# Calibrating densities in mammography

September 2002

Thomas E. Marchant

THE UNIVERSITY
of MANCHESTER

Department of Imaging Science and Biomedical Engineering

A thesis submitted to the University of Manchester for the degree of M.Sc. in the Faculty of Medicine, Dentistry, Nursing and Pharmacy

# Table of contents

# Tables and figures

## Abstract

Breast tissue has both fatty and glandular components. The glandular tissue is denser than fatty tissue, and causes increased x-ray attenuation. The amount of dense tissue apparent on mammographic images has been shown to be an indicator of breast cancer risk. An automatic method of measuring the thickness of dense tissue at each pixel in a digitized mammogram is described here. The method involves radio-opaque magnification markers placed on the breast compression paddle, which allow accurate determination of breast thickness, and a stepwedge calibration device, which is imaged alongside the breast and allows comparison of the amount of attenuation at each point in the image to a set of calibration data. The method is developed to allow use with a wider range of breast sizes, and its performance is assessed using a large set of clinical mammograms. The error on the measurements of dense tissue thickness made using this method is estimated to be 2.4 mm. The measurements at each pixel can be combined to produce measurements of area or volume of dense tissue over the whole breast. Percentage by volume of dense tissue and percentage by area with a significant thickness of dense tissue are found to be the most strongly correlated with an existing density classification scheme, whose relation to cancer risk has been established.

## Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

## Copyright and intellectual property notes

1. Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

2. The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the Department of Imaging Science and Biomedical Engineering.

## Acknowledgements

I would like to thank Sue Astley and Alan Hufton, who have supervised this project, staff at the Nightingale Breast Clinic, where the practical work was carried out, and Christie Hospital NHS trust who provided financial support to the Author.

## The Author

Tom Marchant was educated at Judd School, Tonbridge, Kent, before joining the department of Physics and Astronomy at the University of Manchester as an undergraduate in 1994. He went on to complete a PhD in particle physics at Manchester, before entering employment at Christie Hospital, Manchester, as a medical physics trainee, in September 2001. The M.Sc. project described here was undertaken as part of the NHS Grade A training scheme for clinical scientists.

# 1 Introduction

The most frequently diagnosed form of cancer amongst women in the western world is breast cancer (Lacey 2002). Early detection of breast cancer can significantly improve the chances of successful treatment. X-ray mammography, where x-ray images are taken of the compressed breast, is currently the most effective method of detecting the early signs of breast cancer. Hence many countries, including the UK, have a national breast screening programme. An important part of the screening process is the identification of individuals at risk. This can reduce the total number of examinations required, by focusing on the highest risk groups. Also, groups considered to be at particularly high risk can be screened at more frequent intervals. Commonly used risk factors include age (risk increases with age) and family history (breast cancer has a strong genetic association). This project is mostly concerned with one particular indicator of breast cancer risk – density of breast tissue.

The main aims of this project are to validate a previously developed method for measuring mammographic density, and to extend the method for use with a wider range of breast sizes and for breasts with a higher proportion of dense tissue. This was important because it allowed the method to be used in a study of the effects of lifestyle on the amount of dense breast tissue and breast cancer risk. The study group of mammograms was larger than had previously been analysed using this method, and included many large and highly dense breasts.

The method produces a measurement of the thickness of dense tissue at each pixel in the digitized mammogram. An important part of the method is a breast thickness measurement made using radio-opaque magnification markers placed on the breast compression paddle. Accurate measurements of the marker separations in the image are used to find the degree of magnification, and hence the distance between the compression paddle and the image receptor. The method also uses a calibrated stepwedge device

imaged alongside the breast, allowing comparison between images taken with different imaging parameters. Each pixel in the image is compared to the stepwedge to find the thickness of stepwedge material causing the same degree of attenuation. This thickness is then referred to a set of previously acquired calibration data, to find the combination of dense and fatty tissue thicknesses which cause this amount of attenuation.

## *1.1 Outline*

Chapter 2 explains the physical principles of mammography. The third chapter concentrates on mammographic density, and how it is related to breast cancer risk. Chapter 4 explains the various approaches that have been used to measure mammographic density. Chapter 5 describes in detail the method that is the main focus of this project. Chapter 6 describes the clinical application of the new method. Chapters 5 and 6 contain the majority of the new experimental work carried out by the author. Some discussion of the results is provided in chapter 7.

# 2 Physical principles of mammography

Mammography involves taking an x-ray image of the breast. This chapter describes the physics of the x-ray imaging process, and its application to breast imaging. Mammography makes severe requirements of the imaging system. To identify breast disease, good contrast is required between soft tissues with relatively small differences in absorption coefficient. In addition very good resolution is desirable to detect tiny microcalcifications, commonly associated with breast cancer. It is also necessary to keep radiation doses from the process to an absolute minimum, so that it is suitable for mass screening programs. These often conflicting requirements have lead to mammography equipment being highly specialized. More details can be found in the textbooks, (Bushberg 1994) and (Dendy 1999).

## *2.1 Production of x-rays*

X-rays are produced by accelerating electrons across a high potential difference into a metal target. The electrons are produced at the cathode, which is held at a large negative potential. X-rays are produced at the target by the processes of bremsstrahlung and characteristic radiation.

Bremsstrahlung x-ray photons are produced when electrons are decelerated in the nuclear electric field of the target atoms (see Figure 2.1a). Bremsstrahlung produces a broad energy spectrum with a maximum at the energy of the accelerating potential difference (see Figure 2.1b). This corresponds to electrons which collide head-on with the target nucleus, and lose all their energy to the x-ray photon. Electrons that approach the target nucleus less closely give rise to less energetic x-rays. These more distant interactions are the most probable, but the lower energy x-rays produced are less penetrating, and are reabsorbed in the target itself. This gives rise to a maximum in the energy spectrum at around a third to half of the maximum energy.

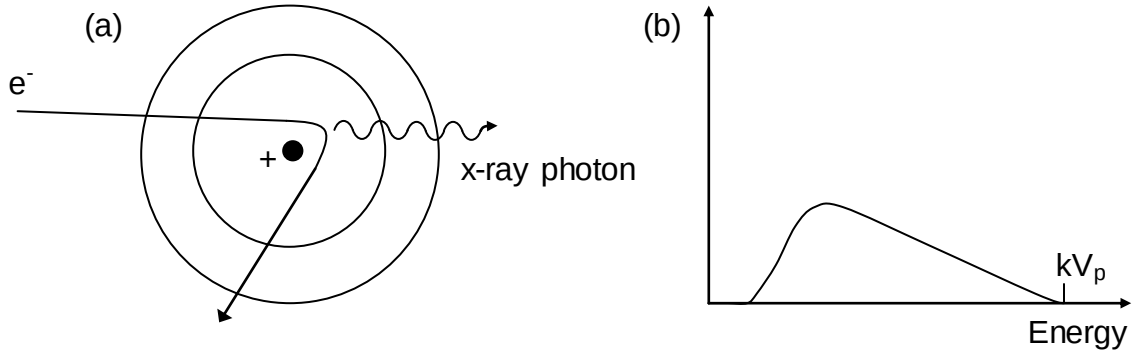


*Figure 2.1: (a) Production of bremsstrahlung x-rays (b) Energy spectrum of bremsstrahlung x-rays.*

Characteristic radiation occurs when the electrons interact with an inner shell electron of a target atom, knocking it out of the atom. This leaves a vacancy in the inner shell, which is filled by the transition of an electron from a higher energy level. An x-ray photon is emitted with energy corresponding to the difference in energy between the two levels (see Figure 2.2a). Hence the energy spectrum has discrete spikes (see Figure 2.2b). Each target material has its own characteristic pattern of energy spikes (giving rise to the name characteristic radiation).

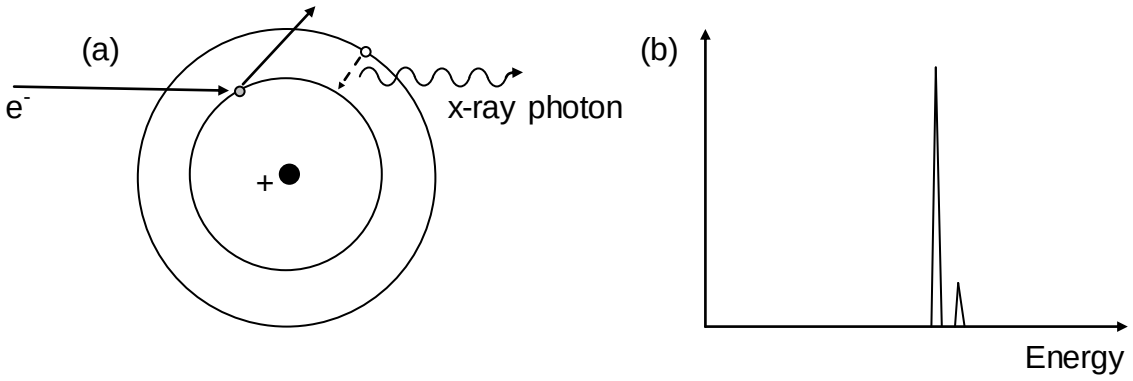


*Figure 2.2: (a) Production of characteristic x-rays (b) Energy spectrum of characteristic x-rays.*

Mammography requires good soft tissue contrast and low dose. The high energy components of the x-ray spectrum give less contrast, while low energies increase dose without contributing to image quality. Hence the best

spectrum for mammography is essentially monoenergetic. The suitable energy range giving maximum contrast with low dose is 17-25 keV for typical breast thicknesses (compressed thickness 3-6 cm). A monoenergetic spectrum in this energy range is approximated by using the characteristic x-rays from target materials such as Molybdenum or Rhodium. The accelerating voltage used is in the range 25-35 kV. This is high enough to give good efficiency for producing the characteristic x-rays, but not so high as to produce a large amount of bremsstrahlung at higher energy.

## *2.2 X-ray target*

Figure 2.3 shows the cathode and target (anode) of a typical mammography x-ray set. The vast majority of the electrons' energy is converted into heat in the anode. A small focal spot is desirable to reduce blurring in the image, but the size is limited by the heat dissipation capacity of the anode material. The anode is formed from a rotating disc, which allows a smaller focal spot to be used. The anode is also tilted at an angle θ. This produces a smaller effective focal spot size from the perspective of the image receptor.

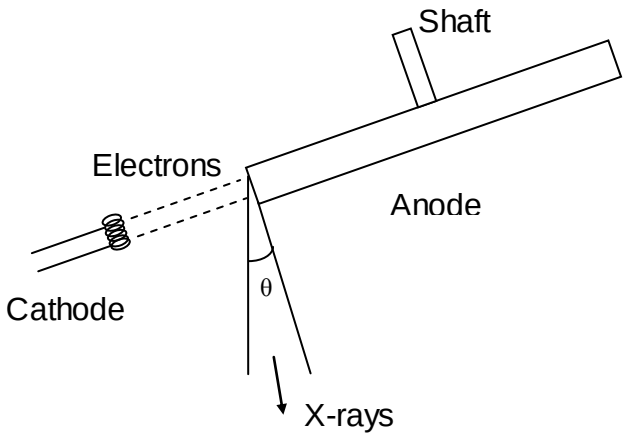


*Figure 2.3 : X-ray tube geometry.*

A consequence of the tilted anode is the "anode-heel effect", which causes the intensity of the x-ray beam to vary across the field of view. This is because x-rays on the anode side of the beam undergo more self filtration in the target material. The anode-heel effect is compensated for in

mammography by arranging the cathode-anode axis of the x-ray set along the chest wall-nipple direction. This evens out the exposure across the image, since the thickness of tissue falls off towards the breast edge – on the same side of the image where the x-ray intensity is also reduced.

## *2.3 X-ray beam filtration*

After the x-ray beam is produced at the target, it is modified to give the optimum energy spectrum using a filter. In mammography, filters are used which are made of the same material as the target. This is to make use of the x-ray attenuation characteristic of the filter, which has a minimum for energies just below the k-edge. X-rays with energy above the k-edge are able to excite k-shell electrons, and hence are absorbed more strongly than x-rays just below the k-edge energy. The minimum in absorption coincides with the energy of the characteristic x-rays, hence this filter is ideal for enhancing these components of the spectrum, while attenuating higher and lower energy bremsstrahlung components (see Figure 2.4).

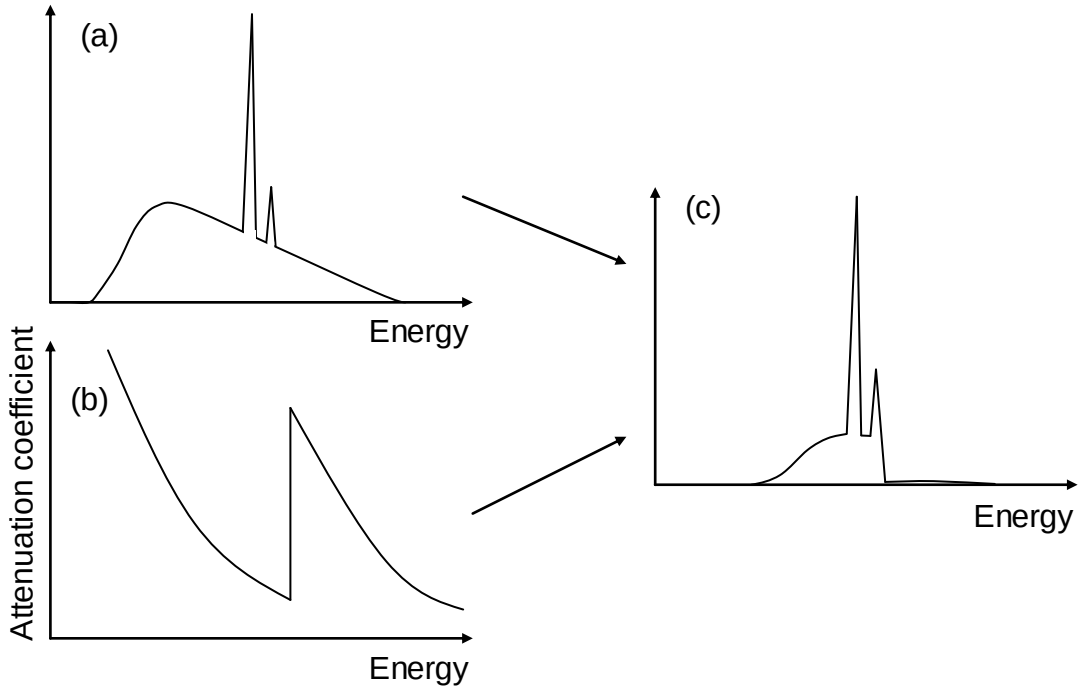


*Figure 2.4 : (a) Unfiltered x-ray spectrum (b) Attenuation coefficient of filter as a function of x-ray energy (c) Filtered x-ray spectrum.*

## *2.4 Interaction of x-rays with tissue*

X-rays interact in tissue by either the photoelectric effect or Compton scattering. Photoelectric interactions occur when the x-ray photon transfers all of its energy to a bound electron, which is then ejected, from the atom. The x-ray is totally absorbed. The probability of photoelectric interactions is approximately proportional to $Z^3/E^3$, where Z is the atomic number of the material and E is the x-ray energy. The photoelectric effect is highly dependent on the atomic number of the absorbing material. This provides contrast between different tissue types (e.g. between bone, which has higher atomic number, and soft tissue). The strongly inverse dependence on E causes photoelectric absorption to dominate at low energy. At higher energies contrast between tissues is reduced because the photoelectric effect becomes less important. Mammography requires good contrast between very similar soft tissues, so lower energies are used than in other x-ray applications.

Compton scattering occurs when the x-ray interacts with weakly bound outer shell electrons, and only transfers part of its energy. The x-ray is not totally absorbed, but continues with lower energy and is deflected through an angle depending on the amount of energy transferred. The probability of Compton scattering increases with x-ray energy, and also depends on the electron density of the material. The electron density in tissue is relatively constant, so Compton scattering does not contribute usefully to tissue contrast. The main effect of Compton scattering is image degradation caused by deflected x-rays reaching the image receptor.

## *2.5 Anti-scatter grid*

Scattered radiation contributes to the image from other parts of the field of view where scattering material is present. This forms a slowly varying background which reduces subject contrast. Anti-scatter grids are used to reduce the amount of scattered radiation reaching the image receptor. The

grid works by stopping any x-rays that are not travelling parallel to the incident x-ray beam direction. The grid also absorbs some of the primary radiation, since it involves placing extra material in front of the image receptor. This means that the dose is higher to achieve the same film exposure.

## *2.6 Breast compression and mammographic projections*

The amount of scattered radiation also depends on the amount of scattering material in the x-ray beam. The thickness of tissue can be reduced by compressing the breast, leading to a reduction in scattered radiation. This increases image contrast without increasing dose. Breast compression also reduces geometrical blurring (due to different magnification for objects at different depths), and spreads out anatomical details.

The mammogram should contain as much breast tissue as possible, and the view used will depend on the purpose of the examination. Most common is the Medio Lateral Oblique (MLO) view (see Figure 2.5a), used in all UK breast screening examinations. The MLO view uses the x-ray beam tilted at about 45-50˚ from vertical, and should include pectoral muscle to the level of the nipple, the intra-mammary fold, and the nipple in profile. Also common is the Cranio Caudal (CC) view (see Figure 2.5b), where the x-ray beam is vertical.

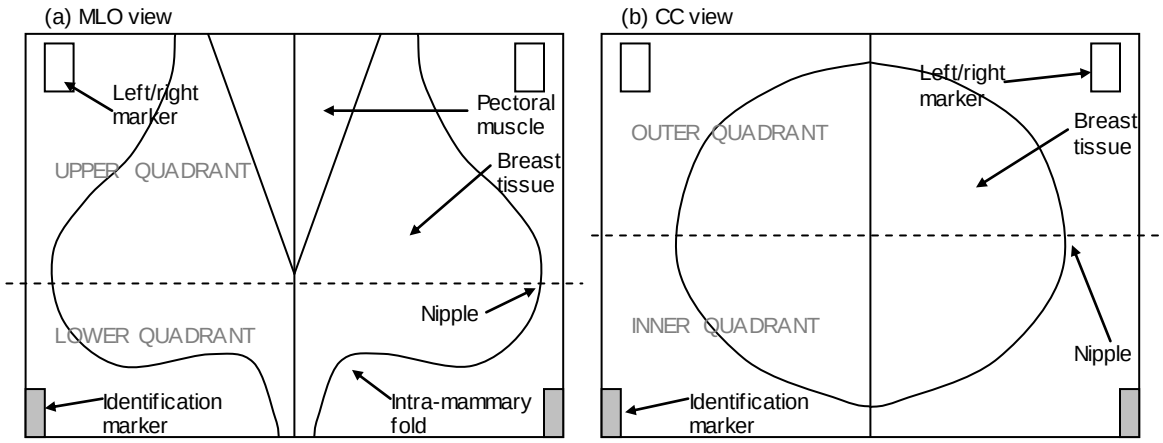


*Figure 2.5 : Schematic diagram of (a) Medio lateral oblique view (b) Cranio caudal view mammograms.*

## *2.7 Image formation*

The image is formed when x-ray photons that have passed through the subject are absorbed by the image receptor. This is usually a cassette containing a sheet of photographic film and an intensifying screen, although digital imaging devices can be used for mammography. Film alone is inefficient at absorbing x-rays, hence a large number of incident x-rays are needed to produce an image. The intensifying screen is used to increase the x-ray absorption efficiency, and hence reduce the dose. The screen is made of a rare-earth phosphor which absorbs the x-ray and converts it to light photons. Thicker screens give better improvements in sensitivity, although resolution is reduced due to the light photons having more room to spread out before reaching the film.

In mammography resolution is very important so single screen systems are used, with the screen placed behind the film. The front surface of the screen, where x-rays are most likely to interact, is thus located as close as possible to the film.

The light photons emitted by the screen are absorbed by the film. The film consists of an emulsion of silver halide backed with a plastic base. The light photons induce a photochemical reaction in the silver halide. This is the latent image, which is not visible at this stage. The film is subsequently developed which causes the film to darken according to how much light it was exposed to. The resultant image is a negative, where more darkening corresponds to greater exposure.

The darkening of radiographic film is measured in units of optical density (OD). This is defined as $OD = \log_{10}\left(\frac{I_0}{I_t}\right)$, where $I_0$ is the incident light intensity and $I_t$ is the light intensity transmitted through the film. The useful range of OD is roughly 0.25-2.0. The response of a film to increasing exposure is described by its characteristic curve (see Figure 2.6). Exposure is a measure of the total number of photons in the x-ray beam. It is usually measured in units of current $\times$ time or mAs, because the number of photons produced is directly proportional to both the tube current (number of electrons incident on the target per unit time), and the time for which the beam is switched on.

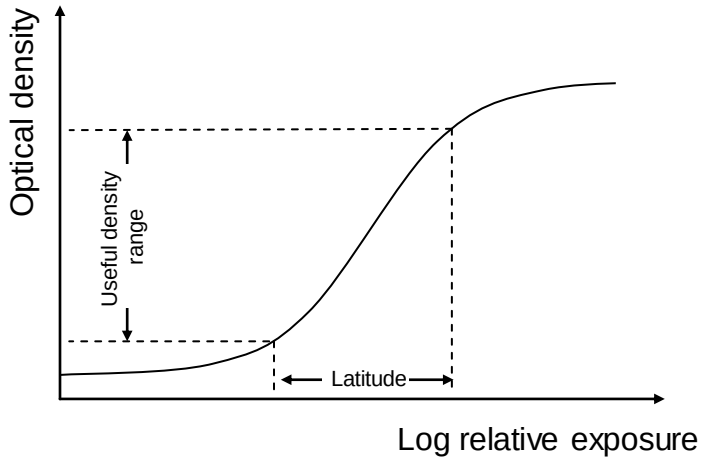


*Figure 2.6: The characteristic curve.*

An important feature of the characteristic curve is its gradient. This determines the film contrast – how quickly the density varies with increasing

exposure. A steeper characteristic curve improves contrast, but decreases exposure latitude. This means the useful range of exposures is smaller, hence it is easier to under- or over-expose the film. Mammography systems use an automatic exposure control (AEC), which aims to terminates the exposure in the centre of the useful range. The AEC only works correctly when positioned underneath the centre of the breast.

## *2.8 Radiation dose*

Mammography involves the use of ionizing radiation, which may cause damage to the exposed tissue. It is important that the risks involved are considered and that doses are kept within acceptable, safe limits. Radiation causes ionization in tissue, which can damage DNA molecules directly or via the formation of highly reactive chemicals called free-radicals. Damaged cells may be unable to reproduce, or may become cancerous, reproducing uncontrollably.

Radiation dose is quantified by the amount of energy absorbed per unit mass of tissue. The unit of absorbed dose is the Gray, which equals 1 J/Kg. The dose due to a mammographic examination varies due to differences in x-ray spectra, exposure level, and in the amount and composition of the breast tissue. The dose delivered by a typical 2-view mammography examination is roughly equivalent to the dose received from natural background radiation in six months.

# 3 Breast density as risk factor

Several factors have been shown to be linked to risk of breast cancer, the most significant being age. Age is the primary criterion by which women are selected for breast screening, with incidence rates being highest for women over the age of 60 (Lacey 2002). Other well established risk factors are later age at first birth, nulliparity (not having borne children), higher family income, and first degree family history (Madigan 1995). The risk of breast cancer is approximately doubled for women with an affected mother or sister (Boyd 1999). The link to family income may be connected to higher surveillance among wealthier women.

Breast density, or the amount of dense tissue present in the breast, is a less widely used risk factor for breast cancer. This factor was first recognized after the observations that fewer cancers occur in very fatty breasts, and that cancers generally arise within the dense tissue component. Several schemes have been developed for classifying the amount of dense tissue, and research has confirmed that denser breasts are associated with a higher risk of developing cancer.

## *3.1 Wolfe classification scheme*

The first scheme for classification of breast density was suggested by Wolfe (Wolfe 1976). His system classified mammograms into four categories with different patterns of tissue (called parenchymal patterns). Wolfe's scheme considers the breast to consist of 3 tissue types: fat, epithelial and connective tissue, and prominent duct. The parenchymal categories are based on the relative amounts of these constituents visible in the mammogram, as follows:

| | |
|---|---|
| *N1 (lowest risk)*: | Primarily fat, little epithelial and connective tissue, no prominent duct. |
| *P1 (low risk)*: | Chiefly fat, prominent ducts occupying no more than a quarter of the breast volume. |
| *P2 (high risk)*: | Prominent duct occupying more than a quarter of the breast volume. |
| *DY (highest risk)*: | Large amount of epithelial and connective tissue, often obscuring underlying prominent duct pattern. |

The Wolfe scheme was significant because it was the first to recognize the link between breast density and cancer risk. However, the category definitions are very subjective, and this leads to a high degree of inter-observer variability. This scheme is therefore not commonly used in the UK.

### *3.2 Boyd classification scheme*

Another scheme was developed by Boyd *et. al*., where mammograms are classified into six classes according to the percentage of dense tissue apparent (Boyd 1982). Dense tissue is defined as anything that is not fatty tissue, which appears as increased density on the mammogram. Mammograms are classified into six categories with 0%, <10%, 10-25%, 25-50%, 50-75%, and 75-100% of the breast containing dense tissue. This scheme is more quantitative than the Wolfe scheme, however the percentages of dense tissue are usually assessed by a visual estimate, rather than an objective measurement technique. It is also unclear whether the percentages should refer to the area of the image with evidence of dense tissue, or to the volume of the breast containing dense tissue. For these reasons the Boyd scheme is more subjective than it at first appears.

## *3.3 Relation between breast density and risk*

Both schemes have been shown to be correlated with breast cancer risk, with the quantitative scheme of Boyd showing stronger correlation (Boyd 1995). Breast density is also strongly correlated with risk within high risk groups, such as women with a family history of breast cancer (Boyd 1999).

Mammographic density has been described by Byrne (Byrne 1997) as the most undervalued and underused breast cancer risk factor. This may be because after the link between breast density and cancer risk was first reported by Wolfe, subsequent studies were inconclusive. This caused people to dismiss the link, although many of the studies which were non-confirmatory have since been shown to have possible bias' due to not blinding radiologists to patient outcome, inadequate radiologist training, and use of low contrast images. The most recent studies have tended to confirm the connection, with risk increasing 4 to 6 times for the highest density category (Byrne 1997).

Byrne also argues that more than 46% of breast cancers are attributable to some measurable density. This makes breast density more significant than many generally accepted risk factors. For example bilateral breast cancer in a mother confers a very high relative risk to the daughter, but is rare and so is only present in a small number of cases. Risk factors such late age at first birth and nulliparity are much more common, but only confer a small increase in breast cancer risk. Breast density is an important risk factor because it combines large relative risk with a high rate of incidence.

The mechanism by which breast density is related to cancer risk is not well understood. However, it has been suggested that if the malignant transformation of certain cells in the breast is a stochastic event, then the probability of this occurring should be proportional to the total number of cells. Most malignancies occur in the dense glandular tissue, so risk may in turn be related to density of breast tissue (Boyd 1999).

Density classification schemes based on subjective assessment by radiologists are inevitably affected by problems with reproducibility. This can be partially addressed by increased training of radiologists, and the link with cancer risk has been shown to increase with more training. It would still be desirable to produce an objective, quantitative measure of breast density, and to establish that it is proportional to risk. This can be done by comparison to classifications where risk is established, or by new studies.

Density is also an important risk factor because of the potential that it can be changed, for example by exercise, diet or hormonal intervention (Boyd 1999). A study by Boyd *et. al*. (Boyd 1997) has shown that changes in diet can have an effect on breast density. Change to a low-fat, high-carbohydrate diet for two years was shown to result in a reduction in the dense area of the breast by about 6%. This was associated with a similar reduction in the overall breast area; hence the percentage of dense breast tissue did not change. It has yet to be established that these changes in density are accompanied by an equivalent change in risk.

# 4 Quantitative measurements of breast density

A number of methods have been suggested for making an objective, quantitative measurement of breast density based on mammographic images. Three methods have used a calibration of the imaging system together with knowledge of the breast thickness to relate film density to thickness of dense tissue at each pixel in the image. These methods are described in sections 4.1 to 4.3. Others have applied computerized image analysis to mammograms, but without calibration of the imaging process they cannot produce the same information about thickness of dense tissue. A summary of some of these methods is given in section 4.4.

## *4.1 Manchester method*

In this approach (Smith 1998), which forms the basis of the work presented in this thesis, a stepwedge is placed on the imaging plate, and imaged as part of the mammogram. Radio-opaque markers are placed on the compression paddle, and these are also imaged as part of the mammogram. The magnification of these markers allows an accurate determination of the compressed breast thickness (to ±0.9 mm). For each pixel in the mammogram, the thickness of the stepwedge that has the same grey level is found, and this is related to the thickness of dense tissue at that point in the breast. To simplify the situation the breast is assumed to consist of only two tissue types – dense or fatty. Calibration data are collected using various thicknesses of (dense and fatty) tissue equivalent materials. Along with the breast thickness measurement, this allows the composition of the breast to be uniquely determined at each point.

This method has the advantage of determining the volume of dense tissue, instead of just the area. If the hypothesis that the amount of dense tissue is important is correct, then this should yield a more sensitive indicator of risk. More work is needed to establish this, though, for example by following large

group of women over a long period of time, and analysing the correlation between the number of cancers and volume of dense tissue.

Another attractive feature of this method is that all of the necessary information is imaged as part of the mammogram, so no external information needs to be recorded by the radiographer.

A disadvantage of this method is that it does not fully take into account the effects of scattered radiation. The film density at each point is due to both primary and scattered radiation. Scattered radiation increases with the thickness and area of the imaged object. A study of scatter to primary ratio (SPR) in mammography (Boone 2000) showed that SPR increases linearly with breast thickness, but is only weakly dependent on tissue composition. The scatter contribution is greater for larger breasts and falls off towards the edge of the breast. Hence there will be variations in film density depending on the breast area, and on position in the breast. These scatter effects are not fully replicated in the image of the stepwedge, so they will introduce spurious variations in thickness of dense tissue calculated by this method. Anti-scatter grids are used for mammography and this is expected to reduce the scatter contribution.

More details on this method, including improvements and evaluation, are given in chapter 5. Chapter 6 describes the application of this method to a sample of clinical mammograms.

## *4.2 Oxford method*

An alternative method of calibrating mammograms based on a model of the physics of the imaging process has been suggested (Highnam 1994, 1996a, 1996b, 1998). This method also considers the breast to consist of only two tissue types, and computes a thickness of dense tissue (termed "$h_{int}$") at each pixel. The first stage of this method is to compute the scatter component of the image, and subtract it to produce a scatter-free version of

the original image. The scatter image is effectively a blurred version of the original image, produced by convolution with a kernel whose properties are determined by the scatter properties of breast tissue and the properties of the anti-scatter grid.

The scatter-free image is then converted into a representation of the thickness of dense tissue using information on the exposure, the X-ray spectrum, the film-screen response, and the thickness of the compressed breast. The exposure and peak tube voltage need to be recorded at the time of imaging. This is potentially inconvenient, but this information is automatically logged by modern mammography systems.

The compressed breast thickness is determined by studying the behaviour of the $h_{int}$ surface near the edge of the breast. The breast thickness is expected to be constant over most of the breast area, and to fall rapidly to zero at the edge. Near the edge, where the true thickness is falling to zero, a region of $h_{int} < 0$ occurs. This is because a thickness $H$ of only fatty tissue would still cause more attenuation than is observed on the film, where $H$ is the compressed breast thickness. The estimate of $H$ can be adjusted to give a well behaved $h_{int} < 0$ region, with a smooth inner edge. This method gives a measure of the thickness accurate to between 2 and 4 mm.

The main differences between the Manchester and Oxford methods are the way that changes in imaging parameters are accounted for (principally exposure), and the treatment of scattered radiation. In the Oxford method, the imaging parameters (such as exposure) are recorded at the time of imaging. The thickness of dense tissue corresponding to each film density is then calculated using a physics based model of the imaging system. This involves complex calculations, requiring knowledge of the shape of the x-ray spectrum, and the transmission functions of the anti-scatter grid and compression paddle as a function of photon energy. The Manchester method does not record the imaging parameters, but records an image of the (previously calibrated) stepwedge on each mammogram. This makes the Manchester method somewhat simpler to implement.

Other differences between the methods are that the Oxford method explicitly attempts to remove the scattered radiation component from the image, and the different methods used to measure the compressed breast thickness.

### *4.3 Calibration approach for digital mammography*

A method suitable for digital mammography systems, based on calibrations of the system using various thicknesses and compositions of tissue equivalent materials has been described by Kaufhold *et. al*. (Kaufhold 2002). The calibration data relate radiation intensity to amount of dense tissue and total breast thickness. The calibration data are parameterized, allowing the percentage of dense tissue to be calculated at each pixel for any breast thickness and radiation intensity value. In this respect the method is similar to the Manchester method described above.

Important differences in the methods arise from the nature of digital mammography. A digital detector is used instead of a film-screen system. This simplifies the analysis because a number of counts is recorded at each pixel, which is linearly proportional to exposure. Measurements can be scaled to a standard mAs value to allow comparison between calibration and clinical data. This eliminates complications due to film characteristic curve and processing, so the stepwedge device used in the Manchester method is not necessary. Digital systems also allow easier correction for geometric and anode-heel effects.

The method described by Kaufhold *et. al*. only makes measurements in the central breast region, where the breast is in contact with the compression paddle. The edge region where the breast thickness is reduced is not considered. Breast thickness measurements are taken from an automatic reading supplied by the mammography unit. Tilt or bending of the compression paddle is not considered. The measurement of breast thickness was considered by the authors to be the dominant source of error

in the percentage dense measurement. An error of 2 mm in the breast thickness measurement was shown to lead to an error of 7 % in the percentage of dense tissue measurement at each pixel (for a 4 cm thick breast with 50 % dense 50 % fatty composition).

### *4.4 Other approaches*

Other researchers have treated measurement of mammographic density as a segmentation problem, with the aim of segmenting the mammogram into dense and non-dense tissue areas. This approach results in a binary segmentation, which ignores information about the amount of dense tissue at each pixel. Hence only area based measurements are obtained, such as percentage of dense tissue by area. This type of measure does, however, allow easy comparison to breast density classification schemes, such as Boyd's SCC scheme, whose relationship to cancer risk has been established.

Byng *et. al*. (Byng 1996) investigated the use of fractal dimension and histogram skewness for assessment of breast density from mammograms. Dense regions of the breast were hypothesized to have lower fractal dimension due to their smoother appearance in mammographic images. Dense areas were also associated with negative histogram skewness due to the high mean grey level and tail at lower values. Both measures were compared to Boyd's six-category density classification scheme, and significant correlations were found (Spearman coefficients –0.76 and –0.88 for fractal dimension and histogram skewness respectively). The greatest correlation was found when considering both variables together. This method was shown to be relatively insensitive to small changes in mammographic imaging parameters such as exposure level and film characteristic curve. It was noted that reproduction of the Boyd scheme was not an end in itself, and that the real aim should be to investigate relative risk associated with the new density assessment based on image parameters.

Another automated approach to density assessment looked at the variance of pixel values in areas containing different tissue types (Heine 2000). Each pixel is labelled as either fatty or dense (non-fatty), and hence a binary separation of the mammogram is achieved. This method was not explicitly compared to expert radiologist classifications, but the authors state that expert readers have visually verified that the segmentation is reasonable.

Another approach based on analysis of the image histogram was used by Zhou *et. al*. (Zhou 2001). Here the mammograms were first filtered to remove low-frequency variations across the image (principally due to variations in breast thickness), and then classified according to the number and width of the peaks in the image histogram. Next the image was segmented into dense/non-dense areas using a grey level threshold chosen by a method dependent on the previous classification. The automatic classification was compared to radiologists' segmentations using an interactive grey level thresholding tool. The tool allowed radiologists to vary the segmentation threshold, and automatically calculated the percentage of dense tissue for the chosen segmentation. This tool restricts the radiologists' segmentation choice by assuming that it is based only on the grey level in the image. However, it did show that the radiologists consistently underestimated the percentage by area of dense tissue, when compared to the automatically calculated area. The automatic measurement of dense area was found to be significantly correlated with the radiologist measurement (correlation coefficients 0.94 and 0.91 for CC and MLO views respectively). However, these results did not include 6% of cases where the automatic method was considered to have failed, with poor resulting segmentation.

# 5 Method for calibrating mammographic densities

This chapter describes in detail the method for calibrating mammographic densities introduced in section 4.1. The two main elements of this method are the magnification markers used to measure compressed breast thickness and a stepwedge used to calibrate optical density against thickness of dense breast tissue. These are described in sections 5.2 and 5.3 respectively. The methods used here were described in (Smith 1998), but part of the author's work involved writing software to apply them. Section 5.4 describes further improvements to the technique and assessment of its performance, which was carried out by the author. Results of applying the method to a large set of mammograms in a study of lifestyle effects on breast density are described in chapter 6.

## *5.1 Equipment used in this project*

The mammography unit used in this project was made by LORAD (Danbury, CT). The films were digitized using a Kodak LS85 digitizer at a pixel size of 50 μm and with 12 bits (4096 grey levels) pixel depth. The grey level is linearly related to OD in the range 0.03-4.1 OD (Kodak 2001). This digitizer produces images with grey level 0 corresponding to an OD of zero, displayed as white, and grey level 4096 corresponding to the maximum OD value, displayed black. This is contrary to the usual convention of low grey values being displayed as black and high values being white. The pixel depth was reduced to 8-bits (256 grey levels) after digitizing, using a window based on the maximum and minimum OD present in the image. This was to reduce file sizes of the stored images. The relationship between the original and reduced grey levels was recorded for each image, to allow consistent comparisons between images. The images were used at maximum spatial resolution for the measurement of compressed breast thickness described in section 5.2. The resolution was reduced to a pixel size of 250 μm for the

thickness of dense tissue measurements described in section 5.3. All image analysis was carried out using Matlab 6 (The Mathworks, Inc.), and data analysis was done using both Matlab and Microsoft Excel.

## *5.2 Measurement of compressed breast thickness*

An essential element in measuring the thickness of dense breast tissue from a mammogram is information about the compressed breast thickness. This allows the x-ray attenuation in the breast to be interpreted in terms of dense tissue thickness. The method used for measuring the compressed breast thickness is based on the use of radio-opaque magnification markers, suggested by Burch and Law (Burch 1995). A transparent acetate sheet with several pairs of circular lead markers is placed on top of the compression paddle. The arrangement of markers is shown in Figure 5.1.

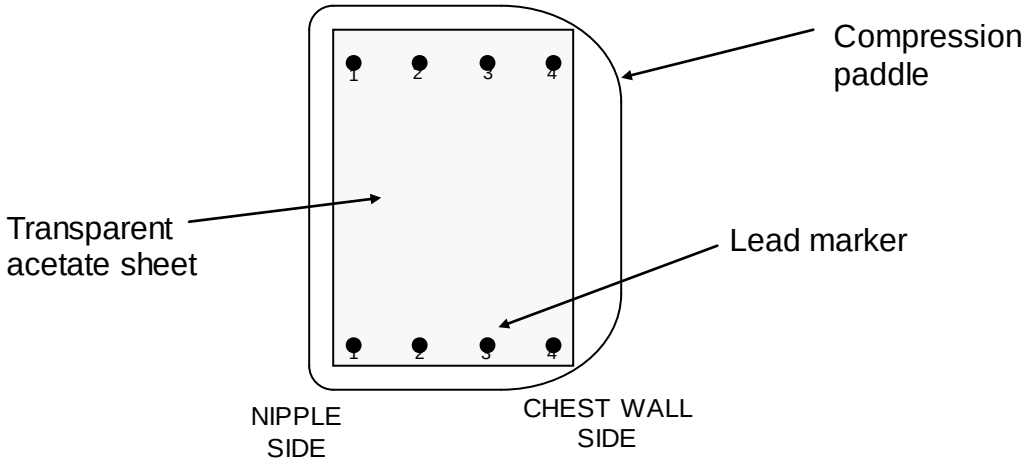


*Figure 5.1: Acetate sheet with lead markers placed on top of compression paddle.*

The marker sheet was made by sticking small lead discs to an acetate sheet using a layer of adhesive backed transparent plastic covering. It is important that the marker sheet is transparent, so that it does not interfere with the radiographer's view of the breast when positioning the subject for the examination.

The magnification markers are imaged as part of the mammogram, and the distance between them on the film is used to find the distance between the compression paddle and the image receptor. This geometry of this principle is shown in Figure 5.2.

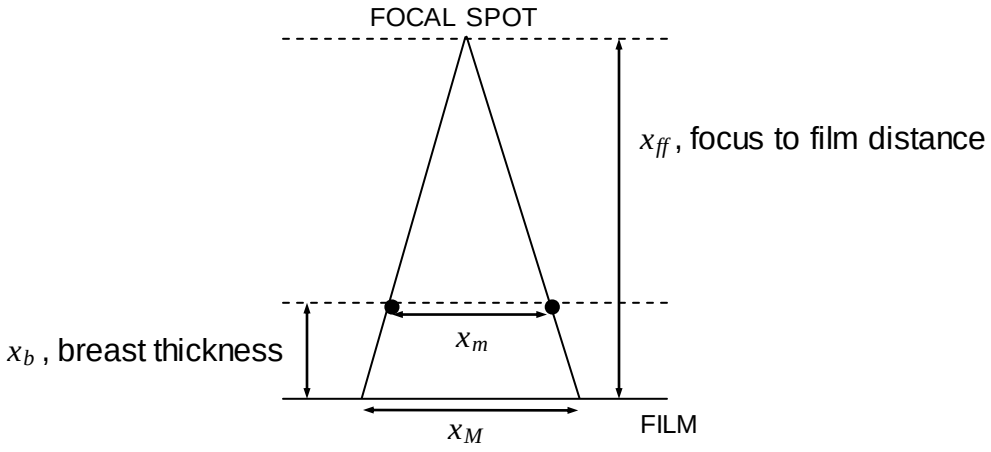


*Figure 5.2: Magnification marker geometry.*

The image of each pair of magnification markers appears magnified on the film, by a factor

$$M = \frac{x_M}{x_m} = \frac{x_{ff} - x_b}{x_{ff}} \quad (1)$$

where $x_m$ is the marker separation on the acetate sheet, $x_M$ is the marker separation in the image, $x_{ff}$ is the focus to film distance, and $x_b$ is the distance between the compression paddle and the film. Hence the breast thickness can be determined as

$$x_b = x_{ff} - \frac{x_M x_{ff}}{x_m} \quad (2)$$

In practice it is difficult to accurately measure the focus to film distance, and the value of $x_b$ measured using equation 2 is not precisely the breast thickness, but also includes the thickness of the compression paddle and film-screen cassette case. These difficulties are overcome by taking calibration films with the compression paddle lowered to until it is just in contact with a series of test objects of known thickness. The separation of the markers on the film is then measured and plotted against the thickness of

the object under the compression paddle. This method has the additional advantage that it is not necessary to know the actual separation of the markers on the acetate sheet. Also, separate calibration plots can be made for each marker pair, so it is not even necessary for all of the marker pairs to have the same separation. Calibration data for one of the pairs of markers are shown in Figure 5.3.

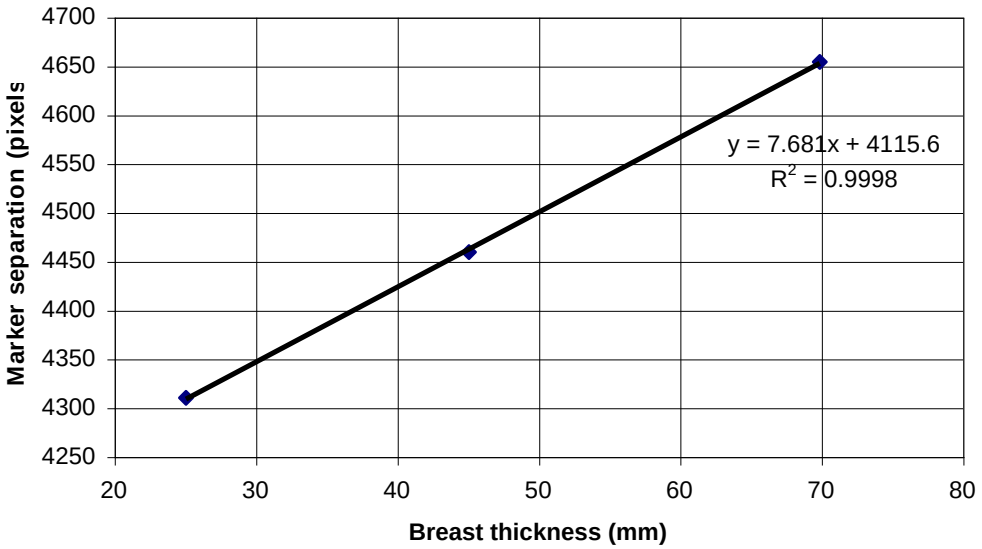


*Figure 5.3: Calibration data for marker pair 1 on small compression paddle marker sheet.*

### 5.2.1 Location of marker positions

The key to accurate determination of breast thickness is finding the position of the magnification markers with high precision. Circular markers are used and the centre of the circle is designated as the reference point on the marker. The centre of the circle can always be accurately located, even when the edges of the marker are blurred, or partially obscured. The marker location is done in four stages, and is semi-automatic. First the user selects a small region around the marker from the main image. This small region is then contrast enhanced so that its contents cover the full range of grey levels from 0 to 255. The third stage is to apply a Canny edge detector[1] to the contrast enhanced image (Canny 1986), which identifies points falling on an

[1] The canny detector parameters used are: Gaussian smoothing filter with sigma = 2 pixels, upper and lower edge thresholds 0.5 and 0.1 respectively.

edge. Finally a circle is fitted to these edge points, and the centre of the fitted circle is taken as the centre of the marker. Figure 5.4 illustrates the stages in finding the marker positions.

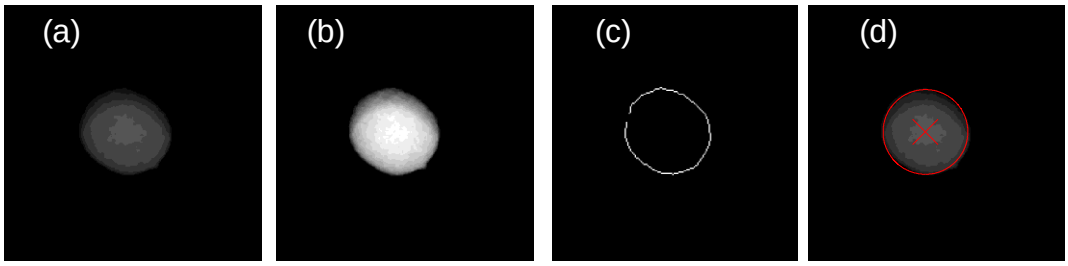


*Figure 5.4: Locating the magnification markers (a) original image of marker (b) contrast enhanced image (c) canny edge detected image (d) fitted circle superimposed onto original image.*

The x-ray films are digitized with spatial resolution of 50 μm per pixel. The magnification markers are located using this high-resolution image. The accuracy of location of the marker centres is estimated to be 5 pixels (see section 5.4.4 for details). The calibration data can be used to convert an error in pixels to an error in the measurement of breast thickness. The error on the breast thickness measurement is thus estimated to be 0.6 mm.

## *5.3 Calibration of thickness of dense tissue*

The aim of the method is to measure the thickness of dense tissue along the projection onto each pixel in the image, using the grey level of that pixel. This is done by imaging a calibrating device called a stepwedge alongside the breast as part of the mammogram. The stepwedge allows an estimate to be made of the thickness of a known material which causes the same x-ray attenuation as the breast tissue. This allows the thickness of dense tissue to be calculated, as explained below.

The intensity of the x-ray beam after passing through a thickness $x_b$ of breast tissue, which is composed of $x_g$ glandular (dense) tissue and $x_f$ fatty tissue is

$$I = I_0 e^{-(\mu_g x_g + \mu_f x_f)}, \quad (3)$$

where $\mu_g$ and $\mu_f$ are the attenuation coefficients of glandular and fatty tissue respectively, and $I_0$ is the incident x-ray intensity. The intensity of radiation after passing through a thickness $x_{sw}$ of the stepwedge material with attenuation coefficient $\mu_{sw}$ is

$$I = I_0 e^{-\mu_{sw} x_{sw}}. \quad (4)$$

If the thickness of stepwedge material causing the same amount of attenuation as a particular point in the breast image is identified, then equations 3 and 4 can be equated to yield the expression

$$\mu_{sw} x_{sw} = \mu_g x_g + \mu_f x_f. \quad (5)$$

Using the assumption that the breast is composed of only glandular and fatty tissue, i.e. $x_b = x_g + x_f$, equation 5 can be rearranged to give

$$x_g = \frac{\mu_{sw} x_{sw} - \mu_f x_b}{\mu_g - \mu_f}, \quad (6)$$

an expression for $x_g$ in terms of only known quantities.

The stepwedge used is composed of PTFE and has 25 steps differing in thickness by 1 mm. The stepwedge is illustrated in Figure 5.5, and the design is discussed fully in (Smith 1998). The material, PTFE, is chosen because of its similar mass attenuation coefficient ($\mu/\rho$) to breast tissue, and because its density is slightly higher than other plastic materials allowing a larger range of attenuations to be achieved without the wedge becoming too thick. The sides of the wedge are shielded with lead to insure that only parts of each step where x-rays have travelled through the whole thickness of the wedge are imaged. Without this precaution the image becomes blurred by x-rays that only pass through part of the wedge, causing the grey level to vary across each step. This is illustrated in Figure 5.6.

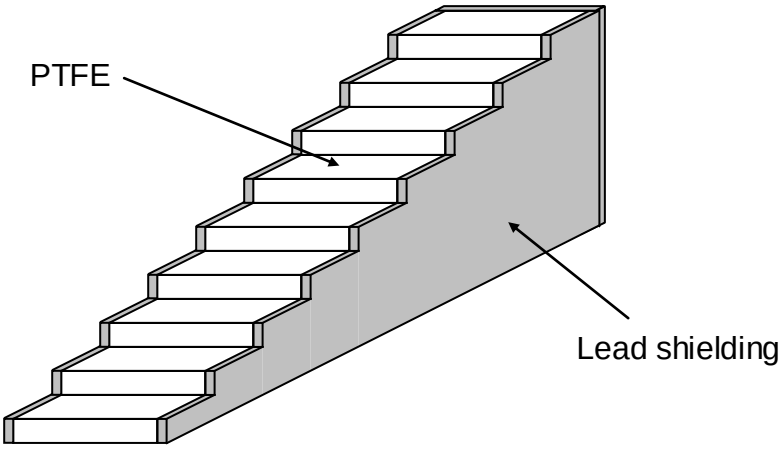


*Figure 5.5: Illustration of stepwedge calibration device.*

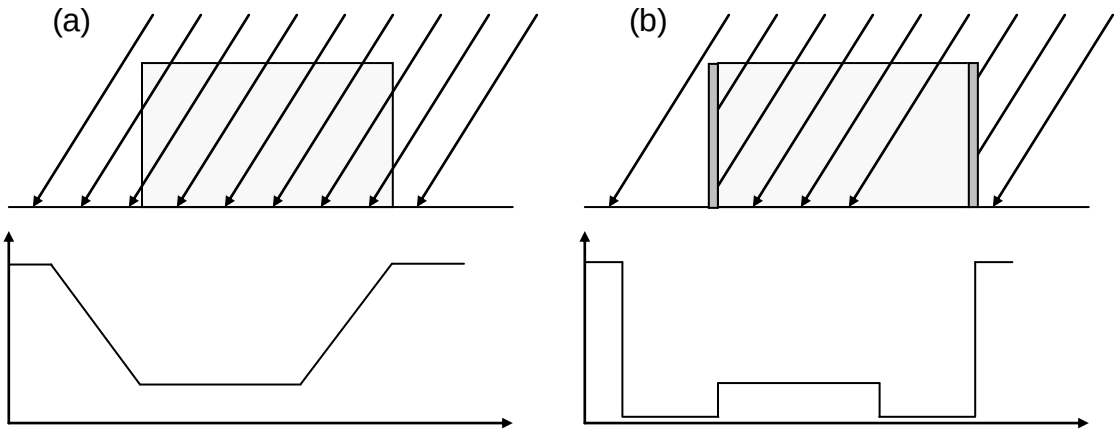


*Figure 5.6: Effect of lead shielding around stepwedge, lower image shows the variation in x-ray intensity across the image of a step. (a) Stepwedge with no lead shielding. The image of the step does not have clearly defined edges. (b) With lead shielding x-rays not passing through entire thickness of step are absorbed by the lead.*

The grey levels in the image depend on several factors other than the thickness and composition of the breast tissue, namely the exposure level of the image, the film-screen response, the processing conditions, the digitizer response, scattered radiation, beam hardening and variations in the x-ray intensity across the image due to inverse square law effects and the anode-heel effect.

The first group of these effects, namely exposure level, film-screen response, processing conditions and digitizer response, can be disregarded because their effect is expected to be the same in the image of the breast and the stepwedge, since the stepwedge is imaged at the same time as the breast.

Beam hardening occurs because the lower energy x-rays are preferentially absorbed from the x-ray beam. This causes the average x-ray energy to increase as the beam passes through an absorber. The attenuation coefficients are dependent on energy, so are expected to change depending on the degree of beam hardening that takes place. It is expected that this effect is small, since a similar degree of beam hardening should occur in the stepwedge material as occurs in the breast tissue.

### 5.3.1 Geometric correction

Variations in the x-ray beam intensity across the image are caused by the inverse square law and by the anode heel effect. The source of radiation, the target, is located directly above a point on the chest-wall side of the film. This results in the distance from the target to the film increasing as a function of distance across the film, as illustrated in Figure 5.7. The increase in distance causes the radiation intensity to fall off due to the inverse square law.

The anode-heel effect also produces higher radiation intensity at the chest-wall side of the film, as explained in section 2.2.

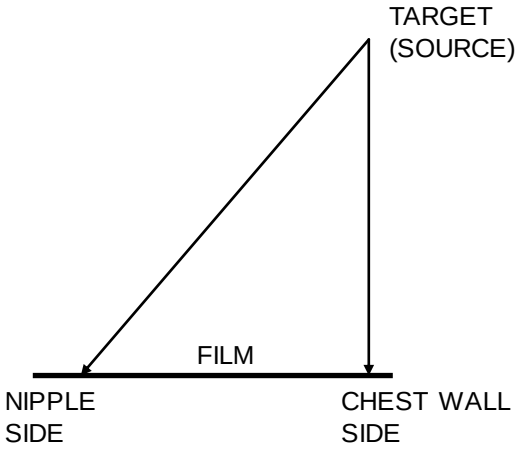


*Figure 5.7: Distance from target to film.*

The variation of radiation intensity across the film is compensated for by means of a correction applied to the grey levels in the digitized image, derived from a blank film taken with very short exposure time. A short exposure time of about 3 to 5 mAs is used so that the variation in intensity is visible across the blank film but without any part of the film being saturated. In this way the optical density remains within the linear part of the characteristic curve.

We can assume that the radiation intensity at each point on the film is some fraction of the maximum intensity occurring directly below the target. The optical density of the film is dependent upon the energy imparted to the film, and this is directly proportional to the radiation intensity. So in order to correct an energy imparted image, $E(x,y)$, for the uneven radiation intensity we must divide it by the fraction of the maximum intensity received at point $(x,y)$. This gives us the corrected image

$$E'(x,y) = \frac{E_B^A}{E_B(x,y)} E(x,y), \qquad (7)$$

where $E_B^A$ is the energy imparted directly beneath the anode in the blank image, and $E_B(x,y)$ is the energy imparted at point $(x,y)$ in the blank image. Provided the optical density is within the linear region of the film's characteristic curve, optical density is related to energy imparted as

$OD = \gamma \log(\beta E)$, where $\gamma$ and $\beta$ are constants related to the film response. Substituting this relationship into equation 7 and rearranging shows that an optical density image $OD(x, y)$ can be corrected for uneven radiation intensity by the addition of a correction according to

$$OD'(x, y) = OD(x, y) + \left(OD_B^A - OD_B(x, y)\right), \quad (8)$$

where $OD'(x, y)$ is the corrected image, $OD_B^A$ is the optical density directly below the anode in the blank film, and $OD_B(x, y)$ is the optical density in the blank film at position $(x,y)$. Since grey level in the digitized image is linearly proportional to optical density, this correction can also be applied in terms of grey levels

$$GL'(x, y) = GL(x, y) + \left(GL_B^A - GL_B(x, y)\right). \quad (9)$$

This correction is valid for all optical densities in the images which are within the linear region of the characteristic curve. This limitation should not be a problem because most optical densities within the breast region are expected to be in this range.

Figure 5.8a shows the blank film used to derive a correction for large format (24x30 cm) films. Film markers are masked off and replaced with intensities copied from adjacent areas on the film. Then the blank film is Gaussian smoothed with standard deviation 15 pixels, found to be sufficient to remove any local or statistical intensity variations, leaving only global intensity variations. The blank image is padded at the edges before smoothing by copying intensity values from the adjacent region on the film. This stops the smoothing from introducing a rapid fall-off to zero at the edges of the film. Figure 5.8b shows the final correction image for the large format film.

This correction makes a significant modification to the image, particularly at the nipple side of the film. The correction increases in size with distance from the point directly below the anode, so the largest correction occurs for the large format film. It is important that the correction is applied to all images (including calibration films) before any further processing is done.

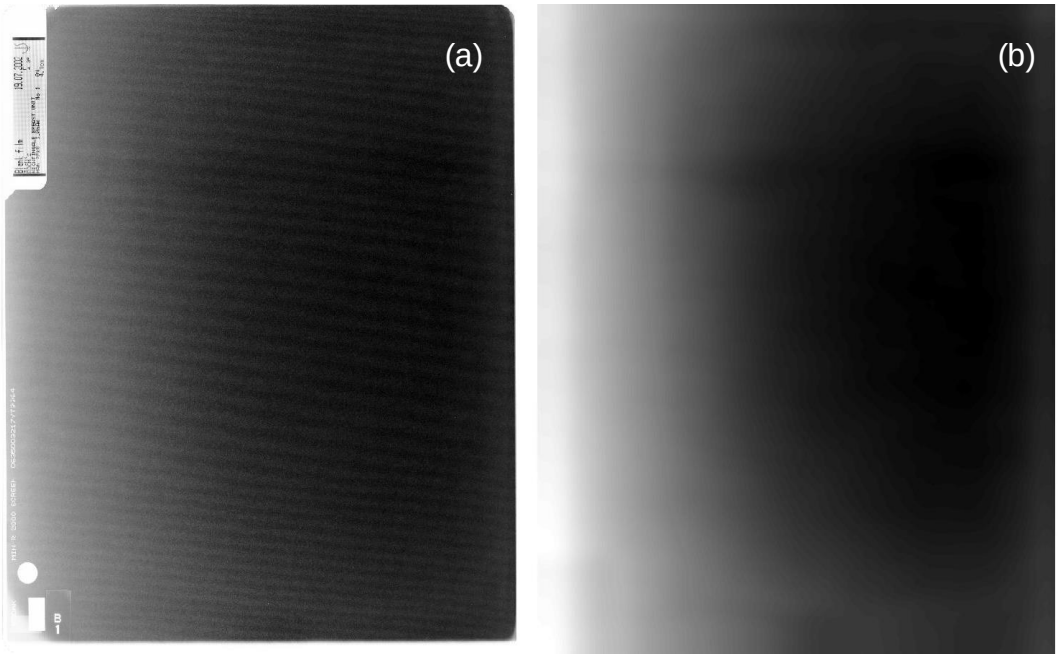


*Figure 5.8: (a) Original blank film (24×30 cm) (b) Correction image made using blank film, after masking film markers and smoothing (note that this image is shown with increased contrast to illustrate variation across film, actual pixel value range in image (b) is 0-92).*

### 5.3.2 Calibration of stepwedge

In practice the stepwedge is calibrated using varying thicknesses of tissue equivalent material. This avoids explicit reliance on published values of the attenuation coefficient of the stepwedge material, PTFE. X-ray images are taken of the stepwedge together with varying thicknesses of glandular and fatty tissue equivalent material. For each image the thickness of stepwedge material producing the same grey level as the tissue equivalent material is found, as illustrated in Figure 5.9. These data are then plotted on a graph of stepwedge thickness against glandular tissue thickness (see Figure 5.10).

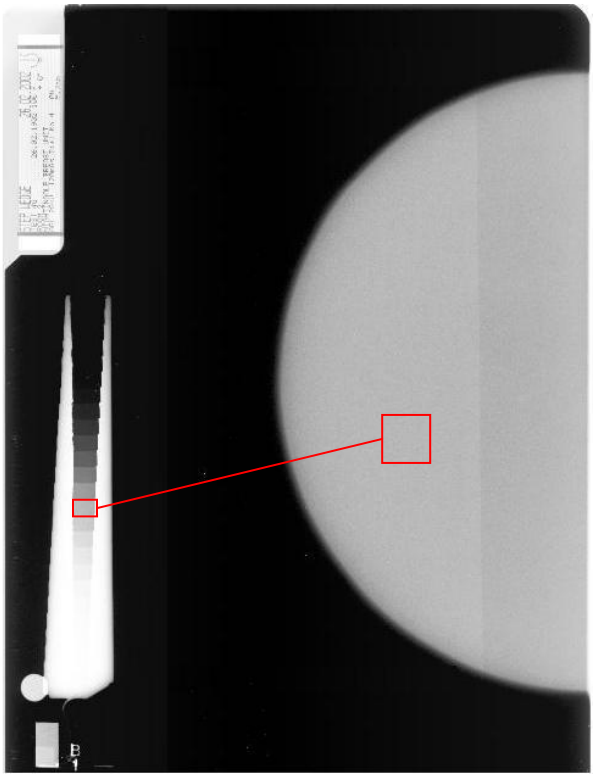

*Figure 5.9: Calibration film showing thickness of stepwedge with same grey level as tissue equivalent material.*

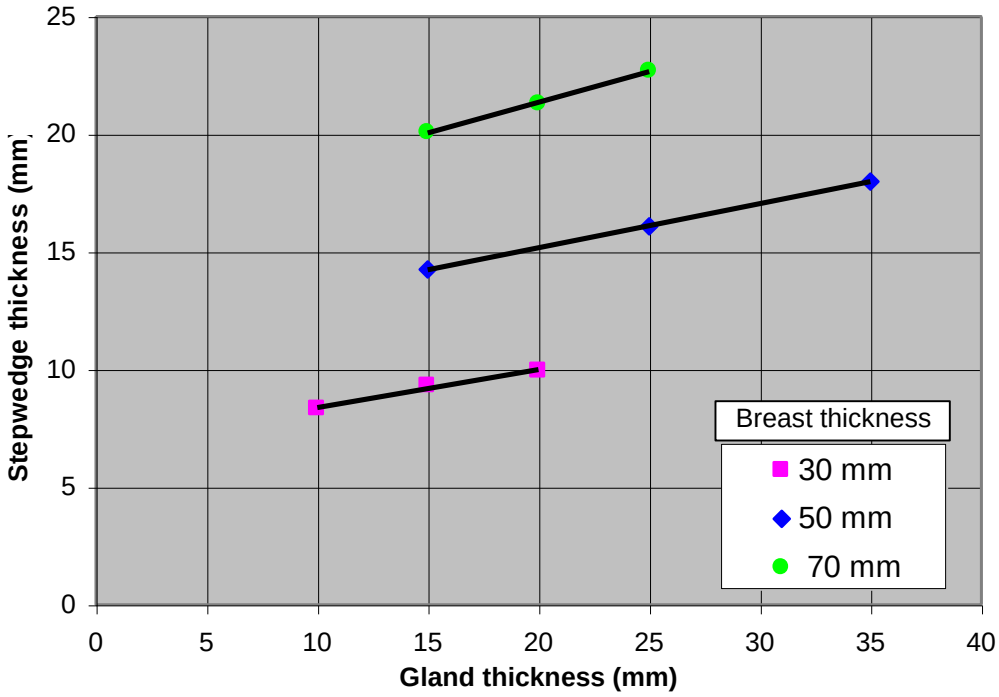


*Figure 5.10: Example calibration data plot.*

The data for each total thickness of breast tissue are found to lie on a straight line, the equivalent stepwedge thickness being proportional to glandular tissue thickness. The intercepts of these straight lines with the stepwedge thickness axis (that is the stepwedge thickness at glandular tissue thickness zero) are also found to form a straight line, with stepwedge thickness proportional to breast thickness. The calibration data can therefore be parameterized with stepwedge thickness being linearly dependent on gland thickness and breast thickness,

$$x_{sw} = ax_g + bx_b + c, \tag{10}$$

where $a$, $b$, and $c$ are constants which can be derived from the calibration data. Equation 10 can be rearranged to give an expression for the glandular tissue thickness,

$$x_g = \frac{x_{sw} - bx_b - c}{a}. \tag{11}$$

It may be noted that equation 11 is similar in form to equation 6, with constant $b$ being related to $\mu_f / \mu_{sw}$, and constant $c$ being related to $(\mu_g - \mu_f)/\mu_{sw}$.

## *5.4 Modifications to the existing method*

Apart from implementing and validating the existing method of breast density assessment using a large sample of clinical mammograms, several modifications have also been incorporated into the method. These are now described and represent new work by the author.

### 5.4.1 Extension to larger film sizes

The existing method for measuring mammographic density was only designed and tested for small format mammography film (18x24 cm) and a small sample of clinical mammograms. It was desired to use the method for a larger set of clinical mammograms, taken for a study of the effects of diet and exercise on breast cancer risk. This group contained a higher than

usual proportion of larger breasts which needed to be imaged using large format mammography film (24x30 cm). This group also contained breasts with a larger amount of dense tissue than is typical in routine screening mammography. It was therefore necessary to assess the method's suitability for use with the large film size and for very dense breasts.

Several test films were taken using large format mammography film, and containing the stepwedge and tissue equivalent material. Films were taken with thicknesses of tissue equivalent material corresponding to the extremes of the range of breast sizes expected to be encountered. These films were then analysed using a densitometer to measure the optical densities of each step on the stepwedge in each exposure. Figure 5.11 shows the stepwedge curves obtained.

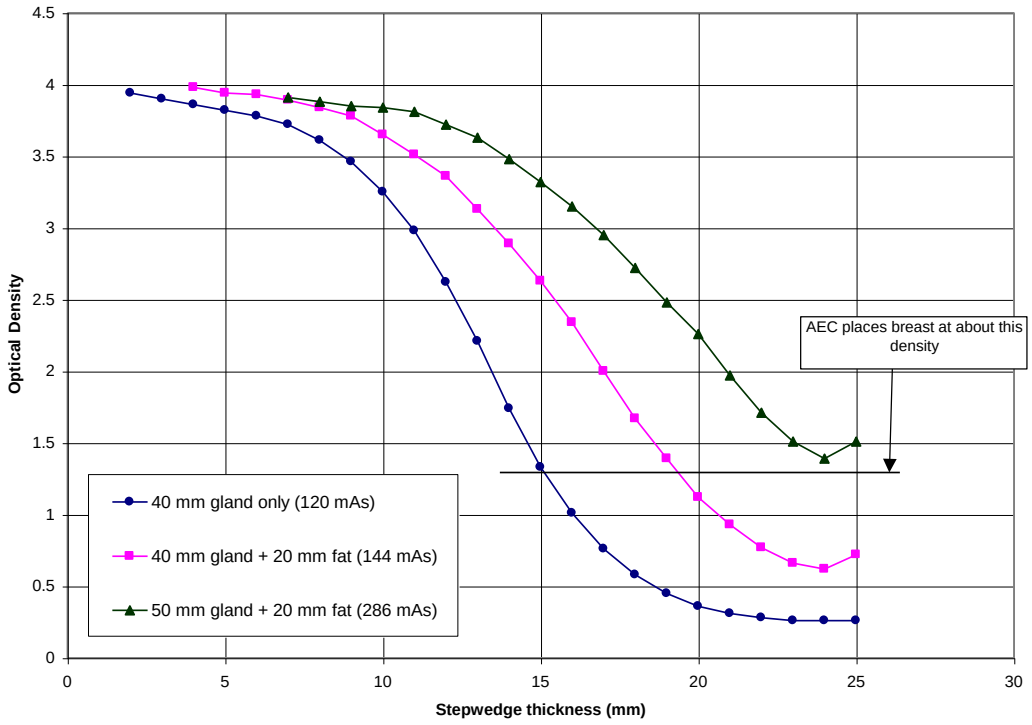


*Figure 5.11: Optical densities on stepwedge for three test films.*

The upturn in the curves in Figure 5.11 at stepwedge thickness of 25 mm is an artefact which occurs on the last step of the wedge. This is because only a very small area of this step is visible for which the x-rays have passed

through the whole stepwedge thickness. The artefact is avoided in the computer analysis of the stepwedge curves by not allowing the curve to change direction in this unphysical manner.

The automatic exposure control (AEC) is located beneath the tissue equivalent material. Hence, as the thickness of tissue equivalent material increases, the exposure also increases. This causes the image of the stepwedge to appear darker – the optical density increases. Ideally, the optical densities in the stepwedge image will cover the full usable range, from 0.2 to 3.9. This would allow any optical density in the breast image to be related to an equivalent thickness of stepwedge material. It can be seen from Figure 5.11 that the film with tissue equivalent of 50 mm gland and 20 mm fat does not meet this requirement. The curve does not even reach an OD value of 1.3, which would be the average OD in the breast region as set by the AEC. The stepwedge is not thick enough to calibrate the high exposure image necessary for a breast with this amount of dense tissue.

The solution to this problem was to design a new thicker stepwedge capable of calibrating higher exposure images. It was estimated from Figure 5.11 that the new stepwedge should have steps with thickness up to 35 mm. This was done by extrapolating the curve corresponding to the higher exposure film along a similar trajectory to the other two curves, to estimate at what stepwedge thickness it would flatten out.

Another issue with using larger films was a reduction in usable area of each step. This is because the stepwedge is placed further away from the point directly below the x-ray source, leading to the x-rays travelling more obliquely to the stepwedge. This reduces the usable width of each step on the stepwedge. Using a new, thicker stepwedge creates a similar problem. This effect can be seen by considering the x-ray intensity profile across the stepwedge shown in Figure 5.6b. If the stepwedge becomes taller, or the x-rays approach at a larger angle to vertical, then the usable width of the stepwedge image is reduced. The usable width of the stepwedge image

must be large enough that there is a sufficient area of the step visible to sample the average pixel value of that step.

The usable width of the stepwedge image was plotted as a function of step thickness for one of the test films where the wedge was placed towards the edge of a large format film, see Figure 5.12. This allows assessment of the usable width for a 35 mm thick step.

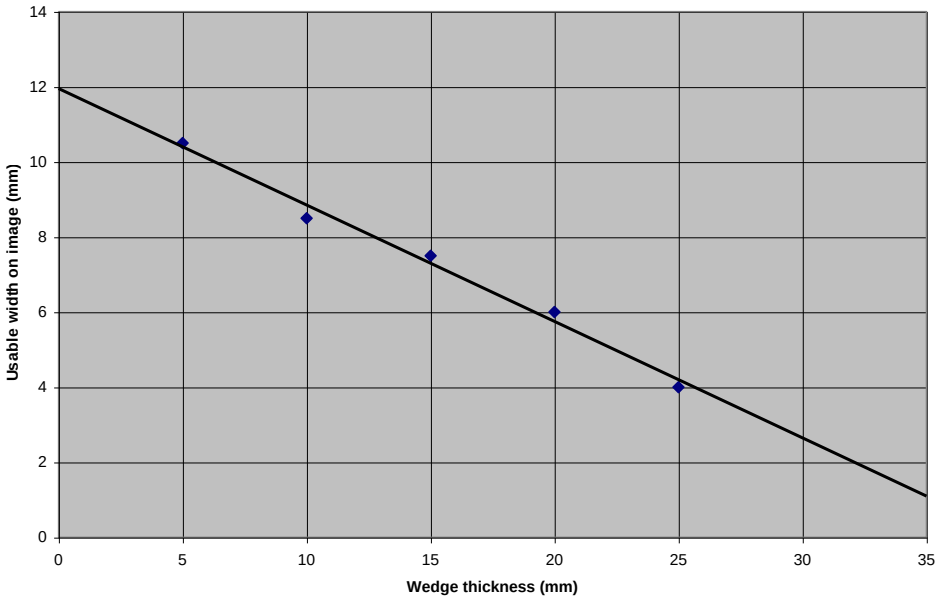


*Figure 5.12: Usable width of stepwedge image as a function of step thickness.*

For wedge thickness zero the line in Figure 5.12 indicates the true width of the stepwedge (the image is not magnified since it sits directly on top of the film-screen cassette) which is 12 mm for the 25 mm stepwedge. Extrapolating the line to wedge thickness of 35 mm shows that the usable width of the image would be only 1 mm. To enable more accurate measurement of the step grey level, the 35 mm stepwedge was designed with an actual width of 15 mm. This results in a usable image width of 4 mm for the thickest step, which gives a reasonable area from which to sample the average pixel value of the step.

### 5.4.2 Improvement to calibration parameterization

It was noticed from the calibration plots (such as Figure 5.10) that the gradient of the gland thickness against stepwedge thickness lines increases with increasing breast thickness. It was found that the agreement between the calibration data and the parameterized values could be improved by modifying equations 10 and 11 so that constant $a$ varies linearly with breast thickness. The new expression for glandular tissue thickness becomes

$$x_g = \frac{x_{sw} - bx_b - c}{dx_b + e}, \tag{12}$$

where $d$ and $e$ are two new constants replacing $a$, and which are derived from the calibration data. Figure 5.13 to Figure 5.16 show the agreement between the calibration data and the parameterized values, illustrating the improvement from allowing the gradient to vary with breast thickness. Note how the old parameterization causes the gradient to be too steep at breast thickness 30 mm and too shallow at breast thickness 70 mm.

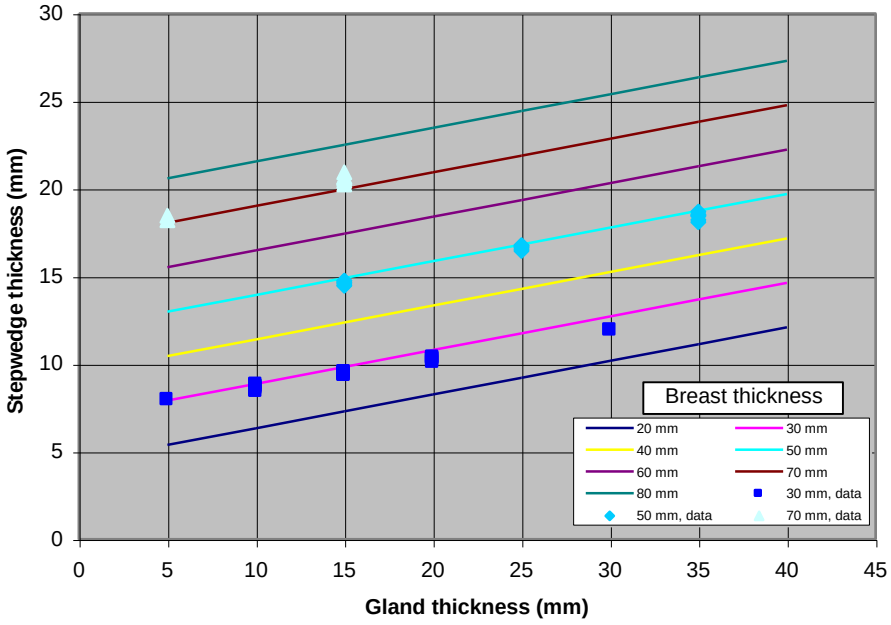


*Figure 5.13: Calibration data (points) and parameterization (lines) for small stepwedge using old parameterization.*

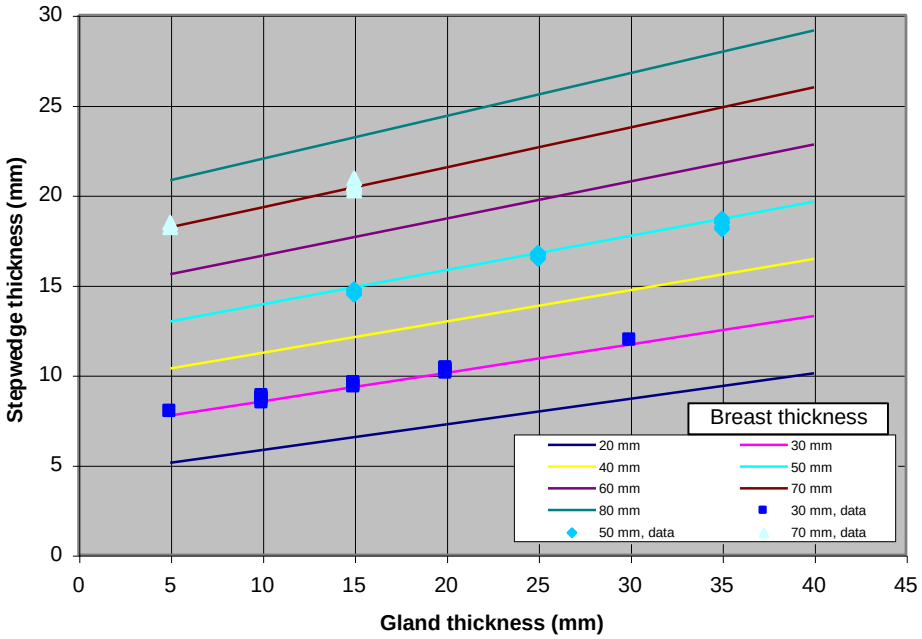


*Figure 5.14: Calibration data (points) and parameterization (lines) for small stepwedge using new parameterization.*

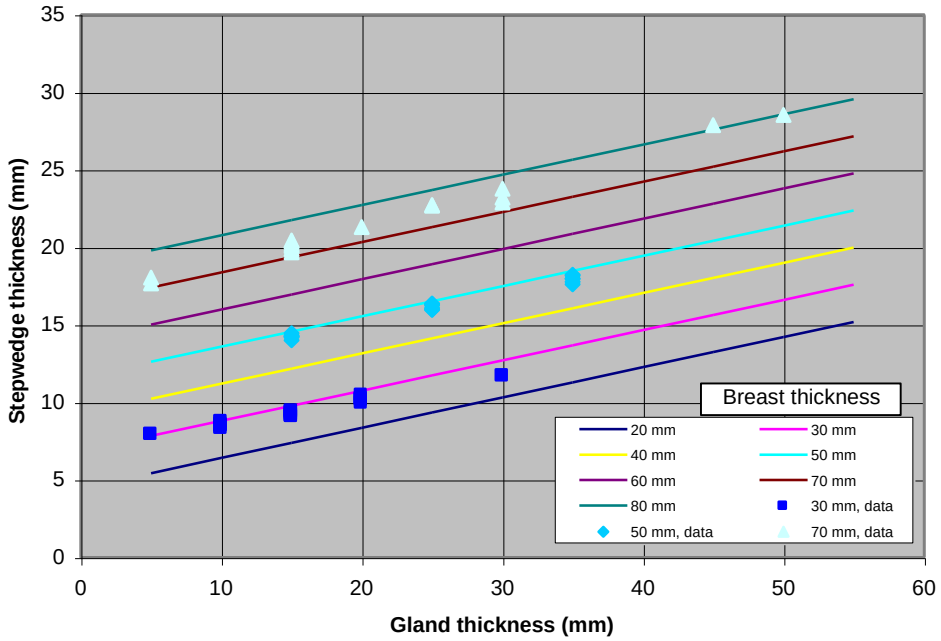


*Figure 5.15: Calibration data (points) and parameterization (lines) for large stepwedge using old parameterization.*

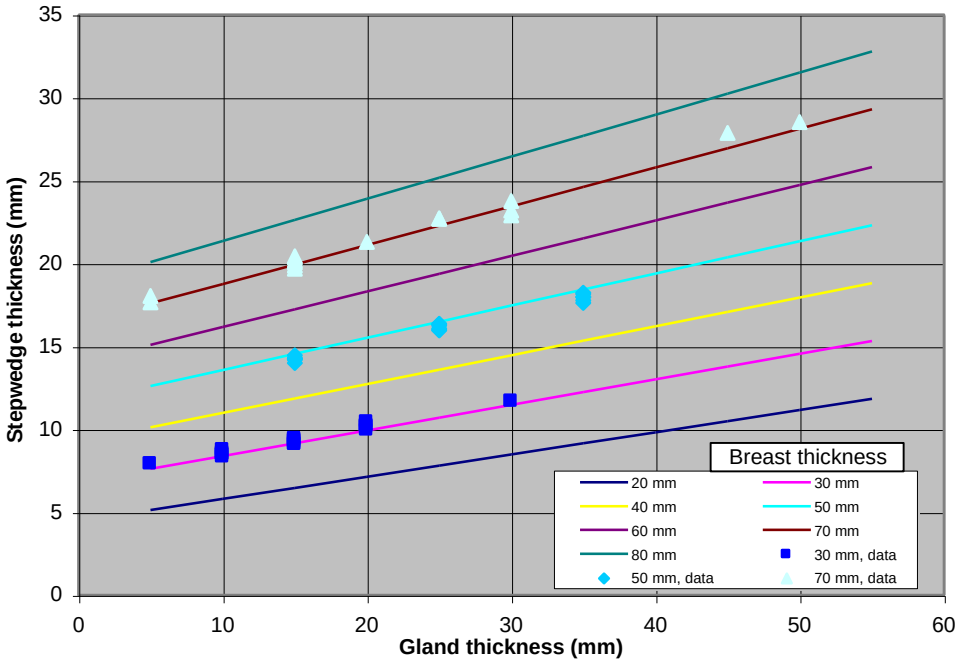


*Figure 5.16: Calibration data (points) and parameterization (lines) for large stepwedge using new parameterization.*

The mammography x-ray unit used in this study selects the tube voltage and target material automatically depending on the degree of attenuation detected in the first few mAs of the exposure. The target material can be either molybdenum or rhodium, and the tube voltage is in the range 27-32 kV. Calibration data were collected separately at each kV value to check for any dependence of the calibration constants on kV. No significant trends were found, so the data at all kV values were combined to calculate the final values of the constants.

A small difference in the calibration data was found between the large and small stepwedge. The equivalent stepwedge thickness of the small stepwedge was consistently slightly larger than the large stepwedge (the difference was typically less than 0.5 mm). The reason for this difference is not known, although it may indicate that the PTFE used for each stepwedge did not have exactly the same x-ray absorption properties. Another possible cause could be the error associated with the anode heel correction (see section 5.4.4), since the calibration films were taken with the large and small stepwedges in different positions on the film. Separate calibration coefficients were calculated for each stepwedge, and they are given in Table 5.1.

| Coefficient | Small stepwedge | Large stepwedge |
|---|---|---|
| $a$ | 0.25 | 0.24 |
| $b$ | -0.66 | -0.33 |
| $d$ | 0.0016 | 0.0020 |
| $e$ | 0.11 | 0.094 |

*Table 5.1: Calibration constants from equation 12, where all thickness are measured in mm.*

### 5.4.3 Compensation for breast edge effects

The method used to measure breast thickness actually measures the separation between the compression paddle and the film cassette. Close to the breast edge the breast loses contact with the compression paddle as the thickness falls off to zero. The possibility of a correction to the breast thickness used in this region has been investigated. The effect of not making a correction is to overestimate the breast thickness, which in turn causes an underestimation of the thickness of dense tissue close to the breast edge.

A method for detecting the region close to the breast edge, called the *breast margin*, where the thickness is less than in the central region, has been described by Byng *et. al.* (Byng 1997). This method uses the general reduction in the grey levels towards the breast edge due to the higher exposure of the film caused by the reduced breast thickness. The aim of Byng *et. al*.'s method was the equalization of grey levels in the breast margin, reducing the dynamic range required for display of the image. A similar method could be used here to identify the width of the breast margin, and to investigate the shape of the breast profile as the thickness falls to zero.

*Location of breast edge and margin*

The first step is to smooth the image to remove local variations in grey level due to dense tissue structures in the breast. This is done using a Gaussian filter with standard deviation 2.5 mm. The remaining grey level variations are caused by global variations in exposure level, due to changes in breast thickness. In the ideal case, the grey level would be constant in the central portion of the breast, where breast occupies the full space between the compression paddle and the film cassette. The grey level then rises smoothly across the breast margin. Figure 5.17 shows a typical image before and after smoothing. A profile is marked for which the grey levels are shown in Figure 5.18.

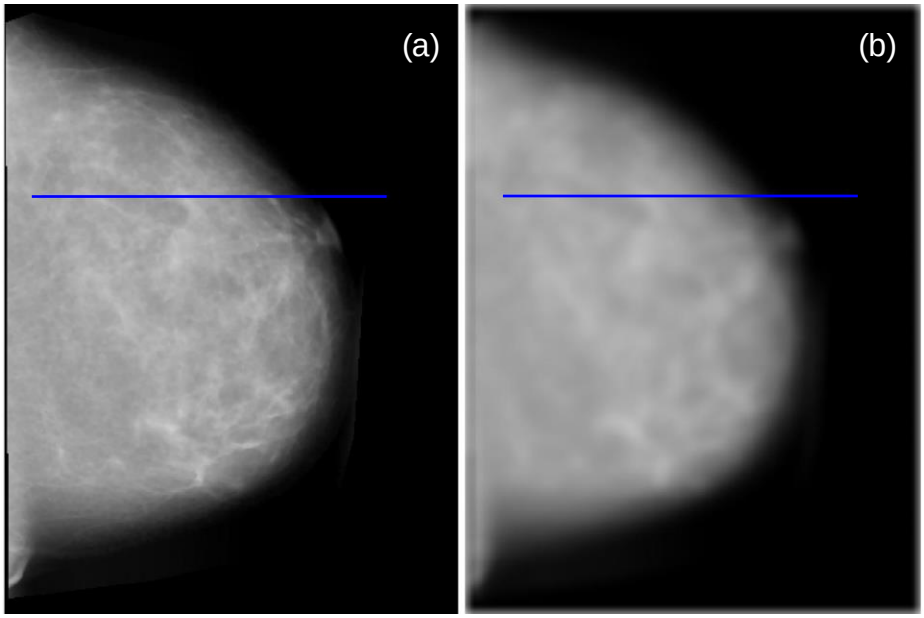


*Figure 5.17: Example image (a) before and (b) after smoothing. A profile is marked in blue which is plotted in Figure 5.18.*

The strategy is to search each line in the smoothed image, looking for the point where the grey level first falls significantly below the background level outside the breast. This point is used to define the breast edge. The search is continued along the same line until the grey level stops falling. This point is used to define the inner edge of the breast margin. This process is repeated for each line in the image, so that the complete breast edge and margin are defined. Typical profiles are illustrated in Figure 5.18, showing the location of the breast edge and margin.

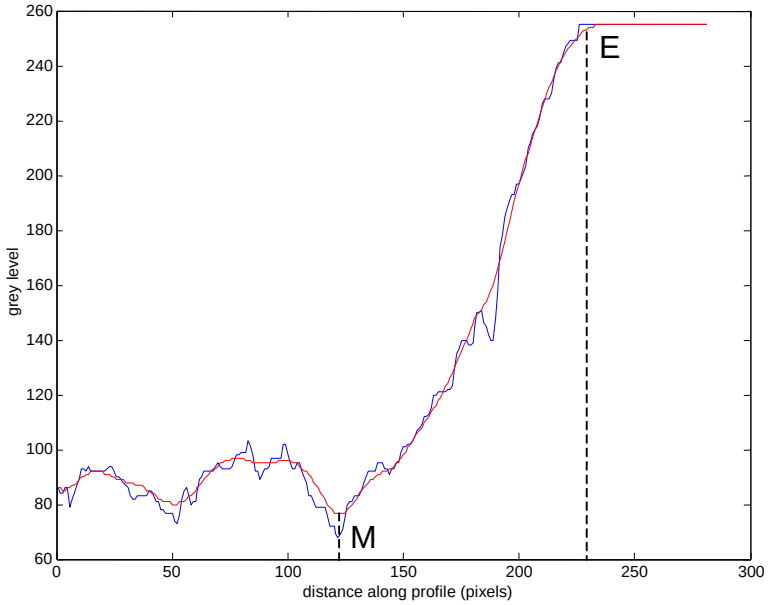


*Figure 5.18: Example profile through breast edge before (blue) and after (red) smoothing. Point E shows the position of the breast edge, and point M marks the inner edge of the breast margin.*

The positions of the breast edge and margin identified by this process tend to be disconnected, as shown in Figure 5.19a. This problem is overcome by smoothing the edge points, and by finding the average width of the breast margin, and using this as a constant width for all rows. Figure 5.19b shows the locations of the breast edge and margin after this smoothing and averaging.

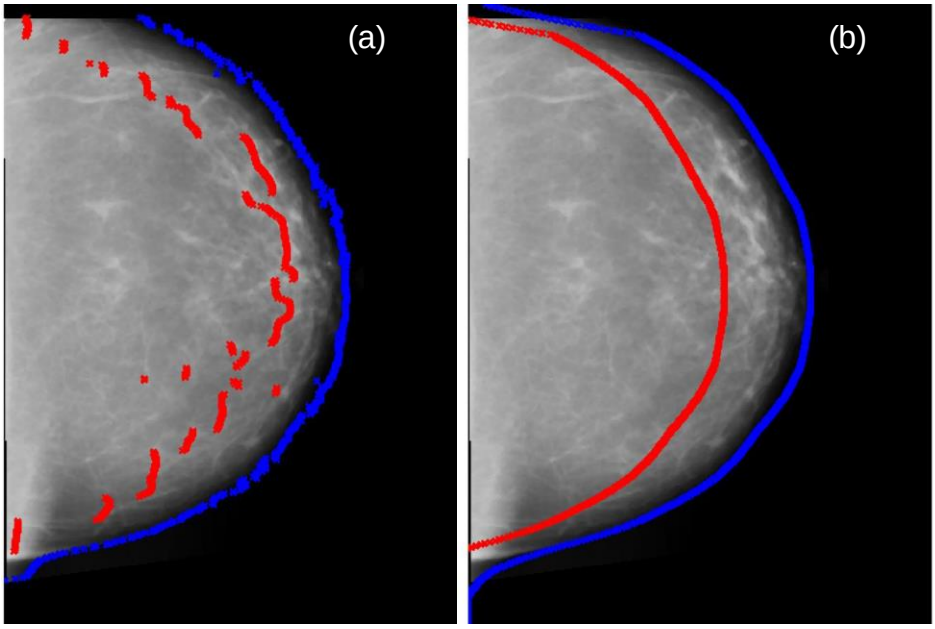


*Figure 5.19: Images showing location of breast edge points (blue) and inner edge of breast margin (red) (a) before and (b) after smoothing.*

*Shape of thickness fall-off*

An extension to the technique described in (Byng 1997) is to use the grey level profiles through the breast margin to model the shape of the breast thickness fall-off in this region. The average grey level at the inner edge of the margin is found (i.e. the grey level at the red points in Figure 5.19a) and all of the profiles through the breast margin are scaled to lie between this value and the average grey level at the breast edge. Next all of the profiles are stretched/squeezed so that their length is the same as the average breast margin width, before being added to find the average grey level profile.

The average grey level profile is converted to a tissue thickness profile by comparison to the image of the stepwedge. A scaling factor, $F$, is applied to the breast thickness inside the margin region, which is calculated as

$$F(x) = \frac{x_{sw}(x)}{x_{sw}(x_i)}, \qquad (13)$$

where $x_{sw}(x)$ is the stepwedge thickness equivalent to the grey level a distance $x$ along the grey level profile, and $x_{sw}(x_i)$ is the stepwedge thickness equivalent to the average grey level at the inner edge of the breast margin. Figure 5.20 shows the average grey level profile, the stepwedge curve, and thickness scale factor profile for the images shown in Figure 5.19. The thickness scale factor falls sharply to zero at the breast edge because there is a minimum thickness that can be detected above the background. This is because the film is highly exposed at the very edge of the breast, resulting in optical densities in the flat region of the characteristic curve.

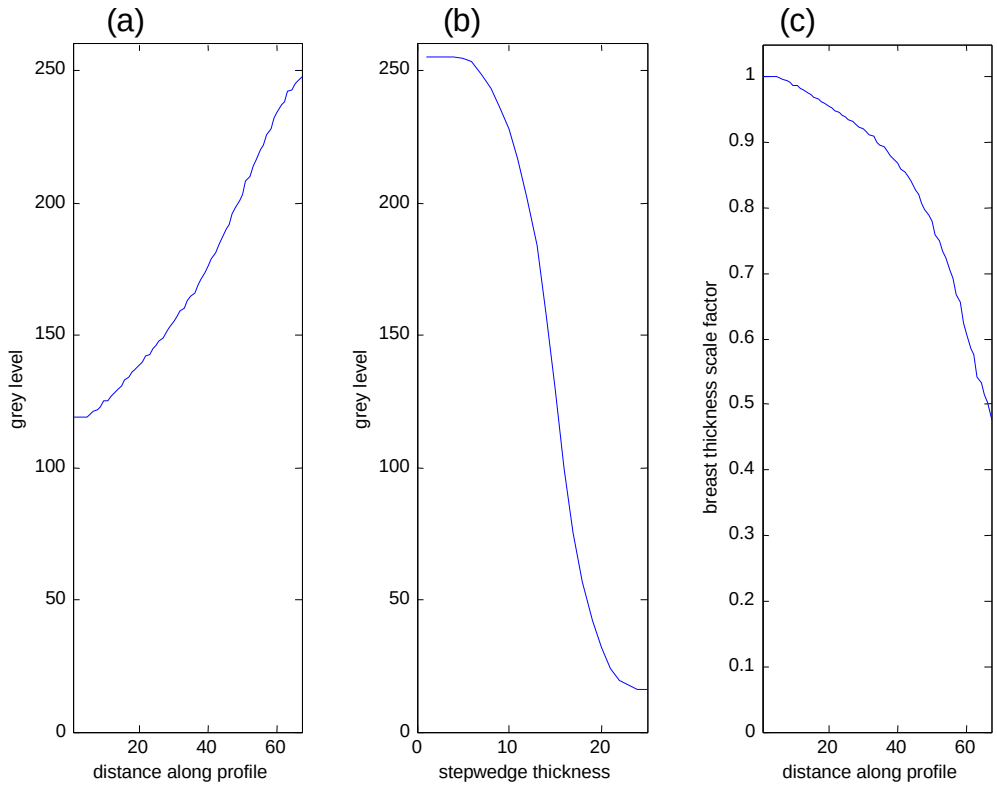


*Figure 5.20: (a) Average grey level profile through breast margin (b) Stepwedge curve (c) Breast thickness scale factor profile through breast margin.*

Figure 5.21 shows the final breast thickness template calculated for the image in Figure 5.19, along with a profile through the breast region. This shows the breast thickness decreasing linearly due to the tilt of the

compression paddle, and then falling off faster at the edge, where the breast loses contact with the compression paddle.

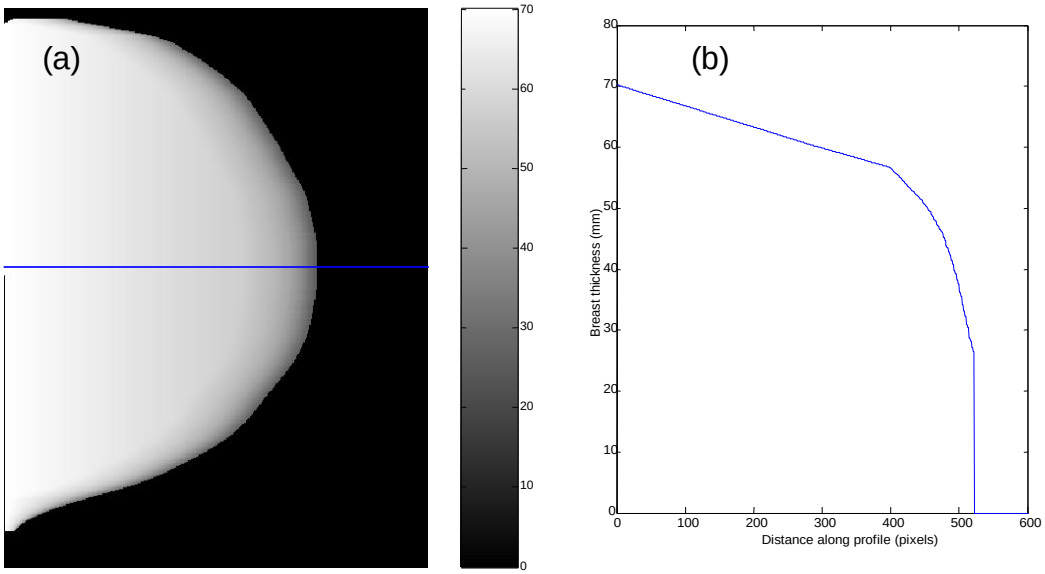


*Figure 5.21: (a) Breast thickness template (b) Breast thickness profile.*

*Problems with breast edge correction*

The breast edge correction described in this section presents several problems when put into practice. The major problem is that using this correction causes significant thicknesses of dense tissue to be measured very close to the breast edge. It is expected that the tissue should be mostly fatty at the edge of the breast, so this is likely to indicate a problem with the method. The most likely cause is that the thickness correction is too large causing the breast thickness to be underestimated. This would cause an overestimation of the dense tissue thickness.

Another problem with the correction is that using a constant margin width in the horizontal direction causes the true margin width to be too small in places where the breast edge is aligned along the horizontal direction. This effect is clearly seen at the top and bottom of Figure 5.19b. This problem could be solved by using a radial, breast cantered coordinate system. However this

would increase both the complexity of the problem and the processing time required.

It is also not clear that this method is correctly identifying the breast margin. Many of the mammograms analyzed have a high proportion of dense tissue, and the smoothing applied to such images is insufficient to eliminate the variations in intensity due to dense tissue structure. The inner edge of the margin is defined by the point where the grey level stops decreasing. Often this occurs at the point where there is most dense tissue. It is difficult to separate the effects of tissue structure from the effect of decreasing breast thickness.

Another problem is that, even in the absence of any dense tissue structure, the grey level fall off is not only affected by the breast thickness. Other factors such as scattered radiation, extra-focal radiation also become important at the breast edge.

The best way to deal with the problems described above is not clear, and the overestimation of dense tissue at the breast edge indicates that their effect is not negligible. Since only a small amount of dense tissue is expected close to the breast edge, it is concluded that the potential errors introduced by the use of the breast edge correction are larger than the error from not correcting for breast edge effects. Hence the breast edge correction described in this section is not used in the analysis of mammograms described in chapter 6.

Other computer assisted measurements of breast density, such as those described in (Highnam 1998) and (Yaffe 1998), do not use a correction to the breast thickness at the breast edge.

### 5.4.4 Assessment of errors associated with breast density measurement

The error on the thickness of dense tissue measurement depends on the uncertainties in the breast thickness measurement and the equivalent stepwedge thickness measurement.

The error on the measurement of breast thickness arises due to uncertainty in the location of the magnification marker centres, and also due to spatial non-uniformity of the film digitizer. The first of these sources is most important, since the second is expected to be negligible. The error associated with the magnification marker locations is assessed by looking at the deviations of the marker positions from the best-fit straight line in the calibration data (see, for example, Figure 5.3). This gives 24 separate measurements of the marker location uncertainty from which the standard deviation is found to be 5.0 pixels. This corresponds to an error in the breast thickness measurement of 0.6 mm. This error is slightly smaller than the 0.9 mm quoted in (Smith 1998) for the same method. This may be because each marker pair is calibrated separately here, whereas Smith used the same calibration for all marker pairs with each pair having the same separation. This introduces additional errors due to slight differences in the separations of the marker pairs, and differing degrees of magnification at each pair due to system geometry (each pair is at a different angle to vertical).

The most significant error in the measurement of equivalent stepwedge thickness is caused by the approximate nature of the anode heel/geometric correction (described in section 5.3.1). Figure 5.22 shows the pixel value against stepwedge thickness curves before and after correction for four films of the same exposure, but with the stepwedge in different positions in the image. If there was no effect due to location in the image then the four stepwedge curves would be identical. The differences before correction are due to the anode heel effect and inverse square law effects.

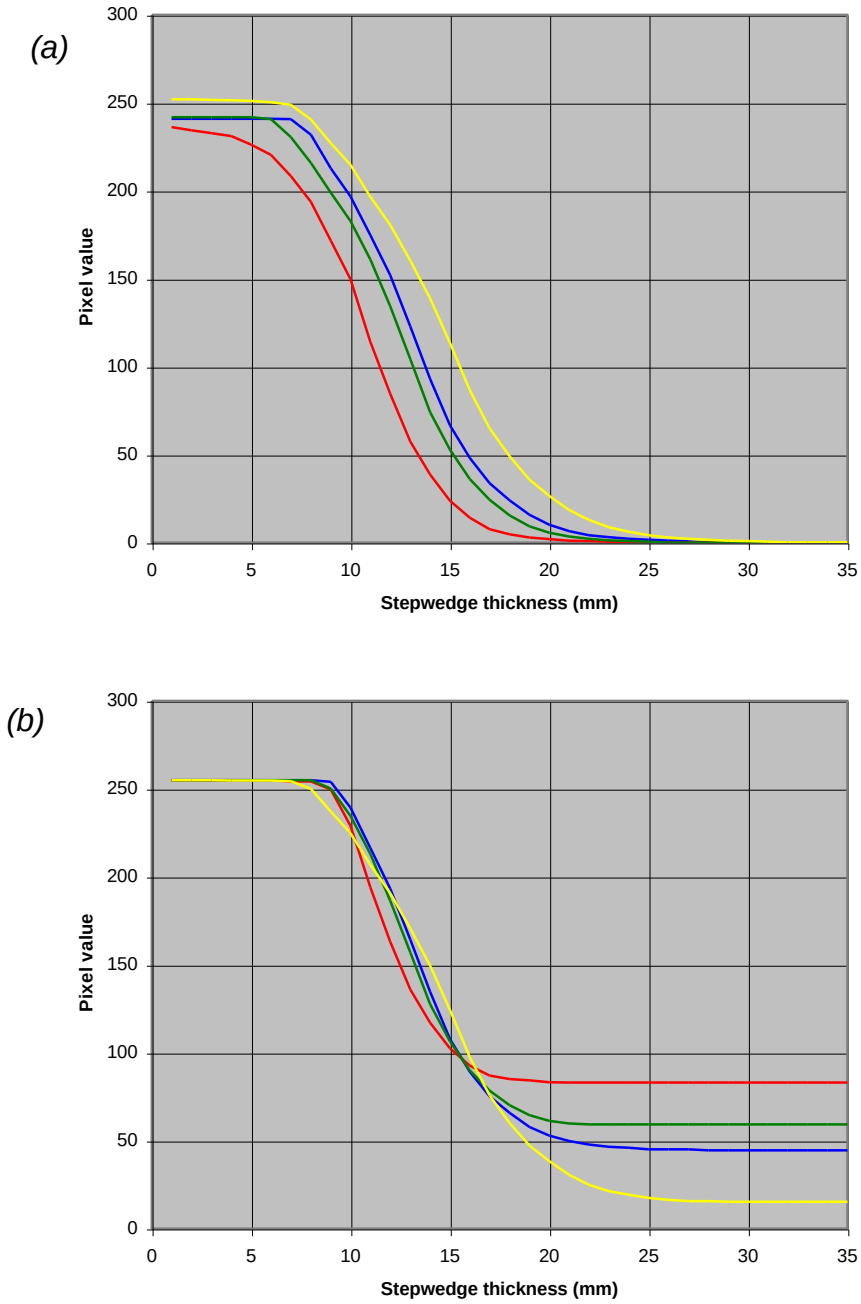


*Figure 5.22: Stepwedge curves (a) before and (b) after correction for four films with the same exposure but with stepwedge in different positions.*

The correction is made by adding an amount to the pixel value at each location in the image. At the edge of the large format films this amount is as much as 90 grey levels. The correction is only correct for the linear portion of the film's characteristic curve, and is least useful for areas of the film in the flat regions of the curve at high and low exposure. This is seen in Figure

5.22b where the correction has failed to bring the curves together at large stepwedge thicknesses. The stepwedge cannot be used to calibrate grey levels that are less than the amount added in the correction. This ranges from 10 to 90 depending on the position of the stepwedge on the film. This effect reduces the usable thickness of the stepwedge, meaning that areas of very dense tissue cannot be accurately calibrated, even though the stepwedge is not saturated by the high exposure.

Apart from this problem of reduced usable stepwedge thickness, it is also clear that the correction is not perfectly accurate for the thinner parts of the stepwedge. The curves are brought into fair agreement after correction, but the differences are still significant. The figure shows that for a grey level of 150 there is a difference of up to 1.5 mm in the corresponding stepwedge thickness, depending on the position of the stepwedge in the image. The average deviation of the corrected stepwedge curves from their mean value is 0.4 mm. This is taken as an estimate of the uncertainty on the stepwedge thickness.

The thickness of dense tissue is calculated from the breast thickness and the stepwedge thickness using equation 12. Errors can be propagated through this equation by the standard relationship:

$$\sigma_{x_g}^2 = \sigma_{x_{sw}}^2 \left( \frac{\partial x_g}{\partial x_{sw}} \right)^2 + \sigma_{x_b}^2 \left( \frac{\partial x_g}{\partial x_b} \right)^2 . \qquad (14)$$

By performing the differentials in equation 14 and using the calibration constants from Table 5.1, it is found that the error on the thickness of glandular (dense) tissue, for an intermediate breast thickness of 50 mm, is

$$\sigma_{x_g}^2 = 27\sigma_{x_{sw}}^2 + 1.6\sigma_{x_b}^2 . \qquad (15)$$

Clearly the measurement of thickness of dense tissue very sensitive to errors in the stepwedge thickness. Substituting in the estimated values of the errors on stepwedge thickness and breast thickness we find that the estimated error on the thickness of dense tissue is 2.2 mm.

The error mentioned above is a systematic error, which would affect all of the measurements on a particular image in the same direction. There is also a random error affecting the measurements at each pixel. This is caused by the random variation in the pixel values on the digitized films. Using the calibration films it is found that this variation causes the glandular tissue thickness measurements to vary randomly with standard deviation 0.9 mm.

# 6 Clinical application

The method for measuring breast density described in chapter 5 has been applied to a set of clinical mammograms, which are part of a study into the effects of diet and exercise on breast cancer risk. The objective of the study is to determine whether the risk of breast cancer can be reduced by changes in diet and exercise. The study follows a group of women who have a family history of breast cancer, and are thus considered to be at high risk. The women have had a number of investigations, including x-ray mammography, to assess various risk factors associated with breast cancer. The tests will be repeated after one year on a prescribed diet and exercise regime. The clinicians require an objective measurement of changes in mammographic density over this period. The study is not yet complete, so only the mammograms taken at the beginning of the study period have been analysed so far.

## *6.1 Statistics of the study sample*

The study sample consists of full 4-view mammograms for 41 cases, which were taken with the stepwedge and magnification markers in place. This gives a total of 164 films on which the method has been tested. A further 10 cases were involved in the study, but these mammograms were taken without the stepwedge and magnification markers in place, due to a misunderstanding at the breast clinic. These 10 cases could not be processed using the stepwedge technique, and are not considered further here. The average age of the women in the sample is 41.4 years and the distribution is shown in Figure 6.1. This is a younger sample of women than those who usually undergo screening mammography, because the study sample is taken from a clinic for women with family history of breast cancer.

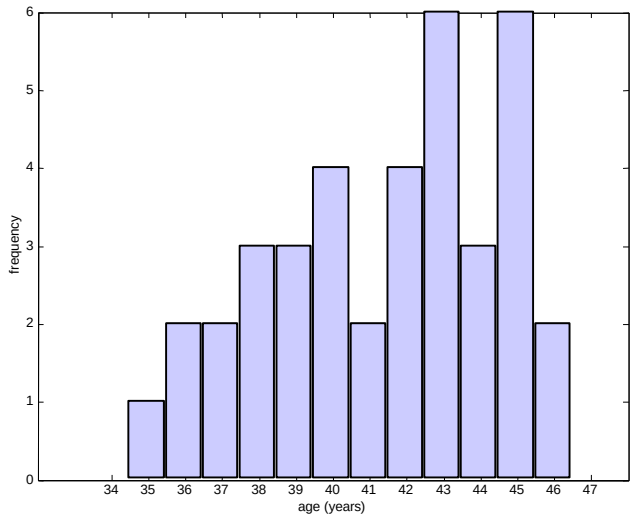


*Figure 6.1: Ages of women in study sample.*

Figure 6.2 shows the range of exposures used for the mammograms in the sample. It also divides the data into mammograms which were taken with the larger, 35 step stepwedge, and those which were taken with the 25 step stepwedge. The radiographers were instructed to use the large stepwedge when using large format (24x30 cm) mammography film, and to use the small stepwedge with standard size film. This was based on the assumption that higher exposures would be used for larger breasts which required the large film. The principal factor in selection of stepwedge is the exposure.

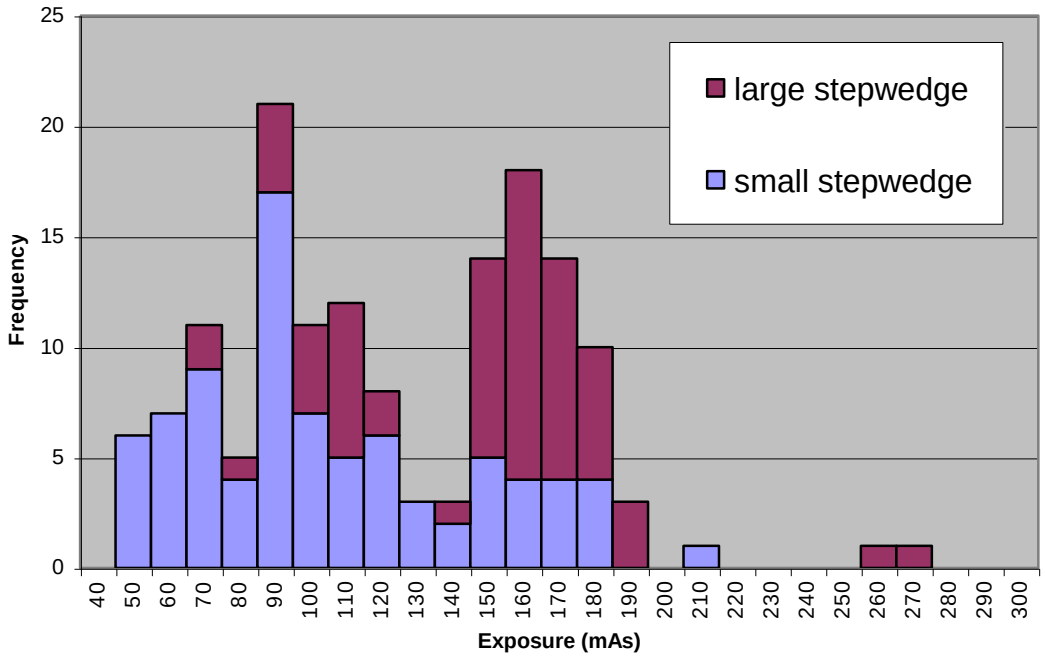


*Figure 6.2: Exposures of mammograms in study sample.*

Figure 6.2 shows that, while there is a general tendency for the mammograms taken on the standard size film to have lower exposure, there is not a clear division. Many mammograms taken using the large size film have relatively low exposures, and a significant number of those taken with standard size film have relatively high exposures. The films with low exposure taken using the large stepwedge are not a problem, since the large stepwedge includes the same range of low stepwedge thickness as the small stepwedge. Mammograms having high exposure taken using the small stepwedge are a problem, though, because the maximum stepwedge thickness is not sufficient to calibrate images of such highly dense breasts.

This evidence shows that it is not sufficient to decide which stepwedge to use based purely on the film size chosen by the radiographer. It is always sufficient to use the large stepwedge for large films, but sometimes the large stepwedge is required for the small films. There are two problems with this. The first is that there is often not enough room for the large stepwedge on the small format film. This problem could be solved by making a new

stepwedge without the lower thickness steps, say omitting the bottom ten steps. This would reduce the dimensions of the wedge, allowing it to fit onto the smaller film. High exposures that require the thicker end of the wedge do not make use of the lower steps, which are not visible in the image.

The second problem is how to decide when to use the larger wedge with the standard size films. One possibility is to decide based on the breast thickness as measured by the mammography unit (a measurement which is available to the radiographer at the time of the examination). The reliability of this measurement is discussed later in section 6.2. Figure 6.3 shows the range of breast thicknesses found in the study sample, as measured by the mammography unit.

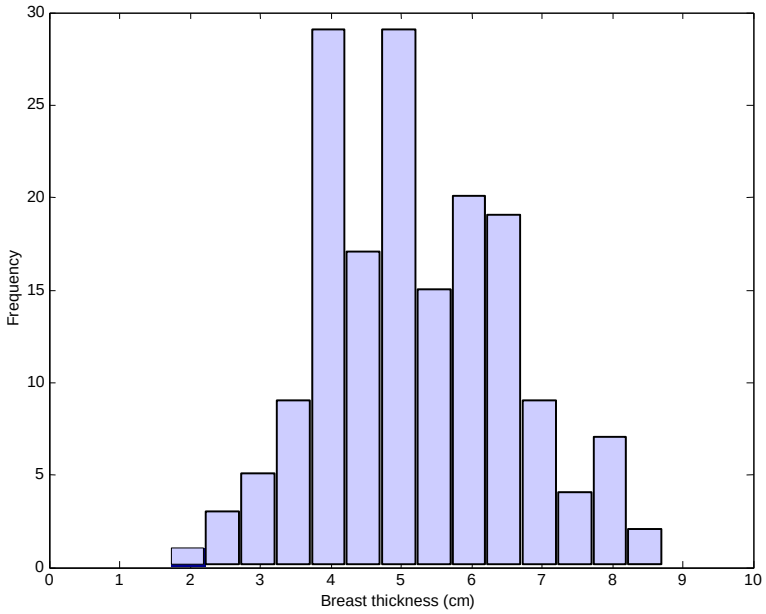


*Figure 6.3: Range of breast thicknesses in study sample.*

Figure 6.4 shows the correlation between breast thickness and exposure. In general thicker breasts lead to higher exposure. Films with exposure greater than 130 mAs require the larger stepwedge. The correct stepwedge can be chosen by using the large stepwedge for breast thicknesses larger than 5.5 cm. This leads to only 9 films where the small stepwedge is used with exposure greater than 130 mAs (these are seen in the upper left quadrant of

Figure 6.4). This is an improvement over using the film size to select the stepwedge, which results in 20 films with exposure greater than 130 mAs using the small stepwedge. However, this would require the women to undergo compression for longer, which might not be acceptable.

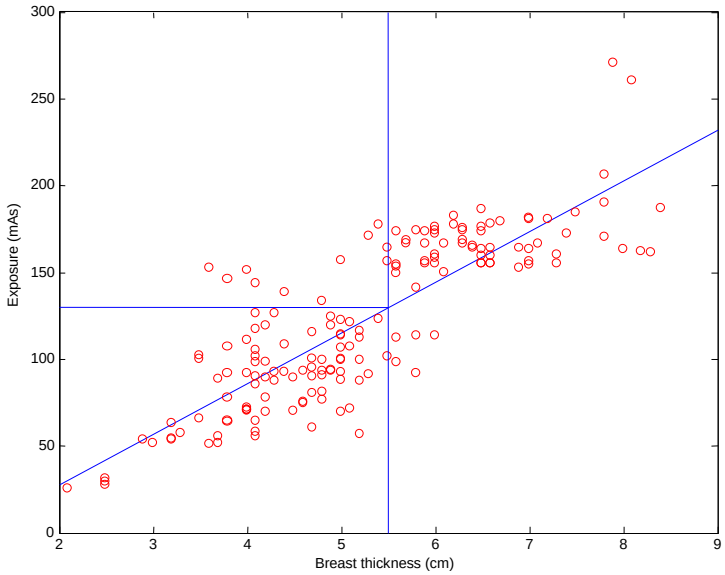


*Figure 6.4: Breast thickness versus exposure for the study sample, showing best fit line.*

## *6.2 Breast thickness measurements*

The set of clinical mammograms was used to assess the performance of the breast thickness measurement provided by the mammography unit. The unit provides a single measurement of the breast thickness along with a measurement of the compression force applied to the breast. Measurements made using the magnification marker method described in section 5.2 show that the compression paddle is able to tilt significantly in the chest-wall to nipple direction, leading to a variation of thickness across the breast. The height of the compression paddle differs by up to 2 cm across the width of the film. This tilting was previously measured (Smith 1998), although the size of the effect is found to be larger with the mammography unit used here (LORAD).

The tilt of the compression paddle means that the single thickness measurement provided by the mammography unit is inadequate. It was found that the thickness measurement coincided best with the magnification marker measurement at a point on the film approximately 7 cm from the chest wall edge of the film for the small films and approximately 11 cm from the chest wall edge for the large films. Figure 6.5 shows the difference between the two measurement methods at this point.

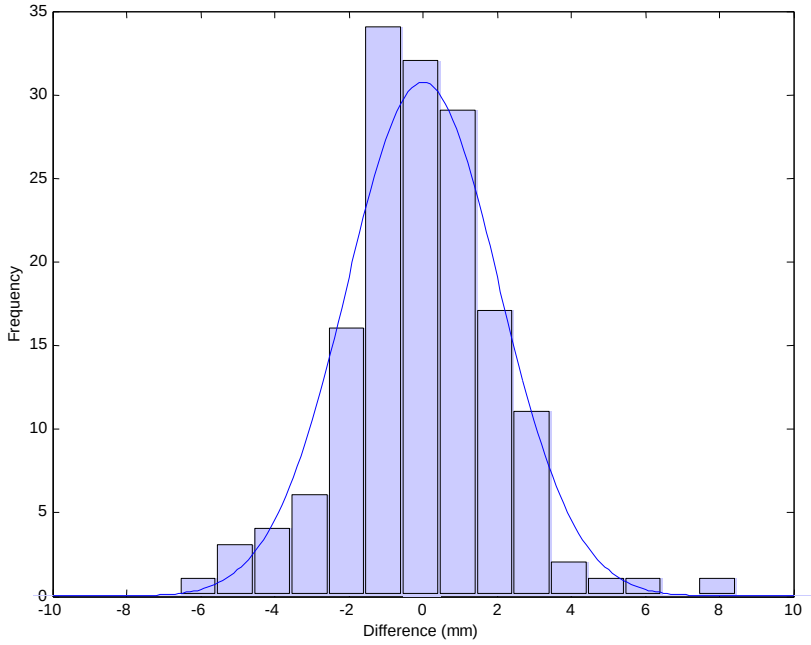


*Figure 6.5: Difference between breast thickness measurement techniques (histogram), and Gaussian curve with mean 0.0 and standard deviation 2.0 mm (line).*

The standard deviation of the difference between the two measurement methods is 2.0 mm. If the measurements are separated into two groups according to film size, then the standard deviation is found to be slightly bigger for the large films (2.3 mm compared to 1.9 mm for the small films). This is significantly larger that the 0.6 mm error expected to be associated with the magnification marker method. Hence the measurement made by the mammography unit is less accurate, in addition to not taking into account the compression paddle tilt.

It was hypothesized that the angle of tilt of the compression paddle may be related to the compression force applied. This idea is supported by the data shown in Figure 6.6, which shows the tilt of the compression paddle plotted against the compression force applied. The tilt increases for larger compression force.

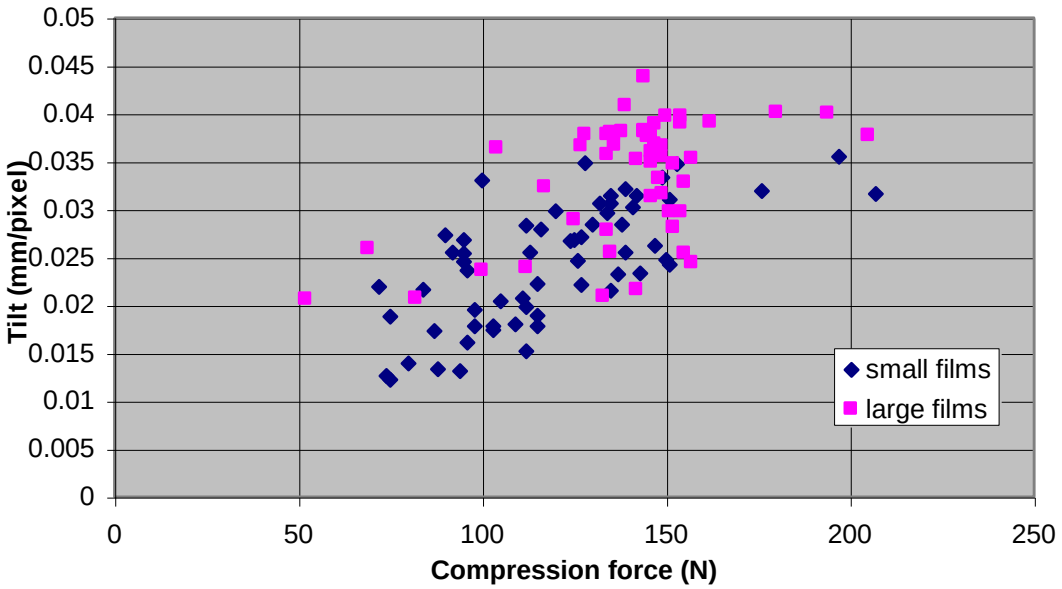


*Figure 6.6: Compression paddle tilt as a function of compression force applied.*

A line could be fitted to the data in Figure 6.6 allowing an estimate of compression paddle tilt to be made based on the compression force reported by the mammography unit. This could be used in conjunction with the unit's thickness measurement to estimate the breast thickness across the film in cases where the magnification markers cannot be used. This would however lead to lower accuracy of the thickness measurement. There were two films in the study sample where no pairs of the magnification markers are sufficiently visible to allow a thickness measurement.

## *6.3 Measurements of breast density*

There are several ways to present the measurements of breast density made. One can express the percentage of the breast area with thickness of dense tissue greater than zero. This measurement corresponds most closely to the Boyd scheme of breast density assessment. Quoting the percentage of dense tissue by area in this way leads to rather high percentages. This is because there are small amounts of dense tissue present throughout the breast, which are not easily distinguishable by eye. Also, the random errors on the thickness of dense tissue measurement cause areas with no dense tissue to appear as a cluster of values around zero. About half of this area will be labelled as containing some dense tissue.

Problems with the measurement of area of dense tissue can be overcome by measuring the area with a significant amount of dense tissue. In this case, the thickness of dense tissue is required to be significantly greater than zero. The significance is judged relative to the expected error on the dense tissue thickness. A lower threshold of 5 mm of dense tissue is used here, corresponding to approximately two standard deviations greater than zero.

Area measurements do not use all of the information available, and measurements of the volume of dense tissue give a more complete picture of the amount of dense tissue. The volume of dense tissue can be expressed either as a percentage of the breast volume or as an absolute measurement.

### 6.3.1 Performance of breast density measurements

The ideal way to test the performance of the breast density measurements would be to compare the results to the true values. Unfortunately it is not possible to find the true amounts of dense tissue in each breast without dissecting the breast tissue, and even this would be difficult. It is clearly not acceptable to do this. An alternative to this would be to compare with the

results of magnetic resonance imaging (MRI). MRI is non invasive and is able to distinguish between breast tissue types with high accuracy (a study by Lee *et. al* (Lee 1997) found that volume of dense tissue could be determined using MRI with an accuracy of 2%). However MRI images of the breast are not available for the present study.

The only practical options are to compare measurements from two views of the same breast, and to compare the results to the density assessments of a trained radiologist.

*Comparison of two measurements from same breast*

Figure 6.7 to Figure 6.10 below show the difference between measurements made from the medio-lateral oblique and cranio-caudal view mammograms of the same breast, for each density indicator. The MLO and CC views are not necessarily expected to give exactly the same results because there may not be exactly the same amount of breast in view, and the different viewpoints will cause changes in the area projection measurements. While it is expected that the differences should be reasonably small, the standard deviations found here should be considered as upper limits on the measurement error.

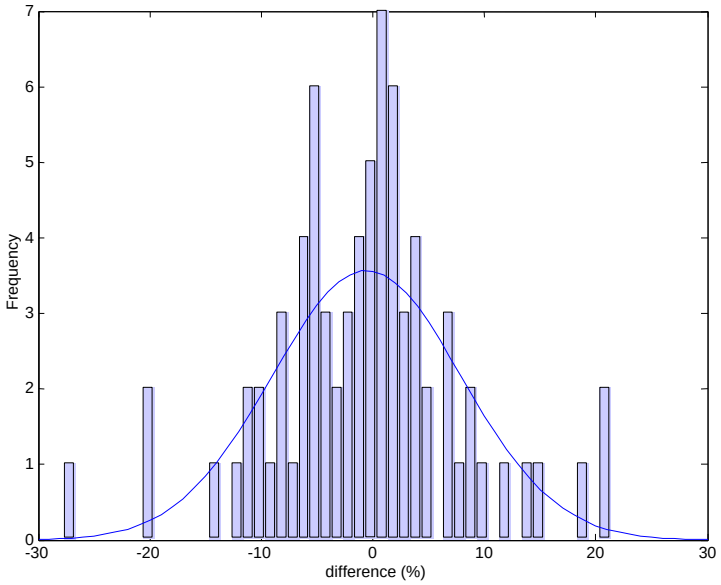


*Figure 6.7: Difference between measurements of percentage of area with thickness of dense tissue greater than zero, taken from two different views of the same breast (histogram), and fitted Gaussian curve (line).*

The difference between measurements of percentage of area with thickness of dense tissue greater than zero has a mean of -0.5 % and a standard deviation of 8.5 % (shown in Figure 6.7). There is no significant difference between the standard deviations for small and large films (8.5 % for the small films compared to 7.5 % for the large films).

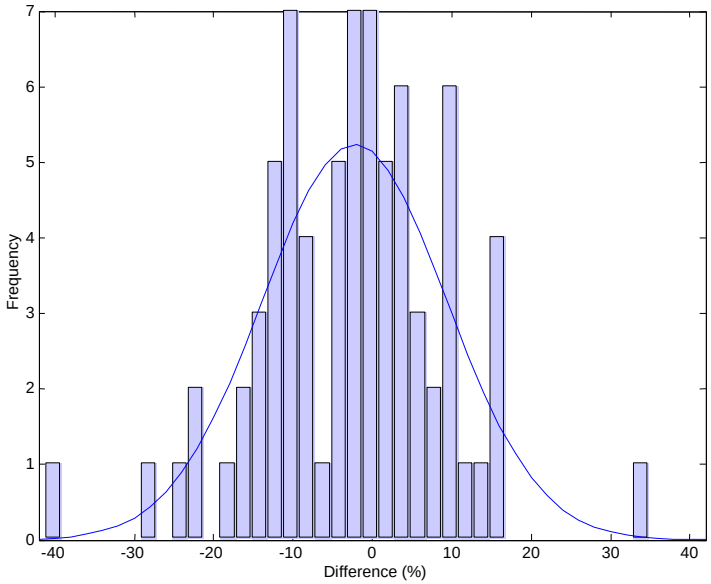


*Figure 6.8: Differences between measurements of percentage of area with thickness of dense tissue greater than 5 mm, taken from two different views of the same breast (histogram), and fitted Gaussian curve (line).*

The difference between measurements of percentage of area with thickness of dense tissue greater than 5 mm has a mean of 2.2 % and a standard deviation of 11.6 % (shown in Figure 6.8). For this measure the standard deviation is smaller for measurements made from small films (7.6 % compared to 14.9 % for large films).

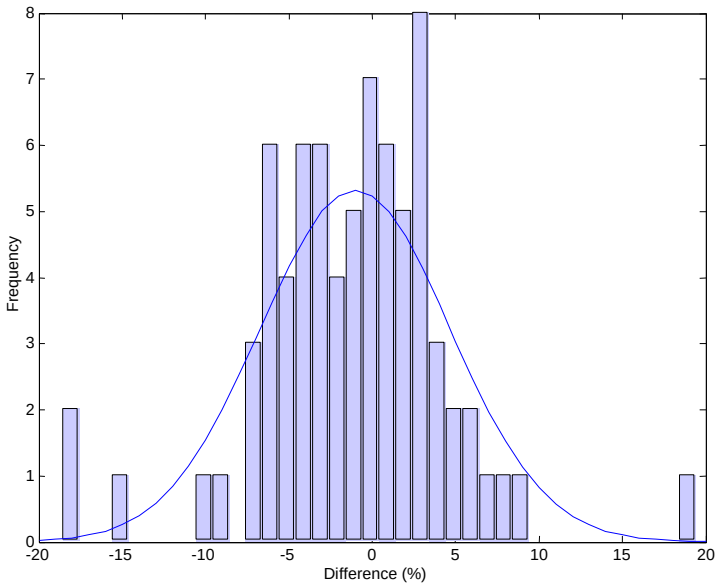


*Figure 6.9: Differences between measurements of percentage of dense tissue by volume, from two different views of the same breast (histogram), and fitted Gaussian curve (line).*

The difference between measurements of percentage of dense tissue by volume has a mean of -1.0 % and a standard deviation of 5.7 % (shown in Figure 6.9). For this measure there is little difference in the resolution between the small and large films (4.8 % for the small films and 5.8 % for the large films).

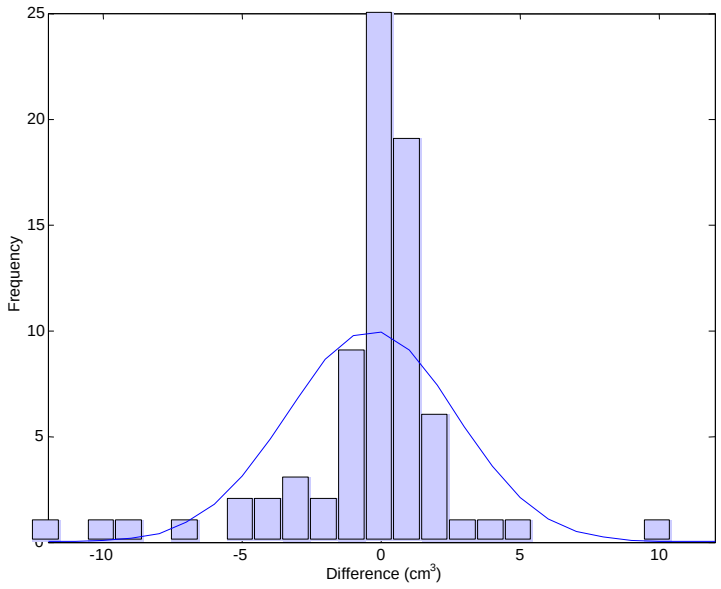


*Figure 6.10: Differences in measurements of volume of dense tissue, made from two different views of the same breast (histogram), and fitted Gaussian curve (line).*

The difference between measurements of volume of dense tissue has a mean of -0.4 $cm^3$ and a standard deviation of 3.0 $cm^3$ (shown in Figure 6.10). The standard deviation of the measurements on small films is found to be significantly smaller than the large films (1.1 $cm^3$ compared to 4.1 $cm^3$). This may be because this is an absolute measurement of the amount of dense tissue (as opposed to a fraction), so the values measured from large films tend to be bigger. The data in Figure 6.10 are not a good fit to the Gaussian curve. This is probably because the true distribution is made up of two Gaussians, one with much smaller standard deviation than the other, corresponding to the two film sizes.

The measurements of breast density can also be used to classify the mammograms according to Boyd's six category scheme. The measurement of breast area with greater than 5 mm of dense tissue was used, because the area with dense tissue thickness greater than zero would result in almost all images classified into the highest density group. Figure 6.11 shows the distribution of the mammograms in the study sample between the six Boyd

categories. This distribution is compared to the classification of the same mammograms by an expert radiologist observer below.

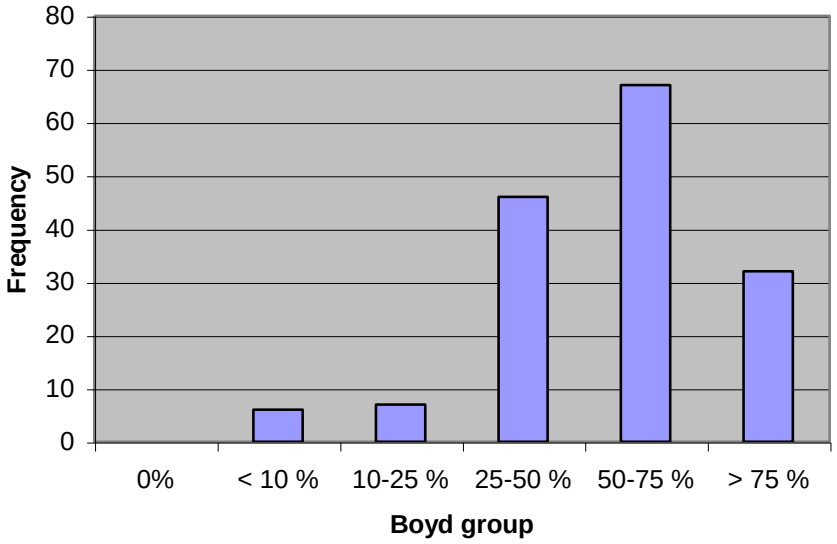


*Figure 6.11: Boyd groups of study mammograms as measured by stepwedge method.*

*Comparison to radiologist's assessment*

In comparing breast density measurements made using the stepwedge method to radiologist's assessment it is important to consider the subjective nature of this method. Two different radiologists may not agree exactly on the percentage of dense tissue (inter-observer agreement), and the same radiologist may not estimate the same percentage on two assessments of the same image (intra-observer agreement). These measures of reliability were previously investigated in (Smith 1998), who found that inter-observer and intra-observer agreement were 91 % and 89 % respectively for classification of mammograms according to the Boyd scheme. This means that different observers, or the same observer on two separate occasions, are expected to disagree about the classification around 10 % of the time. In the cases of disagreement, the differences were only between adjacent percentage bands.

Figure 6.12 shows the Boyd groups of the study set as classified by an expert radiologist observer. The radiologist's assessments tend to be lower than the automated measurements shown in Figure 6.11. This agrees with the previous findings using this and other automatic density measurement techniques, reported in (Smith 1998) and (Zhou 2001).

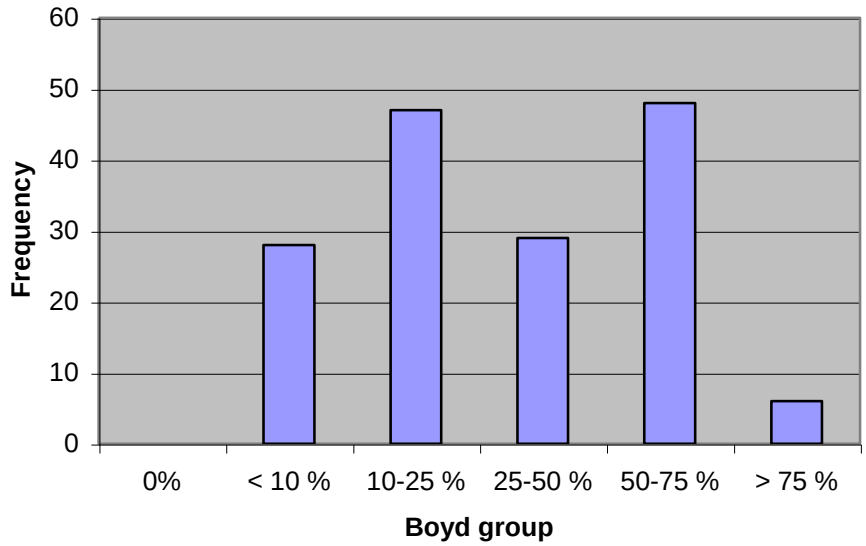


*Figure 6.12: Boyd groups of study mammograms as classified by expert observer.*

Figure 6.13 to Figure 6.14 show the correlation between the various automatic breast density measurements and the Boyd group as classified by an expert observer (the Boyd groups are referred to by number as follows: (1) <10 % (2) 10-25 % (3) 25-50 % (4) 50-75 % (6) >75 %). It should be noted that reproducing the Boyd classification is not an end in itself, but it is a useful guide since Boyd groups have been shown to be correlated with breast cancer risk. The eventual aim should be to establish the relation between the new measures and risk.

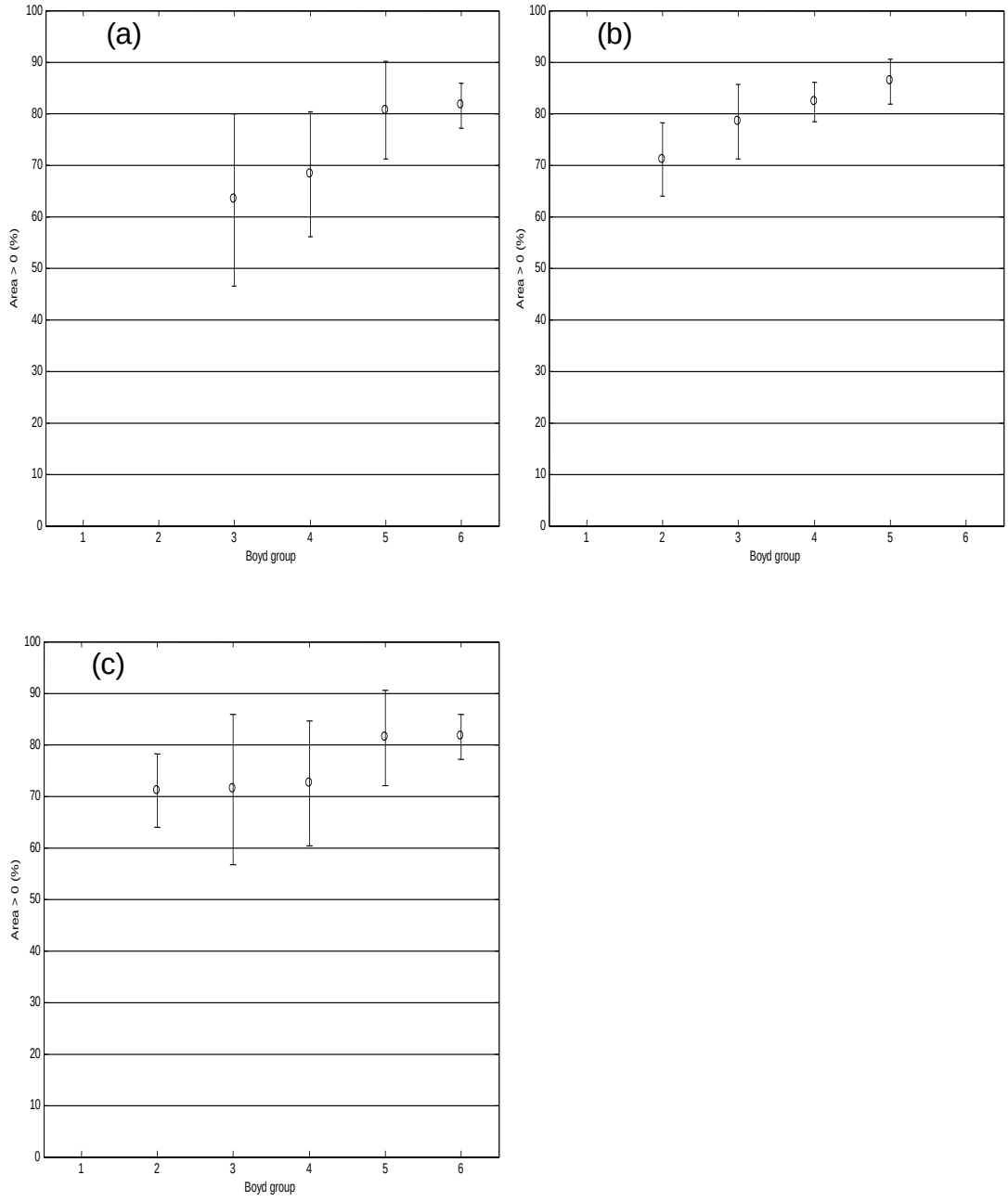


*Figure 6.13: Mean and standard deviation of percentage by area with dense tissue thickness greater than zero against Boyd group (a) small films (b) large films (c) all films.*

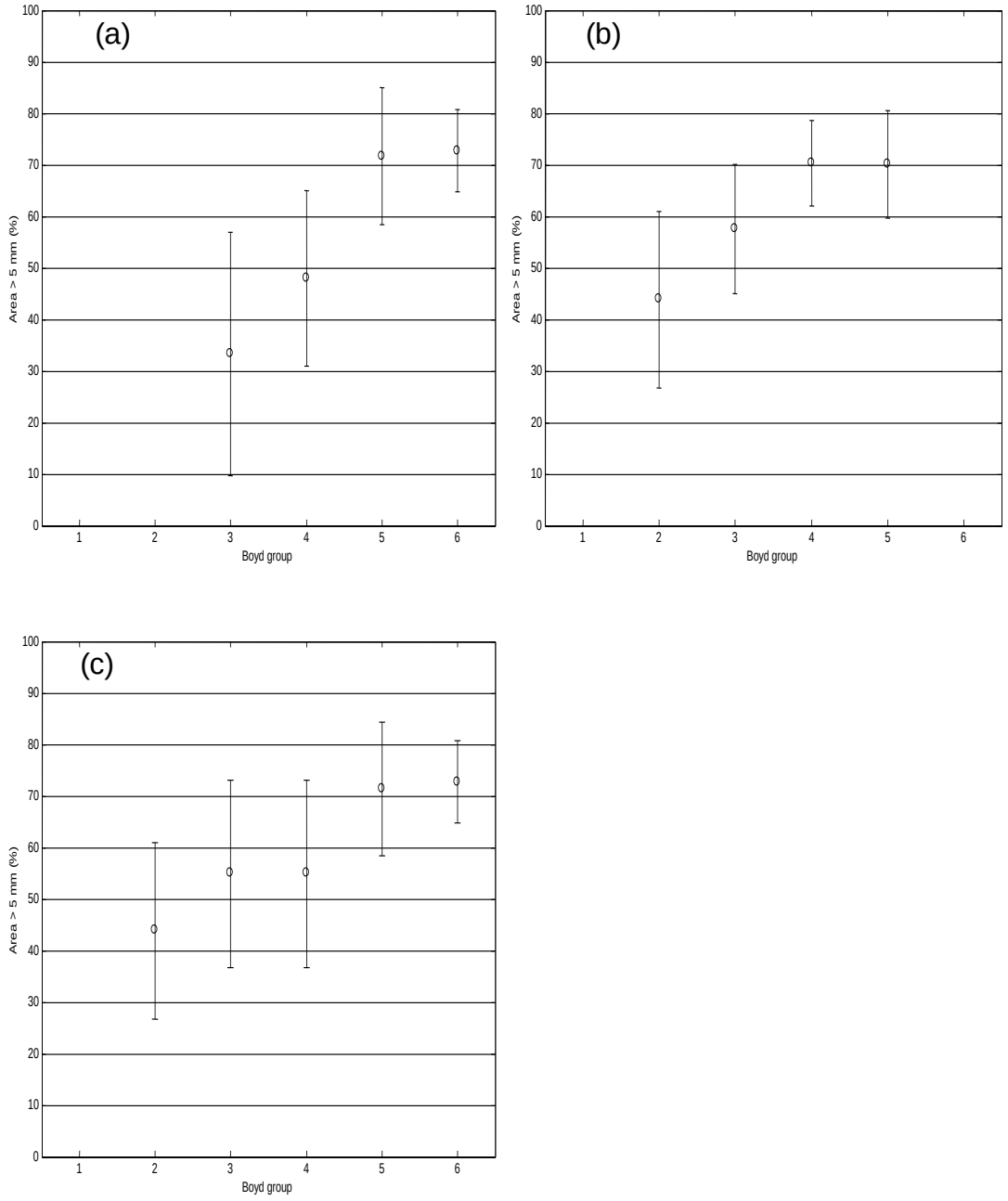


*Figure 6.14: Mean and standard deviation of percentage area with dense tissue thickness greater than 5 mm against Boyd group (a) small films (b) large films (c) all films.*

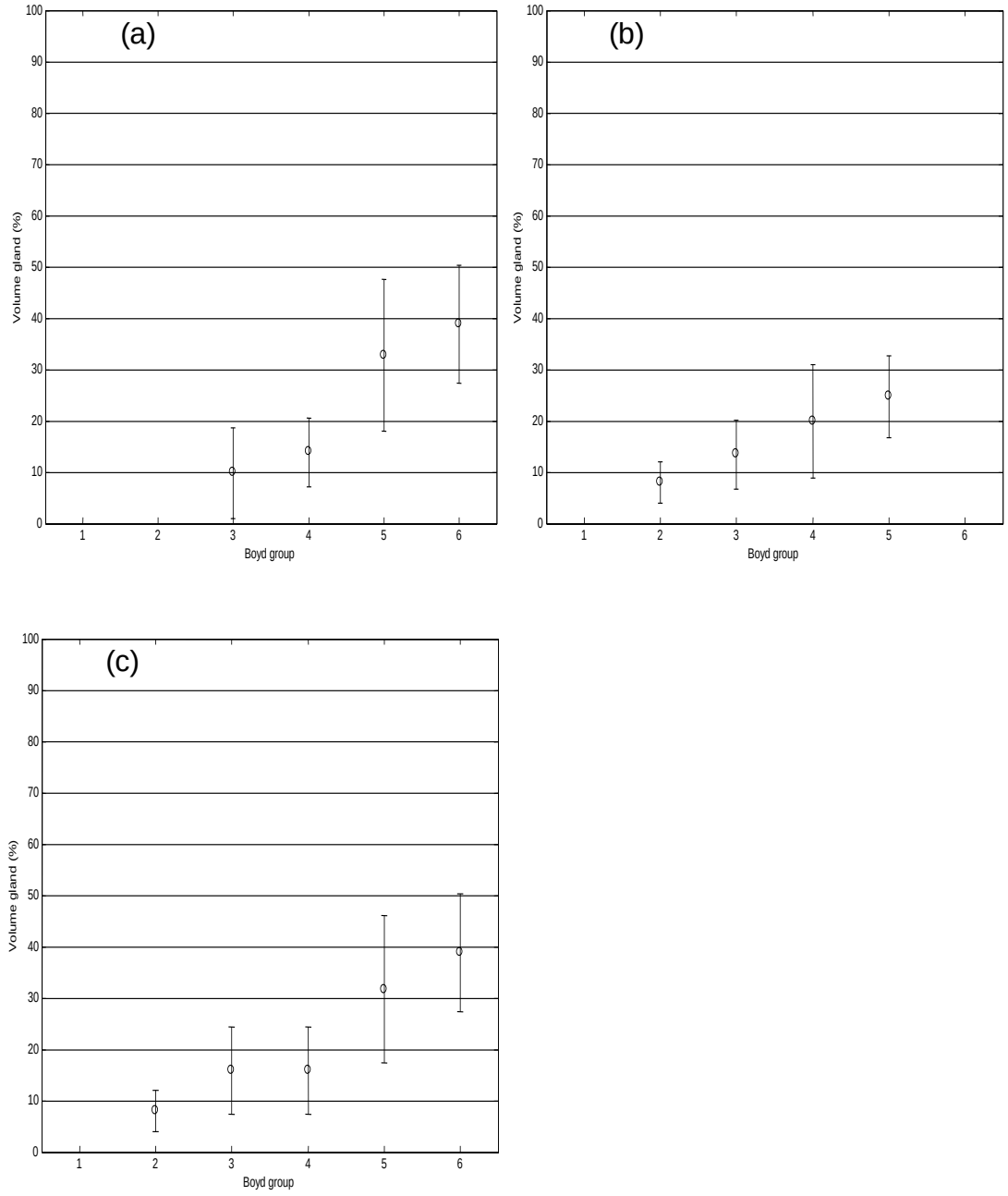


*Figure 6.15: Mean and standard deviation of percentage of glandular tissue by volume against Boyd group (a) small films (b) large films (c) all films.*

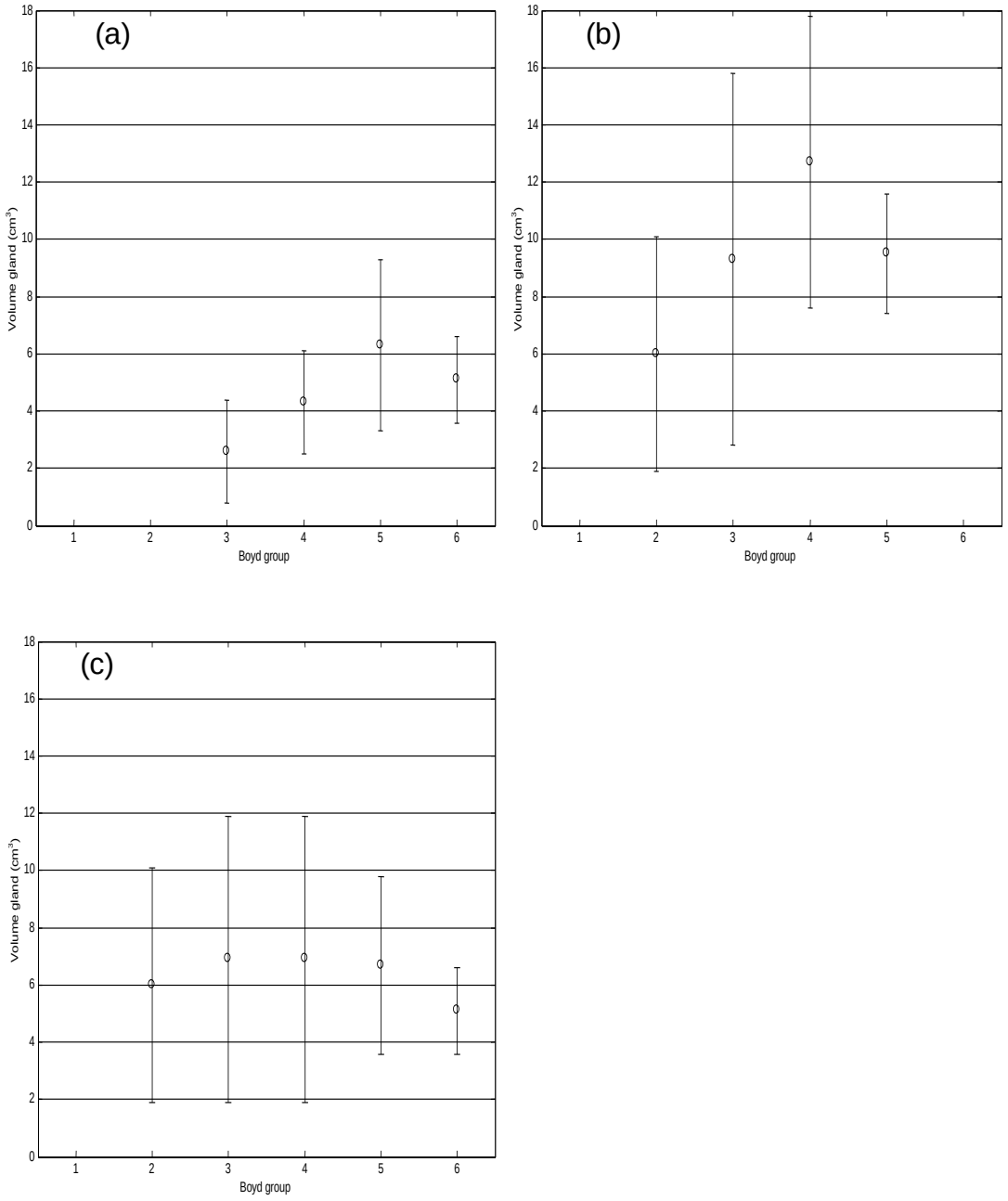


*Figure 6.16: Mean and standard deviation of volume of glandular tissue ($cm^3$) against Boyd group (a) small films (b) large films (c) all films.*

The figures above give a visual impression of relationship between each measure and the expert density classification. For example the absolute measure of dense tissue volume is seen to be rather weakly correlated with Boyd group, particularly when all films are considered together (Figure 6.16c). In order to decide which measure is most useful, i.e. which has most

potential for discriminating between breast cancer risk groups, quantitative measures are needed.

Table 6.1 shows the Spearman correlation coefficients between the various density measures and Boyd group classified by expert observer[2]. The measure with the highest degree of correlation with Boyd group is the percentage by volume of dense tissue. The next most significant correlation is for the percentage by area with thickness of dense tissue greater than 5 mm. The absolute volume of dense tissue is very weakly correlated with Boyd group. This is to be expected because of the wide range of breast sizes in the study group. For all of the measurements, except percentage area with thickness of dense tissue greater than zero, the small films show a stronger correlation with Boyd group than the large films. This is probably because the measurement uncertainty is greater for the large films.

| Measure | Small films | Large films | All films |
|---|---|---|---|
| Area with $x_g > 0$ (%) | 0.55 | 0.69 | 0.43 |
| Area with $x_g > 5$ mm (%) | 0.68 | 0.62 | 0.57 |
| Volume of dense tissue (%) | 0.73 | 0.69 | 0.73 |
| Volume of dense tissue ($cm^3$) | 0.56 | 0.50 | 0.17 |

*Table 6.1: Spearman correlation coefficients between each breast density measure and Boyd groups classified by expert observer.*

Discussion of these results on the performance of the various density measurements is provided in Chapter 7.

[2] Spearman correlation is used because the Boyd classifications are a discrete rather than continuous measure.

### 6.3.2 Examples

Figure 6.17 to Figure 6.21 show an example mammogram from each of the five populated Boyd categories, using the percentage area with greater than 5 mm dense tissue measurement, along with an image representation of the thickness of dense tissue at each point.

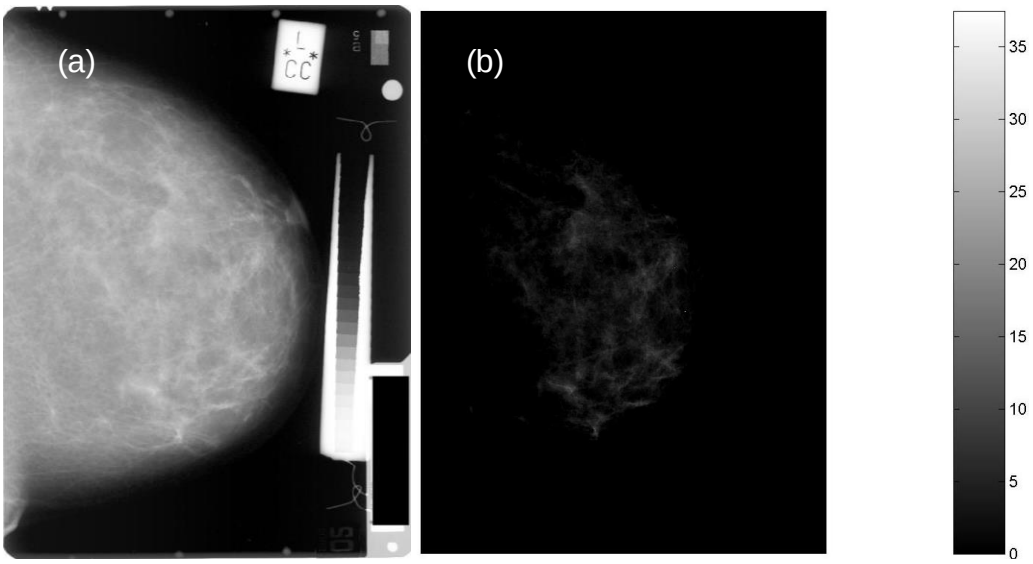


*Figure 6.17: (a) Original mammogram classified into <10% dense Boyd group (b) corresponding representation of thickness of dense tissue (in mm).*

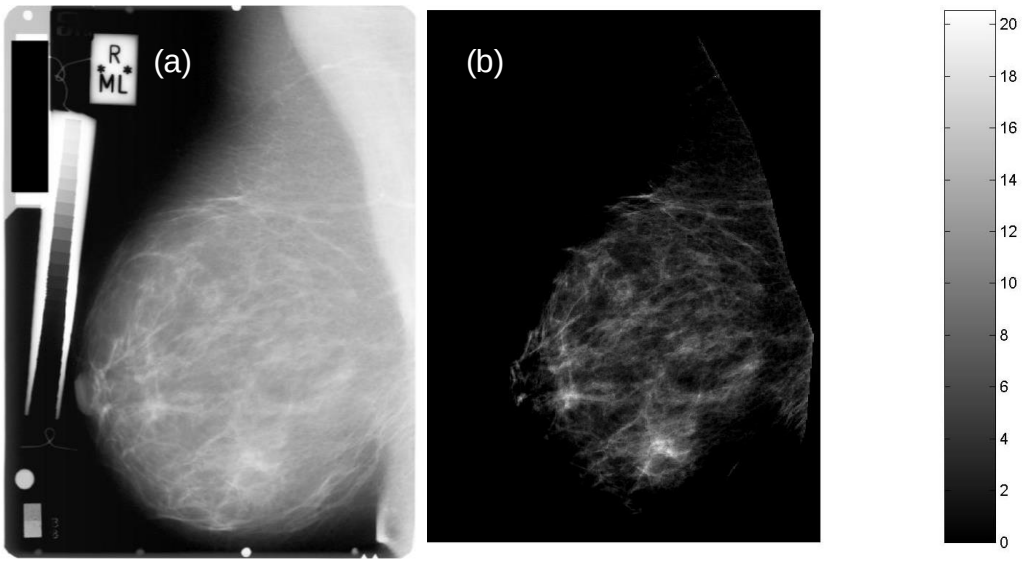


*Figure 6.18: (a) Original mammogram classified into 10-25% dense Boyd group (b) corresponding representation of thickness of dense tissue (in mm).*

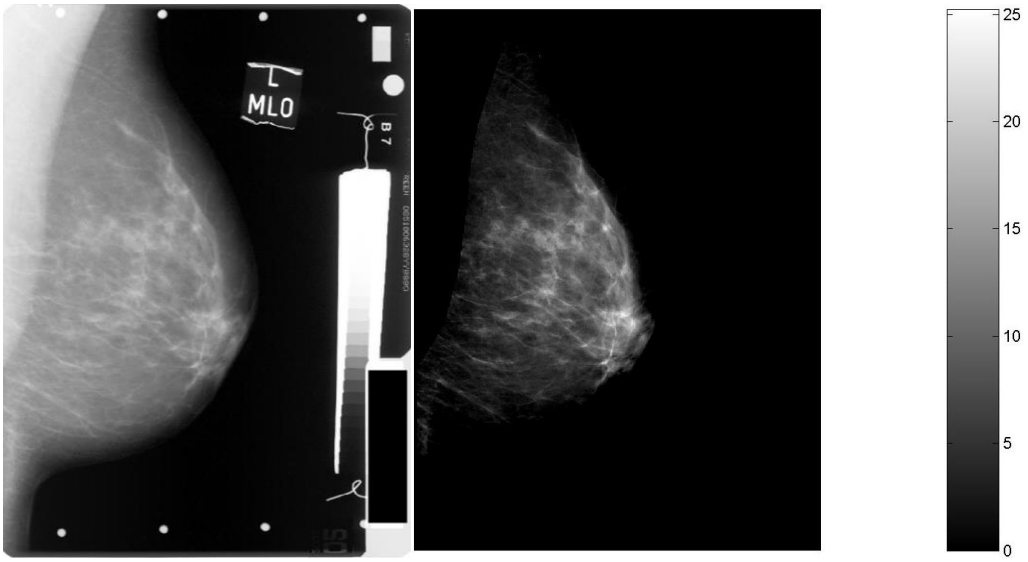


*Figure 6.19: (a) Original mammogram classified into 25-50% dense Boyd group (b) corresponding representation of thickness of dense tissue (in mm).*

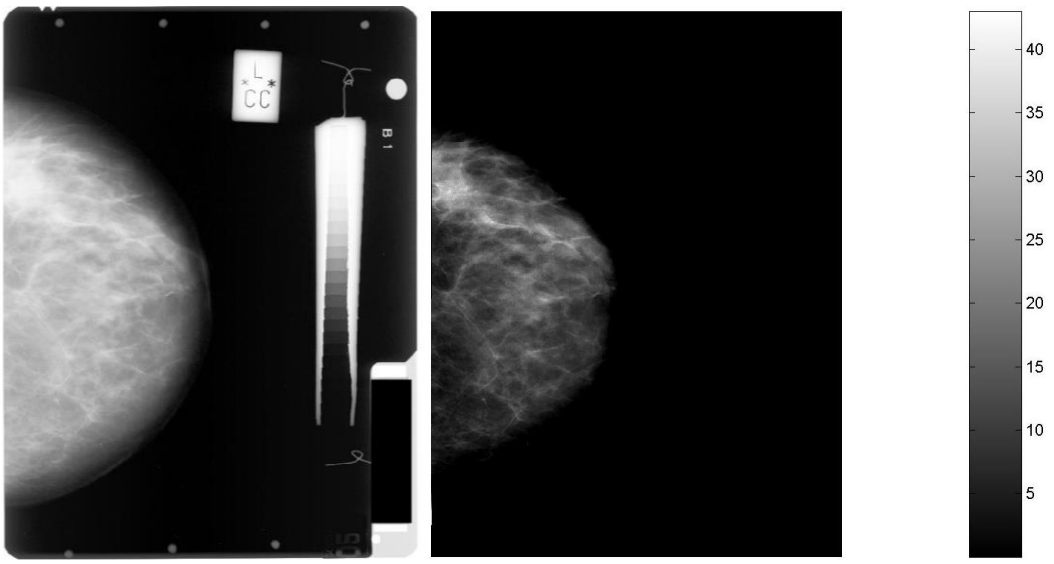


*Figure 6.20: (a) Original mammogram classified into 50-75% dense Boyd group (b) corresponding representation of thickness of dense tissue (in mm).*

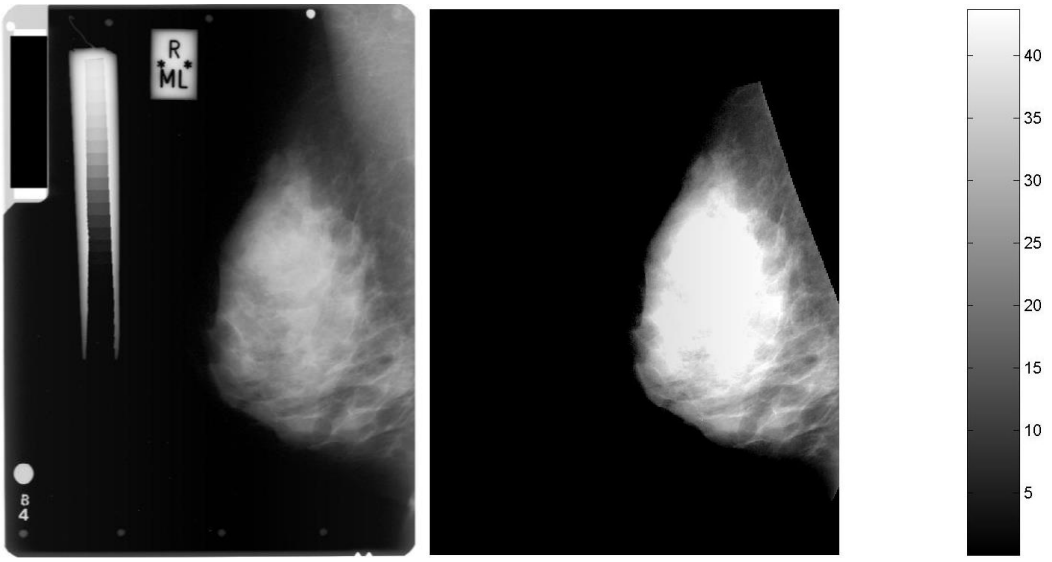


*Figure 6.21: (a) Original mammogram classified into 75-100% dense Boyd group (b) corresponding representation of thickness of dense tissue (in mm).*

# 7 Discussion

This section provides some discussion of the performance of the stepwedge method for measuring breast density.

## *7.1 Breast thickness measurement*

The breast thickness measurement technique, using radio-opaque markers placed on top of the compression paddle, has been shown to be both simple and accurate. Of the 164 clinical films taken with the markers in place, only two were not able to produce a breast thickness measurement because the markers were obscured. The error on the measurement is estimated to be 0.6 mm, significantly less than alternative methods. The error on the breast thickness measurement made by the mammography unit is estimated to be 1.9 mm, and the method used by Highnam *et. al*. has an estimated error of between 2 and 4 mm (see section 4.2).

A further advantage of the breast thickness measurement described here is that it measures the tilt of the compression paddle. The mammography unit used in this study is equipped with a compression paddle that allows a particularly large degree of tilt. This conforms better with the natural breast contour, giving more uniform compression and superior image quality[3]. Tilting can cause a difference in thickness of up to 20 mm from one side of the breast to the other, clearly a significant effect. Other methods of breast thickness measurement do not take into account the tilt of the compression paddle. The correlation between compression force and paddle tilt noted in section 6.2 suggests a way to extend other methods to account for compression paddle tilt, by assuming a value for the tilt based on the reported compression force. However, this would only be possible for machines where compression force measurements are available, and the

[3] Information from LORAD internet site: http://www.loradmedical.com/p225.html

accuracy of such a measurement would still be inferior to the magnification marker method used here.

## *7.2 Dense tissue thickness measurements*

The stepwedge method of measuring dense tissue thickness has been tested on a large sample of clinical films, including many on large format mammography film. Of 164 films there were only 4 where there was not enough room on the film to fit the stepwedge. Dense tissue measurements could not be made on these 4 films, which were spread between 3 cases.

The errors associated with breast density measurements made using this method have been assessed. The error on individual pixel measurements is estimated to be 2.4 mm. Individual pixel measurements can be used to create several different overall density measures. Errors on these have been assessed by comparing measurements from two views of the same breast, and area-based measures were found to have a standard deviation of around 10%. This should be considered as a conservative estimate of the error. The error was found to be larger for large format films than small films. The validity of the 2.4 mm error estimate on the dense tissue thickness measurements was tested by recalculating the area with dense tissue thickness greater than 5 mm after fluctuating the measurements at all pixels up and down by one standard deviation. The change in the area was found to be consistent with the standard deviation estimated from comparing views.

The main source of error in the dense tissue thickness measurements is the high sensitivity to equivalent stepwedge thickness. Measuring equivalent stepwedge thickness involves comparing grey levels from different areas in the image. This is difficult because a large correction is applied to the grey levels to compensate for anode heel and geometric effects. Future development work on the stepwedge method should include further investigation of the anode-heel correction.

It is important to consider whether the measurements presented here are sufficiently sensitive to detect small changes in breast density, which might result from the lifestyle changes being tested in the clinical study described in chapter 6. A previous study of effects of diet on breast density found a reduction in the area of dense tissue of 6.1 %, and reduction in the area of the breast of 2.4 % after 2 years on a low-fat, high-carbohydrate diet (Boyd 1997). If we use these as a guide to the magnitude of changes expected in the present study, then it is clear that our measurements will not be sensitive enough to confidently measure changes in individual cases. However, it should be possible to measure average changes over the study group. The standard error on the sample mean is $\sigma/\sqrt{N}$, where $\sigma$ is the standard deviation and $N$ is the sample size. For the sample size analyzed in section 6.3.1 we find that the errors on the mean are 1.0 % for percentage area with dense tissue thickness greater than zero, 1.3 % for percentage area with dense tissue thickness greater than 5 mm, 0.7 % for the percentage volume of dense tissue, and 0.3 $cm^3$ for the volume of dense tissue. The area measurements should be precise enough to detect changes in dense area of around 6 %.

The results from the correlation tests are rather disappointing, and do not suggest that the stepwedge density measurements provide better discrimination between Boyd's SCC groups than existing methods. For example, the method described by Byng using histogram skewness and fractal dimension (Byng 1996), was reported to achieve Spearman correlation coefficients with SCC category of –0.76 and –0.88 for fractal dimension and histogram skewness, respectively. This is higher than the coefficient of 0.73 measured here for percentage volume of dense tissue. It is important, however, to remember that recreating Boyd groups is not the main aim, only the simplest way to estimate relation to caner risk.

The stepwedge method gives additional information about the volume of dense tissue that cannot be accessed from Boyd type classification alone. It is hypothesized that this volumetric information may prove to be a superior

indicator of cancer risk. Only two other methods provide volume information (Highnam 1996 and Kaufhold 2002). Assessment of the performance of Highnam's method in the clinical situation has not been published, to the best of the Author's knowledge. Kaufhold's estimate of a 7 % error on the dense tissue percentage for a 4 cm thick breast corresponds closely to the present method's 2.4 mm estimated error (6 % for a 4 cm thick breast). It is interesting to note that the largest source of error in Kaufhold's method is the breast thickness measurement. They assume an error of 2 mm on breast thickness. The sensitivity of their method could be significantly improved by adopting a breast thickness measurement technique similar to that used here (estimated error 0.6 mm).

## *7.3 Breast edge effects*

No breast thickness edge correction has been used in the results presented in chapter 6. Attempts to develop such a correction were complicated because of effects other than breast thickness which effect the grey level fall-off at the breast edge. One such effect is scattered radiation. The method used by Highnam *et. al*. attempts to explicitly remove the scattered radiation component from the image. A similar approach may be useful here. Other automatic breast density measurement techniques do not make a thickness correction at the breast edge. The effect of neglecting this correction is to underestimate the amount of dense tissue near the breast edge. This is expected to be a small effect.

The use of a reliable edge correction would allow accurate measurements of the total breast volume. Methods of breast volume determination based on measurements made from mammograms have been compared by Kalbhen (Kalbhen 1999). Volume measurements made using thickness information from magnification markers should prove more accurate than the methods described by Kalbhen. Breast volume measurements are useful to plastic surgeons and physicians who examine cancer patients being considered for

breast conservation surgery, in addition to their use for breast density measurements.

# 8 Conclusion

Density of breast tissue has been shown to be significantly correlated with risk of developing breast cancer. This factor has not yet been fully utilized in the identification of women at risk. Breast density is particularly interesting because of the possibility to modify it using changes in diet, exercise, or treatment with hormones. Existing methods of assessing breast density have been subjective, and at best only semi-quantitative. These methods suffer from poor reproducibility and inter-observer variability. Objective and quantitative methods of measuring breast density are being developed using computer analysis of mammographic images. Calibrated methods, which compute the thickness of dense tissue at each point in the breast, make it possible to estimate the volume of dense tissue. This is expected to be a superior indicator of breast cancer risk.

A method of calibrating the thickness of dense tissue at each point in the image, using a stepwedge calibration device, has been developed and tested on a set of clinical mammograms. The resolution of measurements made using this method has been assessed, and is expected to be sufficient to detect changes in breast density during a study of diet and exercise effects on breast density.

A significant correlation is found between measurements made using this method and the Boyd SCC classification scheme which has an established relation with breast cancer risk. However, the discrimination between categories provided by these measurements is inferior to other automatic density measures.

More research is needed to investigate how the new methods of measuring breast density, particularly volumetric measurements, are correlated with cancer risk.

## Bibliography

Boone, J. M., K. K. Lindfors, et al. (2000). "Scatter/primary in mammography: comprehensive results." Medical Physics **27**(10): 2408-16.

Boyd, N. F., B. O'Sullivan, et al. (1982). "Mammographic signs as risk factors for breast cancer." Br. J. Cancer **45**(2): 185-193.

Boyd, N. F., J. W. Byng, et al. (1995). "Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study." Journal of the National Cancer Institute **87**(9): 670-5.

Boyd, N. F., C. Greenberg, et al. (1997). "Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: results from a randomized trial." Journal of the National Cancer Institute **89**(7): 488-96.

Boyd, N. F., G. A. Lockwood, et al. (1999). "Mammographic densities and risk of breast cancer among subjects with a family history of this disease." Journal of the National Cancer Institute **91**(16): 1404-8.

Burch, A. and J. Law (1995). "A method for estimating compressed breast thickness during mammography." British Journal of Radiology **68**(808): 394-399.

Bushberg, J. T., J. A. Seibert, et al. (1994). The essential physics of medical imaging. Baltimore, Williams & Wilkins.

Byng, J. W., N. F. Boyd, et al. (1996). "Automated analysis of mammographic densities." Physics in Medicine and Biology **41**: 909-923.

Byng, J. W., J. P. Critten, et al. (1997). "Thickness-equalization processing for mammographic images." Radiology **203**: 564-568.

Byrne, C. (1997). "Studying mammographic density: implications for understanding breast cancer." Journal of the National Cancer Institute **89**(8): 531-533.

Canny, J. (1986). “A Computational Approach to Edge Detection.” IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6): 679-698.

Dendy, P. P. and B. Heaton (1999). Physics for diagnostic radiology. Bristol, Institute of Physics Publishing.

Heine, J. J. and R. P. Velthuizen (2000). "A statistical methodology for mammographic density detection." Medical Physics **27**(12): 2644-51.

Highnam, R. P., J. M. Brady, et al. (1994). "Computing the scatter component of mammographic images." IEEE Transactions on Medical Imaging **13**(2): 301-313.

Highnam, R. P., J. M. Brady, et al. (1996). "Mammographic image analysis." European Journal of Radiology **24**: 20-32.

Highnam, R. P., J. M. Brady, et al. (1996). "A representation for mammographic image processing." Medical Image Analysis **1**(1): 1-18.

Highnam, R. P., J. M. Brady, et al. (1998). "Estimation of compressed breast thickness during mammography." British Journal of Radiology **71**: 646-653.

Kalbhen, C. L., J. J. McGill, et al. (1999). "Mammographic determination of breast volume: comparing different methods." Am J Roentgenol **173**(6): 1643-1649.

Kaufhold, J., J. A. Thomas, et al. (2002). "A calibration approach to glandular tissue composition estimation in digital mammography." Medical Physics **29**(8): 1867-1880.

Kodak (2001), “Kodak LS85 film digitizer product specifications.”, Eastman Kodak Company, Rochester, NY. Cat. No. 165 2908.

Lacey, J. V., S. S. Devesa, et al. (2002). "Recent trends in breast cancer incidence and mortality." Environmental and molecular mutagenesis **39**: 82-88.

Lee, N. A., H. Rusinek, et al. (1997). "Fatty and fibroglandular tissue volumes in the breasts of women 20-83 years old: comparison of X-ray mammography and computer-assisted MR imaging." Am. J. Roentgenol. **168**(2): 501-6.

Madigan, M. P., R. G. Ziegler, et al. (1995). "Proportion of breast cancer cases in the United States explained by well-established risk factors." Journal of the National Cancer Institute **87**(22): 1681-5.

Smith, J. H. (1998). “Prediction of the risk of breast cancer using computer vision techniques.” Ph.D. thesis, Department of Medical Biophysics, University of Manchester, UK.

Wolfe, J. N. (1976). "Breast patterns as an index of risk for developing breast cancer." Am. J. Roengenol. **126**: 1130-1139.

Zhou, C., H. Chan, et al. (2001). "Computerized image analysis: estimation of breast density on mammograms." Medical Physics **28**(6): 1056-1069.