuation changes with the type of material and with the photon energy, as seen in Figure 1.2. In breast tissue at mammographic energies, three main types of interaction occur: the photoelectric effect, Compton scatter and coherent scatter. They account for approximately 75%, 15% and 10% of the interactions, respectively [59]. In photoelectric events, the x-ray is absorbed and no scatter occurs, as opposed to Compton and coherent scatter events, in which the photon changes direction and can scatter multiple times.

In assessing attenuation, the x-ray beam that is not removed by attenuation is measured. Therefore, the Beer-Lambert law only holds for the primary photons, that is the photons that don’t interact and thus travel in a straight line from the source to the detector. In most x-ray imaging systems, the patient is subjected to a broad beam of x-ray that exposes the whole organ, and the image receptor generally has no or only imperfect collimation, so that both the primary radiation and scattered radiation are received. The scattered radiation causes degradation of the image quality. Scattered photons deviate from their initial straight line trajectory, adding a layer of diffuse and noisy signal on the image. In screen-film mammography, the detrimental effect of scatter is in a reduction of the contrast. In digital mammography, the contrast and brightness can be arbitrarily enhanced by adjusting the window and level on the viewing system. Thus scatter causes a reduction in the signal-to-noise ratio and consumes part of the dynamic range.

In mammography, the signal arising from scattered radiation is comparable in magnitude to the signal arising from primary radiation; the scatter to primary ratio is approximately 25%, 40%, 60% and 75% for breasts 2, 4, 6 and 8 cm thick, respectively [60]. Mammography systems employ anti-scatter grids (described in Section 1.4.4) which discriminate against scattered radiation.

### 1.4.2 X-ray Production

X-rays are produced in an evacuated tube. It contains a cathode that generates an electron beam that is accelerated with a high voltage towards an anode. The cathode is composed of a filament which emits electrons when heated using an electric current (thermionic emission). The filament is placed into a focusing cup that is negatively charged and which directs the electron beam to a
small area of the anode, the focal spot. In mammography, the anode is typically molybdenum or rhodium, and the focal spot is typically 0.3×0.3 or 0.1×0.1 mm². The interaction of the electrons with the anode material generates x-rays at various energies, which radiate towards the image receptor. With an applied potential of 25-34 kV and a molybdenum or rhodium filter (25-30 µm thick) to further attenuate the low energy x-rays, the tube produces an x-ray spectrum at the proper energy (16-20 keV) to maximize attenuation contrast. Typically, the following combinations of target and filter material are used in an attempt to optimize the energy spectrum: Mo/Mo, Mo/Rh, Rh/Rh, W/Ag... etc. See Figure 1.5 for examples of x-ray spectra used in mammography.

![Figure 1.5: Example of spectra in mammography. The mean energy for the Mo/Mo and Rh/Rh spectra are 17 and 19 keV, respectively.](image)

The x-ray spectrum varies in intensity and energy across the image field, due to the anode heel effect and the inverse square law. Usually, the anode is angled with respect to the electron beam, and the x-rays generated at some depth within the target go through a longer path length when emitted downwards directly under the anode then when emitted forward towards the cathode side. This differential self-filtration makes the spectrum more penetrating (higher energy) on the anode (chest wall) side, but less intense compared to the spectrum on the cathode (nipple) side. In addition, the x-ray beam traverses a longer distance to reach a point \((x, y)\) on the receptor a distance \(R = \sqrt{x^2 + y^2}\) from the focal point directly under the focal spot. Due to the inverse square law, the intensity of the beam decreases such that \(I(x, y) = I_0 \left(1 + \frac{R^2}{SID^2}\right)^{-1}\), where \(I_0\) is
the focal point intensity and SID the source to image distance, typically 650 or 660 mm. Since $SID \gg f$, the focal spot size, the x-ray source can be considered a point source. In digital mammography, most of the variations in x-ray intensity and energy can be corrected by flat-fielding the image.

Some radiation can also originate from outside the focal spot. Off-focus radiation occurs when stray electrons deviate from the main focused beam or bounce off the focal spot to be absorbed elsewhere in the anode. The off-focus radiation generally does not follow the same direction as the primary x-ray beam and thus behaves as scattered radiation, degrading the image quality while increasing the dose to the patient. As we will see in Section 2.3.4.1, it accounts for approximately 3% of the emitted radiation.

1.4.3 Breast Compression

As mentioned above, in mammography the breast is compressed to reduce tissue superposition, scatter-to-primary ratio, and dose. The compression device is typically a 2.5 mm thick piece of rigid plastic. The compression force varies between 60 and 120 N (6-12 kg), bringing the thickness down to 5-6 cm on average. In digital mammography, the compression force and compressed breast thickness are reported in the image header. The type of paddle has a strong influence on the thickness profile of the breast [61,62]. As will be discussed in Sections 1.5.2 and 2.5, accurate knowledge of the thickness of the breast is of great importance to make accurate volumetric breast density measurements. Therefore, it is important to characterize how the compression paddle responds to the compressive forces and to determine the accuracy of the thickness readout in the image header. For instance, some paddles are designed to remain flat under compression, while some have a built-in hinge to reduce the compression on the chest wall side of the breast, which slants the compression paddle forward. Moreover, some mammography units have a relatively accurate thickness readout (within 1-2 mm), while some are grossly inaccurate with 8-10 mm errors [63]. The characterization of the compression paddle is described in Section 4.2.1.

1.4.4 Anti-Scatter Grid

Mammography systems have an anti scatter grid that geometrically discriminates against scattered radiation. A linear grid consists of a series of strips of a highly attenuating material
(such as lead) interleaved with strips of a radiolucent material, typically paper (cellulose). The strips are focused towards the x-ray source (in a single direction, not spherically) to maximize the transmission of primary radiation. See Figure 1.6. The scattered radiation, which is generally at an angle with respect to the primary radiation, is thus partly blocked by the attenuating strips. Some grids use a honeycomb hexagonal pattern of attenuating material. In addition, grids undergo a reciprocating or linear movement during exposure to blur the potentially distracting shadow of the grid lines. The grid is contained in the “bucky”, which includes a breast support plate and a sliding mechanism to take the grid in and out from the top of the image receptor.

![Figure 1.6: Schematic of an anti-scatter grid.](image)

Linear grids are characterized by the grid ratio $h/D$ and the frequency $(t + D)^{-1}$, where $h$, $D$ and $t$ are the height of the strips, the space between the strips and the strip width, respectively. Typical values are $h = 1.5$ mm, $D = 0.3$ mm and $t = 0.02$ mm, so that the grid ratio is 5:1 to 4:1, and the frequency is at 30 lines/cm. The grid ratio is a good indicator of the selectivity of primary to scatter transmission; a tall and narrow grid has a lower acceptance angle than a short and wide grid. However, for a given height a high ratio implies a high frequency, which increases the density of the grid and results in a larger loss in primary radiation. As seen in Section 2.3.4, the transmission of primary radiation $T_P$ is approximately 70%, while the scatter transmission $T_S$ is approximately 15%. As a result, for a 5 cm thick breast, around 50% of the total radiation transmitted by the breast is blocked by the anti-scatter grid before reaching the detector. Thus, in
order to maintain an adequate exposure on the image when using a grid, the exposure, and dose, to the patient must be increased by a factor of 2. A compromise must therefore be found between the image quality improvement from the scatter removal, and the increase in dose from the loss of primary signal and overall exposure. Some studies have shown that breasts under 4 cm thick do not benefit from using a grid; the dose is lower without the grid while maintaining the same image quality [64,65].

1.4.5 Screen-Film Mammography

Screen-film mammography is a valuable tool for the detection of breast cancer, and will continue to be so: most of the long-term randomized studies showing the reduction of breast cancer mortality from screening programs have been done using screen-film technology. A recent study has shown that screen-film and digital mammography perform similarly in a screening population, although digital technology performed better for women with high breast density or for younger or premenopausal women [6].

A conventional screen-film system has similarities to the indirect digital detectors discussed below. It uses a cassette that contains a single phosphor screen and single emulsion film. The screen absorbs x-rays and emits visible light, in an amount proportional to the intensity of incident x-rays. The light exposes the film, and the cassette is removed to extract the film for processing. The type of film, screen, cassette and processing all affect the image quality and dose.

Screen-film mammography has advantages. The technology is relatively inexpensive, fine-grained film can achieve a very high resolution, and the processed films are conveniently displayed on view-boxes. Moreover, the film has a logarithmic response to exposure that inherently performs a compression of the transmitted exposure to optical density (the processed film’s blackness). See Figure 1.7. This compression often provides an acceptable representation of the radiographic structures in the breast in a single image; the compressed range of intensities is within the eye’s dynamic range. In comparison, the “raw” digital mammograms contain a broader range of intensities than can be accommodated by the observer. Therefore, it is necessary to adjust the display parameters (brightness and contrast) during viewing to visualize the full range of the information entirely. To facilitate visualization, the “raw” mammograms are processed to compress the intensity range into a “for presentation” mammogram.
One of the limitations of film imaging is that the attenuation contrast from structures in the breast is compounded with the film’s characteristic response. As seen in Figure 1.7, there are regions in the toe and shoulder of the curve where large changes in exposure lead to only very small changes in the film’s optical density: the tissue contrast is poorly rendered in those regions. Thus, it is critical that the film be properly exposed, i.e., the mean exposure transmitted by the breast should be chosen such that it lies within the linear region of the film’s response. Even when this is done, it is often the case that the displayed contrast will be sub-optimal in the more opaque and lucent regions of the breast, even though there is significant attenuation contrast in those regions. This can be problematic if lesions are present in those regions, or in breasts with a high amount of dense tissue. Lowering the film gradient would increase the width of the linear part of the response and lower the noise, but would require a higher dose and have lower display contrast.

1.4.6 Digital Mammography

Digital mammography provides several advantages compared to screen-film mammography. First, the detector produces an electronic signal that is proportional to the intensity of the x-rays transmitted by the breast, so that the attenuation contrast is faithfully recorded over a large range of exposure. Moreover, in digital systems the image acquisition function is separated from the storage and display functions, so that each can be optimized independently. In screen-film mammography, the film acts as both the image receptor and the storage and display device, and
thus certain compromises result; once the film is exposed, the display contrast and brightness characteristics are fixed. If those characteristics are inadequate, the patient should be re-exposed.

There are two main types of x-ray detectors used in digital mammography: phosphor-based “indirect” detectors and photoconductor-based “direct” detectors.

In indirect detectors, the phosphor material, typically Cesium Iodide doped with Thallium, absorbs x-rays and produces light in the visible spectrum, which is then absorbed by an array of photodetectors, which convert the optical energy into an electrical signal. CsI is grown in a columnar structure on top of a photosensitive array (amorphous Silicon) of detector elements, in order to reduce the spreading of light from the point of x-ray absorption to the photodetector directly below, which causes blurring. This blurring is exacerbated in a thick phosphor screen, so a compromise must be made between x-ray absorption efficiency and sharpness. See Figure 1.8. General Electric uses this type of detector in mammography, with a 100 µm detector element.

![Indirect detector diagram](image)

**Figure 1.8:** Indirect detector. The light is emitted isotropically. The closer the x-ray absorption is from the detector array, the less distance the light will travel laterally.

In direct detectors, x-rays interact in the photoconductor (amorphous Selenium) and release electron-holes pairs, which are guided directly to the surfaces by an applied electric field. See Figure 1.9. The charge is then directly collected on an electrode coupled to each detector element. Compared to indirect detectors, the direct conversion of x-ray energy into an electrical signal is more efficient. Moreover, the charge doesn’t spread from the point of interaction to the electrode, so the photoconductor can be made thick to increase the absorption efficiency. This increase requires an increase in the electric field, which is typically 10-30 V/µm. Hologic uses this type of detector in mammography, with a 75 µm detector element.
In both types of detectors the detector elements in the array are each coupled to a readout device, a thin-film transistor (TFT). During readout, the TFTs of an entire row are simultaneously activated with control lines and read to low-noise charge amplifiers on each column through signal lines. The signals are then selected in a multiplexer and digitized, and the readout is completed by repeating the process on all the rows of the detector array.

Since each element on the array must contain the image sensor, the TFT switch, and control and signal lines, only a fraction (the fill factor) of the element is actually sensitive to x-ray exposure. The fill factor is typically 96-93% depending on the element size.
1.4.7 Image Quality

1.4.7.1 Quantum Efficiency and Gain

In order to produce a signal, the x-ray photons must interact and be absorbed in the detector material. The quantum efficiency of the detector is given by

$$\eta(E) = \left(1 - e^{-\mu(E)T}\right). \tag{1-3}$$

The quantum efficiency is derived from the Beer-Lambert law, and is the probability that a photon of energy $E$ interacts in the detector of linear attenuation $\mu(E)$ and thickness $T$. Often, the quantum efficiency $\eta(E)$ is multiplied by the factor $\tau = \mu_{en}(E)/\mu(E)$: the fraction of the interacting photons that are absorbed in the detector. Some photons interact in the material but only transfer part of their energy before scattering away from the material. $\mu_{en}$ is the linear absorption coefficient. In this dissertation we make the substitution $\eta \rightarrow \eta\tau$. For $N$ photons impinging on the detector, $N\eta$ will be absorbed.

The quantum efficiency can be increased by making the detector thicker, and the efficiency generally decreases with increasing x-ray energy. For a 100 $\mu$m layer of CsI at mammographic energies (15-20 keV), $\eta = 0.88-0.65 \ (\mu_{en}/\mu = 0.95-0.93)$. Doubling the thickness to 200 $\mu$m brings $\eta$ to 0.95-0.85.

The energy of the $N\eta$ photons that interact in the detector must be converted to an electrical signal $q$. This conversion efficiency is expressed by the total gain factor $\Gamma$, so that

$$q = N\eta\Gamma. \tag{1-4}$$

There are various gain stages in a detector. The first gain factor $\Gamma_1$ is expressed by the number of visible light photons or electron hole pairs generated by the absorption of an x-ray photon of a given energy, for indirect or direct detectors, respectively. Later gain stages express the probability of the visible light or electron hole pairs escaping the detector, the efficiency of the photodetector or electrode in generating an electrical charge, etc. The larger the gain and quantum efficiency, the more sensitive the detector is. In a good detector, $\eta$ is the primary
inefficiency; the gain $\Gamma_1$ is sufficiently large so that the subsequent gain factors do not bring the level of signal lower than $N\eta$.

For CsI, it takes approximately 19 eV of energy to create one light photon [66]. Thus the gain is approximately 1000 for a 20 keV x-ray. We note that the light photons are 2-3 eV; the energy carried by a 20 keV x-ray would correspond to 8000 light photons, but the conversion is only about 12% efficient.

### 1.4.7.2 Noise

In the x-ray source, x-rays are produced at a constant rate, but each x-ray is produced independently of the other. This characteristic implies that the x-ray beam follows Poisson statistics; the number of produced x-rays will fluctuate randomly about the mean rate. For a mean number $N$ of x-rays, the standard deviation is $\sigma_N = \sqrt{N}$, so that the signal-to-noise ratio

$$\text{SNR} = \frac{N}{\sqrt{N}}.$$  

This deviation ultimately causes unwanted signal fluctuations, i.e. noise, in the image.

The conversion processes in detectors will increase the noise from the quantum Poisson noise. Each gain factor is susceptible to fluctuations about its mean value. From the signal $q$ in equation (1-4), it has been shown [67] that the variance is

$$\sigma_q^2 = N\eta(\Gamma^2 + \sigma_{\Gamma}^2),$$  \hspace{1cm} (1-5)

where $\sigma_{\Gamma}^2$ is the variance in the gain factor. Normally there are multiple gain stages, and each contributes to an increase in signal variance. In this simple analysis, the SNR is

$$\text{SNR} = \sqrt{N\eta} \left(1 + \frac{\sigma_{\Gamma}^2}{\Gamma^2}\right)^{-1/2}. \hspace{1cm} (1-6)$$

In a good detector, the variance in $\Gamma$ should be small compared to $\Gamma$, so that the noise in the system is limited by the incident x-ray fluence and by $\eta$. 
1.4.7.3 Resolution

One important factor limiting resolution in digital detector is the size $d$ of the detector element: single structures of minimum size $d$ can be resolved. Moreover, the sampling theorem states that only the spatial frequencies below the Nyquist frequency $(2d)^{-1}$ can be faithfully imaged. Patterns of a higher frequency will lead to artifacts (aliasing) on the image. For $d = 100 \, \mu \text{m}$, the Nyquist frequency is 5 cycles/mm. In comparison, screen-film systems can have up to 20 cycles/mm resolution. In most digital systems, the blurring associated with the screen or photoconductor added to the resolution loss from the finite focal spot size or motion are the limiting factors in the resolution of the detector.

In indirect detectors, a significant loss of resolution is due to the light photons spreading laterally in the screen. See Figure 1.8. The light can travel a few millimetres laterally in the phosphor and causes a drop of resolution at the low frequencies resolution in the detector. This effect, called glare, is discussed in Section 2.3.3. There are a few ways to reduce the blurring. In screen-film systems, the film is placed on top of the phosphor. Since the attenuation (or absorption) of x-rays depends exponentially on thickness, the light is created preferentially near the top, and it has less length to travel and diffuse to reach a detector on top than at the bottom. In digital systems, the detector is placed below the phosphor. Thus, a thicker screen will suffer from more blurring, but a thicker screen increases $\eta$, so that a compromise between resolution and efficiency must be made. CsI can be grown in vertical columns that channel light like optic fibres to reduce the lateral spreading of light, which improves the resolution at a given $\eta$.

It should be noted that there always is a compromise between image quality and dose in radiography. A lower dose means fewer photons and a smaller Poisson SNR. A smaller detector element increases the resolution, but the smaller area receives fewer photons, resulting in a lower SNR at the same exposure. Thus, it is important to compare the image quality of different systems at a given patient dose. In other words, a better imaging system has the potential to either improve on the image quality while preserving the dose, or to lower the dose while preserving the image quality.
1.4.7.4 Anatomical Noise

The overall signal, resolution and noise performance of imaging detectors versus spatial frequency $f$ is described by the detective quantum efficiency DQE($f$). The DQE describes the efficiency in transferring the SNR$^2$ contained in the incident x-ray beam to the detector output; ideally, DQE($f$)=$\eta$. By multiplying the DQE by the number of incident quanta, one obtains the noise equivalent quanta (NEQ), or SNR$^2_{\text{out}}(f)$, that is, the number of incident quanta necessary to achieve a given SNR on the image.

The DQE or NEQ of a detector outlines the baseline performance of an imaging system. However, in practice, the limiting factor in imaging performance is the anatomical noise. The anatomical noise refers to the normal structures in the image that can prevent an observer from detecting a pathological lesion (the “signal”). The diagnostic performance of an imaging system is thus ultimately assessed by determining its sensitivity and specificity characteristics. As discussed in Section 1.3, the amount of normal dense tissue on a mammogram has a strong effect on the sensitivity of mammography. Novel imaging methods such as tomosynthesis, contrast-enhanced mammography, and dual-energy mammography are being developed on a digital mammography platform to reduce the detrimental effect of tissue superposition (anatomical noise) that is present in planar mammography.

1.5 Volumetric Breast Density

1.5.1 Limitations of Mammographic Breast Density

As we have seen in Section 1.3, mammographic breast density is an important risk factor for breast cancer. The methods to estimate breast density have some limitations. We have seen that breast cancer risk should be related to the amount, or volume, of epithelial and connective (i.e. dense) tissue in the breast. However, the traditional breast density estimation methods, including the Wolfe categories, BIRADS classes, computer-thresholding, etc. are based on the projected area of dense tissue on the image. The categorical methods typically only have a few classes, and thus can only yield a coarse estimate of the amount of dense tissue. The computer-assisted percent area density method intends to be more quantitative, but remains susceptible to errors due to unaccounted variations in the breast thickness or dense tissue thickness. Typically, a global grey-level threshold is set that best delineates the dense tissue on the image. It is
unavoidable that thin dense regions will be visible on the image but fall below the threshold. Similarly, regions where the breast is thinner (such as near the periphery) can also be omitted even though they contain dense tissue. In other words, the method does not allow a column of tissue to have a mixture of fat and fibroglandular tissue: a pixel is considered to represent a path through either entirely dense or entirely fatty tissue. Moreover, a small difference in the threshold level can have large effect on the measured area density. See Figure 1.11. Finally, the method is manual and requires a trained observer, making it relatively time-consuming and subjective (see Section 5.2).

![Figure 1.11: Thresholding method to estimate mammographic density. The left image is the mammogram (negative logarithm of the raw image). The middle and right images represent two thresholds levels leading to an area density of 28% and 39%, respectively. As we can see, both thresholds lead to a reasonable delineation of the dense tissue, yet the area density measure differs by an absolute value of 11%. The thresholds levels differed by 5 units on a scale of 1 to 256 (2%).](image)

The aforementioned limitations in the area-based measurements are likely to attenuate the association between breast cancer risk and breast density. Volumetric breast density measurements can in principle circumvent those limitations by providing an objective (and automatic) measure representative of the breast composition with a degree of accuracy. That is, volumetric breast density measurements have the potential to be more strongly linked with breast cancer risk.

There have been some methods developed to estimate the volume of dense tissue from mammograms, which will be discussed in the following sections. To date, there have been four studies examining the relation between breast cancer risk and volumetric breast density from
mammograms. Shepherd et al. [68] has found an odds ratio for breast cancer risk of 4.1 between upper and lower quintiles of percent volumetric breast density, versus an odds ratio of 2.5 for the area measures. However, Ding et al. [69] found that the percent volumetric density measure was associated with risk, but less strongly than percent area. And after adjusting for the percent area measurement, the volume was no longer associated with breast cancer risk. Similarly, Boyd [70] found that the percent volumetric density did not improve on the risk prediction made by area density, although the volumetric measure was an independent risk factor. Those three studies were done on film screen mammograms. Using digital mammograms, Lokate et al. [71] found a similar association of breast cancer risk with the percent area and percent volume methods.

Volumetric breast density (VBD) can be measured through other imaging modalities. Shepherd et al. [72] have developed a dual-energy absorptiometry technique, normally used for bone densitometry, but modified for use in the breast, that offers precise measurements of VBD at very low radiographic doses. Ducote et al. [73] have developed a method using dual-energy mammography that has also shown promising results in measuring the VBD. Blackmore et al. [74] have developed a spectroscopy system to measure the breast composition and have shown good correlation with mammographic density. A CT system dedicated for breast imaging [75-77] has been developed, which has been used to measure VBD directly by segmentation [78,79]. MR can also be used to measure breast density. In MR, the breast tissue contrast arises from the difference in the resonance frequency between protons in water and in adipose tissue [80]. Thus, the volumetric water content fraction is a good measure of the volumetric fibroglandular fraction, if the fact that adipose tissue contains 5-15% water is taken into account. Studies have shown a strong correlation between mammographic percent area density and MR percent water-content [81]. Another method extracts the fibroglandular fraction in MR images through an adaptive fuzzy C-means classification [82]. It is also possible to quantify density using ultrasound in a tomographic ultrasound system dedicated for breast imaging has been developed [83,84]. In ultrasound imaging, the breast tissue contrast arises from the difference in sound speed between fatty and fibroglandular tissue. The speed of sound \( v = \sqrt{E/\rho} \), where \( E \) is the tissue bulk modulus (stiffness) and \( \rho \) the density. It has been shown [85,86] that for most tissues, \( E \propto \rho^3 \), so that \( v \propto \rho \). Therefore, breast density can be investigated by measuring the speed of sound in the breast. In the tomographic ultrasound system, a map of sound speed is reconstructed for each slice and segmented using k-means clustering, so that the ultrasound breast density can be
computed. The ultrasound volume density has been shown to be strongly correlated with mammographic density [47].

1.5.2 Volumetric Breast density in Mammography

The most practical way to obtain the VBD is from mammograms, as women frequently receive this examination for breast cancer screening. The basic principle behind the measurement of volumetric breast density in mammography is as follows. Suppose a narrow beam of monoenergetic x-rays of intensity $I_0$ are impinging on a breast of thickness $T$. Assuming the detector perfectly records the transmitted intensity $I$, and that the scattered radiation is rejected, $I$ is given by Beer’s law:

$$I = I_0 e^{-\mu T}. \quad (1-7)$$

We can further assume that the breast is composed of two types of tissue, fatty tissue and fibroglandular tissue, so that the attenuation of the breast $\mu = m\mu_{fs} + (1-m)\mu_f$, where $m$ is the fraction of fibroglandular tissue in the column of tissue above the pixel and $\mu_{fs}$ and $\mu_f$ the attenuation of fibroglandular and fatty tissue, respectively. We can thus extract the fractional density as

$$m = \frac{\log(I_0/I) - T\mu_f}{T(\mu_{fs} - \mu_f)} \quad (1-8)$$

That is, the breast composition is deduced from the image signal, $I$ and $I_0$, the breast thickness $T$ and the attenuation coefficients of breast tissue. If we assume that the breast is of constant thickness and composition, the dense volume is given by $V_D = ATm$, and the total breast volume is $V = AT$, where $A$ is the breast area. The volumetric breast density is then:

$$VBD = \frac{V_D}{V} = m. \quad (1-9)$$

In that simplistic case we can see that the volume density is inversely proportional to the thickness $T$. The difficulty in measuring VBD from mammograms is in the relatively high sensitivity of the measurement with respect to the breast thickness. This high sensitivity is expected; since the measurement is done on a 2D image, the volume can only be accurately
deduced from an accurate thickness. Say the true thickness and density are $T$ and $m$, respectively, but we erroneously assume the thickness is $T + \Delta T$. The attenuation $\ln(I_0/I) = \mu T$, so that the corresponding error in density is

$$\Delta m = m_T - m_{T+\Delta T} = \frac{\mu}{\Delta \mu} \frac{\Delta T}{T + \Delta T} = \frac{\mu}{\Delta \mu} \frac{\alpha}{1 + \alpha} = \frac{\mu}{\Delta \mu} \alpha(1 - \alpha),$$

(1-10)

where $\alpha = \frac{\Delta T}{T}$. The approximation is valid for $|\alpha| \leq 0.1$. The value $\frac{\mu}{\Delta \mu} = m + \frac{\mu_I}{\Delta \mu}$ increases with x-ray energy and density $m$, as seen in Table 1-2. The error in thickness is generally amplified: for a breast 25% dense, imaged at 18 keV, a relative error in thickness of $\pm 3\%$ (1.5 mm over 50 mm) will lead to an error in density of $\pm 4$ percentage points (%VBD). That is, the breast will be attributed a density of $21\%$ or $29\%$, a relative error of $16\%$.

<table>
<thead>
<tr>
<th>keV</th>
<th>16</th>
<th>18</th>
<th>20</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.08</td>
<td>1.19</td>
<td>1.31</td>
<td>1.47</td>
</tr>
<tr>
<td>25</td>
<td>1.33</td>
<td>1.39</td>
<td>1.56</td>
<td>1.72</td>
</tr>
<tr>
<td>50</td>
<td>1.58</td>
<td>1.69</td>
<td>1.81</td>
<td>1.97</td>
</tr>
<tr>
<td>75</td>
<td>1.83</td>
<td>1.79</td>
<td>2.06</td>
<td>2.22</td>
</tr>
<tr>
<td>100</td>
<td>2.08</td>
<td>2.19</td>
<td>2.31</td>
<td>2.47</td>
</tr>
</tbody>
</table>

Table 1-2: Values of the parameter $\mu/\Delta \mu$ as a function of density $m$ and x-ray energy.

As we will see in Section 2.5, the sensitivity of the VBD measurement with respect to thickness remains similar when beam polychromaticity, scatter and detector efficiency are taken into account. It thus should be emphasized that accurate knowledge of the breast thickness is important to make accurate volumetric density measurement. In some systems the breast is compressed to a relatively uniform breast thickness. However, mammography systems aren’t designed to accurately report the breast thickness. Moreover, the thickness will invariably decrease to zero in the peripheral region where the breast has lost contact with the compression plate. Finally, some compression plates are deliberately slanted, so that the compression thickness is not uniform. The density measurement is also sensitive with respect to the attenuation coefficients of breast tissue. This issue is discussed in Section 2.2.2.3.

As discussed in Section 1.4, the relation between the image signal and the breast thickness and composition is more complex than the simple Beer’s law of attenuation. The intensity of primary
x-rays transmitted by the breast is an integral function of Beer’s law over the x-ray energies of the x-ray spectrum. It also includes the scattered radiation, which itself depends on the x-ray energy, breast thickness and composition. Moreover, the primary and scattered radiations are filtered through the anti-scatter grid, and the absorption in the detector also depends on the x-ray energy. Despite these complications, the principle behind the estimation of density remains the same: the breast composition is deduced from the image signal, the breast thickness, as well as the tissue attenuation and x-ray energy. However, since the relation is an integral function, it cannot be readily inverted as in equation (1-8), and a lookup table of signal value versus density, thickness, attenuation and x-ray energy must be used.

As we saw in Sections 1.4.4 and 1.4.5, digital mammography has some advantages compared to screen-film mammography, and those advantages remain if one attempts to make quantitative volumetric breast density measurements. The non-linear signal-exposure relation in screen-film systems, which depends on the film processing characteristics, adds a layer of complexity in the relation between breast density and image signal. In comparison, the pixel value of un-processed digital images is linear with respect to exposure, making mathematical modeling more accurate and experimental calibration data more reliable. In addition, screen-film images have a limited dynamic range, so that the thinner or thicker regions of the breast are poorly rendered and cannot be analyzed accurately. Finally, screen-film images must be digitized for analysis, and the exposure characteristics, breast thickness and compression force information, which are necessary for the density measurement must be obtained separately. Digital mammography images contain a digital header with all the necessary information that can be extracted automatically.

1.6 Overview of Thesis

1.6.1 Hypothesis and Aims

This thesis presents a theoretical and experimental investigation of the following general hypothesis:

\[ \text{The volumetric breast density can be measured from digital mammograms.} \]

The work was carried out in terms of three specific aims. These include:
Aim 1) Development of a method to estimate the volumetric breast density from digital mammograms;

Aim 2) Validation of the method for measuring the volumetric breast density using breast CT

Aim 3) Measurement of the volumetric breast density characteristics on a population of women.

The completion of those aims resulted in published work as described in Table 1-3.

1.6.2 Outline of Thesis

The core chapters of this thesis are adapted versions of published work shown in Table 1-3. Specifically, Chapter 2 provides a description of the method used to measure the volumetric breast density from digital mammograms. Chapter 3 describes the validation of the method, where the calculated volumetric breast density was compared to the volumetric breast density measured with breast CT. Chapter 4 shows the results of the volumetric breast density calculations done on a large set of patients, including the method used to determine the compressed breast thickness and the thickness in the peripheral region. Finally, Chapter 5 provides preliminary results validating the density algorithm with MRI, a comparison between the area and volumetric density calculation and their association with breast cancer risk, and provides a summary of results with concluding remarks.

<table>
<thead>
<tr>
<th>Published work reference</th>
<th>Chapter:</th>
<th>Aim:</th>
</tr>
</thead>
</table>

Table 1-3: List of published work relevant to each aim and chapter.
Chapter 2
A Method to Estimate the Volumetric Breast Density from Digital Mammograms

2.1 Introduction

As we have seen in Chapter 1, the traditional measures of breast density (BIRADS classes, percent mammographic density) are strongly associated with breast cancer risk, in that women with high mammographic breast density are at a considerably higher risk of developing breast cancer compared to women with low breast density. However, those methods suffer from some limitations that may attenuate their association with breast cancer risk: a) they are subjective and prone to variability in the users’ performances; b) the resulting measurement is based on the projected area of dense tissue and does not accurately represent the tissue composition of the breast. In principle, the measurement of volumetric breast density overcomes those limitations, and thus could potentially be a stronger risk factor for breast cancer.

In this chapter we present a detailed description of the method for measuring the volumetric breast density from a mammogram. Our method is partly based on an empirical calibration of the imaging system that yields attenuation measurements of plastic materials representative of fat and breast fibroglandular tissue. The measurements are made for a range of tissue thicknesses and for various x-ray beams (tube kilovoltage and target/filter combination). However, we recognize that the x-ray attenuation properties of these materials are slightly different than those of actual tissues. Therefore, the calibration is complemented and corrected by a simulation model of x-ray transmission that predicts the attenuation of actual breast tissue as a function of tissue thickness, tube kV and target/filter. The density algorithm then operates as follows: the (corrected) attenuation references for fat and fibroglandular tissue are fitted to polynomial functions of tissue thickness and kV, for each specific target/filter combination. Next, the composition of the breast (of a known thickness, and imaged with a known kV and target/filter) is deduced at every location in the image by comparing the measured attenuation at that point to the reference attenuation data. This allows for the computation of the volume of fibroglandular
tissue, which is divided by the total breast volume in order to obtain the volumetric breast density, the VBD.

In Section 2.2 we present the details of the empirical calibration procedure, of the polynomial fitting method and its limitations, as well as the method to compute the dense tissue volume, total breast volume and VBD from the image. We also discuss the problems due to the difference in attenuation between the plastic phantoms and breast tissue. In Section 2.3 we present the details of the simulation model, as well as the results of the characterization of detector glare, anti-scatter grids and off-focus radiation. In addition, we investigate on the sensitivity of the volumetric breast density measurement with respect to thickness.

2.2 Calibration Approach

2.2.1 Phantom Calibration

As mentioned in Section 1.5.2, the relation between the image signal and the breast thickness and composition is complex. Therefore, a common approach to estimate the volumetric breast density is to measure the response of the imaging system versus breast thickness, breast composition and x-ray spectrum using phantoms representative of breast tissue. This approach has been tested by researchers using screen-film [87] or digital systems [88], and the method proposed here is a refinement of the method developed by Pawluckzyk et al. [87]. However, as we will see in Section 2.2.2.3, the calibration method has a significant flaw: the attenuation of the plastic material does not match the attenuation of breast tissue. As described in Section 2.3, the calibration method proposed is supplemented by modeling in order to compensate for that attenuation difference.

For a given target/filter combination, slabs of breast plastic (by CIRS, Norfolk, VA, USA) of thicknesses $T$ ranging between 1 cm and 8 cm (in 1 cm increments) containing cylindrical inserts 2.5 cm in diameter, representing a breast fat and breast fibroglandular tissue composition, are imaged under different x-ray spectra $\phi$ (different kV and anode/filter combination). To keep the contribution from scatter constant, the slabs have a uniform density of 50% adipose and 50% fibroglandular, and are surrounded by other slabs of uniform density, so that most of the imaging field-of-view is covered. See Figure 2.1 for an image of the calibration phantom.
The images are first divided by the corresponding mAs exposure. This is done to normalize the image values to the same relative exposure level, and this normalization is possible thanks to the linear signal-exposure relation of digital systems. The detector signal value for fat $I_f(kV,T)$ and fibroglandular $I_{fg}(kV,T)$ are measured as the average signal over a approximately 20×20 mm region within the projected image of the cylindrical inserts, for each anode/filter combination.

At the same kV and target/filter combination, the image signal $I_0(kV)$ with no object on the field is obtained at the same locations as $I_f$ and $I_t$, and are also normalized per unit mAs. The calibration is completed by computing two empirical relations: $f_0$ and $f_1$ for the fat and fibroglandular attenuation respectively, obtained by fitting a two-dimensional second-degree polynomial surface to the measured fraction of transmitted signal:

\[
\ln \left( \frac{I_f(kV,T)}{I_0(kV)} \right) = f_0(kV,T) = \sum_{i,j=0}^{2} a_{ij} kV^i T^j \quad (2-1)
\]

\[
\ln \left( \frac{I_{fg}(kV,T)}{I_0(kV)} \right) = f_1(kV,T) = \sum_{i,j=0}^{2} b_{ij} kV^i T^j \quad (2-2)
\]

The parameters $a_{ij}$ and $b_{ij}$ are the coefficients of the polynomial least square fit. Equations (2-1) and (2-2) are analogous to Beer’s law. The functions $f_0$ and $f_1$ represent $-\mu_fT$ or $-\mu_{fg}T$, but include the effects of scatter, the polychromatic x-ray beam and the detector efficiency; the functions
correspond to the effective attenuation times the thickness. The polynomial fitting is necessary in order to interpolate the calibration data over a continuous thickness range and at kV values between calibration points. See Figure 2.2 and Figure 2.3 for an example of the experimental calibration data and its corresponding polynomial fit. As we can see, the functions are close to being linear, so that a second-order polynomial with a single inflexion point is appropriate.

Figure 2.2: Effective attenuation for \( f_0 \) (fat, dashed line) and \( f_1 \) (fibroglandular, solid line) as a function of thickness, for the Rh/Rh anode/filter combination on a GE Essential unit. The dashed and solid lines represent the polynomial fits for \( f_0 \) and \( f_1 \), respectively, and the data markers represent the experimental calibration data. For each of the two groups of curves, the kVs (increasing from bottom to top) are 28, 30, 32 and 34.

Figure 2.3: Effective attenuation for fat (left) and fibroglandular (right) as a function of kV, for the Rh/Rh anode/target combination on a GE essential unit. The dashed and solid lines represent the polynomial fits and the data markers represent the experimental calibration data. For the two groups of curves, the thickness increases from top to bottom in 1 cm steps up to 8 cm.

In addition to the effective attenuation, the mAs-normalized \( I_0(\text{kV}) \) values are recorded separately and a second-order polynomial fit is computed for every target/filter combination:
\[ \ln(I_0(\text{kV})) = S_0(\text{kV}) = \sum_{i=0}^{2} c_i \text{kV}^i. \]  

(2-3)

In this case, the signal over a region as large as the phantom containing the inserts (see Figure 2.1), is averaged to measure \( S_0 \). See Figure 2.4.

![Figure 2.4: \( S_0 \) for a GE essential unit, as a function of kV for the Rh/Rh and Mo/Rh anode filter combination. The lines represent the polynomial fit and the data markers represent the experimental data points.](image)

With the image of a breast or phantom \( I(x,y) \), of known thickness \( T(x,y) \) perpendicular to the \( x,y \) plane of the detector, and obtained at a known kV, mAs and anode/filter combination, the effective attenuation of the column of tissue at the location \( (x,y) \) is computed as:

\[ f_m(x,y) = \ln\left( \frac{I(x,y)}{\text{mAs} \cdot I_0(\text{kV})} \right) = \ln\left( \frac{I(x,y)}{\text{mAs}} \right) - S_0(\text{kV}), \]

(2-4)

where \( S_0 \) is computed using the appropriate polynomial coefficients (equation (2-3)). The calibration procedure is performed for all mammography machines, and yields coefficients specific to every unit.

Next, the oblique path length, corresponding to the path length traversed by the x-ray beam through the object, \( T' \), is computed such that

\[ T'(x,y) = T(x,y) \cdot \sqrt{1 + \frac{x^2 + y^2}{\text{SID}^2}}, \]

(2-5)
where $SID$ is the source to image distance. Typically $SID = 65$-$66$ cm, and the origin of the $x$, $y$ plane is located at the focal point, at the centre of the chest wall line ($x$ axis). For a 19x23 cm detector, the ratio $T'/T$ can reach a maximum of 1.06. The values $f_0(kV, T')$ and $f_1(kV, T')$ are computed, representing the effective attenuation of pure fat and fibroglandular tissue, respectively, at the thickness $T'$. The fractional amount of density $m(x, y)$ corresponding to each image pixel is determined by linear interpolation between the values $f_0$ and $f_1$, so that

$$m(x, y) = \frac{f_m (x, y) - f_0 (x, y)}{f_1 (x, y) - f_0 (x, y)}.$$  

(2-6)

In equation (2-6), the $x, y$ dependence of $f_0$ and $f_1$ is from the variation in thickness $T'(x, y)$, and the dependence on the x-ray spectrum ($kV$ and anode/filter combination) is implicit. Finally, the volumetric breast density (VBD) is computed as the ratio of the dense volume $V_D$ and total volume $V$:

$$VBD = \frac{\sum_{x,y} m(x, y) \cdot V_c(x, y)}{\sum_{x,y} V_c(x, y)},$$  

(2-7)

where $V_c$ is the volume of tissue above the pixel $(x, y)$ traversed by the x-ray beam. $V_c$ corresponds to the volume of a truncated slanted pyramid of height $T$, such that

$$V_c = \left( \frac{L}{SID} \right)^2 \cdot a^2 T \cdot \left( 1 - \frac{T}{L} + \frac{T^2}{3L^2} \right),$$  

(2-8)

where $a$ is the (square) pixel’s linear dimension and $L$ is the distance between the x-ray source and the breast support plate. That is, $L = SID - d$, where $d$, typically 20 mm, is the height of the bucky. It is important to compute the volumes in the projected geometry: for flat objects 30, 50 and 70 mm thick, the ratio between the slanted volume $V_c$ and the perpendicular volume $a^2 T$ is 0.90, 0.87 and 0.84, respectively.

The VBD without the skin can also be estimated. We can assume that the breast is surrounded by a shell of skin, of thickness $T_{sk}$. A study done on CT has shown that the skin thickness is
approximately 1.5 mm [89]. The skin volume $V_{sk}$ can be calculated from the two truncated slanted pyramids located at the bottom and top surfaces of the breast:

$$V_{sk} = \left( \frac{L}{SID} \right)^2 \cdot 2T_{sk} \cdot a^2 \cdot \left( 1 - \frac{T}{L} - \frac{T_{sk} T}{2L^2} + \frac{T^2}{2L^2} + \frac{T_{sk}^2}{3L^2} \right)$$

(2-9)

We note that this equation is only valid for the uniformly thick region of the breast. In the peripheral region where the breast has a rounded shape, the x-rays travel through a total thickness of skin larger than $2T_{sk}$. Moreover, skin is approximately 3% more attenuating than fibroglandular tissue at the energies used in mammography [90]. To compensate for that increased attenuation, we can artificially increase the thickness of the skin, so that $T_{sk} \rightarrow T_{sk} \times 1.03$. The VBD without the skin can be then calculated as follows:

$$\text{VBD}_{msk} = \frac{V_{b} - V_{sk}}{V}.$$  

(2-10)

2.2.2 Limitations of the Calibration Method

2.2.2.1 Fitting and Interpolation Errors

The fitting process introduces small errors: the fitted function does not coincide exactly with the experimental data points. Also, it is assumed in equation (2-6) that the effective attenuation $f_m$ can be linearly decomposed into its fatty and fibroglandular components, such that $f_m = mf_f + (1-m)f_f$. While this interpolation is true for linear attenuation, in the absence of scatter, with a monoenergetic beam and a perfectly efficient detector, the assumption will not hold perfectly in this case. However, the contribution from scatter is relatively constant for a uniformly thick object; it varies little with breast density.

To estimate the errors associated with the fitting and interpolation, we can analyze the images from which the calibration was extracted. As illustrated in Figure 2.1, there are additional inserts in the phantoms, with intermediate density values, so that equation (2-6) can be tested. When the phantoms are superposed to increment the thickness, the density in those inserts spans the density range between 0% and 100% with an approximate increment of 10% density.
Figure 2.5 agglomerates all the data for the three anode/filter combinations, as well as the entire kV and thickness range. Figure 2.6 shows the average absolute difference in density between the truth and the calculation, as a function of the anode/target combination, kV, thickness and density. As we can see, the calibration method is generally accurate within 3 %VBD (volumetric density percentage points), and the accuracy is similar between the intermediate density values and the calibrated 0% and 100% density values.

![Figure 2.5: Histogram of the true phantom density minus the calculated density, for the entire calibration image set. The mean difference is -0.9 %VBD, with a standard deviation of 3.1 %VBD.](image)

### 2.2.2.2 Field Variations

The $S_0$ values (equation (2-3)) are measured over specific regions on the calibration images. However, there are signal variations over the image plane due to the anode heel effect and inverse square law (see Section 1.4.2). The results shown in Figure 2.5 and Figure 2.6 were obtained from phantoms that occupied a large field of view on the image plane. That is, the images were analyzed over different regions than where the calibration was performed.

In order to investigate the potential errors resulting from the variation of $S_0$ over the image field, we can analyze the images differently. Instead of using a fitted function for $\ln(I_o) = S_0$, we can use the full $I_0(x,y)$ image in equation (2-4), so that $S_0$ is specific to every point on the image. Figure 2.7 shows the difference in percent density between the true density of the phantoms and the calculation. Using the location-specific $S_0$ improves the accuracy of the algorithm by
approximately 1 %VBD; the mean and standard deviation of the difference are -0.2 and 2.2 %VBD, respectively.

Figure 2.6: Mean density difference (calculation minus truth), as a function of anode/filter combination (top left), kV (top right), thickness (bottom left) and density (bottom right). For each data point, the absolute density error is averaged for all the parameters (kV, target/filter, thickness and density) except for the parameter on the x-axis. The errors bars correspond to ± the standard deviation in the density difference.
3.7

Figure 2.7: Histogram of the true phantom density minus the calculated density (using the location-specific $S_0$), for the entire calibration image set. The mean difference is -0.2%VBD, with a standard deviation of 2.2%VBD.

2.2.2.3 Difference in Attenuation between Phantoms and Breast Tissue

The biggest limitation of the calibration method lies in the phantoms themselves: the attenuation coefficients of the plastic materials from which the phantoms are made is different than the attenuation coefficients of breast tissue [91]. See Figure 2.8. The breast tissue attenuation was measured directly [28], while the plastic material mixture was chosen to match the measured elemental composition of breast tissue. As shown by Byng et al. [91], the plastic material’s elemental composition doesn’t exactly match the elemental composition of tissue measured by Hammerstein et al. In addition, the calculated attenuation, based on the measured mixture of elements doesn’t exactly match the measured attenuation of Johns and Yaffe [28]. It should be noted that a range of attenuation for breast tissue (fat and fibroglandular), from different tissue samples, was measured by Johns and Yaffe. At mammographic energies, the attenuation of “plastic fat” lies just above the upper range of the measured fat attenuation, and the attenuation of “plastic fibroglandular” lies just below the lower range of the measured attenuation of fibroglandular tissue [91].
The difference in attenuation between breast tissue and the plastic material will induce a
significant error in measuring breast density. Following the simplified analysis seen in Section
1.5.2, it can be shown that the measured fractional density $m'$, using the plastic attenuation ($\mu'_f$
and $\mu'_g$ for fat and fibroglandular, respectively), is related to the true fractional density $m$ (using
the correct breast tissue attenuation, $\mu_f$ and $\mu_g$) as follows:

\[
   m' = m \cdot \frac{\mu_g - \mu_f}{\mu_g - \mu'_f} + \frac{\mu_f - \mu'_f}{\mu_g - \mu'_f}. \tag{2-11}
\]

For 18 keV x-ray photons, the slope and intercept of this relation are 1.36 and -0.22,
respectively. That is, for a breast that is truly 0%, 25%, 50%, 75% and 100% dense, using the
plastic attenuation values will yield densities of -22%, 12%, 46%, 80% and 114%. The results
are similar between 16 keV and 24 keV.

Thus, the phantom calibration unfortunately cannot be used directly to compute the breast
density. In the next section, we describe a method in which the effective attenuation is modeled
using the proper attenuation coefficients in order to measure the breast density.
2.3 Modeling Approach

As we can see in the previous section, the calibration method for measuring density is fairly straightforward to implement and use. It has an excellent accuracy (less than 1 %VBD mean error) and a precision of 2-3 %VBD when analyzing breast phantoms. However, it will not be accurate in the analysis of breast images, because the attenuation of the plastic material is significantly different than the attenuation of breast tissue. In order to overcome that issue, we propose the following method. First, develop a simulation model that accurately reproduces the measured effective attenuation of plastic phantoms. Next, calculate the effective attenuation, this time using the proper breast tissue attenuation. Finally, use the modeled effective attenuation and follow the methodology outlined in Section 2.2.1 to measure the breast density.

2.3.1 X-Ray Propagation

The simulation model must take into account all the steps in the propagation of x-ray photons that lead to the formation of the image. As described in Section 1.4, the x-rays propagate as follows: 1) the x-ray spectrum is generated; 2) the radiation is transmitted through the imaged object with primary and scattered components; 3) the transferred radiation is filtered by the anti-scatter grid; 4) the filtered radiation is absorbed in the detector to form the radiographic image; 5) the image is blurred by the detector glare. The steps 3 and 5 require specific parameters that need to be directly measured.

The spectrum simulation is based on the data provided in the work of Boone et al. [92]. The unfiltered spectrum $\phi(E)$ is a function of the x-ray energy $E$, and will vary depending on the anode material and the kV. The filtered photon fluence spectrum $\phi(E)$, in photons per mm$^2$, is given by

$$\phi(E) = \phi'(E) \cdot e^{-\mu_r T_F},$$  

where $\mu_r(E)$ and $T_F$ are the attenuation and the thickness of the filter, respectively. $T_F$ is 30 $\mu$m and 25 $\mu$m for Mo and Rh, respectively. The modeled photon fluence spectrum is not scaled to any specific mAs or Roentgen exposure level, and the spectrum is thus normalized to a unit area, and then scaled appropriately. As mentioned in Section 1.4.2, the x-ray spectrum varies in
intensity and energy across the image field, due to the anode heel effect and the inverse square law.

When no imaging object is present, the x-ray spectrum is absorbed in the detector, with the quantum efficiency \( \eta(E) \), where

\[
\eta(E) = \left(1 - e^{-\mu_d(E)T_d}\right) \frac{\mu^d_{en}(E)}{\mu^d(E)}. 
\]  

(2-13)

\( T_d \), \( \mu^d \) and \( \mu^d_{en} \) are the thickness, linear attenuation and linear absorption attenuation of the detector material. See Section 1.4.7.1. The x-ray energy available for the formation of the image is thus given by \( \phi(E)\eta(E)E \), and this energy is converted into a digital signal through the multiplicative detector gain \( \Gamma \) in units of digital signal per absorbed photon energy. There are multiple gain stages in a detector (e.g. the generation of scintillation photons, followed by the generation of charge in the photosensitive element in an indirect detector). Here the gain is expressed as the compound total gain, and we make the assumption that \( \Gamma \) is independent of x-ray energy.

The gain might vary for each detector element on the array. The gain map can be obtained by acquiring an open-field image \( I_0(x, y) \) with no object in the beam. That is, the open field image, per unit exposure (i.e. normalized by the mAs), and specific to the x-ray spectrum in question, is such that

\[
I_0(x, y) = \overline{I} = \frac{\int \phi(E)\eta(E) \cdot \Gamma E \cdot dE}{\int \phi(E)\eta(E)dE}. 
\]  

(2-14)

Using the empirical open-field image has the advantage of including the spatial inhomogeneities due to the anode heel effect and inverse-square law.

The object to be imaged is described by a 3D map of attenuation coefficients, \( \mu(r; E) \). The radiographic projection of the object on the image plane \( (x,y) \) is described by the line integral of the attenuation along the x-ray path \( C(x,y) \), such that the attenuation times path length product is
\[
\bar{\mu}(x, y; E) = \int_{C(x, y)} \mu(r; E) \mathrm{d}l.
\]

(2-15)

In the coordinate system with the origin at the focal point, the x-ray source is located at (0,0,SID), and the path \(C(x, y)\) represented by the vector \((x, y, -SID)\), where SID is the source-to-image distance.

The image signal due to the transmitted primary radiation, relative to the incident radiation \(I_0(x, y)\), is given by

\[
P(x, y) = \frac{\int \phi(E) \exp[-\bar{\mu}(x, y; E)] \eta(E) \Gamma EdE}{I_0(x, y) \cdot \int \phi(E) \eta(E) dE} = \frac{\int \phi(E) \exp[-\bar{\mu}(x, y; E)] \eta(E) EdE}{\int \phi(E) \eta(E) EdE}.
\]

(2-16)

We can express the image signal due to the transmitted scattered radiation, relative to the incident radiation \(I_0(x, y)\), as \(S(x, y)\). The scatter signal is covered in Section 2.3.2. The anti-scatter grid, if present, filters the transmitted radiation with a primary transmission ratio \(T_p(T, \phi)\) and a scatter transmission ratio \(T_s(T, \phi)\), which depends on the thickness \(T\) of the imaged object and the imaging spectrum \(\phi\) (kV and anode/filter combination). The anti-scatter grid characteristics are discussed in more details in Section 2.3.4. Therefore, considering that \(I_0(x, y)\) is measured with a grid present, the image \(I'(x, y)\) per unit mAs is given by

\[
I'(x, y) = \frac{I_0(x, y)}{T_p(0)} \cdot \left( T_p(T) \cdot P(x, y) + T_s(T) \cdot S(x, y) \right),
\]

(2-17)

where the dependence on the x-ray spectrum is implicit. Finally, the effects of the detector glare, discussed in Section 2.3.3, are included by convolving the image \(I'\) by the glare point-spread function \(G_{psf}(x, y)\). The simulated image \(I(x, y)\) per unit mAs is thus computed as

\[
I(x, y) = \left[ \frac{I_0(x, y)}{T_p(0)} \cdot \left( T_p(T) \cdot P(x, y) + T_s(T) \cdot S(x, y) \right) \right] \otimes G_{psf}(x, y).
\]

(2-18)
2.3.2 Scatter Simulation

Scatter occurs when x-rays interact with matter and are deflected away from the primary beam. Most of the scattered radiation is directed in a forward angle and thus contributes to the signal in the image. The amount of scatter in mammography is significant: the scatter-to-primary ratio is 25%, 40%, 60% and 75% for breasts 2, 4, 6 and 8 cm thick, respectively. The magnitude of the scattered radiation varies on the x-ray spectrum, the breast thickness, the shape of the breast, the dimensions of the x-ray field, the air gap between the breast and the detector, and the anti-scatter grid characteristics.

In this work, we use the scatter point-spread functions ($S_{PSF}$) calculated by Boone et al. [93]. The $S_{PSF}$ were calculated with a Monte Carlo procedure, and the results were validated against experimental data with good accuracy. The $S_{PSF}$, in units of scattered energy per mm$^2$ per incident primary photon$^1$, was available as a function of radial distance from the input x-ray path (0 to 200 mm every 1 mm), of x-ray energy (5 to 120 keV every 1 keV), of breast thickness (2 to 8 cm every 1 cm), of breast composition (0%, 50% and 100% density) and air gap (0 to 30 mm every 1 mm). The $S_{PSF}$ at the air gap value $d = 10$ mm was used, and the $S_{PSF}$ was interpolated to finer thickness and density intervals, to 2 mm and 10% density increments, respectively. Moreover, the $S_{PSF}$ was truncated to a 22x22 cm square region, the size of a mammography image. This truncation had a negligible effect on the accuracy: at mammographic energies, the total scatter energy from 11 cm to 20 cm was only 0.6% of the total scatter energy from 0 cm to 11 cm.

Thus, for a given object, of density $m$ and of thickness $T$ (respectively rounded to the nearest 10% density and 2 mm thickness), the scatter energy at each energy $E$ is given by the convolution:

$$S'(x, y; E, T, m) = \text{FOV}(x, y) \otimes S_{PSF}(x, y; E, T, m) \cdot a^2.$$  \hspace{1cm} (2-19)

Where $\text{FOV}(x, y)$ is the field-of-view, defined as unity where the breast is present and zero elsewhere, and $a$ is the pixel linear dimension. $S'$ is expressed per incident primary photon, that is

---

$^1$ The scatter PSF is actually expressed in the units described above, but per radial area $A = \pi(r_n^2 - r_{n-1}^2)$: the area of the annulus between successive radial distances [93].
per primary photon at the plane defined by the top of the breast. The scatter energy per primary
photon incident on the detector plane is obtained using the inverse-square law:

\[ S'(x, y; E, T, m) \rightarrow \frac{S'(x, y; E, T, m)}{(1 - (T + d)/\text{SID})^2}. \tag{2-20} \]

In a breast, the density is not constant, but the average \( \overline{m} \) density is used. Similarly, the thickness \( T(x, y) \) will vary, but the compression thickness \( T_c \) is used.

Following the analysis of Section 2.3.1, the scatter signal, relative to \( I_0(x, y) \), is given by

\[ S(x, y) = \frac{\int \phi(E)S'(x, y; E, \overline{m}, T_c)\eta(E)dE}{\int \phi(E)\eta(E)dE}. \tag{2-21} \]

The value for \( S(x, y) \) is inserted in equation (2-18) to compute the simulated image \( I(x, y) \).

2.3.3 Glare Characterization

Glare occurs in indirect-type scintillation detectors. When an x-ray is absorbed in the phosphor, a
number of visible light photons are generated, and they are emitted isotropically. Therefore, only
a fraction of those photons are emitted towards the photosensitive detector element located
directly below the point of absorption; some of the photons will be absorbed in neighbouring
elements. This lateral propagation of visible photons is the glare, and it induces a loss in
resolution in the image. The further the light travels laterally, and the greater the amount (relative
to the primary light) of the scattered light is, the larger the blurring effect due to glare becomes.
The blurring effect due to glare occurs at low spatial frequencies: the light spreads laterally to
distances on the order of a few mm. Thus, glare may have a significant effect on low-frequency
structures, such as density patterns on mammograms.

Modern CsI indirect detectors are designed to minimize the glare: the phosphor material is grown
into columnar structures that act as optical fibres, channeling the scintillation photons vertically
towards the detector array. Nevertheless, glare is still present and contributes to the characteristic
low-frequency drop in the modulation transfer function (MTF) of indirect-type detectors [94,95].
See Figure 2.9.
Glare (also named veiling glare or flare) was initially characterized in image-intensifier tubes [96,97,98], and on modern detectors [99,100]. Seiberg et al. [96] proposed a theoretical description of the phenomenon based on the energy diffusion of the scattered light. We can assume that the phosphor screen is two-dimensional (it is much thinner at ~0.1 mm, than it is wide at 200 mm), and that a fraction $F_0$ of the light is scattered radially in the screen. At a distance $r$ from the point of x-ray absorption and light emission, the light is distributed along the circumference $2\pi r$. Furthermore, it is attenuated exponentially in the material with a characteristic decay length $r_0$. Thus, the fraction of the scattered light energy per area is given by

$$E(r) = \frac{F_0}{2\pi rr_0} e^{-r/r_0}. \quad (2-22)$$

The term $r_0$ appears in the denominator to normalize the energy function, such that the integral of $E(r)$ over the (infinite) 2D plane of the detector is equal to the total light energy fraction $F_0$.

The common way to measure the effect of glare is to image a series of radio-opaque disks, typically made of lead. The signal at the centre of the disk, of radius $r$, will be due to the glare originating from the exposed detector outside of the disk. The disk transfer function (DTF) [96,98], or the glare PSF, $G_{PSF}$, is thus given by
DTF\(r\) = G_{PSF}(r) = \int_{0}^{2\pi} \int_{0}^{r_{0}} \frac{F_{0}}{2\pi r' r_{0}} e^{-r'/r_{0}} r' dr' d\theta = F_{0} e^{-r/r_{0}}. \quad (2-23)

The integration is done assuming the detector is infinite. The assumption is valid given that the detector size is much larger than \(r_{0}\). The spatial frequency \(f\) representation of the \(G_{PSF}\) is given by its Fourier transform [97]:

\[ \Im[DTF] = \frac{F_{0}}{\sqrt{1 + (2\pi r_{0} f)^2}}. \quad (2-24) \]

### 2.3.3.1 Measurement of Glare

The glare DTF was measured on three GE machines that use similar structured CsI detectors: a Senograph 2000D, a Senograph DS and a Senograph Essential. The disks were machined from lead sheet of 1.5 mm thickness, with a radius varying between 35 mm and 1 mm. The disk’s outline was detected by automatic thresholding and the radius of the disk was measured on the thresholded shape. The eccentricity in the disk’s shape was measured to estimate the error on the radius; the error was less than 1% for the disks larger than 2 mm, and on average 4% for the smaller disks. A circular region, scaled to 40% of the disk’s radius was selected around the centre, where the average and standard deviation of the image signal was measured. Finally, the DTF was computed as the measured signal divided by the image signal in the absence of a disk, which was obtained from a separate image obtained at the same exposure level (normalized by the mAs). See Figure 2.10 for an example of the DTF.

As we can see in Figure 2.10, the DTF in a semilog scale has an inflexion point, and a sum of two exponentials with distinct decay lengths signal fractions, as observed by Lutha et al. [98], offers a better fit to the data compared to a single exponential. That is, the DTF and its Fourier transform can be described by:

\[ \text{DTF}(r) = F_{1} e^{-r/r_{1}} + F_{2} e^{-r/r_{2}}, \quad \Im[DTF] = \frac{F_{1}}{\sqrt{1 + (2\pi r_{1} f)^2}} + \frac{F_{2}}{\sqrt{1 + (2\pi r_{2} f)^2}}. \quad (2-25) \]
The sum $F_1 + F_2$ would correspond to the low-frequency drop in the MTF. The results are summarized in Table 2-1. The larger amplitude glare occurs at a shorter range while the longer range glare has smaller amplitude. The latter is only apparent when imaging disks larger than ~15 mm. Very large disks cannot be imaged due to the finite detector size, which may explain the larger errors in $r_2$. Seibert et al. [96] and Shen et al. [99] didn’t investigate large disk sizes, which might explain why the longer range glare component was not detected.

As we can see in Figure 2.11, the frequency response of the glare corresponds very well to the low-frequency drop in the MTF. The DTF or $G_{PSF}$ can thus be used to deconvolve the glare from images in order to restore the contrast in the lead disk images.

As we can see in Table 2-1, the glare was characterized for three x-ray spectra for the DS system, spanning the clinical range of beam energy. The changes in the glare parameters for different spectra were within the fit errors. This is further illustrated in Figure 2.12, which shows the variation in the glare parameters for multiple beams. Thus, the changes were not considered significant, so that a single set of glare parameters can be used to perform the deconvolution on an image acquired under any spectrum.
<table>
<thead>
<tr>
<th>Machine</th>
<th>$F_1$</th>
<th>$r_1$ [mm]</th>
<th>$F_2$</th>
<th>$r_2$ [mm]</th>
<th>$F_1 + F_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000D (28kV Rh/Rh)</td>
<td>0.120 (0.008)</td>
<td>2.3 (0.3)</td>
<td>0.094 (0.007)</td>
<td>10 (3)</td>
<td>0.214</td>
</tr>
<tr>
<td>DS (29 kV Rh/Rh)</td>
<td>0.085 (0.005)</td>
<td>4.1 (0.3)</td>
<td>0.020 (0.005)</td>
<td>18 (3)</td>
<td>0.105</td>
</tr>
<tr>
<td>DS (27 kV Mo/Rh)</td>
<td>0.089 (0.007)</td>
<td>4.0 (0.4)</td>
<td>0.018 (0.008)</td>
<td>18 (6)</td>
<td>0.107</td>
</tr>
<tr>
<td>DS (26kV Mo/Mo)</td>
<td>0.089 (0.005)</td>
<td>4.1 (0.3)</td>
<td>0.017 (0.005)</td>
<td>18 (4)</td>
<td>0.106</td>
</tr>
<tr>
<td>Essential (32kV Rh/Rh)</td>
<td>0.050 (0.003)</td>
<td>3.9 (0.2)</td>
<td>0.016 (0.003)</td>
<td>16 (2)</td>
<td>0.066</td>
</tr>
</tbody>
</table>

Table 2.1: glare parameters for three similar CsI detectors. The values in parenthesis correspond to the errors in the fitted parameters.

Figure 2.11: Frequency response of the glare on a 2000D system (left); effect of glare deconvolution on a lead disk image on a Essential system (right).
Figure 2.12: Parameters of the DTF (see equation (2.25) and Table 2.1) as a function of kV for the Senographe DS. The circles, squares and dots are for the Mo/Mo, Mo/Rh and Rh/Rh beams, respectively. A single error bar is shown for clarity; all the data points had similar errors.

2.3.4 Anti-Scatter Grid Characteristics

As mentioned in Section 1.4.4 and 2.3.1, the anti-scatter grid is characterized by its primary transmission and scatter transmission ratios, $T_P$ and $T_S$ respectively. $T_P$ mostly depends on the density of the anti-scatter grid, while $T_S$ depends mostly on the grid ratio. Moreover, the transmission ratios will increase with an increase in beam energy incident on the grid. A harder beam that results from a high kV or that is transmitted by a thick breast has more penetrating power and will be less attenuated in the grid.

Carton et al. [101], Shen et al. [99] and Savagnini et al. [102] have investigated the transmission ratios of clinical anti-scatter grids, but using different beam energies. Therefore, it was necessary to measure the transmission ratios directly, as follows. Consider two images of a phantom of
thickness $T$, $I_g$ and $I_{ng}$, obtained with and without the anti-scatter grid, respectively. We can express the images as $I_g = P_g + S_g$ and $I_{ng} = P_{ng} + S_{ng}$, where $P_g$ and $P_{ng}$ are the amount of signal due to the primary radiation with a grid present and with no grid, respectively, and where $S_g$ and $S_{ng}$ are the amount of signal due to the scattered radiation with a grid present and with no grid, respectively. The primary and scatter transmission ratios are given by

$$T_p = \frac{P_g}{P_{ng}}, \quad T_s = \frac{S_g}{S_{ng}}.$$

(2-26)

It is possible to isolate the signal due to the primary radiation by collimating the broad cone beam to a narrow pencil beam using a small circular aperture near the x-ray tube. In doing so, the $P_g$ and $P_{ng}$ values can be measured in order to deduce $T_p$. To obtain the primary transmission at $T = 0$, no aperture image was necessary since the image is free from scatter (we neglect the scatter occurring in air). Then another image $I$ is obtained without the aperture, and the scatter can be isolated as $S = I - P$ and the corresponding $T_s$ is computed. As we will see in the next section, choosing the right aperture size will require determining the effects of the off-focus radiation.

### 2.3.4.1 Off-Focus Radiation

In order to find the optimum aperture that allows the primary radiation through, we measured the signal transmitted by a series of apertures of different diameters. The aperture was surrounded by a piece of lead in order to block the rest of the detector from exposure, and was placed approximately 15 cm below the x-ray source. Figure 2.13 shows the measurements, with apertures 0.34, 0.50, 0.74, 1.00, 1.50, 2.0, 3.2, 4.8, 7.2, and 10.8 mm in diameter, drilled in a 0.4 mm sheet of Tantalum. An image with the entire detector exposed (no aperture) was also obtained, but was assigned a 51 mm diameter aperture: the aperture size that would expose the entire detector at the same magnification factor.
As we can see in Figure 2.13, the signal variation displays an inflection point around a 1 mm aperture diameter. This inflection originates from the off-focus radiation [99]. At a small aperture size, the radiation coming from the focal spot is partly blocked, until the critical size (in this case 0.74 mm) where the entire focal spot radiation is transmitted. Increasing the aperture size from this point gradually allows the off-focus radiation through. As we can also see, there is a large difference between the raw signal and the signal after glare has been deconvolved. Glare has a strong effect in high contrast images. The projected sizes of the apertures had radii ranging between approximately 0.6 and 20 mm, which is similar to the lead disks images discussed in Section 2.3.3. The amount off-focus radiation relative to the primary signal, $OF/I$, can be obtained as follows:

$$\frac{I_0}{I_c} = \frac{I + OF}{I} \Rightarrow \frac{OF}{I} = \frac{I_0}{I_c} - 1,$$

(2-27)

Where $I_0$ is the open-field signal (corresponding to the 51 mm aperture size in figure Figure 2.13) and $I_c$ is the signal at the critical aperture. The experiment was repeated for multiple x-ray beam, and the contribution from the off focus is shown in Figure 2.14.
A small increase of $OF/I$ with kV was found, as by Shen et al. [99]. However, the total variation over the range of kV was close to the uncertainty in the measurement. Without deconvolving the glare, the average $OF/I$ (over all beam energies) was $0.124 \pm 0.008$, while deconvolving the glare leads to a significantly smaller contribution from the off-focus radiation: $OF/I = 0.032 \pm 0.008$. Shen et al. found a similar result, with values ranging between 0.035 and 0.055.

### 2.3.4.2 Grid Transmission Measurement

As we have seen in the above section, glare must be taken into consideration when the aperture is used to measure the primary signal $P$, and the contribution of the off-focus radiation to the signal must also be accounted for. In other words, the true primary signal $P$ is obtained from the aperture image $P'$, such that $P = P \otimes G_{PSF}^{-1}$. For consistency, the glare was also deconvolved from the image without the aperture. The total signal can be expressed as $I = (P + S) \cdot (1 + OF/I)$, such that the scatter signal

$$S = \frac{I}{1 + OF/I} - P.$$  \hspace{1cm} (2-28)
The measurements of $P$ and $S$ were done over a range kV for the usual anode/filter combination, with and without the grid, and with phantoms 3, 5 and 7 cm thick, of 30% density, in order to obtain the grid transmission factors $T_P$ and $T_S$. See Figure 2.15.

Figure 2.15: Primary transmission factor (top row and bottom left) and scatter transmission factor (bottom right) for a Senograph DS system, as a function of thickness kV and anode/filter. A single set is shown with error bars in each figure for clarity; the errors were similar for the other points.

As we can see in Figure 2.15, there is an important change in $T_P$ from 0 to 3 cm thickness. However, the changes between 3 to 7 cm, between different kVs or between the different anode/filter combination were smaller and close to the experimental error in the measurement. See Table 2-2. The experimental error in $T_S$ was proportionally more important, so that the differences for different kV and thickness were not significant. The values for $T_S$ at the other anode/filter combination were similar; the average was $T_S = 0.15 \pm 0.02$. Carton et al. [101] and Shen et al. [99] and Salvagnini et al. [102] found higher values of $T_P \sim 0.75-0.80$. All used different x-ray beams, and the methods to account for the glare were different. The $T_S$ values
found here were similar to those found by Carton et al. (0.13-0.22 ± 0.03) and Salvagnini et al. (0.12-0.14). Shen et al. found higher values of $T_S$ between 0.19 and 0.30. Shen et al. used a slot scan system with the slot perpendicular to the grid’s focused orientation. Therefore, the lower scatter rejection rate is expected.

<table>
<thead>
<tr>
<th>Anode/filter</th>
<th>Rh/Rh</th>
<th>Mo/Rh</th>
<th>Mo/Mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>kV</td>
<td>28</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>0 cm</td>
<td>0.64</td>
<td>0.65</td>
<td>0.66</td>
</tr>
<tr>
<td>3 cm</td>
<td>0.69</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>5 cm</td>
<td>0.69</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>7 cm</td>
<td>0.69</td>
<td>0.69</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Table 2-2: Primary transmission factor $T_P$ for a Senograph DS as a function of thickness, kV and anode/filter for a plastic breast phantom of 30% density. The experimental error was on average ± 0.01.

Preliminary measurements were also performed on a high transmission cellular anti-scatter grid (4:1 ratio, copper septa, 23 lines/cm) on a Hologic Selenia machine. The results for $T_P$ at 28 kV Mo/Rh were 0.73, 0.74, 0.75, 0.75 and 0.76 at 0, 2, 4, 6 and 7 cm thickness, with an average error of ± 0.02. The scatter transmission $T_S = 0.08 ± 0.04$. Salvagnini et al. found similar values for $T_P$, but found lower values of $T_S$ ranging between 0.02 and 0.04.

### 2.4 Simulation Results

Using equation (2-18), we can compute the effective attenuation for the plastic breast material. The simulation of images is also described in Section 3.2.3.2. We simulated 20×20 cm images, each containing a 20×12 cm slab of breast plastic, filling the chest-wall side of the image plane. The slabs had a uniform density of 0 and 100% density, with the thickness ranging from 1 to 8 cm in 1 cm increments. The simulations were done using the Mo/Mo, Mo/Rh and Rh/Rh beam with the kV ranging between 25 and 33 kV in 1 kV increments. A ray-tracing algorithm adapted from Siddon [103] by Dr. James Mainprize was used to compute the transmission of an obliquely incident x-ray beam though the phantom volume in order to compute the attenuation times path length product (equation (2-15)). The $\mu T$ maps were interpolated into a simulated detector array
with a coarse detector element 1×1 mm in size. The geometry of the projection mimicked that of a mammography system: the SID was set to 660 mm and the phantom slabs were offset from the detector by an amount close to the bucky thickness at \( d = 10 \) mm. See Figure 2.16. The computation was done using the anti-scatter grid parameters outlined in Section 2.3.4.2 and the glare parameters of the Essential machine (Table 2-1). The CsI detector thickness used in the simulation was 160 \( \mu \)m. Figure 2.17 shows the comparison between the measured attenuation from the calibration, and Figure 2.18 shows the corresponding error in density between the simulation and calibration.

As we can see in Figure 2.17, the simulated attenuation curves are very close to the experimental ones. However, minor differences do lead to an appreciable difference in density, as seen in Figure 2.18. The difference in percent density \( \Delta m \) was computed using the following formula, based on equation (2-6):

\[
\Delta m = m_p - \left( \frac{f_s - f_0}{f_1 - f_0} \right) \times 100 ,
\]

(2-29)
where \( m_p \) is 0 or 100 (the true density), \( f_s \) the simulated effective attenuation of fat or fibroglandular tissue, and \( f_0 \) and \( f_1 \) the calibrated effective attenuation values for fat and fibroglandular tissue, respectively. As we can see in Figure 2.18, the largest errors occur at low thicknesses, close to the converging point in the effective attenuation for fat and fibroglandular tissue. The average root mean square error in density between the simulation and the calibration was 3.2 % (Mo/Mo, T = 3-7 cm, kV = 25-29), 4.5 % (Mo/Rh, T = 2-7 cm, kV = 26-32) and 2.7 % (Rh/Rh, T = 2-7 cm, kV = 27-34). The errors in the simulation are thus similar in magnitude to the intrinsic errors in the calibration, as seen in Section 2.2.2. We note that the model performed best in the Rh/Rh beam, the beam used for the majority of clinical images. Since the accuracy of the simulation is acceptable, the model was used to simulate the effective attenuation of breast tissue, using the linear attenuation coefficients measured by Johns and Yaffe [28]. See Figure 2.19.

The difference between the plastic and breast effective attenuation is significant. The errors are similar to those estimated in Section 2.2.2.3. For a 5 cm breast, imaged at 29 kV Rh/Rh that has 0, 25, 50, 75 and 100 %VBD, using the plastic effective attenuation would lead to calculated density of -18, 15, 48, 81 and 115 %VBD. The errors are similar at different thicknesses and imaging kVs. Breast images can thus be analyzed using the method described in Section 2.2.1, but using the simulated effective attenuation. The results of the analysis of clinical breast images are shown in Chapter 4. The model was also used to simulate images for a GE DS unit in order to validate the algorithm, as described in Chapter 3.
Figure 2.18: Density errors resulting from the differences in the simulated and calibrated effective attenuation as a function of thickness and kV. Top figures show the density difference for the Rh/Rh beam, for the fat attenuation (left) and fibroglandular attenuation (right). The bottom figures show the differences for the Mo/Mo beam, for the fat attenuation (left) and fibroglandular attenuation (right).
2.5 Thickness Sensitivity

There can be some uncertainty in determining the breast thickness, especially when analyzing clinical mammograms. In this section we can calculate the sensitivity of the density measurement with respect to thickness, as was done in Section 1.5.2 with linear attenuation at discrete energy values. The calculation is done using the modeled effective attenuation of breast tissue shown in Figure 2.19. We can consider an object of constant thickness \( T \) and fractional density \( m \), giving rise to an effective attenuation \( f \), so that

\[
f = mf_i + (1 - m)f_o,
\]

where \( f_o \) and \( f_i \) are the calibrated or modeled effective attenuation values for fat and fibroglandular tissue at the thickness \( T \), respectively. The same attenuation \( f \) can be due to an object with a different but erroneous thickness and composition \( T' \) and \( m' \), respectively. Then, the error in the breast density from assuming the erroneous thickness \( T' \) is given by

\[
\Delta VBD = m - m' = m - \frac{f - f_o'}{f_i' - f_o'},
\]

where \( f_o' \) and \( f_i' \) are the calibrated attenuation values for fat and fibroglandular tissue at the thickness \( T' \), respectively. In Figure 2.20 we show the resulting absolute error in percent density.
when \((T - T')/T = \pm 0.03\). We averaged the density errors due to the positive and negative thickness errors; the difference between those was on average less than 0.3 %VBD.

Thus, for a 5 cm breast imaged at 29 kV with a Rh/Rh anode/filter, a 3 % (or 1.5 mm) thickness error will induce a error in density of 4.5, 5.0, 5.6, 6.1 and 6.7 %VBD when the breast is actually 0, 25, 50, 75 and 100 % dense, respectively. The results are similar to those obtained in Section 1.5.2 (a 29 kV Rh/Rh beam has a mean energy of approximately 19 kV).

Similarly, we can investigate the density error as a function of the thickness error. As we can see in Figure 2.21 (left) for a 5 cm breast, the density error is approximately linear with respect to the thickness error. In contrast, the response is non-linear for a thinner 2 cm attenuator (Figure 2.21 right). We can compute the average sensitivity \(\Delta VBD/\Delta T\) as the slope of the calculated relations: the value is only approximate for a thickness lower than 3 cm. Figure 2.22 shows the average sensitivity for the thickness ranges 0-3 cm and 3-8 cm. Table 2-3 summarizes the sensitivity values as a function of breast thickness as breast density. The sensitivity of the density measurement with respect to the thickness is high: for breasts 5-6 cm thick, the sensitivity ranges between 2.5 and 4.5 %VBD per mm of thickness error. For thin attenuators under 1 cm (such as near the breast edge), the sensitivity to thickness is extremely high.
Figure 2.21: Density error $\Delta V_{BD}$ as a function of the thickness error $\Delta T$ for different density values, for a 5 cm thick breast (left) and for a 2 cm breast (right), both imaged at 29 kV with a Rh/Rh anode/filter. A thin attenuator has a non-linear sensitivity response.

Figure 2.22: Average sensitivity of the volumetric breast density (VBD) measurement per mm of thickness error, for an error range $\Delta T = [-5, 5]$ mm. The computations were done assuming a beam of 29 kV with a Rh/Rh anode/filter.
Table 2-3: sensitivity in percent density per mm as a function of thickness, for a uniformly thick breast imaged at 29 kV with a Rh/Rh beam. The sensitivity is computed as the slope of the relation between $\Delta m$ and $\Delta T$, where $\Delta T$ ranges from – 5 to 5 mm.

<table>
<thead>
<tr>
<th>%</th>
<th>1 cm</th>
<th>2 cm</th>
<th>3 cm</th>
<th>4 cm</th>
<th>5 cm</th>
<th>6 cm</th>
<th>7 cm</th>
<th>8 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>15.1</td>
<td>7.1</td>
<td>4.8</td>
<td>3.6</td>
<td>3.0</td>
<td>2.5</td>
<td>2.2</td>
<td>2.0</td>
</tr>
<tr>
<td>25%</td>
<td>17.7</td>
<td>8.2</td>
<td>5.5</td>
<td>4.1</td>
<td>3.3</td>
<td>2.8</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>50%</td>
<td>20.2</td>
<td>9.4</td>
<td>6.2</td>
<td>4.7</td>
<td>3.7</td>
<td>3.1</td>
<td>2.7</td>
<td>2.3</td>
</tr>
<tr>
<td>75%</td>
<td>22.8</td>
<td>10.5</td>
<td>6.9</td>
<td>5.2</td>
<td>4.1</td>
<td>3.4</td>
<td>2.9</td>
<td>2.5</td>
</tr>
<tr>
<td>100%</td>
<td>25.4</td>
<td>11.7</td>
<td>7.7</td>
<td>5.7</td>
<td>4.5</td>
<td>3.7</td>
<td>3.1</td>
<td>2.7</td>
</tr>
</tbody>
</table>

2.6 Summary

In this chapter we have described a method to measure the VBD from digital mammograms. The method can be summarized as follows: 1) the effective attenuation of plastic breast phantoms, representing pure fat and pure fibroglandular tissue, is measured from phantom images; 2) the same attenuation is simulated, using empirical or modeled data describing the x-ray spectrum, x-ray scatter, anti-scatter grid and x-ray detector; 3) once the simulation adequately models the x-ray transmission of breast plastic, the simulation is repeated to yield the effective attenuation of breast tissue; 4) the effective attenuation is fitted to a second degree polynomial as a function of tissue thickness and tube kV; 5) the fractional density at a given point on a mammogram is deduced by comparing its corresponding effective attenuation to the fitted effective attenuation of pure fat and pure fibroglandular tissue; 6) the VBD is obtained by summing the fractional density over the image, taking into account the obliquity of x-rays and of the volume elements above each pixels.

Based on the work presented in Section 2.2.2 and Section 2.4, the breast density method is accurate to approximately 3 %VBD. This represents the baseline accuracy; the images analyzed in these sections were homogeneous blocks of tissue of uniform thickness. In practice, breasts are heterogeneous with a non-uniform thickness, which may affect the algorithm’s accuracy. In
Chapter 3, we present a validation of the algorithm on breast images using CT. In Chapter 4, we use the algorithm to measure the VBD on a large sample of clinical mammograms.
Chapter 3
Validation of a Method for Measuring the Volumetric Breast Density using Computed Tomography

3.1 Introduction

As we have seen in the previous chapters, there is some interest in measuring the volumetric breast density from mammograms. Compared to the traditional methods of measuring breast density, which are area-based and subjective, volumetric breast density has the potential to be a stronger risk factor for breast cancer. The method to measure the volumetric density described in Chapter 2 is relatively simple. Phantom images are obtained and/or simulated in order to calibrate the relation between the imaging signal versus the phantom thickness and composition. Then, this relation is inverted to deduce the composition from the image signal and breast (or phantom) thickness. We have seen in Sections 2.2 and 2.4 that the method (using an empirical or simulated calibration) is accurate to within approximately 3 %VBD, when tested on phantom images. However, there are some important differences between the phantom and breasts. As we have seen in Sections 2.2.2.3 and 2.4, the x-ray attenuation coefficients of the breast plastics phantoms differ from those of human breast tissue, hence the need for the simulated effective attenuation. In addition, the plastic slabs are perfectly homogeneous, with a constant thickness. In comparison, breasts are inhomogeneous, and their thickness varies: the compressed region might bulge or slant due to the compression paddle, and the thickness drops smoothly to zero in the periphery of the breast. Thus the x-ray scatter and the overall effective attenuation of the phantom slabs will differ from those occurring in an actual breast being imaged, and the accuracy of the algorithm might suffer. Thus, it is important to test the accuracy of the method on breast images. In this chapter, we use volumetric attenuation data obtained with a prototype dedicated breast computed tomography (DBCT) system developed at the University of California, Davis [75,76,77] to perform such a validation of the volumetric breast density algorithm. The work presented in this chapter was published in Phys. Med. Biol. 55 (2010) 3027-44 [79] and some of the text and figures are reproduced with permission from IOP Science.
The approach used to validate the algorithm is illustrated in Figure 3.1. Women were imaged on the DBCT scanner, and the image slices were segmented according to the major tissue types: fibroglandular, fat and skin. The CT data were then “deformed” using finite-element analysis in order to simulate the mechanical compression of the breast that occurs in mammography. The deformed CT breast volumes were converted to x-ray attenuation coefficients, according to the tissue type, at the energies corresponding to those used in mammography, and simulated projection digital mammograms were created, following the method described in Section 2.3. Finally, the simulated mammograms were analyzed with the volumetric breast density (VBD) algorithm, as described in Section 2.2, and the resulting volumetric breast density measure was compared to the reference “true” density from the deformed CT volume. Twenty-six cases were analyzed in total. For the density algorithm, we used the empirical $I_0$ maps as opposed to the fitted $S_0$ function, as we have seen in Section 2.2.2.2 that it leads to an increase in the precision of the algorithm.

Figure 3.1: Flow chart summary of the validation process. From top to bottom: sagittal slice from a segmented CT volume, sagittal slice of the deformed volume, negative logarithm of the simulated digital mammogram and calculated density map of the mammogram. Used with permission from IOP Science [79].
3.2 Material and Methods

3.2.1 Breast CT

The prototype DBCT scanner is described in detail in the references mentioned above. In short, a patient lies prone on a table with the breast pendant through an aperture during imaging. An x-ray tube and a flat-panel detector array rotate about the breast to acquire the projection images. The cone-beam projections are acquired at 80 kV with a Tungsten target and a 0.3 mm Copper filter. The CT images are reconstructed from the projection data as a volumetric array of effective attenuation coefficients, representing breast tissue. These measurements reflect the x-ray energy fluence transmitted by the breast and absorbed in the flat-panel detector, and they include the effects of x-ray scattering and beam hardening. Following the image reconstruction, a “cupping” correction was performed to correct the scatter and hardening artifacts. The reconstructed volume elements (voxels) had dimensions ranging between 0.21 and 0.41 mm (with a 512×512 field-of-view) in the coronal plane, while their axial length varied between 0.25 and 0.41 mm.

The voxels in the CT image were classified into four components: air (the breast exterior), skin, adipose and fibroglandular tissue. The frequency distribution of the reconstructed attenuation coefficients was computed and the threshold between the breast and the air, which has low attenuation values, was determined manually. See Figure 3.2. Because skin and fibroglandular tissue have similar attenuation values, we initially used a two-compartment Gaussian fit to segment the histogram between fibrogladular (plus skin) tissue and adipose tissue. This allowed the determination of an approximate threshold attenuation value at the intersection of the two Gaussian distributions. In most cases, the Gaussian functions were not well separated, and in order to verify whether the threshold was satisfactory, images of central slices in the three orthogonal main planes were displayed. The threshold was then adjusted manually until the most satisfactory separation between skin/fibroglandular tissue and fat was obtained. See Figure 3.2.
Figure 3.2: Histogram of the CT attenuation for one case in the data set. We can easily identify the threshold at low CT numbers that separates the air from the breast. We can see that the peaks in attenuation corresponding to the fat and fibroglandular tissue are poorly separated. An initial two-compartment Gaussian fit gave a good estimate of the threshold between the two tissues, but it was adjusted manually to a lower value to obtain a better separation between the two tissue types.

To identify the voxels representing skin, the periphery of the breast was determined for each axial slice, and a morphological “dilation” operator was applied to contract the periphery inwards towards the breast centre by an amount similar to the thickness of the skin (approximately 1.5 mm [89]). This operation allowed the identification of a ring of tissue just inside and around the breast. The voxels that were in the peripheral ring and also previously segmented in the fibroglandular class (which included the skin) were classified as skin voxels.

Finally, a median filter of size $3 \times 3$ was applied to the segmented slices. The segmentation reduced the noise in the image but removed some detail in the parenchyma. See Figure 3.3. As discussed in Section 3.4, we note that an accurate segmentation of the original CT data is not critical in this study, since the segmented data is directly used to generate the simulated mammograms: the segmented volumetric density is thus used as the ground truth, as opposed to the actual anatomic volume density.
The volumetric breast density (VBD) in the segmented CT images was calculated as

$$VBD = \frac{V_D + V_{sk}}{V_D + V_F + V_{sk}},$$  \hspace{1cm} (3-1)$$

where $V_D$, $V_F$ and $V_{sk}$ are the volume of dense (fibroglandular), fatty (adipose) and skin tissues, respectively. We also computed the VBD excluding the skin as

$$VBD_{sk} = \frac{V_D}{V_D + V_F + V_{sk}}.$$  \hspace{1cm} (3-2)$$
3.2.2 Deformation

3.2.2.1 General

A number of finite-element studies have been conducted to simulate the deformation of breast tissue, in order to investigate the effects of material properties [104] and mechanical compression [105, 106]. These tools were used in our study to simulate the mechanical displacements that would occur under compression of the breast in mammography.

From the segmented data obtained from the CT images of the volunteers, 3D finite-element representations of the breast were constructed, and the linear elastic material properties [106] as listed in Table 3-1 were attributed correspondingly to the skin, fat and fibroglandular tissue classes. The modulus of elasticity is defined as the slope of the stress-stain curve in the elastic (linear) deformation regime. Stress corresponds to the force or pressure applied to the material, and the strain is the ratio of the resulting change in length of the material to the original length of the material. Poisson’s ratio is the negative ratio of the transverse strain to axial strain. If a material is stretched (or compressed) along one direction it will likely contract (or stretch) in the two directions perpendicular to the stretching, and Poisson’s ratio is a measure of the contraction relative to the stretching (or vice-versa).

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Modulus of elasticity [kPa]</th>
<th>Poisson’s ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fat</td>
<td>1.0</td>
<td>0.49</td>
</tr>
<tr>
<td>Fibroglanular</td>
<td>10.0</td>
<td>0.49</td>
</tr>
<tr>
<td>Skin</td>
<td>88.0</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Table 3-1: Elastic properties of breast tissue. From ref. [105].

To simulate the compression of the breast, the finite-element breast model is placed between two rigid plates with rounded corners (with a 0.5 cm radius). The bottom plate is initially raised to a fixed position under the breast to provide support, and the top plate moves downwards to achieve the specified target compression thickness. See Figure 3.4. The plates were offset towards the nipple (i.e. away from the chest wall) by approximately 2 cm to facilitate the modeling of the
compression. The target compression thickness was selected as a function of breast diameter at the chest wall (on the first few axial CT slices), using the relation determined by Boone et al. [107] on patients that had both the DBCT scan and a mammography exam.

3.2.2.2 Boundary Conditions

The surface of the breast at the chest wall (the first axial CT slice) was allowed to be displaced in the superior-inferior direction only. The breast was allowed to slide on the plates as compression was applied, using a contact-surface model consisting of the external surface of the skin and the surface of the plates facing the breast. A coefficient of friction of 0.1 was used to simulate the movement restriction of the breast in contact with the plates. Similar values have been used by Pathmanathan et al. [108] and Yin et al. [109], but to the best of our knowledge, experimental coefficients of friction between the breast and compression plate are not available. In order to accommodate large deformations in the breast, non-linear geometry was used in the analysis.

3.2.3 Mammogram Simulation

3.2.3.1 Deformation

The deformation algorithm produced a displacement vector \( \mathbf{s}(\mathbf{r}) = [s_x(\mathbf{r}), s_y(\mathbf{r}), s_z(\mathbf{r})] \) for each point \( \mathbf{r} = (x, y, z) \) in the initial segmented CT volume. We can denote the segmented volume by \( m_0(\mathbf{r}) \), where \( m_0 \) represents the tissue composition, having the two possibilities of \( m_0 = 1 \) for fibroglandular or skin tissue and \( m_0 = 0 \) for adipose tissue. The displacement vector was used to compute the displaced density image \( m(\mathbf{r}) \), so that

\[
m(\mathbf{r} + \mathbf{s}(\mathbf{r})) = m(x + s_x(\mathbf{r}), y + s_y(\mathbf{r}), z + s_z(\mathbf{r})) = m_0(x, y, z) = m_0(\mathbf{r}).
\] (3-3)

The displacement values generally had a finer resolution than the voxel dimensions (0.2-0.4 mm) in the CT image, and the “destination” points \( \mathbf{r} + \mathbf{s}(\mathbf{r}) \) fell between voxels of the CT image. Thus, the density at the destination points was interpolated onto a regular Cartesian grid with the same dimensions and voxel size as the original density image \( m_0(\mathbf{r}) \).
The interpolation was performed using a three-dimensional 3×3×3 Gaussian kernel scaled to the voxel dimensions. That is, for the point \( r' \) in the grid closest to \( r + s(r) \), the interpolated density is

\[
m(r') = m_0(r) \times A(r') \times \exp[-(\Delta x^2 + \Delta y^2 + \alpha^2 \Delta z^2)/(2 \cdot \sigma^2)],
\]

(3-4)
where \( \mathbf{r} + \mathbf{s}(\mathbf{r}) - \mathbf{r}' = (\Delta x, \Delta y, \Delta z) \). The factor \( A(\mathbf{r}') \) is used to normalize the Gaussian kernel to unit volume, and \( \sigma \) is set to half a voxel length (0.5). The grid was normalized to the coronal plane dimensions, so that \( \alpha = d_{x,y}/d_z \), the ratio of the voxel dimensions in the \((x,y)\) plane to the voxel length in the axial \((z)\) direction. The operation was repeated for the 3×3×3 region around \( \mathbf{r}' \), i.e. for \( \mathbf{r}' + (i,j,k) \) where the each index spans the values \(-1,0,1\). The points in the destination grid were populated multiple times by different displacement vectors through the interpolation operation. Thus, the Gaussian weights \( A(\mathbf{r}') \) were tallied for every point in the destination grid, and the summation of the weights was used to normalize the resulting displaced map \( m(\mathbf{r}) \) at every point, which now had fractional density values ranging between 0 and 1.

A second displacement and interpolation operation was performed to displace the skin voxels only, that is, the map \( m_0(\mathbf{r}) \) was converted such that skin had a value of 1 and both fibro glandular and fatty tissue had a value of 0. In this case, the displaced fractional skin values were rounded to the nearest integer, and the two displaced maps where merged. This allowed for the labeling of the displaced skin, while avoiding the confusion between partial skin and partial fibro glandular volumes; fractional skin voxels were treated as fractional fibro glandular voxels.

### 3.2.3.2 Projection

The displaced density map \( m(\mathbf{r}) \) was converted into x-ray attenuation values using

\[
\mu(\mathbf{r}; E) = m(\mathbf{r}) \cdot \mu_{fg}(E) + (1 - m(\mathbf{r})) \cdot \mu_f(E),
\]

where \( \mu_{fg}(E) \) and \( \mu_f(E) \) are the attenuation of fibro glandular and fatty tissue at the x-ray energy \( E \), respectively. The attenuation values were obtained from Byng et al. [91] for breast plastic phantoms. The skin voxels were also converted to attenuation, using the data of Hammerstein et al. [90]. The simulated projection was similar to the projections done in Section 2.4. The x-ray tracing algorithm adapted from Siddon [103] was used to model the transmission of oblique x-rays though the volume, and for each x-ray path incident on the detector plane \((x,y)\), the product of the attenuation and thickness and at each energy, \( \mu T(x,y; E) \) was computed. The \( \mu T \) maps were then interpolated onto the simulated detector array at a 0.25 mm detector element size. The projection geometry was the same as in Figure 2.16: the lower surface of the
breast was assumed to lie \( d = 10 \) mm above the detector plane (corresponding to the upper surface of the anti-scatter bucky), with a SID of 660 mm. The source was located directly above the centre of the chest-wall edge of the detector. The projected thickness of the breast \( T(x,y) \), including the obliquity effects, was determined by projecting a volume of unity attenuation. The compression thickness was calculated as an average over the relatively flat portion of the thickness map. The projections were done at the energies that corresponded to a spectrum chosen as a function of the compression thickness. Three sets of projections were performed for the three target/filter combinations. See Table 3-2.

<table>
<thead>
<tr>
<th>Compression thickness [mm]</th>
<th>kV (Rh/Rh)</th>
<th>kV (Mo/Rh)</th>
<th>kV (Mo/Mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-50</td>
<td>28</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>50-70</td>
<td>29</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>70-80</td>
<td>30</td>
<td>29</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 3-2: Correspondence between the spectrum and breast thickness

The simulated images were calculated following the method described in Section 2.3, for a Senograph DS mammography system. We used a CsI thickness of 120 \( \mu \)m. The relative signal due to the directly transmitted primary x-rays was computed using equation (2-16). The relative signal due to the transmitted scattered x-rays was calculated using equation (2-21). The images were computed using equation (2-18), using the glare and anti-scatter grid coefficients for the DS machine found in Sections 2.3.3 and 2.3.4, respectively. The values for \( T_p \) and \( T_s \) were linearly interpolated at the compression thickness of the breast. An experimental open field image for the machine, \( I_0(x,y) \), expressed in analog-to-digital units (ADU), was obtained for all of the exposure conditions used in the simulations.

In order to test the accuracy of the simulation, an analysis similar to that shown in Section 2.4 was done. Images of slabs of breast plastic phantoms of 0 % density and of dimensions 10×12.5 cm (oriented with the short axis parallel to the chest wall), with thicknesses of 3, 5 and 7 cm, were simulated at 26, 28, 30 and 32 kV with a Rh/Rh beam and compared with the
corresponding experimental images. The differences in signal \( \Delta I = I_{\text{exp}} - I_{\text{sim}} \) were observed, where \( I_{\text{exp}} \) and \( I_{\text{sim}} \) are the mean signal values within a central region of the phantom images for the experiment and simulation, respectively. \( \Delta I \) was also converted to its corresponding difference in breast density, such that \( \Delta \text{VBD} = \Delta I \cdot S_I \). The sensitivity \( S_I \) was calculated as

\[
S_I = \left. \frac{dVBD}{dI} \right|_{\text{VBD}=0}.
\] (3-6)

For this calculation, a series of simulated phantom images were generated as described above, with a density ranging from 0 to 100 % in 1% increments. \( I \) corresponds to the image pixel value in a region at the centre of a simulated phantom image of a given VBD.

3.3 Results

3.3.1 CT Analysis and Deformation

As a first test of the deformation procedure, we compared the volumes and volumetric density of the CT data and of the displaced CT data. We found that the displacement conserved the total breast volume within an average of 1.5 %, whereas the volumetric density was conserved within 4.4 % on average, relative to the original data. As mentioned in Section 3.2.2.2, the mechanical deformation procedure required offsetting the compression plate toward the nipple for the application of boundary conditions. This resulted in approximately 33 % of the total breast volume, located in a region proximal to the chest wall, not being compressed when the deformation was applied. For the purposes of this study, only the portion of the breast uniformly compressed was used for the mammogram simulation, and the remainder of the breast was cropped. Therefore, the absolute volumes and dense tissue volume of the deformed and cropped breast will be substantially lower than those of the entire breast. See Figure 3.5 for an illustration of the deformation and cropping process.

For the purposes of comparison with the VBD algorithm, the VBD determined by CT, \( \text{VBD}_{\text{CT}} \), was computed as the mean of the displaced fractional density volume \( m(r) \) within the cropped region. The total breast volume determined by CT, \( V_{\text{CT}} \), was also calculated on the deformed and cropped region. The average \( \text{VBD}_{\text{CT}} \), \( V_{\text{CT}} \) and compression thickness of the 26 cases were
23.6%, 558 cm³ and 62 mm, respectively. The mean relative difference between $V_{BD_{CT}}$ and the un-cropped original CT image was 9%. The difference is explained by the presence of pectoralis muscle, segmented as fibroglandular tissue, often present near the chest wall but removed by the cropping.

Figure 3.5: Illustration of the finite-element deformation. Left: original segmented CT image; right: displaced fractional density image. From top to bottom: sagittal, transverse and coronal slices. The white vertical line represents the cropping to remove the non-compressed part of the breast. Used with permission from IOP Science [79].
3.3.2 Mammogram Simulation

Figure 3.6 shows the distribution of values, in ADU/mAs, in the simulated and experimental phantom images. The simulation and experiment were done for a thickness of 3, 5 and 7 cm and at a kV of 26, 28, 29, 30 and 32, with a Rh/Rh anode/filter. The distributions of signal were obtained over a central region of the phantom image, covering approximately 85% of the projected surface of the phantom. Figure 3.7 shows the comparison of the means of the signal distributions, and the corresponding difference in density $\Delta$VBD. The sensitivity $S_i$ in percent VBD ($\%$VBD) is shown in Figure 3.8 and in Table 3-3.

The distributions of values between the experimental and simulated images were remarkably similar over the kV and thickness range. The root mean square (rms) $\text{rms}(\Delta I)$ was 0.72 ADU/mAs, which corresponded to a rms error of 2.4% in the signal. Those errors in the signal resulted in the rms difference of 2.4 PVBD between the experiment and simulation.

![Figure 3.6: Distribution of ADU/mAs values on simulated and experimental images of a plastic breast phantom of thickness 5 cm (a), and 7 cm (b), imaged at 28 and 30 kV respectively, with a Rh/Rh anode/filter. The distributions are computed over an area of approximately 107 cm$^2$ on the projected area of the phantom. Used with permission from IOP Science [79].](image-url)
Figure 3.7: (a) Comparison between the means of the experimental (solid line) and simulated (dashed line) phantom images as a function of thickness and kV. The marker size is representative of the error in the experimental value. (b) Corresponding difference $\Delta V_{BD}$ between experiment and simulation. Used with permission from IOP Science [79].

Figure 3.8: Image signal in ADU/mAs as a function of density and thickness, for a 29kV Rh/Rh beam. The sensitivity $S_f$ for each thickness is the inverse of the slopes of the dashed lines.
### Table 3-3: Sensitivity $S_I$ in %VBD/ADU/mAs.

<table>
<thead>
<tr>
<th>T [cm]</th>
<th>26</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-4.07</td>
<td>-2.73</td>
<td>-2.30</td>
<td>-1.97</td>
<td>-1.50</td>
</tr>
<tr>
<td>5</td>
<td>-7.61</td>
<td>-4.81</td>
<td>-4.00</td>
<td>-3.38</td>
<td>-2.48</td>
</tr>
<tr>
<td>7</td>
<td>-15.01</td>
<td>-9.15</td>
<td>-7.48</td>
<td>-6.23</td>
<td>-4.50</td>
</tr>
</tbody>
</table>

#### 3.3.3 Comparison of the VBD Algorithm with the CT Data

The 26 CT volumes were segmented, deformed, projected and used to simulated and image using equation (2-18). The images were then analyzed with the mammographic VBD algorithm to obtain VBD$_{mammo}$ (with and without skin) and the total breast volume $V_{mammo}$. The algorithm was tested for self-consistency by analyzing images of the calibration phantoms, as in Section 2.2.2.2. Figure 3.9 shows the frequency distribution of the difference between the true density VBD$_{exp}$ and the calculated density VBD$_{calc}$, for phantoms 3-8 cm imaged at 26, 28, 30 and 32 kV with a Rh/Rh beam. The accuracy and precision of the algorithm was determined to be 1.25 %VBD and 2.3 %VBD, respectively.

![Figure 3.9: Frequency distribution of the VBD difference between the true density of the phantom and the calculated density using the algorithm. The results show are for the Rh/Rh beam. Used with permission from IOP Science [79.](#)
For the simulated breast images, we restricted the computation of the VBD to points where the density was in the physically acceptable range from 0 % to 100 %. On average, for the 26 cases simulated with a Rh/Rh beam, only 6.4 % of the total number of pixels in the breast images were excluded for this reason. These pixels were close to the outside edge of the breast where it is thin: they corresponded to only 2.3 % of the breast volume on average. For the Mo/Rh and Mo/Mo sets, these pixels corresponded to 1.9 % of the breast volume. The likely cause of these non-physical results is discussed in the next section.

Good agreement was observed between VBD_{CT} and VBD_{mammo}, with and without the skin. See Figure 3.10 and Figure 3.11. Table 3-4 shows the results of the linear regression analysis, and also the rms difference rms(ΔVBD), were ΔVBD = VBD_{CT} - VBD_{mammo}. The average VBD_{mammo} for the Mo/Mo, Mo/Rh and Rh/Rh target/filter combinations were 25.4 %, 26.0 % and 22.0 % respectively, compared to 23.6 % for VBD_{CT}.

Excellent agreement between V_{CT} and V_{mammo} was observed. See Figure 3.12 and Table 3-4. The results for the skin volumes are also shown in Figure 3.12 and Table 3-4.

Figure 3.10: Comparison between the VBD from CT and from the mammographic breast density algorithm, for the set of images simulated with a Rh/Rh beam. The dashed line is the identity function. Used with permission from IOP Science [79].
Figure 3.11: Comparison between the VBD from CT and from the mammographic breast density algorithm, for the set of images simulated with a Mo/Rh beam (a) and a Mo/Mo beam (b). The dashed line is the identity function. Used with permission from IOP Science [79].

Figure 3.12: Comparison between the total breast volume (left) and skin volume (right) from the CT image and the mammographic density algorithm. The dashed line is the identity function.
Table 3-4: Linear regression and rms analysis between $V_{BDCT}$ and $V_{BDmammo}$ (VBD with and without the skin) and between $V_{CT}$ and $V_{mammo}$ (total breast volume $V$ and skin volume $V_{sk}$).

### 3.4 Discussion and Conclusion

The purpose of this study was to investigate the performance of an algorithm that measures the VBD from digital mammograms. We observe that the algorithm agrees very well with the values obtained from breast CT, both on the total volume of the breast and on the VBD with or without including the skin, as shown in Figure 3.10, Figure 3.11, Figure 3.12 and Table 3-4. To the best of our knowledge, only the algorithm of Van Engeland et al. [110] also proposed a validation by means of comparison with a 3D dataset. Van Engeland et al. used a simple polychromatic model of primary attenuation, without considering the effects of scatter or detector efficiency. Internal calibration was achieved by finding a reference fatty pixel value on the image. The validation was done by comparison with segment MR images of the same 22 patients for which 4 digital mammographic views (2 for each breast) were obtained. The total breast volumes from the MR images and mammograms were not compared, but a correlation of 0.97 with a relative error of
13.6% was observed between the dense tissue from MR and average dense tissue from the four mammographic views of each patient.

As discussed in Section 2.2, the volumetric density algorithm is similar to those of Pawluczyk et al. [87] and Kaufhold et al. [88], in that it uses an empirical calibration of the detector signal as a function of thickness and composition from plastic breast phantoms to determine the density of a breast image acquired under known conditions (beam quality and breast thickness). As seen in Section 2.2.2, the algorithm is susceptible to various sources of error, and the baseline accuracy of the algorithm is approximately 2%VBD when tested on the calibration phantoms (see Figure 3.9 and Section 2.2.2.2). Similarly, we have seen in Figure 3.7 and in Section 2.4 that the simulation model is accurate to within approximately -4 to 3%VBD. Thus, the rms error between the ground truth from CT and the results of the algorithm, of approximately 2.5%VBD for the VBD including the skin, and for the 3 target/filter combination, suggests that the algorithm performs in line with expectations when tested on the simulated breast images; the main sources of error are thus likely due to errors in the calibration and in the simulation.

We can analyze the errors associated both with the simulation and the algorithm in more detail by separating the image set (with the Rh/Rh beam) into kV subsets. Table 3-5 indicates the the average $\Delta VBD = VBD_{CT} - VBD_{mammo}$ and the predicted average density error $\Delta M$ due to the model, for the different kVs. $\Delta M$ was linearly interpolated or extrapolated from the results of Figure 3.7 to the appropriate thickness of each set. We can add the $\Delta M$ error from the model with the systematic 1.25 error arising from the algorithm (Figure 3.9), and we obtain the predicted VBD error shown in Table 3-5. The predicted error agrees with $\Delta VBD$ within 1.5%VBD. The slight discrepancy may arise from the calculation of $\Delta M$, which was done at 0% density, while the breasts in this study had a higher density. As we can see in Table 3-3 and Figure 3.8, the sensitivity $S_I$ increases with an increase in density.

Another type of error occurs when the algorithm yields a density value outside of the acceptable range between 0% and 100%. The average thickness where those errors occurred was 18 mm, indicating that the erroneous point tended to be located mostly in the peripheral region. There are several factors that may be responsible for these errors. The calibration curves for fat and fibroglandular tissue converge to the same value at 0 thickness, and thus are close together for thin attenuators, causing the density estimation to be more sensitive to small errors in measuring
or simulating the image signal at lower thickness. Moreover, the thickness of the breast decreases in the peripheral region, and the scattered radiation profile will differ between the calibration exposures, which were obtained on phantoms that have uniform thickness, and actual mammograms, causing further errors.

<table>
<thead>
<tr>
<th></th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cases</td>
<td>3</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Mean VBD [%]</td>
<td>32 %</td>
<td>26 %</td>
<td>17 %</td>
</tr>
<tr>
<td>Mean T [cm]</td>
<td>3.9</td>
<td>5.9</td>
<td>7.4</td>
</tr>
<tr>
<td>Observed error</td>
<td>-3.0</td>
<td>2.0</td>
<td>2.5</td>
</tr>
<tr>
<td>ΔVBD [%VBD]</td>
<td>-2.7</td>
<td>0.8</td>
<td>2.9</td>
</tr>
<tr>
<td>ΔM [%VBD]</td>
<td>-1.4</td>
<td>2.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Predicted error</td>
<td>-1.4</td>
<td>2.0</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Table 3-5: Observed and predicted difference as a function of kV for the Rh/Rh image set.

Another source of error may arise from the assumption in the algorithm of a strictly two-compartment model for the breast. That is, the algorithm assumes that for a given tissue thickness and corresponding image signal, there is a unique combination of fat and fibroglandular tissue possible. Certainly, the presence of skin (and of blood vessels) in the breast is a departure from the two-compartment model. The skin has attenuation very close to that of fibroglandular tissue, and thus such errors are likely to be noticeable only for pixels where the density is already close to 100 % or 0 %, causing the density to fall outside the expected range.

We note that the performance of the algorithm in estimating VBD_{sk} was not as good as VBD_{fk}. This is likely due to the discrepancies observed between the calculated and true skin volumes (Figure 3.12), which arises from the assumption of a constant total skin thickness of 3 mm (1.5×2 mm). In actuality, the skin thickness in the CT images varied between patients and within the same breast. Moreover, equation (2-9) for the skin volume is only valid for the uniformly thick portion of the breast. In the peripheral region, the breast has a rounded shape, so that the
skin layer thickness effectively increases towards the edge of the breast in the projection. This would cause the skin volume to be underestimated, as we have observed.

Using actual mammograms with the corresponding CT from the same volunteers for purposes of validation will remove any errors due to the simulation of images. On the other hand, there are some advantages in using simulated mammograms. This approach allows for a self-consistent comparison between CT and the density algorithm. The validation is, therefore, insensitive to potential errors in the absolute anatomical VBD due to segmentation or the finite-element deformation since the VBD is preserved in the simulated projection mammogram. Furthermore, we compare a breast CT volume with its direct projection with no potential discrepancy in positioning between the two. And finally, the thickness of the breast is known exactly, and potential errors due to a thickness measurement or estimation are avoided. Thus, this validation offers an estimate of the baseline error of the algorithm when tested on breast images. We have found that the performance of the algorithm is similar to its performance when assessed on phantom images; it yields a measure of the percent volumetric breast density within approximately 3 percentage points.

In Chapter 4, we submit the VBD algorithm to a further test. A large sample of clinical mammograms, for which there is some uncertainty in the breast thickness, and for which there is no ground truth to validate the measurement, was analyzed with the algorithm. In Chapter 5, we present another validation of the algorithm, this time comparing the VBD values from a subset of clinical mammograms to the VBD measurement done on MR images of the same patients.
Chapter 4
Volumetric Breast Density Characteristics as Determined by
Digital Mammograms

4.1 Introduction

In Chapter 1, we have discussed about the interest in using volumetric breast density (VBD) measurements as a risk factor for breast cancer. In Chapter 2 we have described a method to measure the volumetric breast density (VBD) from digital mammograms, and in Chapter 3 we have presented a validation of the method using CT images. In this chapter, we present the results of the VBD analysis done on 55,087 digital images, from 15,351 individuals imaged at the Sunnybrook Health Sciences Centre (SHSC) breast imaging clinic. The relationship of density versus age and versus time (for a given individual) were investigated, as well as the general distributions of the VBD, the VBD with no skin, the breast volume and the comparison between the left and right breast density and volume.

Analyzing mammograms presented a new challenge, that of accurately determining the breast thickness. In the previous chapters, the thickness of the phantoms or of the simulated mammograms was known. In mammography the breast is compressed via a relatively flexible plastic plate, and the mammography system reports the thickness in the header of the digital image. The accuracy of the readout thickness varies as a function of the compression force, with the type of paddle, and with mammography unit. Moreover, the thickness in the compressed region might vary due to the flexibility of the paddle or by design: some compression paddles are hinged in order to alleviate some of the pain of compression, and as a result the plate slants forward. In addition, in the periphery of the breast (outside of the compressed region) the breast will assume a rounded shape, and the thickness will smoothly decrease to zero. Accurately determining the compression thickness is essential in making an accurate VBD measurement; as we have seen in Section 2.5, a 1.5 mm error in assessing the thickness of a breast that is actually 50 mm and 25% dense will lead to an error of 5% VBD. In Section 4.2 we describe the
characterization of the thickness response of the mammography machines used in this study, as well as the algorithm used to detect the peripheral region of the breast.

The work presented in this chapter was published in ref. 112 and ref. 111.

4.2 Thickness Estimation

4.2.1 Compression Thickness Correction

The density algorithm includes a model based on the work of Hauge et al. [63] and Mawdsley et al. [61] to estimate the compressed breast thickness. A flexible breast phantom was compressed at different forces, and its thickness was measured at multiple points on the compressed region. The flexible phantom is made of gel and is approximately semi-circular in shape, with a radius of approximately 20 cm. It has a small ledge near the chest wall side so that it can be positioned with consistency in relation to the support plate. The thickness measurement was done using a stand resting on the ground, to which a plate is attached and parallel to the breast support table. The plate contains a series of holes, regularly positioned in a grid spaced 20 mm apart. See Figure 4.1. A digital height gauge is inserted in the holes in order to measure the height of the plate with respect to the breast support plate and the compressed breast phantom (plus the compression paddle), denoted by $T_0$ and $T'$ respectively. The hole’s diameter on the plate is matched to the gauge size to reduce location errors. The height readout of the gauge was zeroed to the bottom of the plate in order to exclude its thickness. Thus, the thickness $T$ of the compressed breast on a point on the grid is given by:

$$T = T_0 - T' - t_p, \quad (4-1)$$

where $t_p$ is the compression paddle thickness, which was measured separately to 2.5 mm. The measurements were done at the compression forces of 5, 10 and 15 dN, for four paddle types (two hinged paddles, 19×23 and 24×30 cm; and two semi-rigid paddles, also 19×23 and 24×30 cm), and for the three GE essential machine of the clinic at the Sunnybrook Health Sciences Centre. The measurement was done on six points: on three points along a line parallel and as close as possible to the chest wall ($x$ axis), spaced by 40 mm; at three points along a line parallel to the first line but 40 mm forwards towards the nipple. See Figure 4.1.
For the semi-rigid paddles, the thickness across the compressed region was uniform within a standard deviation of 0.5-0.8 mm, and the thickness for all the grid points was averaged at a given compression force. The hinged paddles slant forward by design, and the percent slope of the slant was characterized as a function of compression force. To determine the slope, the thickness was averaged for the two lines \( x = 0 \) and \( x = 40 \) mm, so that the slope \( S = 100\% \times \frac{\overline{T}_{40} - \overline{T}_{0}}{40} \). The thickness along a given line was uniform within 0.2-1.0 mm. See Figure 4.2 for an illustration of the thickness measurement.
Using this method, the thickness and the slant of the phantom can be determined as a function of the compression force. Figure 4.3 shows the thickness and the percent slope as a function of compression force $P$ of two paddles for the three mammography units. A linear function was fitted to the slant data, such that $S = m + t \times P$, see Table 4-1.

![Figure 4.3: Thickness (left) and percent slope (right) of the phantom as a function of compression force. The left figure is for the 19×23 cm semi-rigid paddle; the right figure is for the 24×30 cm hinged paddle. The dashed lines represent a linear least-square fit to the data. The error bars originate from the standard deviation in the thickness measurement.](image)

Table 4-1: Coefficients of the force response for the hinged paddles.

<table>
<thead>
<tr>
<th></th>
<th>Unit 1</th>
<th></th>
<th>Unit 2</th>
<th></th>
<th>Unit 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t$ [%/N]</td>
<td>$m$ [%]</td>
<td>$t$ [%/N]</td>
<td>$m$ [%]</td>
<td>$t$ [%/N]</td>
</tr>
<tr>
<td>19×23 hinged</td>
<td>-0.045</td>
<td>-6.4</td>
<td>-0.039</td>
<td>-3.8</td>
<td>-0.035</td>
</tr>
<tr>
<td>24×30 hinged</td>
<td>-0.028</td>
<td>-4.3</td>
<td>-0.009</td>
<td>-5.0</td>
<td>-0.023</td>
</tr>
</tbody>
</table>

In order to transfer those results to breasts of different sizes than the phantom, we characterized the difference between the measured thickness $T$ of the phantom and the thickness readout $T_{ro}$ of the mammography machine. That is, we measured $T - T_{ro}$ as a function of compression force $P$, for the different paddles and machines, and computed a linear least-square fit $T - T_{ro} = a + b \times P$.

Figure 4.4 shows that relation for two paddles, and Table 4-2 summarizes the results. For the hinged paddles, the measured thickness at the chest wall ($x = 0$) was used. When analyzing a breast image, the paddle type, machine type, compression force and compressed thickness are
extracted from the digital image header, and the coefficients found in Table 4-2 are used to estimate the breast thickness. For the hinged paddles, the slant in the paddle is also extracted using the values of Table 4-1, and is applied on the thickness map of the image. \( T_{ro} \) and \( P \) are given to the nearest mm and dN, respectively.

The relations between slant and thickness different and compression force are well modeled by a linear function. The experimental accuracy in \( T - T_{ro} \) is approximately \( \pm 1 \) mm, and the percent slant measurement is accurate to within 2.5 percent. By comparing Figure 4.3 and Figure 4.4 we can see that the thickness readout accuracy for the small semi-rigid paddle varies between mammography units: at 80 N of compression force it is very accurate for the units 1 and 3,

<table>
<thead>
<tr>
<th></th>
<th>19×23 rigid</th>
<th>19×23 hinged</th>
<th>24×30 rigid</th>
<th>24×30 hinged</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a ) [mm]</td>
<td>2.2</td>
<td>1.6</td>
<td>0.0</td>
<td>5.6</td>
</tr>
<tr>
<td>( b ) [mm/N]</td>
<td>-0.033</td>
<td>-0.014</td>
<td>0.018</td>
<td>0.009</td>
</tr>
<tr>
<td>Unit 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a ) [mm]</td>
<td>0.3</td>
<td>2.9</td>
<td>-1.2</td>
<td>4.9</td>
</tr>
<tr>
<td>( b ) [mm/N]</td>
<td>-0.035</td>
<td>-0.036</td>
<td>0.012</td>
<td>-0.0175</td>
</tr>
<tr>
<td>Unit 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a ) [mm]</td>
<td>2.4</td>
<td>0.9</td>
<td>0.8</td>
<td>6.0</td>
</tr>
<tr>
<td>( b ) [mm/N]</td>
<td>-0.031</td>
<td>-0.024</td>
<td>0.0175</td>
<td>-0.004</td>
</tr>
<tr>
<td>Unit 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-2: linear least-square fit coefficients for the thickness correction parameter.

Figure 4.4: Measured thickness of the phantom minus the readout thickness of the mammography machine versus compression force for the semi-rigid 19×23 cm paddle (left) and for the hinged 24×30 cm paddle (right). The dashed lines represent the linear least square fit.
whereas unit 2 overestimates the thickness by 2.5 mm. As we can see in Figure 4.4 and Table 4-2, the readout thickness is largely inaccurate for the 24×30 cm hinged paddle: the thickness is underestimated by 3 to 6 mm at 80 N. Moreover there is more variability in the response as a function of mammography units for the hinged paddles.

The thickness correction was applied to the 55 087 images of the clinical cohort discussed in the Section 4.3. Table 4-3 presents the distribution in the type of paddle and in the mammography unit used in this study. Unit 1 was used less frequently than the two other machines: it was in a room typically used for biopsy procedures.

<table>
<thead>
<tr>
<th>19×23 rigid</th>
<th>19×23 hinged</th>
<th>24×30 rigid</th>
<th>24×30 hinged</th>
<th>Sum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 1</td>
<td>5801</td>
<td>114</td>
<td>362</td>
<td>2922</td>
</tr>
<tr>
<td>Unit 2</td>
<td>14 797</td>
<td>106</td>
<td>991</td>
<td>5564</td>
</tr>
<tr>
<td>Unit 3</td>
<td>16 836</td>
<td>141</td>
<td>488</td>
<td>6965</td>
</tr>
<tr>
<td>Sum (%)</td>
<td>37 434 (68)</td>
<td>361 (0.6)</td>
<td>1841 (3.4)</td>
<td>15 451 (28)</td>
</tr>
</tbody>
</table>

Table 4-3: Distribution in the paddle and machine for the clinical cohort.

Table 4-4 summarizes the results from the distribution in $T - T_{ro}$ and $P$. As we can see, the correction equally increased or decreased $T_{ro}$ within ± 3 mm. The mean and standard deviation of $T$ was 55.7 mm and 13.2 mm, respectively.

<table>
<thead>
<tr>
<th>$T-T_{ro}$ [mm]</th>
<th>Mean (std)</th>
<th>Median</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6 (3.1)</td>
<td>-0.1</td>
<td>-1.8, 3.1</td>
</tr>
<tr>
<td>$P$ [N]</td>
<td>84 (24)</td>
<td>80</td>
<td>70, 100</td>
</tr>
</tbody>
</table>

Table 4-4: thickness correction and compression force data.

In order to investigate whether the thickness correction induced a bias in the density, breast volume or thickness measurements, we separated those measures by mammography units and by paddle type. See Figure 4.5, Figure 4.6 and Table 4-5. The VBD and breast volume results are presented in more details in Section 4.3. The density, volume and thickness measurements were
reasonably consistent across units for the 19×23 cm semi-rigid paddle and the two 20×34 cm paddles. We note a higher density measurement for unit 2 and the 19×23 cm semi-rigid paddle. This might be a result of the thickness correction, which is different to units 1 and 3, and would give a lower thickness (thus higher density) at the same force. However, the measured volume and thickness were similar to unit 3. Similarly, for that paddle the volume & thickness was higher for unit 1, but the density was similar to unit 3. Likewise, the density was higher for unit 1 on the 24×30 cm hinged paddle, but the volume and thickness were all similar. Thus, despite of the differences in the observed responses of the compression paddles vs. force for different machines, the consistency in the breast volume, breast thickness and density measurements indicates that the thickness correction is reasonably accurate. We note that the density and volume measurement were less consistent for the hinged 19×23 cm paddle. It is possible that in this case the thickness correction might not be as accurate. Fortunately, only 0.6 % of the images were obtained with that paddle.

Figure 4.5: VBD as a function of paddle type and mammography unit. Paddles 1, 2, 3 and 4 represent the 19×23 cm semi-rigid paddle, the hinged 19×23 cm paddle, the semi-rigid 24×30 cm paddle and the hinged 24×30 cm paddle, respectively.
4.2.2 Peripheral Thickness Correction

In this section we present a method to identify the peripheral region of the breast (the region where the breast has lost contact with the compression paddle), and to model the breast shape and thickness in that region. We tested the accuracy of the thickness prediction and volumetric breast density (VBD) calculation on the 26 simulated digital mammograms presented in 0 for which the true thickness and VBD were available. The method was also tested on 55 087 digital mammograms, and an estimate of the thickness error from the algorithm was calculated.

4.2.2.1 Methods

The algorithm is similar to those presented by other researchers [113,114]. The inner edge of the peripheral region is deduced from the variation of signal intensity along radial lines. The negative logarithm of the pixel value of the digital mammogram \( L(x, y) = -\ln I(x, y) \) is taken.
The variation in $L$ represents changes in attenuation due to change in thickness and/or composition. The outer edge of the breast, $E$, is determined by automatic thresholding (the signal in the outside region of the breast is typically saturated). Next, radial lines are generated, centred on the point $M$ that lies midway between the breast edges along the chest wall edge of the image. See Figure 4.7. For each radial line at the angle $\phi \in [0, \pi]$, the intensity range $\Delta_{\phi} = \max L(r, \phi) - \min L(r, \phi)$ is determined from the maximum and minimum signal intensities on the image along the line. See Figure 4.7 for an example of a line $L(r)$. Typically, on the compressed region the variations in intensity are relatively small, presumably due to small changes in tissue composition rather than changes in thickness. Further from $M$, we note a sudden change in intensity, which arises from the reduction in thickness that occurs in the periphery. To identify the onset of that sharp change, for every $\phi$ various linear fits are performed on $L(r, \phi)$. The fits are done on subsets of the line from $r = 0$ (at $M$) to $r = R_c$. The distance $R_c$ was determined empirically, and varied as a function of breast size. It was chosen such that $R_c = \alpha \cdot R_\phi$, where $R_\phi$ is the distance from $M$ to $E$ at the angle $\phi$, with $\alpha = 0.6, 0.7$ and 0.8 for breast with a respective projected area of <100, 100-225 and >225 cm$^2$. Typically, smaller breasts have a larger peripheral region proportionally to the breast size. Thus, for each $\phi$, linear fits are performed on the 10 regions $[0, 0.25] \cdot R_c$, $[0, 0.5] \cdot R_c$, ..., $[0, 1] \cdot R_c$, $[0.25, 0.5] \cdot R_c$, ..., $[0.25, 1] \cdot R_c$, ..., $[0.75, 1] \cdot R_c$, and the one with the smallest absolute slope is recorded, yielding the linear function $a(\phi)r + b(\phi)$, where $a$ and $b$ are the slope and intercept at the angle $\phi$, respectively. The multiple fits are done to obtain the best estimate of the baseline constant-thickness intensity in the image, by avoiding localized pockets of adipose or fibroglandular tissue. Then, the maximum distance $r_0$ for which $[ar + b - L(r)] < \beta \Delta_{\phi}$ is determined. The factor $\beta$ was chosen empirically to be 0.07: if the breast was homogeneous in composition, $r_0$ would correspond to the point were the thickness has decreased by 7% from the constant thickness region. By repeating the procedure for all angles $\phi$, the approximate periphery contour $C'$ is found, and is then smoothed using a low-pass filter. See Figure 4.7.
Figure 4.7: Illustration of the peripheral detection algorithm. On the left we see the breast image, the outer breast edge, the points $r_0$ (dots), the smoothed approximate periphery $C'$ (solid line), the eroded periphery $C$ (dotted line), the central point $M$ and the radial lines in white. On the right we see an example of a radial line, along with the baseline linear fit, the detected periphery point $r_0$, the fitting limit $R_c$ and the intensity range $\Delta$.

The location of the inner periphery contour is then determined using the thickness profile to be applied in the periphery. The empirical thickness profile was obtained from Rico et al. [114] and was extracted from the image of a plastic phantom with a machined semi-circular peripheral region. See Figure 4.8. The thickness profile $T(r)$ was scaled such that $T(0) = 1$ and $T(1) = 0$, where $r$ is the normalized distance from the inner periphery edge to the outer breast edge. Assuming that the thickness at $C'$ corresponds to $(1 - \beta)$ of the compressed breast thickness $T_0$, the inner periphery contour $C$ is obtained by applying a morphological operation of erosion to $C'$, by a distance $r' = \delta \cdot T_0/2$, with $\delta$ such that $T(\delta) = 1 - \beta$ (the semi-circular periphery extends a distance $T_0/2$ forward). For $\beta = 0.07$, $\delta = 0.49$. Finally, the thickness map $t$ is applied on the breast image: the thickness profile is interpolated on every radial line $L(r)$ in the image such that $t(C) = 1$ and $t(E) = 0$, and the thickness map is multiplied by the compression thickness $T_0$. 
4.2.2.2 Results

The peripheral detection algorithm was tested on the 26 simulated mammograms described in 0. For those images, we the true thickness $T(x,y)$ was available, and the peripheral detection algorithm was used on $T(x,y)$ (as opposed to the image) in order to determine the true peripheral region. Thus, the true volumes $V$ and $V_P$ in the entire breast and in the peripheral region were computed, respectively. Similarly, the dense volumes in the entire breast and in the periphery, $V_D$ and $V_{PD}$, were also computed, as well as the peripheral area $A_P$ and the VBD. The peripheral region is defined by the region where the thickness is lower than 93% of the compression thickness. The simulated mammograms were then analyzed with the peripheral detection algorithm outlined above, and the thickness map $T'(x,y)$ was computed. The total breast volume, the dense volume, the VBD, the peripheral volume, the peripheral area, as well as the dense peripheral volume were computed. See Table 4-6 for a comparison of the results. We also show the mean thickness within the respective peripheral regions, $T_P$. Figure 4.9 (left) shows the comparison between the true VBD and VBD’ as determined using the algorithm. Figure 4.9 (right) shows the comparison between the total breast volumes and the peripheral region volume. See for a description of the differences between the true and calculated values.
To evaluate the thickness errors, the rms of \((T - T')\) was computed over the peripheral region determined by the algorithm for every image. Figure 4.10 shows the histogram of the rms difference in thickness: the average was 3.3 mm. We also computed the number of points with “extrema errors”, that is the point with a density below 0 % or above 100 %. On average, the fraction of points showing extrema errors was 6 % when using the true thickness \(T\), and 9 % in when using \(T'\).

Table 4-6: Comparison between the true volumes and densities and those determined using the peripheral detection algorithm.

<table>
<thead>
<tr>
<th></th>
<th>(V) [cm(^3)]</th>
<th>(V_D) [cm(^3)]</th>
<th>VBD [%]</th>
<th>(V_P) [cm(^3)]</th>
<th>(V_{PD}) [cm(^3)]</th>
<th>VBD(_P) [%]</th>
<th>(A_P) [cm(^2)]</th>
<th>(T_P) [cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Truth (CT)</strong></td>
<td>566</td>
<td>108</td>
<td>23.1</td>
<td>158</td>
<td>39</td>
<td>27.9</td>
<td>43</td>
<td>3.8</td>
</tr>
<tr>
<td><strong>Periphery algorithm</strong></td>
<td>564</td>
<td>110</td>
<td>24.5</td>
<td>175</td>
<td>44.5</td>
<td>30.6</td>
<td>47</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Table 4-7: Differences between the true VBD and volumes and those calculated using the peripheral correction algorithm.

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Intercept</th>
<th>Pearson corr.</th>
<th>Rms diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD [%]</td>
<td>1.01</td>
<td>1.1</td>
<td>0.94</td>
<td>4.3</td>
</tr>
<tr>
<td>(V) [cm(^3)]</td>
<td>1.02</td>
<td>-12.4</td>
<td>&gt;0.99</td>
<td>13</td>
</tr>
<tr>
<td>(V_P) [cm(^3)]</td>
<td>0.95</td>
<td>24.5</td>
<td>0.98</td>
<td>27</td>
</tr>
</tbody>
</table>
Figure 4.9: Comparison between the true VBD (left) and volumes (right) with those obtained using the peripheral detection algorithm (left). The dashed line is the identity function.

Figure 4.10: Histogram of the rms difference in thickness within the peripheral region between the true and the calculated thickness.

The peripheral detection algorithm was also used on the 55 087 images from the clinical cohort presented in the Section 4.3. The method outlined in Section 4.2.1 is used to estimate the compression thickness. Table 4-8 provides a summary of the density measurement in the entire breast and in the periphery only. The peripheral region is again defined as the region where the thickness is lower than 93 % of the compression thickness.

<table>
<thead>
<tr>
<th>$V$ [cm$^3$]</th>
<th>$V_D$ [cm$^3$]</th>
<th>VBD [%]</th>
<th>$V_P$ [cm$^3$]</th>
<th>$V_{PD}$ [cm$^3$]</th>
<th>VBD$_P$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>687</td>
<td>182</td>
<td>30.5</td>
<td>147</td>
<td>53</td>
<td>38.5</td>
</tr>
</tbody>
</table>

Table 4-8: density measurement in the breast and in the periphery.
In this study, the extrema errors were used as a surrogate measure of the inaccuracy of the thickness prediction. For the density calculations, the extrema errors were clipped to the nearest upper or lower bound. That is, for the points for which we calculated a density above 100 % of below 0 %, we assumed the true density was 100 % and 0 % respectively. Figure 4.11 shows an example of the density map and regions where the extrema error occurred. Figure 4.12 shows the distribution of the thickness at which the errors occurred for the 55 087 cases. Negative and positive values represent thickness overestimation (density below 0 %) and underestimation (density above 100 %), respectively.

![Figure 4.11: Example of a density map calculated on a mammogram (left). The window and level was adjusted so that density values above 70 % are saturated. The right image shows where the points above 100 % and below 0 % are located, in white and light gray, respectively. The periphery contour at 93 % compression thickness is shown by the line in a darker gray.](image)

On average, extrema errors occurred on 7.5 % of the image area, or 2.9 % of the breast volume. Most of those errors (87 %) were due to a density above 100 %. The over-100 % error occurred on all the images, while under-0 % errors occurred on 33 676 (61 %) of the images, at the respective thicknesses of 11.7 (std = 2.5) mm and 46.0 (std = 13.8) mm, or at 22 % (std = 7 %) and 79 % (std = 14 %) of the compression thickness. Based on the thickness profile in the periphery, the over-100 % density errors and under-0 % thus occurred within a range of 0.3-1.1 mm and 5-15 mm of the breast’s outer edge, respectively.
4.2.2.3 Discussion

In comparing the detected peripheral volume and thickness with the truth from the 26 simulated images, the thickness was well predicted, with the rms \((T - T') = 3.3\) mm in the peripheral region. Since the thickness errors were partially compensated by correcting the density to the nearest acceptable value, the mean VBD values were within 1.4 % with rms(VBD – VBD') = 4.1 %. The peripheral detection algorithm on average found slightly larger peripheral region (in area and in volume) compared to the truth.

This validation study had limitations. The breast deformation was simulated and thus it is possible that real breasts would be compressed in a different manner, yielding a differently-shaped outer bulge in the periphery for which the thickness model might not apply. Moreover, the simulation of the mammograms and the general VBD algorithm are also susceptible to error. For example, even when the true breast thickness profiles were used from the simulated mammograms, an average of 6 % of the area had extrema errors (density outside the 0% to 100% range).

The method was also used on 55 087 clinical digital mammograms. We note that the peripheral region was on average denser compared to the density in the entire breast. This is due to the presence of the skin: in the thinner peripheral region, the skin occupies a proportionally greater volume. Moreover, near the outside edge, the breast is 100 % dense due to the skin. On average, the fraction of points showing extrema errors was 7.5 %, which is similar to what is observed.
with the simulated mammograms. Most of the errors were due to thickness underestimation, causing density values to be above 100%. The majority of these occurred near the outside edge of the breast at the average thickness of 11.7 mm, and thus only accounted for 2.9% of the breast volume. Errors are to be expected in the outer edge region. As we can see in Figure 4.8, the thickness profile that is applied drops sharply to zero at the outer edge. Thus, small errors in the position and shape of the thickness profile, as well as interpolation errors, can lead to large errors in thickness near the edge. As discussed in Section 4.3.3, the net effect of these errors was small. Reversing the correction, that is attributing a density of 0% to the points above 100%, and attributing a density of 100% to the points below 0%, only caused an average absolute change in the VBD of 2 percentage points. We note that regions were extrema errors occurred are likely surrounded by a region where there also are density errors, but remain in the physical range between 0% and 100%. We did not account for these regions since they cannot be directly detected.

Another limitation of the peripheral detection algorithm is in the use of radial lines to detect the periphery edge and to apply the thickness profile. This method works well for breast images that have an approximately semi-circular shape, since the radial lines then are approximately orthogonal to the breast edge. For breasts with a more elongated or contracted shape, the radial lines, for small angles with respect to the chest wall, were not orthogonal to the breast edge. For those images and at those locations, the largest errors in the VBD or thickness estimation occurred. It would be possible to improve the algorithm by using conformal lines that intersect the breast edge perpendicularly. Finally, the algorithm didn’t perform well in the nipple region and where skin folds occurred. Generally, those regions accounted for only a small area of the mammogram.

4.3 Volumetric Breast Density Characteristics of a Large Clinical Image Set

4.3.1 Methods

The method used to measure the VBD is described in Chapter 2, and is complemented by the thickness estimation method described in Section 4.2. As discussed, the effective attenuation of plastic breast phantoms is characterized as a function of phantom thickness, phantom composition, imaging kV and target/filter composition. The same relation is mathematically
simulated using a complete model of the x-ray propagation in the mammography system. Seeing that the simulation accurately reproduces the effective attenuation of the plastic phantoms, a second simulation is performed, this time using the attenuation coefficients of breast tissue [28], which differ significantly from the plastic attenuation [91], and this simulated effective attenuation (see Figure 2.19) is used as to analyze the breast images. For each pixel in the breast image, of a known thickness and imaging parameters (kV and target/filter), the signal intensity is compared to the corresponding calibrated effective attenuation of fat and fibroglandular tissues, and the fractional composition of the tissue column is determined. The total volume of the column is then computed using equation (2-8), taking into account magnification and path obliquity, and the dense volume of the column is also computed. The VBD is taken as the ratio of the total dense volume to the total breast volume, both obtained by summing the individual volume values for every pixel on the breast image. The VBD with the skin excluded, $VBD_{nsk}$, was also calculated, assuming a constant 1.5 mm skin thickness.

It occurred that some pixels on a given image fall outside the range predicted by the calibration data, yielding a percent density outside of the physical range between 0 and 100%. In such cases, the density was set to the nearest “extremum” value (e.g. 105% is set to 100%). As we have seen in Section 4.2.1, the algorithm includes a model to estimate the breast thickness from the thickness and compression readout of the mammography machine, and the correction differs for different paddle types and the difference mammography units. Despite the thickness correction, there sometime remained a number of points with “extrema errors” within the uniformly compressed region of the breast. When that number exceeded 5% of the total number of points in that region, we further modified the breast thickness in increments of 1 mm, until the number of erroneous points fell under 5%. When the required absolute shift was above 10 mm, we estimated that it was indicative of an algorithm failure, and the images were rejected. The effect of those extrema errors is investigated in the Discussion section. As discussed in Section 4.2.2, the density algorithm contains an algorithm to predict the breast thickness in the peripheral region.

For this study, we used the digital mammograms from women who were clients of the SHSC clinic in Toronto, Ontario. Approval was obtained from the Sunnybrook Research Ethics Board to make density measurement on de-identified images. We used the raw (for processing) images, acquired from three General Electric Senograph Essential mammography units. The necessary
parameters for the VBD analysis (x-ray beam exposure and quality information, compression thickness, compression force and paddle type) were obtained from the DICOM (Digital Imaging and Communication in Medicine) headers. The examination date, patient age, series description, as well as the mammographic view and laterality were extracted from the DICOM header. Any information related to the patient’s identity was anonymized: the patient name and medical record number were encrypted systematically, and the birth date was randomized within ± 90 days. Unfortunately, no other epidemiological parameters (such as ethnicity, body-mass index, menopausal status, etc.) other than age were available for this cohort of patients.

The images were acquired from consecutive examinations of all women who presented for mammography between September 2008 to July 2011 and the images were analyzed retrospectively. After some exclusions, described below, we analyzed a total of 55,087 digital images from 15,351 women, of the crano-caudal (CC) mammographic view only. The medio-lateral oblique (MLO) images were excluded, as they often include a significant amount of pectoralis muscle, which would be mistakenly accounted as fibroglandular tissue by the algorithm. Moreover, the presence of the stiffer muscle in the breast might influence the response of the thickness readout with compression force: the response was tested with a uniformly stiff breast phantom (see Section 4.2.1). 13,481 individuals had at least one pair of left and right CC images taken on the same day, and 6,838 individuals has pairs of images taken on at least two separate dates. A minority of the images contained a tag in the DICOM header information (Series Description), which is entered manually, identifying 2,279 women as clients of the Ontario Breast Screening Program, with 1,380 of those women with exclusively such a tag in their images. However, it is likely that the screening clients were not always identified as such; we identified a total of 11,911 women who underwent at least one standard 4-view mammogram on a given day, typical of a screening examination, with 7,132 of those that exclusively had the 4-view examination. For those women with had multiple 4-view mammograms, the time between the examinations peaked at 1 and 2 years intervals. It is thus likely that those women were screening patients, the remainder being diagnostic or recall patients. Abnormal images such as ones obtained for needle localization, ductograms, magnification images, tissue specimen images, etc. were excluded whenever it was identified in the DICOM header (Series Description). There were 3,467 images that were taken without a compression paddle. Those were excluded, since the breast thickness cannot be determined. In total, there were 19,972
exclusions, leaving 57 547 MLO and 57 533 CC images. On 2446 of the CC images (4.2 %), the additional correction in the compression thickness due to unphysical extrema errors was above 10 mm, and those images were excluded from the analysis.

4.3.2 Results

The distribution in age, VBD, VBD\textsubscript{nsk}, breast volume and dense volume are shown in Figure 4.13, Figure 4.14 and Figure 4.15. Table 4-9 summarizes the data shown in the figures, as well as the skin volume and normalized skin volume information. Figure 4.16 shows the relation between VBD\textsubscript{nsk} and VBD. A linear fit between the two yielded VBD = VBD\textsubscript{nsk}×1.07 + 8.2 %, with a Pearson correlation > 0.99. As seen in Table 4-9, the average difference between the two is 9.6 %.

![Figure 4.13: Age distribution for the image set. With permission from IOP science [112.](image)]
Figure 4.14: VBD (left) and VBD with no skin (right) distributions for the image set. With permission from IOP science [112].

Figure 4.15: Total breast volume (left) and dense volume (right) distributions for the image set. With permission from IOP science [112].

Figure 4.17 shows the relation between the VBD and breast volume, respectively, for left and right breasts of the women in the data set. 13 481 individuals had one suitable left-right image pair, with the addition of 9310 pairs of individuals that had pairs of images taken on multiple dates. We identified 74 outliers for which the absolute VBD difference between the left and right breast was greater than 30 %. Similarly, we identified 124 outliers for which the absolute difference between the right and left breast volumes were greater than 700 cm$^3$. See the Discussion section for a description of the outliers. Table 4-10 shows a summary of the comparison between the left and right VBD and volume.
<table>
<thead>
<tr>
<th></th>
<th>Mean (std)</th>
<th>Median</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age [years]</strong></td>
<td>57.2 (12)</td>
<td>57</td>
<td>48-65</td>
</tr>
<tr>
<td><strong>VBD [%]</strong></td>
<td>30.5 (15.8)</td>
<td>26.3</td>
<td>19.4-37.4</td>
</tr>
<tr>
<td><strong>VBD no skin [%]</strong></td>
<td>20.9 (14.6)</td>
<td>16.9</td>
<td>10.7-27.2</td>
</tr>
<tr>
<td><strong>Breast volume [cm³]</strong></td>
<td>687 (402)</td>
<td>614</td>
<td>396-888</td>
</tr>
<tr>
<td><strong>Dense volume [cm³]</strong></td>
<td>182 (108)</td>
<td>157</td>
<td>108-229</td>
</tr>
<tr>
<td><strong>Skin volume [cm³]</strong></td>
<td>61 (29)</td>
<td>54</td>
<td>40-75</td>
</tr>
<tr>
<td><strong>Norm. skin volume [%]</strong></td>
<td>9.6 (2.1)</td>
<td>9.1</td>
<td>8.2-10.4</td>
</tr>
</tbody>
</table>

Table 4-9: Summary of the age, VBD, VBD with skin excluded, breast volume, dense volume, skin volume and normalized skin volume (skin volume/breast volume × 100) data from the image set.

Figure 4.16: Scatter density plot of the relation between VBDnsk and VBD (including the skin). The dashed line represents the identity function, and the solid line the linear least-square fit. The gray scale (light to dark) represents a decrease in the density of data points. With permission from IOP science [112].
Table 4-10: Values of the mean, median, rms difference and linear fit coefficients describing the difference between the left and right VBD and breast volume.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>RMS diff.</th>
<th>Slope</th>
<th>Offset</th>
<th>Pearson corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD left [%]</td>
<td>30.3</td>
<td>26.2</td>
<td>6.4</td>
<td>0.93</td>
<td>2.3</td>
<td>0.92</td>
</tr>
<tr>
<td>VBD right [%]</td>
<td>30.5</td>
<td>26.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume left [cm³]</td>
<td>700</td>
<td>627</td>
<td>164</td>
<td>0.91</td>
<td>58</td>
<td>0.915</td>
</tr>
<tr>
<td>Volume right [cm³]</td>
<td>694</td>
<td>625</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.18 shows a boxplot illustrating the relationship between breast volume and VBD, and Table 4-11 shows the mean percent VBD and number of images in the successive groups. For all the boxplots presented in this study, the whisker’s end represent 1.5 time the inter-quartile range, the outliers (dots) are the values that fall outside of the values bounded by the whiskers, and the notches represent the 95 % confidence interval in the median.
Figure 4.18: Boxplot of the VBD vs. breast volume. The volume groups include the upper bound but not the lower. The mean VBD for each volume group is shown by the dots joined with a line. With permission from IOP science [112].

<table>
<thead>
<tr>
<th>Volume [cm³]</th>
<th>0-250</th>
<th>250-500</th>
<th>500-750</th>
<th>750-1000</th>
<th>1000-1250</th>
<th>1250-1500</th>
<th>1500-1750</th>
<th>1750-3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD [%]</td>
<td>48.9</td>
<td>36.6</td>
<td>27.4</td>
<td>24.4</td>
<td>23.3</td>
<td>22.2</td>
<td>20.8</td>
<td>18.9</td>
</tr>
<tr>
<td>Image count</td>
<td>5382</td>
<td>14976</td>
<td>14705</td>
<td>10063</td>
<td>5174</td>
<td>2494</td>
<td>1168</td>
<td>1125</td>
</tr>
</tbody>
</table>

Table 4-11: Mean VBD and number of images in the volume groups.

Figure 4.19 shows a boxplot of the variation in VBD and breast volume for different age groups. Table 4-12 shows the mean, median of the VBD, dense volume and breast volume, as well as the number of images for each age group. The body-mass index (BMI) for Ontario women [115]² is also shown (see the Discussion section).

² The data in ref. presented the BMI in groups of under 18.5, 18.5-25, 25-30, 30-35, 35-40 and over 40. For the purposes of this comparison, those groups were assigned the average BMI value of 16, 22, 27.5, 32.5, 37.5 and 42, respectively.
Figure 4.19: Age dependence of VBD (left) and breast volume (right). The age groups include the upper bound but not the lower. The mean VBD or breast volume of each age group is shown by the dots joined with a line. See Table 4-12. With permission from IOP science [112].

Figure 4.20 show boxplots of the difference in VBD and breast volume for the 6838 women who had images taken on multiple dates, as a function of time between the mammography examinations. The average VBD and breast volume of the left and right breasts of a given date were used, and the difference of the value at the latest minus the value at the earliest date was taken. The difference in breast volume is shown relative to the volume value at the latest date. Table 4-12 gives the number of cases, the mean VBD and breast volume difference in each time group. We note that the total number of cases exceeds 6838; 2444 women had more than two separate mammography exams. The average rate of change was computed as -2 %VBD per year and a 2 % relative increase in breast volume per year. The rate of change in VBD was similar for all age groups, but the relative rate of change in breast volume per year was higher for women in their 40’s (4 %), and lower for women in their 60’s (1%).
<table>
<thead>
<tr>
<th>Age group [years]</th>
<th>20-35</th>
<th>35-40</th>
<th>40-45</th>
<th>45-50</th>
<th>50-55</th>
<th>55-60</th>
<th>60-65</th>
<th>65-70</th>
<th>70-75</th>
<th>75-80</th>
<th>80-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image count</td>
<td>1378</td>
<td>2591</td>
<td>5562</td>
<td>7660</td>
<td>8691</td>
<td>7999</td>
<td>7650</td>
<td>5522</td>
<td>3622</td>
<td>2441</td>
<td>1971</td>
</tr>
<tr>
<td>Mean VBD [%]</td>
<td>45.5</td>
<td>42.9</td>
<td>39.3</td>
<td>35.1</td>
<td>30.2</td>
<td>26.4</td>
<td>25.1</td>
<td>24.9</td>
<td>25.5</td>
<td>26.9</td>
<td>29.8</td>
</tr>
<tr>
<td>Med. VBD [%]</td>
<td>42.7</td>
<td>39.3</td>
<td>35.9</td>
<td>30.9</td>
<td>26.5</td>
<td>23.3</td>
<td>22.7</td>
<td>22.4</td>
<td>22.9</td>
<td>24.1</td>
<td>26.6</td>
</tr>
<tr>
<td>Mean dense vol. [cm³]</td>
<td>231</td>
<td>213</td>
<td>212</td>
<td>198</td>
<td>183</td>
<td>171</td>
<td>168</td>
<td>166</td>
<td>165</td>
<td>162</td>
<td>156</td>
</tr>
<tr>
<td>Med. dense vol. [cm³]</td>
<td>204</td>
<td>186</td>
<td>186</td>
<td>174</td>
<td>160</td>
<td>145</td>
<td>146</td>
<td>142</td>
<td>137</td>
<td>136</td>
<td>132</td>
</tr>
<tr>
<td>Mean vol. [cm³]</td>
<td>599</td>
<td>591</td>
<td>630</td>
<td>665</td>
<td>703</td>
<td>734</td>
<td>738</td>
<td>724</td>
<td>701</td>
<td>657</td>
<td>569</td>
</tr>
<tr>
<td>Med. dense vol. [cm³]</td>
<td>518</td>
<td>484</td>
<td>544</td>
<td>575</td>
<td>624</td>
<td>665</td>
<td>680</td>
<td>668</td>
<td>653</td>
<td>611</td>
<td>523</td>
</tr>
<tr>
<td>BMI</td>
<td>24.4</td>
<td>24.8</td>
<td>27.8</td>
<td>27.6</td>
<td>26.3</td>
<td>26.3</td>
<td>26.6</td>
<td>26.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-12: Mean and median VBD, dense volume and breast volume, and number of images for each age groups. The BMI was obtained from ref. 115.
Figure 4.20: Boxplot of the VBD difference (left) and relative volume difference (right) versus time between examinations. The mean difference for each time interval group is shown by the dots joined with a line. See Table 4-12. With permission from IOP science [112].

<table>
<thead>
<tr>
<th>Days between examinations</th>
<th>1-300</th>
<th>301-400</th>
<th>401-500</th>
<th>501-600</th>
<th>601-700</th>
<th>701-800</th>
<th>801-1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cases</td>
<td>283</td>
<td>5646</td>
<td>2016</td>
<td>602</td>
<td>426</td>
<td>2120</td>
<td>718</td>
</tr>
<tr>
<td>Mean % VBD difference</td>
<td>-1.9</td>
<td>-1.2</td>
<td>-1.7</td>
<td>-1.8</td>
<td>-3.4</td>
<td>-3.3</td>
<td>-5.1</td>
</tr>
<tr>
<td>Mean relative vol. difference</td>
<td>3.3 %</td>
<td>2.6 %</td>
<td>2.7 %</td>
<td>3.4 %</td>
<td>5.0 %</td>
<td>4.6 %</td>
<td>6.9 %</td>
</tr>
</tbody>
</table>

Table 4-13: Number of cases, mean VBD difference and mean relative breast volume difference for each time bin.

On average, points, the points were extrema errors occurred accounted for 7.5 % of the total breast area, or for 2.9 % of the total breast volume. As we have seen in Section 4.2.2, the errors for the most part occurred near the periphery of the breast where the thickness changes fairly rapidly. The imaging techniques used in this study are summarized in Table 4-14.
### 4.3.3 Discussion

In this study, we present the results of an automatic and reproducible algorithm that measures the volumetric density, i.e. the volume of dense tissue divided by the total breast volume. The algorithm was used on 55,087 digital mammograms (CC view) corresponding to 15,351 individuals. The measured values of VBD, VBD with skin excluded and breast volume are in reasonable agreement with studies done with 3D breast imaging methods. See Table 4-15. In Chapter 5 we present density and volume measurements determined by MR imaging and compare them to the measurement done on mammograms for the same women.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Modality</th>
<th>Age (range)</th>
<th>Number of cases</th>
<th>VBD (std) [%]</th>
<th>VBD_nsk (std) [%]</th>
<th>Volume [cm³]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current work</td>
<td>Mammo.</td>
<td>57 (20-98)</td>
<td>55,087</td>
<td>30 (16)</td>
<td>21 (15)</td>
<td>687</td>
</tr>
<tr>
<td>Yaffe et al. [78]</td>
<td>CT</td>
<td>54 (35-82)</td>
<td>191</td>
<td>26 (13)</td>
<td>14 (10)</td>
<td>769</td>
</tr>
<tr>
<td>Glide et al. [83]</td>
<td>3D US</td>
<td>48 (21-85)</td>
<td>93</td>
<td>31 (12)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Nie et al. [82]</td>
<td>MRI</td>
<td>50 ± 11</td>
<td>230</td>
<td>n/a</td>
<td>16 (10)</td>
<td>676</td>
</tr>
</tbody>
</table>

Table 4-15: Comparison of the percent breast density and breast volume between this work and those measured with 3D imaging techniques. The values are the mean ± standard deviation.

As a sign of good internal consistency, the algorithm provided similar similar values and excellent correlation of both VBD and total breast volume between the right and left breasts of a given woman. See Figure 4.17. We investigated 74 outliers for which the absolute difference between the left and right VBD was greater than 30 %. For 44 of those pairs, the appearance of the two images was markedly different, due to density patterns, abnormalities (the data set
included some post-operation and post-radiation patients), poor exposure or poor compression. For the remaining 30 outliers the left and right images had a similar appearance, and thus the VBD mismatch is presumed to be due either to poor performance of the algorithm or to superposition of fibroglandular tissues. Similarly, we found 124 outliers in the relation between the left and right breast volumes. As an independent way to estimate the volume of the breast, we calculated an approximate volume as the product of the projected breast area with the reported readout breast thickness. We found an excellent Pearson correlation $> 0.99$ between the approximate and calculated volumes, and the outliers also followed that trend. Moreover, the relation between the two was relatively narrow; the standard deviation of the difference was 125 cm$^3$. As conformed by an inspection of the images, the outliers were thus the result of a difference in the positioning (breast area) or in the actual size (thickness) between the left and right breasts. As expected, the VBD vs. volume (Figure 4.18) shows a steady decline of VBD with increasing breast size, presumably because of an association between breast size and body fat.

We also found the VBD, dense volume and breast volume had an expected variation with age, decreasing with increasing age (Figure 4.19). Similar results were found by Huang et al. [116], who used breast CT to measure the VBD (without skin) and dense volume. We note that the mean VBD increases slightly beyond age 80. It is possible that the women in that older age group continue to present for mammography because of a higher breast cancer risk, which may imply a higher breast density. Moreover, the number of samples in those age groups was smaller than for the 60-80 age groups (see Table 4-12). The change in VBD as a function of time over a period of 1 to 1000 days for a given women (Figure 4.20) were also expected, with a mean rate of decline of 2 percentage points per year. The variation of breast volume with age showed a small increase from ages 35 to 55, followed by a plateau (55 to 70 years) and by a small decrease for later ages. It is reasonable to speculate that the variation in breast volume follows that of the weight of the women. For a rough comparison, we looked at BMI data for Ontario women [115], and found that it had a similar variation with age. See Table 4-12. Finally, the changes in breast volume as a function of time of a period of 1-1000 days were small and expected, again showing good reliability in the algorithm.

We note that there is a fair amount of variability of VBD and breast volume for each age group (Figure 4.19), indicative of the heterogeneity of those characteristics in the studied population. A
similar distribution was observed by Checka et al. [117], who used BIRADS classification on 7007 images. We thresholded our VBD values by quartiles to mimic the BIRADS classes and combined the two top and bottoms as dense and fatty classes. We found that 66, 45, 35 and 38 % of women, respectively in their 40’s, 50’s, 60’s and 70’s had dense breasts, a result similar to Checka et al. In Chapter 5, we investigate the relation between manual area density measurements similar to the BIRADS method and the automatic VBD measure.

The algorithm has been shown to be accurate within ± 3 %VBD, in phantom studies and in a validation study with CT images. See Chapters 2, 3 and ref. 79. In this study, the accuracy of the algorithm cannot be precisely determined, since no measure of the true VBD was available. It is expected that the error will be larger than ± 3 %, because of natural variation in the x-ray attenuation of breast tissue, because of variation in the breast thickness, and because of errors in the thickness estimation in the compressed region and in the peripheral region. We estimate that errors in thickness are the largest contributing factor for the algorithm’s inaccuracies: as seen in Section 2.5, the sensitivity of the VBD measurement with respect to thickness is approximately 3-4 %VBD per mm of thickness error, with an underestimation in breast thickness leading to an overestimation in VBD. We note that we observed higher VBD values when compared to the 3D methods (Table 4-15). Moreover, the population in our study was of older age, which should normally lead to lower density values. However, our studied population was much larger and thus likely more heterogeneous in ethnicity and cancer risk status. In Chapter 5, we present another validation of the method, by comparing the VBD measurement with the density measurement done on MR images of a small subset of women in this cohort. The results show a reasonable agreement between the two methods.

In this work, the breast thickness was estimated by characterizing the response of the thickness readout as a function of paddle type and compression force using a flexible breast phantom. It is possible that breast have different deformation characteristics under compression, and those characteristics may depend on breast size, density, etc. As seen in the previous section, the rms of the thickness correction was 3.1 mm. In order to verify our estimate of breast thickness, we calculated an “inferred” breast thickness for each image. This was done by finding the lowest intensity point on the image within the compressed region and assuming the corresponding column of tissue was comprised of only fatty tissue with two bounding layers of skin (totaling 3 mm). From the attenuation calibration, we can determine the thickness of fatty tissue (plus skin)
for that column. The mean and median of the inferred thickness minus the thickness estimate using the thickness calibration was 2.5 mm, with an inter-quartile range of 4.0 to 0.8. The standard deviation and the rms of the difference were 4.0 and 4.7 mm, respectively. We believe that this method tends to overestimate the thickness, since it is likely that at least some fibroglandular tissue will reside in that column. From the small difference, we can expect a reasonable accuracy from the thickness algorithm, yet this indicates that the thickness is possibly underestimated, which may explain the higher density values when compared to the 3D methods.

We note that the thickness correction algorithm was in some cases inaccurate; for 2088 images (4 % of the total), it required a further correction to bring the VBD within realistic bounds (i.e. between 0 and 100 %), with an average absolute shift of 3.2 mm. The mean positive and negative shifts were 2.8 and -3.7 mm, respectively, with 1135 and 953 cases. For an additional 2446 images, the required absolute shift was above 10 mm, and those images were rejected from the analysis. The breasts in that rejected subset were particularly small and dense: the mean breast volume and density were 495 cm$^3$ and 79.4 %, respectively (the thickness shift was positive in 2147 cases). It is likely that the thickness correction is less accurate for these types of breasts. The high density and small size might also interfere with the peripheral region detection, of which there were 285 failures in this subset.

To estimate the worst-case effect that extrema error (local density estimates below 0 % or above 100 %) could have, we set the densities for those columns to be 100 % and 0 %, respectively. That is, a point with density above 100 % has likely a true density near 100 %, but in theory could have any density value, the most unlikely case being 0 %. Even with this gross adjustment the overall effect in VBD for the breast is small, producing a mean absolute VBD difference of 2.0 %.

We also expect our estimates of $\text{VBD}_{\text{nsk}}$ to be less accurate. We assumed a uniform layer of 1.5 mm of skin [89] surrounding the breast. Variations of the skin thickness between patients and internal variation of the skin thickness were neglected. In 235 images a negative $\text{VBD}_{\text{nsk}}$ was computed. For those images the $\text{VBD}_{\text{nsk}}$ values were set to zero. Nie et al. [118] investigated the effect of skin thickness in VBD measurements using MRI on 50 women. Their measured values for total skin volume were similar to ours. Their total breast volume was higher, so that the mean relative volume of skin was slightly lower than ours (9 % vs. 10 %). Moreover, they observed a larger contribution of the skin to the VBD with increasing VBD: in their work the slope and
offset of the linear fit between VBD_{nsk} and VBD was 1.23 and 7 %, respectively. We note that there were only a few cases with high breast density in that study.

The methods used in evaluating breast density from mammograms vary between researchers, and the VBD values found in the literature differ from those found in this work, for similar populations. See Table 4-16.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of cases</th>
<th>Age (std)</th>
<th>VBD (std) [%]</th>
<th>VBD_{nsk} (std) [%]</th>
<th>Volume [cm^3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current work</td>
<td>55 087</td>
<td>57 (12)</td>
<td>30 (16)</td>
<td>21 (15)</td>
<td>687</td>
</tr>
<tr>
<td>Shepherd et al. [68]</td>
<td>1100</td>
<td>57 (11)</td>
<td>45 (17)</td>
<td>n/a</td>
<td>451</td>
</tr>
<tr>
<td>Ding et al. [69]</td>
<td>2508</td>
<td>50-75 range</td>
<td>24 (7)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Boyd et al. [70]</td>
<td>1020</td>
<td>59 (11)</td>
<td>10 (15)</td>
<td>n/a</td>
<td>740</td>
</tr>
<tr>
<td>Lokate et al. [71]</td>
<td>370</td>
<td>54 (4)</td>
<td>n/a</td>
<td>7</td>
<td>886</td>
</tr>
<tr>
<td>Schroeder et al. [121]</td>
<td>79 978</td>
<td>57 (12)</td>
<td>n/a</td>
<td>10 (6)</td>
<td>755</td>
</tr>
</tbody>
</table>

Table 4-16: Mean density and volume measurement done on mammograms found in the literature. Some values are estimated from the published data.

Shepherd et al. use a method based on Malkov et al. [119], and applies a correction to shift the attenuation of reference of fatty tissue to that of pure fat. Adipose tissue contains approximately 15 % water, which may explain the observed higher density values. They also use a reference attenuation phantom on each image, which also allows for the estimation of the compression thickness. Moreover, the studied population had approximately 27 % of Asians, which may explain the lower breast volume. Ding et al. use a method based on Highnam et al. [120] and is based on the simulation of x-ray propagation in mammography. Schroeder et al. use the Volpara software based on [122,123]. The work of Schroeder et al. was done on a subset of the cohort used in this work, but included the MLO views; detailed and updated results were obtained from the authors. This method presents the VBD without the contribution of skin, and uses an internal reference for the attenuation of fat. We have found that using an internal attenuation reference lead to an estimate of thickness 2.5 mm higher than the compression thickness. Based on the
sensitivity values found in Section 2.5 at the average thickness of 5.5 cm, this would decrease the VBDisk by approximately 7.5 percentage points to 13.5%, which is similar to what was found by Schroeder et al. Lokate et al. use a method based on van Engeland et al [110], and also uses an internal reference for the attenuation of fat. Finally, the method of Boyd et al is similar to ours, but was applied on screen-film mammograms without a machine-specific thickness correction model.

All of those studies except the work of Schroeder et al. conducted a case-control risk analysis. Shepherd et al. found that the percent volumetric density measure was a stronger risk factor for developing breast cancer compared to traditional percent area density (odds ratio of 4.1 versus 2.5). The work of Ding et al. and Boyd et al. found that the volumetric measure strongly associated with breast cancer risk, similarly to the area method. However the percent volume measurement lost significance when adjusted with the area measurement, indicating that the volume measurement did not improve on the area method. Finally, Lokate et al. found that the volume measurement was associated to known breast cancer risk factors, but less strongly that the area measurement. In Chapter 5 we present a preliminary comparison between the risk associations of our volumetric measure with the traditional percent area method. In addition, we present a preliminary validation of the VBD measurement done a subset of the clinical images by comparing it with the VBD measurement done on MR images of the same patients.
Chapter 5
Future Work and Summary

5.1 Introduction
In the previous chapters we have outlined the motivation in making accurate volumetric breast density measurements from mammograms. In principle, VBD should be a better predictor of breast cancer risk compared to the traditional breast density estimation methods, which are subjective and based on the projected area of dense tissue. In Section 5.3 we address directly that question in a preliminary study (365 cancer cases, 2188 controls) in which we compare the odds of breast cancer between VBD and area density categories. In Chapters 2 and 3 we have described and validated the method to measure VBD and determined its baseline accuracy. In Chapter 4 we tested the VBD algorithm on a large set of clinical images. In that study, there was no ground truth available to test the accuracy of the method. In Section 5.2, we perform a preliminary validation of that study using MRI.

5.2 Comparison of the Density Algorithm with Magnetic Resonance Imaging
In this section we present some preliminary results comparing the density and volume measurements obtained from mammograms with those obtained from MR images. We also show the comparison from both volumetric methods to the area density method. The MR data set consisted of 105 examinations conducted between 2004 and 2009 on women who were part of multiple studies conducted at the Sunnybrook Health Sciences Centre (SHSC) [124-127]. The purpose of these studies was to investigate the effectiveness of contrast-enhanced MRI in detecting breast cancer on high-risk women. Although a greater number of women were imaged, we excluded the most recent cases, as the MR imaging sequence was different from the remainder of the images, requiring a modification in the segmentation algorithm. We also excluded the older images because of the large time difference between the MR and mammography examinations. Of the 105 patients, 80 had had digital mammograms (no more that 4.5 years after the MR exam) that were part the set of images discussed and analyzed in
Chapter 4. The MR images were de-identified, and approval was granted from the Research Ethics board to perform density measurement on the images. The area breast density (ABD) was measured by Dr. Jennifer Harvey of the University of Virginia, an experienced radiologist and one of our research collaborators, using the Cumulus software program [37].

The MR images used were the pre-Gadolinium contrast agent injection series. They were obtained on a General Electric 1.5 T scanner using a T1-weighted scan protocol. Forty-three images were fat-suppressed (fatty tissue appears dark), and the remaining 37 images were not fat-suppressed (fatty tissue appears bright). The scan was in the sagittal plane, and the breasts were mildly compressed in that direction. The pixel size varied between 0.70 mm and 0.82 mm (mean 0.76 mm) and the slice thickness varied between 2 mm and 4 mm (mean 2.8 mm). The image matrix size of all the images was 256×256×28 for both the left and right breasts.

5.2.1 Segmentation of Dense Tissue on MR Images

The algorithm to measure the breast density on MR images was developed by Dr. James Mainprize. It operates in four stages: correction of the field inhomogeneities, delineation of the breast tissue, partial volume correction and thresholding of the dense tissue.

Because of inhomogeneities in the breast coil sensitivity, the signal intensity in the slice (coronal plane) was not uniform. A crude estimate of the field map was created by subdividing each slice into 8×8 subregions. In each subregion and within the breast, the maximum signal intensity and its position were recorded. A thin-plate smoothing spline algorithm was used (Matlab function `tpaps`) to fit those points into a smooth and interpolated surface for each slice. This generated a smoothed map of the low spatial frequency signal variation in the image, i.e an estimate of the inhomogeneity map. The correction was achieved by dividing the image by the map for the points inside the breast. Slices were also compared visually, where a user could manually adjust the signal level for any slices that appeared to be overcorrected. This allowed for the density structures to be properly visualized across the entire field of view. The slices requiring user correction were generally limited to those at the edges of the volume containing little tissue.

Next, the breast-air interface was delineated by manually selecting an intensity threshold on every image slice of the breast. If needed, a polygonal region was drawn on the thresholded region in order to add breast tissue that was missed by the threshold, or to remove non-breast
tissue that was included by the threshold. Similarly, a polygonal region was selected to extract the breast tissue from the tissue posterior to the chest wall. See Figure 5.3

Because of the relatively thick slices (2-4 mm), partial volume effects can distort the density and volume estimates at the breast edge. To correct this, an estimate of the tissue-occupying fraction (the partial volume) for edge voxels was determined. This was accomplished by creating an isosurface of the segmented breast tissue and by smoothing the surface. A new binary voxel model of the breast (where 1 and 0 represent the inside and outside of the breast, respectively) was extracted from the isosurface at 10 times the resolution along the slice (coronal) direction through interpolation. The high-resolution image was then averaged at the original voxel resolution in order to compute the partial volume fraction. See Figure 5.1.

![Figure 5.1: the left image shows original breast in black and the oversampled and smoothed breast edge in as the line. The right image shows the calculated partial volume fraction in greyscale.](image)

In the final step, the user selected a region of interest (ROI) in the central slice of the breast that contained both fatty and fibroglandular tissue, and selected the signal intensity peaks corresponding to the two types of tissue on a histogram of the ROI. The peaks are denoted by $I_f$ and $I_{fg}$ for fatty and fibroglandular tissue, respectively. Then, a soft-threshold sigmoid function was applied to transform the signal intensity $I$ into fractional fibroglandular content $m$, such that

$$m = -\tanh(\alpha \cdot (I - \bar{I}))/2 + 0.5,$$

where $\bar{I} = (I_f + I_{fg})/2$ and $\alpha = 2 \cdot \tanh^{-1}(-0.90/(I_{fg} - I_f))$. The factor $\alpha$ is chosen such that the fractional density $m \approx 0.1$ and 0.9 for $I = I_f$ and $I_{fg}$, respectively. See Figure 5.2
The dense volume was computed by summing the fractional fibroglandular content map over the breast and multiplying by the voxel volume map. The average volumetric density was obtained by dividing the dense volume by the total breast volume. See Figure 5.3 for an example of a segmented slice.

The MR segmentation algorithm has not been validated. However, the results of the segmentation compared favourably to the mammography measurements shown in the next section.
5.2.2 Results

The MR images of 80 patients were analyzed using the algorithm described above. The breast volume, the volumetric breast density and the dense volume of the left and right breast were computed. The matching mammograms (CC view only) that were obtained at a date closest to that of the MR exams were selected, and the measured values of breast volume, dense volume and volumetric breast density (VBD) and ABD were extracted for the comparison.

Figure 5.4 shows the age distributions of the patients at the time of the mammography and MR exams, as well as the time difference between the exams. The average time difference was 3 years, the MR exams having systematically occurred before the mammography exam. See Table 5-1. Figure 5.5 and Figure 5.6 shows the comparison between the VBD, breast volume and dense volume measured by MR and with mammography. Figure 5.7 shows the comparison between the VBD from the mammograms and MR with the ABD. Table 5-2 shows the statistical parameters in the distributions of the VBD, breast volume and dense volume measured by MR and mammography. Table 5-3 shows the results of the comparison between MR and mammography. Table 5-4 shows the results of the comparison between the VBD and the ABD values.

![Figure 5.4: Left: age distribution of the patients at the time of the mammography and MR exams. Right: distribution of the time difference between the MR and mammography exams.](image-url)
Table 5-1: Summary of the patients’ age and time difference between the MR and mammography exams.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Median</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at MR</td>
<td>42.1</td>
<td>8.6</td>
<td>42</td>
<td>34.5-47.5</td>
</tr>
<tr>
<td>Age at mammo</td>
<td>45.1</td>
<td>8.5</td>
<td>44</td>
<td>38.5-51</td>
</tr>
<tr>
<td>Time difference [months]</td>
<td>36.4</td>
<td>10.8</td>
<td>36.3</td>
<td>25.3-47.5</td>
</tr>
</tbody>
</table>

Figure 5.5: Comparison between the VBD as measured by MR and by mammography. The dashed line represents the identity function.

Figure 5.6: Comparison between the breast volume (left) and dense volume (right) as measured with MR and mammography. The dashed line represents the identity function.
Figure 5.7: Comparison of the VBD from the mammograms (left) and from the MR (right) with the ABD measurement.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Median</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD MR [%]</td>
<td>32.3</td>
<td>12.8</td>
<td>29.1</td>
<td>23.2-40.2</td>
</tr>
<tr>
<td>VBD mammo [%]</td>
<td>32.3</td>
<td>16.8</td>
<td>27.2</td>
<td>20.3-38.5</td>
</tr>
<tr>
<td>ABD mammo [%]</td>
<td>31.2</td>
<td>22.0</td>
<td>30.6</td>
<td>11.5-49.1</td>
</tr>
<tr>
<td>Vol. MR [cm³]</td>
<td>645</td>
<td>338</td>
<td>584</td>
<td>386-835</td>
</tr>
<tr>
<td>Vol. mammo [cm³]</td>
<td>677</td>
<td>411</td>
<td>581</td>
<td>385-881</td>
</tr>
<tr>
<td>Dense vol. MR [cm³]</td>
<td>185</td>
<td>85</td>
<td>170</td>
<td>121-230</td>
</tr>
<tr>
<td>Dense volume mammo [cm³]</td>
<td>183</td>
<td>94</td>
<td>158</td>
<td>118-216</td>
</tr>
</tbody>
</table>

Table 5-2: Comparison of the measured density and volume between MR and mammography.
Table 5-3: RMS difference, correlation and the linear least square fit coefficients between the MR and mammography measurements.

<table>
<thead>
<tr>
<th></th>
<th>RMS diff.</th>
<th>Pearson correlation</th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD</td>
<td>10.1 %</td>
<td>0.801</td>
<td>1.05</td>
<td>-1.6 %</td>
</tr>
<tr>
<td>Volume</td>
<td>157 cm³</td>
<td>0.934</td>
<td>1.14</td>
<td>-55 cm³</td>
</tr>
<tr>
<td>Dense volume</td>
<td>66 cm³</td>
<td>0.733</td>
<td>0.81</td>
<td>33 cm³</td>
</tr>
</tbody>
</table>

Table 5-4: RMS difference, correlation and linear least-square fit coefficients between the MR and mammography VBD values and the ABD values.

<table>
<thead>
<tr>
<th></th>
<th>RMS diff.</th>
<th>Pearson correlation</th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBD mammo</td>
<td>12.2 %</td>
<td>0.837</td>
<td>0.64</td>
<td>12.4 %</td>
</tr>
<tr>
<td>VBD MR</td>
<td>15.3 %</td>
<td>0.738</td>
<td>0.43</td>
<td>18.9 %</td>
</tr>
</tbody>
</table>

5.2.3 Discussion

As we can see in Table 5-2, the volumes and densities measured by MR and mammography agree reasonably well. The mean VBD and dense volume values were almost identical, but the mammography breast volume was 5% higher than the MR volume. This goes counter to the expectation that density increases with age and that breast volume increases with age; the MR images were obtained 1-4 years before the mammogram. Moreover, the values measures using the mammographic methods had more variance: the standard deviation and inter-quartile range were higher for the mammography measurements.

As we can see in Figure 5.5 and Figure 5.6, there is some variability between the density and volume measurements: the RMS difference is 10 VBD percentage points or 160 cm³. The Pearson correlation is high at 0.8 for the VBD, and excellent for the volume measurement. This variability arises from multiple factors: the experimental error in the measurements, a difference in the imaging field of view, and a change in the breast density and volume due to the time
difference between the exams. We have seen in Chapters 2 and 3 that the baseline accuracy of the density algorithm is 3 percentage points. Considering the added source of error arising from the estimation of the compression thickness, the method should be accurate within 4-6 percentage points. Similarly, there is some error in the segmentation of the MR images. We observed a variation of 2-3 % VBD by testing the repeatability of the manual segmentation. Moreover, some of the MR images were noticeably non-uniform due to coil-sensitivity effects. In some cases, parts of the image of the breast had to be cropped out due to spurious high or low intensity regions. See Figure 5.8. We generally favoured the accuracy of the VBD measure vs. the accuracy of the volume measure, whenever possible. Considering those factors, we estimate that the MR segmentation is accurate to within approximately 2-5 % percent VBD. The error in the difference between the MR and mammography measurements combines in quadrature, and thus should be approximately 4.5-8 percent VBD, which is close to the observed RMS difference. To put those differences in perspective, we can investigate the difference between the left and right measurements. See Table 5-5.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS diff.</td>
<td>5.2 %</td>
<td>89 cm³</td>
<td>5.0 %</td>
<td>110 cm³</td>
</tr>
<tr>
<td>Peason corr.</td>
<td>0.919</td>
<td>0.966</td>
<td>0.957</td>
<td>0.964</td>
</tr>
</tbody>
</table>

Table 5-5. RMS difference and Pearson correlation of breast density and breast volume, between the left and right breast for the MR and mammography results.
We can notice some outliers in the relations shown in Figure 5.5 and Figure 5.6. We found 8 points (3 pairs of left-right images of a given patient, and 2 images of two other patients) for which the absolute difference between the MR and mammography VBD is larger than 20 % (twice the RMS difference). On two of the pairs, the mammogram looked very dense, with a mammographic VBD ranging between 65 % and 90 %, and the MR density ranged between 35 % and 60 %. Approximately 25 months had passed between the exams. See Figure 5.9 for an illustration of one of these outliers. The cause of the discrepancy is uncertain: the results were reasonable for both modalities. For the other pair of images, the mammogram had a density of 17 %, while the MR density was 37-41 %. Again, it is not certain why the discrepancy occurred, but it is possible that there were some changes in the breast: the time elapsed between the exams was 50 months. For the remaining 2 outliers, the mammograms were poorly exposed, which likely increased the error in the VBD measurement. Removing the outliers lead to a correlation of 0.839 between the measurements.
Similarly, we found 12 points (4 pairs of images of a given patient, and 4 individual images) for which the absolute difference between the MR and mammography volume is larger than 320 cm$^3$ (twice the RMS difference). There was no overlap between the VBD and volume outliers. The MR volume measurement should be robust, since it mostly depends on the accuracy in the manual selection of the breast-air and breast-chest interfaces, both of which are easy to perform. Nevertheless, as explained above, coil sensitivity effects prevented the selection of the entire breast volume in some cases. Such effects were observed in 3 of the outliers, but with only mild consequences on the volume estimate. Thus, no clear reason behind the volume discrepancies was found; the results from both modalities seemed reasonable. The breasts in that set had either a high compression thickness when the mammography volume was higher (average of 73.5 mm for 8 images) or a low compression thickness (average of 33.8 mm for 4 images) when the mammography volume was lower. Thus one can assume that the difference is due to an actual change in breast size in the time between the exams or from a difference in the scanned volume. Removing the outliers led to a correlation of 0.950 between the measurements.

As we have seen in Section 4.3.2 (page 101), the VBD decreases by approximately 2 percentage point per year. We applied a corresponding shift to increase the mammographic VBD as a function of the elapsed time between the two exams. The mean VBD increased to 38.4 %, the Pearson correlation between the shifted density and the MR density increased slightly to 0.822, and the RMS difference increased to 11.3 %. It is likely that the VBD change rate is different for
this population than for the large population studied in Chapter 4. However, we would still expect a lower density from the mammography results because of greater age of the woman when the mammogram was performed; in Figure 5.5 we can notice a cluster of points below the identity line. We isolated this cluster by excluding the points with mammographic or MR VBD over 50% (15% of the data), the mean mammographic and MR VBD were 27% and 29%, respectively. Overall, 58% percent of the points had a higher MR VBD compared to the mammographic VBD.

Van Engeland et al. [110] also have proposed a validation of a VBD algorithm using MR images. See Section 3.4. They compared 88 images from 22 patients, with an average of two months between the mammogram and MR images. The average relative error between the dense volumes, per patient, was 13.6%, while in our case it was 25.8%. Moreover, they found a correlation of 0.94 and 0.97 between the MR and mammography dense volumes, when determined per image and per patient, respectively. This is a significantly higher than the corresponding correlation coefficients of 0.73 and 0.75 in our study. The variability between the measurements should be less because of the much smaller time difference between the exams. In addition, the authors acknowledge that they did not accurately segment the breast tissue in the MR images: they used a single plane to separate the breast from the chest wall. In other words, they did not compare the total breast volume determined by MR with that determined by mammography. Thus, it is possible that the good correspondence in the dense volumes hides a mismatch in the total breast volume. Moreover, they excluded any MR case where glandular tissue was present behind the single separating plane, and thus may have specifically selected the cases where the MR and mammographic field of view were well-matched. Finally, they also found a good correspondence between the VBD measurements. However, they used the total breast volume obtained from mammography in calculating the MR VBD, making the comparison essentially identical to the dense volume comparison.

The results of the comparison between the MR VBD and the ABD are similar to results found in the literature. Kazhem et al. [128] found a correlation of 0.78, with a slope and intercept of 0.56 and 2%, respectively, in a study of 264 breasts from 133 women. The researchers used a similar MR segmenting method. Graham et al. [129] found a correlation of 0.79 in a study of 42 women. In their study, a novel and quantitative imaging sequence was used to determine the MR density. In both cases, the same Cumulus algorithm was used for the ABD computation. Lee et al. [130]
found a correlation of 0.63 between the two measures, done on 40 women. Finally, Wei et al. [131] have found a significantly higher correlation of 0.91 between MR and ABD measurements in a study of 67 women. It is not clear why such a strong correlation was found. We note that we observed a better correlation between the VBD obtained from the mammograms and the ABD. This is expected because the measurements are obtained from the same image: there is not variability arising from a difference in the field of view between the MR image and the mammogram or from temporal changes. See Section 5.3.2 for additional results of the comparison between the mammographic VBD and ABD.

We believe the results of the comparison between the MR and mammography measurements of VBD and volume are encouraging. The results validate the accuracy of the mammography measurement with a 3D technique. Certainly, the excellent correlation between the measured volumes indicates a good accuracy of the thickness estimation in the VBD algorithm. As we have seen, both methods suffer from measurement errors, and thus a deviation of at least 5% VBD is to be expected. Moreover, we can expect further differences to arise because of differences in the positioning of the breast in the mammography exam versus the full field of view of the MR scan, or because of changes in the breast in the time interval between the examinations.

5.3 Comparison between the Risk Factors Associated with VBD and with the Manual Area Density Method

In this section, we present preliminary results of the breast cancer risk factor associated with the automatic VBD calculation method. We also measured the area breast density (ABD) using the Cumulus software [37], which allowed for the calculation of the breast cancer risk factor associated with the traditional area method. In addition, we compared the VBD values with the corresponding ABD values. The risk analysis presented here had a limited scope: a single confounding factor for breast cancer, age, was used. Moreover, the manual ABD measurements were done by an inexperienced reader.

5.3.1 Area Density Measurement

The Culumus algorithm operates in three steps. First, a user delineates the outside edge of the breast using a gray-level threshold slider and/or by drawing a polygonal shape, which defines the
breast area. Next, the user delineates the pectoralis tissue (if present) by drawing a polygonal shape; the pectoralis tissue is excluded from the density analysis. Finally, the user selects a gray-level threshold that best delineates the fibroglandular tissue in the image. The area encompassed by that threshold is the dense area, and the area breast density (ABD) is defined as the ratio of the dense area by the breast area. The user can also adjust the window and level of the image to facilitate the visualization of density structures. Generally, the mammograms are analysed in batches (referred to as reads), with repeat images to test for consistency in the density measurement.

The author (O. Alonzo-Proulx) was the reader for most the work presented in the following sections. The reader was trained on a set of 40 digitized screen-film images, half of the images being repeats. The reader was unaware of the proportion of repeats. There was good reliability between the repeats, and the results were similar to the ones from an experienced reader (Dr. N. Boyd). See Table 5-6.

<table>
<thead>
<tr>
<th>Comparison with:</th>
<th>Slope</th>
<th>Offset [%]</th>
<th>ICC</th>
<th>Pearson corr.</th>
<th>RMS diff. [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeats</td>
<td>0.968</td>
<td>-1.4</td>
<td>0.948</td>
<td>0.935</td>
<td>5.5</td>
</tr>
<tr>
<td>N. Boyd</td>
<td>0.825</td>
<td>3.2</td>
<td>0.890</td>
<td>0.885</td>
<td>7.2</td>
</tr>
<tr>
<td>N. Boyd*</td>
<td>0.812</td>
<td>2.9</td>
<td>0.924</td>
<td>0.927</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Table 5-6: Results of the comparison of the ABD measurement between the repeats and with an experienced reader. ICC stands for intraclass correlation. *For the last row, we removed 4 outliers in the data.

The reader then analysed 2688 images, in batches of 100-105. The images were digital mammograms consisting of cancer cases and screening controls, which were randomized within the reads. The case and control selection is described in Section 5.3.3.1. The set of images contained 320 repeats from 300 randomly selected images. Most of the repeats were within a given read to test for the intra-read variability, and 10 of the repeat images were repeated in two extra reads to test for the inter-read variability. From the various possible combinations of the repeats between reads, there were 50 images to test the inter-read variability. The reader had excellent repeatability for the area density measurements. See Figure 5.10 and Table 5-7.
Figure 5.10: Plot of the dense area (left) and area density (right) for the repeat images. The identity function and the linear fit are represented by the dashed and solid lines, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Offset</th>
<th>ICC</th>
<th>Pearson corr.</th>
<th>RMS diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total area</strong></td>
<td>0.999</td>
<td>0.06 cm²</td>
<td>1.000</td>
<td>1.000</td>
<td>2.0 cm²</td>
</tr>
<tr>
<td><strong>Dense area</strong></td>
<td>0.972</td>
<td>1.0 cm²</td>
<td>0.949</td>
<td>0.948</td>
<td>5.3 cm²</td>
</tr>
<tr>
<td><strong>ABD</strong></td>
<td>0.943</td>
<td>1.4 %</td>
<td>0.953</td>
<td>0.952</td>
<td>3.7 %</td>
</tr>
<tr>
<td><strong>ABD (N=300)</strong></td>
<td>0.940</td>
<td>1.2 %</td>
<td>0.959</td>
<td>0.957</td>
<td>3.5 %</td>
</tr>
<tr>
<td><strong>ABD (N=50)</strong></td>
<td>0.951</td>
<td>2.8 %</td>
<td>0.924</td>
<td>0.926</td>
<td>4.7 %</td>
</tr>
</tbody>
</table>

Table 5-7: Results of the comparison between the repeat images. The first three rows are for the entire repeat set of 350 images. The last two rows are for the intra-read (N=300) and inter-read (N=50) repeat images.

5.3.1.1 Film-Like Conversion

Normally, the Cumulus algorithm is used on digitized screen-film images or on processed (for presentation) digital images. The raw (for processing) digital mammograms cannot be used for the area analysis: the range of image intensities is very large, so that the density structures in the
image cannot be easily visualized without repeatedly adjusting the window and level. For the risk study presented in this section, the processed images were unavailable, and we used an algorithm to convert the raw images into a screen-film like appearance. The algorithm was developed by C. Peressotti in our laboratory and is briefly described here. See Figure 5.11. It uses a reference sensitometric curve (relationship between digitized optical density and log-exposure) that was previously obtained from a digitized film sensitometric strip [132]. A region of interest (ROI) is automatically selected near the centre of the breast in the raw image, and the average logarithm of the signal or the exposure value in the ROI is computed. Next, the sensitometric curve is shifted so that its speed point (the log-exposure value corresponding to the maximum slope in the curve) matches the ROI log-exposure. Finally, the digital log-exposure values of the mammogram are converted into a film-like optical density (OD) using the shifted reference curve.

Figure 5.11: Film-like conversion algorithm. The left image is the negative logarithm of the raw mammogram, which is converted into a film-like mammogram (right) using a reference sensitometric curve (bottom). The speed point is indicated by the dot. A low exposure leads to a high digitized value, i.e. a low OD.

In an internal study, we tested if the film-like conversion algorithm was adequate to measure area density. This was done by comparing the area measurements from 200 converted images to a)
processed images obtained from the same raw image and b) with digitized screen-film images from the same patient. The anonymized images were from participants of the DMIST study [6], and we had approval from the Sunnybrook Health Sciences Centre Research Ethics board to perform the density measurements. The analysis was performed by A. Gunasekara in our laboratory and results are reproduced here. Figure 5.12 shows the comparisons for the percent area density, as measured by an experienced reader (Dr. M. Yaffe), and Table 5-8 and Table 5-9 summarize the results. Three other readers performed the analysis with similar results. The conclusion of the study was that the film-like conversion algorithm was adequate to measure area density: it produced results similar to the ones using processed images. In addition, the differences between the measurements from film-like and processed images were similar to the difference arising from the inter and intra user variability. There was a greater difference between the measurements from film-like or processed images and the screen-film images. This was expected since the screen-film image was obtained on a different imaging system in a distinct examination. See Figure 5.13.

Figure 5.12: Film-like ABD vs. the processed ABD (left) and screen-film ABD (right). The dashed line is the identity function and the solid line a linear fit. See table Table 5-8.
<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Offset [%]</th>
<th>Pearson corr.</th>
<th>RMS diff. [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processed vs.</td>
<td>0.895</td>
<td>2.2</td>
<td>0.90</td>
<td>7.4</td>
</tr>
<tr>
<td>film-like</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screen film vs.</td>
<td>0.755</td>
<td>0.3</td>
<td>0.84</td>
<td>11.7</td>
</tr>
<tr>
<td>Film-like</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-8: Linear fit coefficients, correlation value and RMS difference between the ABD measurement from the different images.

<table>
<thead>
<tr>
<th></th>
<th>Total area [cm²]</th>
<th>Dense area [cm²]</th>
<th>ABD [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
</tr>
<tr>
<td>Screen-Film</td>
<td>158.4</td>
<td>69.0</td>
<td>35.6</td>
</tr>
<tr>
<td>Processed</td>
<td>162.5</td>
<td>64.1</td>
<td>27.3</td>
</tr>
<tr>
<td>Film-like</td>
<td>160.8</td>
<td>63.6</td>
<td>27.1</td>
</tr>
</tbody>
</table>

Table 5-9: Mean and standard deviation of the total area, dense area and area density for the three image types.
5.3.2 Comparison between the Area and Volume Density

Figure 5.14 shows the comparison between the area and volume density. While we could expect a power law relationship between the volume and area measures (area = volume$^{2/3}$), we found that a linear fit provided better results. The slope and intercept were 0.796 and 12.3 %, respectively. The Pearson correlation and RMS difference were 0.683 and 13.3 %, respectively. We identified 15 outliers for which the absolute difference between the measurements was larger than 50 percentage points, and the corresponding images were reviewed. For the 10 most extreme outliers (top left in Figure 5.14), the VBD value was erroneous. The corresponding volumetric density maps had a background density between 80-90 % with large regions at 100 % density. For those images the estimate of compression thickness was likely erroneous; the VBD algorithm required a shift in the compression thickness of 7-10 mm (mean 9.1 mm). The density algorithm has a cut-off shift of 10 mm; images that necessitate a shift in the compression
thickness above 10 mm are rejected (see Section 4.3.1). Those images narrowly passed the cut-off limit. Only 2 of those outliers were cancer cases. The remaining 5 outliers had no shift in the thickness, but the background density was relatively high between 40-50%. It’s possible there was a small error in the compression thickness for those cases. See Figure 5.15 for an illustration. Excluding the outliers, the Pearson correlation was 0.744.

Figure 5.14: Comparison between the volumetric and area breast density. The outliers are shown with the circles.

Figure 5.15: Example of an outlier in the ABD vs. VBD relation. From left to right: film-like converted digital image; thresholded image (ABD = 22.4%, the pectoralis muscle was excluded); volumetric density map (VBD = 72.8%). In the density map, the background density ranged between 50 and 60%.
5.3.3 Risk Calculation

5.3.3.1 Case and Control Selection

The selection of cases was done using the Sunnybrook Breast Research Biomatrix. It is a large repository containing de-identified data on patients with a breast cancer diagnosis confirmed by biopsy and linkages to their de-identified mammograms. The date of the diagnosis, patient’s age at diagnosis and the laterality of the cancer are known.

We also needed to identify controls to perform the analysis. It is common to select screening patients as controls, since they are presumed asymptomatic. As discussed in Section 4.3.1, we selected women with a specific tag in the image header that identified them as screening patients, and there were 2279 women with such a tag in their images. It is likely that this method of identification underestimates the true number of screening patients in our set of images, since the tag is entered manually. Thirteen of those women were in the cancer case group, and were thus not considered as controls. We further removed 78 controls for which we did not have a valid VBD measurement of both the left and right breast. 381 of the remaining 2188 women had more than one separate screening examination. For those women, we randomly selected one set of images from a single examination. In addition, we randomly selected the laterality of the density measurement for each control. Finally, 112 controls had repeat images of a given laterality, due to positioning or exposure errors. We manually selected the image with the largest amount of tissue exposed and/or with the best exposure. For three controls, the film-like conversion algorithm failed, and the area density was not computed. The set of control images was saved in order to be analysed with the area density algorithm. The age of controls ranged from 32 and 87 years (see Table 5-10 and Figure 5.16). We note that 54 controls were aged between 32 and 49 years. Organized mammography screening officially starts at 50 years in Ontario, and those younger patients had a tag on their image distinct from the other screening patients.

There were originally 486 breast cancer cases extracted from the Biomatrix. A total of 121 cases were excluded, leaving 365 cancer cases. Two cases had unspecified cancer laterality and 43 cases had a bilateral cancer. Since the breast density of the unaffected breast must be used in the risk analysis, those cases were excluded. For 4 cases there was no match for the patient in our database of images: it is possible that those women were referred from another medical centre. For 64 cases, no contralateral CC images were obtained during the patient work-up, and in 8
cases, the algorithm failed on the available contralateral CC image, so that no VBD value was available. The failures were all due to an excessive number of unphysical density values, which required a correction in the compression thickness in excess of 10 mm. See Sections 4.3.1 and 4.3.3. As with the controls, there were 25 cases with repeat images due to exposure or positioning errors, and we manually selected the properly acquired images. The contralateral images for the cancer cases were saved in order to perform the area density analysis. See Table 5-10 and Figure 5.16 for a distribution of the cancer cases’ ages. There is a significant difference in the age distribution of the cases and controls. There are 106 cancer cases aged below 50 years, versus 54 for the controls. There are 30 cancer cases above 79 years, versus 10 in the control group.

We note that there were 127 cases (35 %) with a non-zero time difference between the mammography imaging and the cancer diagnosis. See Figure 5.16. The average absolute time difference, excluding the nil values, was 5.9 months. We took the age at the time of imaging as the age of the cancer case. We note that in 93 cases, the imaging occurred with an average of 9.2 months after the diagnosis. For those cases it is possible that the patient underwent treatment that might affect the density in the contralateral breast.

<table>
<thead>
<tr>
<th></th>
<th>Mean (Std)</th>
<th>Median</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control group</strong></td>
<td>58.9 (6.8)</td>
<td>58</td>
<td>54-64</td>
</tr>
<tr>
<td>(N = 2188)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cancer case group</strong></td>
<td>59.2 (13.6)</td>
<td>59</td>
<td>48-70</td>
</tr>
<tr>
<td>(N = 365)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-10: Age (in years) distribution of the case and control groups
5.3.3.2 Matched Analysis

We performed an aged-matched case-control analysis to compare the odds of developing cancer between different VBD or ABD classes. We randomly matched every cancer case to a control of the same age plus or minus one year difference. Once a control was selected as a match, it was removed from the control pool. When there were multiple cancer cases of a given age, they were randomly permuted before finding the matches. The ABD and VBD from the entire control group was separated into quintiles (see Table 5-11), and the odds of cancer for a given density quintile was computed as the number of cancers over the number of controls within the density quintile. The odds ratio (OR) was computed between the top and bottom quintiles of VBD or ABD.

<table>
<thead>
<tr>
<th>Quintile range</th>
<th>VBD</th>
<th>ABD</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.4 – 34.4</td>
<td>8.5 – 28.1</td>
<td></td>
</tr>
</tbody>
</table>

Table 5-11: Lower and upper quintiles for the VBD and ABD for the control population (N = 2188).

Because of the random matching, the process was iterated 10 000 times. As expected, multiple cancer cases did not have a control match; an average of 63.2 ± 0.8 (range 61-65) cases out of 365 was excluded. See Figure 5.17 for a typical distribution of the unmatched cases’ age. Figure 5.18 shows the histograms in the OR for this analysis. The OR for VBD was 2.2 ± 0.3 and the OR for ABD was 1.2 ± 0.2, with respective 95% confidence intervals of [1.7, 2.8] and [0.9, 1.6].
5.3.3.3 Discussion

The results presented in the previous section are unexpected; the OR for ABD was not significant at the 5% level. These results cast doubts on the validity of the method of processing the digital images into a film-like appearance. In addition, it is possible that the inexperienced reader (OAP) interpreted the images inadequately, although he had good reliability. We compared his ABD results to those from two experienced readers (Dr. M. Yaffe and Dr. J. Harvey, MY and JH respectively) on a subset of 105 randomly selected images within the case-control dataset. See Figure 5.19. The inexperienced reader’s assessment of density using...
Cumulus had a good linear correlation between the experienced readers, but the intraclass correlation (ICC) was weaker; the experienced readers systematically read a higher ABD, and agreed strongly with each other. An ICC of 0.9 is usually observed for Cumulus [47]. See Table 5-12. Nevertheless, the strong Pearson correlation between the two readers potentially indicates that the risk results would be similar if the experienced reader had performed the ABD measurements.

Figure 5.19: Comparison between the inexperienced (\(ABD_1\)) and experienced (\(ABD_2, ABD_3\)) readers in measuring area density on a subset of 105 images from the case-control dataset. The dashed and solid lines represent the identity function and the linear fit, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Intercept [%]</th>
<th>Pearson Corr.</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAP vs. MY</td>
<td>1.574</td>
<td>-7.1</td>
<td>0.902</td>
<td>0.759</td>
</tr>
<tr>
<td>OAP vs. JH</td>
<td>1.666</td>
<td>-7.6</td>
<td>0.922</td>
<td>0.746</td>
</tr>
<tr>
<td>MY vs. JH</td>
<td>0.949</td>
<td>2.4</td>
<td>0.917</td>
<td>0.919</td>
</tr>
</tbody>
</table>

Table 5-12: Linear parameters and correlations between readers of ABD

Those results may serve to demonstrate the challenges in using the subjective ABD measure as a risk factor; the method is sensitive to reader experience and to image processing. In comparison, the VBD measure is objective, and was found to be a stronger risk factor for breast cancer. It should be emphasized that the risk analysis presented here had a limited scope: without including confounding, potentially independent risk factors such as BMI, parity, menopausal status, family history, etc., the values of the ORs for the two density measures cannot be interpreted separately.
Specifically, adjusting for BMI can increase the OR or relative risk (RR) \([48,49]\). Moreover, the study population came from a diagnostic centre, and thus was a mix of symptomatic and screening women. This has been shown to reduce the RR \([49]\). Similarly, we considered prevalent cancers; the density measurement was obtained from the contralateral breast at diagnosis. This may lead to an underestimation of the RR \([49]\), since since sensitivity in mammography is higher for fatty breasts. Conversely, the RR is higher when considering incident cancers, \(i.e.\) when using the density from a previous, negative mammogram. Therefore, the OR values should not be trusted in and of themselves. Rather, the purpose of this analysis was to compare the ORs for VBD and ABD, since the potentially confounding factors are then kept constant. With that perspective, and keeping in mind that the reader was inexperienced, we can see that the VBD measure is a stronger predictor of risk compared to the ABD measure.

**5.4 Summary and Final Words**

Mammographic breast density is an important metric in breast cancer research. As we have seen in Chapter 1, it would be logical that the association between breast cancer risk and breast density would be due to tissue volumes. Area-based breast density algorithms are certainly correlated to the volumetric amount of dense tissue in the breast, but they are subjective and sensitive to user performance. Therefore, there is some interest in measuring volumetric breast density accurately and automatically. While 3D imaging modalities such as CT and MR can provide such a measure readily, mammography will remain, at least in the near future, the most common method of imaging the breast.

This dissertation presents a theoretical and experimental framework for estimating accurately the volumetric breast density from digital mammograms. Chapter 2 provides the theoretical foundation, validated with experimental data, underlying the volumetric breast density algorithm. That is, we showed how the imaging signal produced by a digital detector relates to the composition and thickness of breast-like phantoms and breast tissue, and how that relation can be used to deduce the breast density from the image signal. The algorithm incorporates the effects of beam polychromacity, beam obliquity, x-ray scatter, anti-scatter grid effects as well as detector glare and efficiency. Chapter 3 presents a validation of the algorithm using simulated breast images obtained from deformed CT data. This allowed us to test the density measurement in a controlled setting: the 3D configuration of the compressed breast was known from the CT
images. It was found that the algorithm performed equally well with the breast images than with the phantom images, with a baseline accuracy of 3 density percentage points. Chapter 4 presents the results of the volumetric density measurement from a large set of 55,087 clinical mammograms. The algorithm yielded results comparable to previously published values using 3D imaging methods. Chapter 4 also presents additions in the algorithm necessary to analyze clinical images: the modeling of the thickness in the peripheral region of the breast, and the characterization of the readout thickness of the mammography machine as a function of compression force. As expected, the volumetric density measurement is sensitive to errors in the breast thickness: for a 25% dense breast, 5 cm thick, this sensitivity is of 3.3 density percentage points per mm of thickness error. Thus, for clinical images where thickness errors on the order of 1 mm are expected, the accuracy of the algorithm should range between 4 and 6 density percentage points. We note that the method was tested rigorously a specific mammography machine, the GE Senographe Essential. Preliminary tests have been done on GE Senographe 2000D and on Hologic Selenia machines. While the algorithm performed well using breast-like phantoms, the thickness readout systems on those machines tend to be inaccurate, and we expect the accuracy of the density measurement to be reduced. Chapter 5 presents preliminary results validating the mammographic density measurements against MRI, as well as a comparison of the volumetric measurement with the traditional area-based density measure.

The volumetric breast density algorithm described in this work is accurate and validated. By design, it is reliable and reproducible, and because it is fully automatic, it can be used to analyze a large number of images in a reasonable time. We have shown that volumetric breast density is a potentially better risk factor compared to the standard are-threshold method. The method can in principle be extended to any mammography machine, provided that the compression thickness of the breast can be accurately determined. Thus, ultimately, we hope to incorporate this automatic measurement of density in a personalized risk model.
References


35. American College of Radiology 1993 ACR Breast imaging reporting and data system Reston, VA

36. American College of Radiology 2003 ACR Breast imaging reporting and data system Reston, VA


90. Hammerstein GR, Miller DW, White DR, Masterson ME, Woodard HQ and Laughlin JS Absorbed radiation dose in mammography. Radiology 1979; 130: 485-91


115. Statistics Canada Table 105-2001 *Measured adult body-mass index, by age group and sex, household population aged 18 and over excluding pregnant females.* Canadian Community Health survey cycle 2.1, Canada and provinces, occasional. 2004 CANSIM database.


134. Statistics Canada, Table 051-0001. Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual (persons unless otherwise noted), from CANSIM database 2012.