TOWARDS OPTIMISATION OF COMPRESSION FORCE IN MAMMOGRAPHY

Katherine SZCZEPURA

SCHOOL OF HEALTH SCIENCES, UNIVERSITY OF SALFORD, SALFORD, UK

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy

October 2018

To Steve and Robin

I finally did this for you

Acknowledgements

During my career I have been extremely lucky to work with some incredibly inspiring scientists and clinical practitioners who have supported my development throughout my professional life.

I would like to thank my supervisors Professor David Manning, Professor Peter Hogg and Dr. Gillian Crofts for their unrelenting motivation.

David, your insight and support throughout the many iterations of this thesis have been invaluable. You are truly an inspiration and I have enjoyed the opportunity to work with you over the past few years, I have thoroughly enjoyed our numerous and wide ranging discussions, and I promise to bore you with more craft talk in the future!

Peter, thank you for stepping into the breech at such a crucial time and for getting me through with your continued advice, critique and wisdom. The opportunities you have given me since joining The University of Salford have allowed me to complete this important step in my research journey.

Gillian, thank you for starting this process with me, your kindness and support got me through the important stages, and I am grateful for your input.

Table of Contents

List of Figures

Glossary of Terms

AEC – Automatic Exposure Control

BSP – breast screening programme

CC – cranial caudal

FFDM – full field digital mammography

MGD –mean glandular dose

MLO – medio lateral oblique

QA – quality assurance

QC – quality control

SFM – screen film mammography

VBD – volumetric breast density

TMD – thickness measuring device

MTF – modulation transfer function

Abstract

This work focusses on the compression force applied to the breast during mammographic imaging. The lack of guidance and evidence base has led to variability in practice, and the implications of this will be discussed. All research was carried out on mammography units used within the U.K. National Health Service Breast Screening Programme (NHSBSP), and as such, emphasis has been placed on U.K. protocols and guidance. It is important to note however that this work does have implications for international breast screening programmes, as the impact considers technological, client and practitioner- based issues.

The aim of this thesis is to discuss the implications of the lack of standardisation of compression force in terms of image quality and dose. The first section provides a contextual background using literature. The second section includes the published papers for consideration, 10 papers are included which have all been published in peer reviewed journals. The author's contribution for each paper has been demonstrated and confirmed by co-authors, the main input to the work presented for this thesis is to ensure the approach taken has validity, reliability and reproducibility and the data is presented in a novel and appropriate manor.

The systematic acquisition and distribution of the acquired and novel knowledge are discussed in the critical review section. This section addresses each objective, critically reviewing the published research, together with key findings and the contribution of each of the papers.

Additionally, further work will be discussed, demonstration the author's independent research building on the previous work, and future direction in the field.

Aims and Objectives

Aim of the Critical Review

To establish, with reference to the published works submitted, the effect, extent and impact of variation in applied compression force in mammographic imaging

Objectives of the Critical Review

1. Establish the impact of inaccuracy in breast thickness measurements
2. Establish the extent of practitioner variation in application of compression force
3. Establish the impact on positioning technique and applied compression force on the breast area on the image receptor
4. Establish the extent of breast thickness reduction with applied compression force
5. Demonstrate the application of measuring Conspicuity Index to demonstrate the impact of compression on lesion detection in mammography

Introduction

Mammography is an imaging technique that utilises X-rays to image the breast. It is unique in the imaging field in that its purpose is to only image one organ.

Mammography is one of the most technically challenging imaging modalities. As it is used as a screening tool to detect small subclinical lesions it requires high spatial resolution, and additionally excellent soft tissue contrast is needed due to the similarities between the lesions and the surrounding tissues. However, as it is used as a screening tool, it needs to deliver these high-quality images at a relatively low radiation dose, as higher doses cannot always be justified in a normal asymptomatic population.

A typical setup for a mammography machine can be seen in Figure 1 below

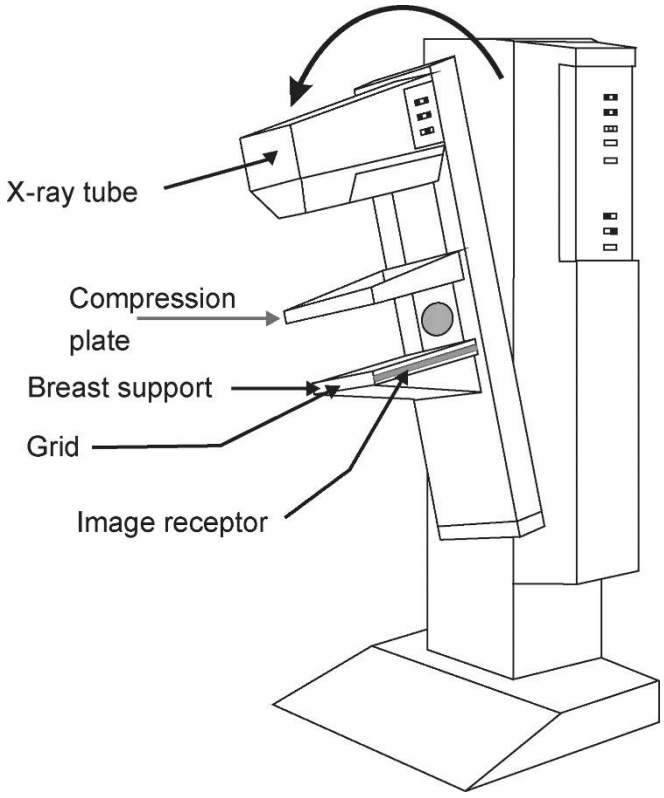


Figure 1 mammography unit (Yaffe, Bunch et al., 2009)

During imaging the breast is placed on the breast support and the paddle is then used for compression. Generally, in clinical practice, automated exposure control (AEC) is used to terminate the exposure, where all exposure factors are controlled by the unit. The factors chosen by the machine are based on the compressed thickness of the breast, this means that appropriate

positioning and compression are essential to ensure the AECs are able to function appropriately to optimise the exposure for the patient.

The compression force that is applied to the breast during mammographic imaging is one of the most important parts of the image acquisition. There are many reasons that have been stated for the justification of compression: it reduces the thickness of breast tissue, which in turn reduces the radiation dose that the patient receives and it brings the breast tissue closer to the image receptor and so reduces geometric unsharpness (Pisano, Gatsonis et al., 2005, Taplin, Rutter et al., 2002). Additionally, there is a suggestion that the applied compression force displaces the glandular tissues, reducing the impact of overlying tissues on the resultant mammograms which improves contrast and lesion visibility due to the visual similarity of lesions and glandular tissue. Finally, compressing the breast reduces movement unsharpness by stabilising the breast.

The compression is applied by the practitioners undertaking the mammogram, and although the importance of appropriate compression force has been recognised by the Breast Screening Service standards (National Quality Assurance Coordinating Group for Radiology, 2006), where compression has been advised to be applied gently and slowly until the breast is held firmly in position, there is no guidance on appropriate compression except that it should not exceed 200 Newton (20 kg). Additionally, there is no European or international guidance, with only broad guidance given such as "compress until the skin is taut to touch", which can lead to subjective interpretation (Perry, Broeders et al., 2008, Poulos and McLean, 2004)

Literature Review

Breast Cancer: burden, risks and control

Breast cancer is the most common cancers in women in the world. It is estimated that worldwide over 508 000 women died in 2011 due to breast cancer (Global Health Estimates, WHO 2013).

Incidence rates vary greatly worldwide from 19.3 per 100,000 women in Eastern Africa to 89.7 per 100,000 women in Western Europe, this difference in incidence rates might be explained by many factors, including increased early detection in Weston Europe, as well as genetic and environmental influences (Coleman, Quaresma et al., 2008). In most of the developing regions the incidence rates are below 40 per 100,000 (GLOBOCAN 2008) but the incidence rate is increasing.

Breast cancer survival rates vary greatly worldwide, a worldwide population-based study conducted in 2008 found that 5-year survival rates ranged from 80% or over in North America, Sweden and Japan to around 60% in middle-income countries and below 40% in low-income countries (Coleman

et al., 2008). Within this publication it was suggested the low survival rates in less developed countries can be explained mainly by the lack of early detection programmes, resulting in a high proportion of women presenting with late-stage disease, as well as by the lack of adequate diagnosis and treatment facilities.

In the UK breast cancer makes up 15% of all cancers, with 53,696 news cases reported in 2013, it is the most common cancer in the UK. The majority of diagnoses are in women, with around 340 cases diagnosed in men in that year. Almost half of breast cancers are diagnosed in women over 65, with men half are diagnosed in over 70-year olds.

Over the past 10 years, breast cancer incidence has increased by 4% in the UK, and the incidence rate is the sixth highest in Europe. Worldwide, 1 in 8 women and 1 in 870 men will be diagnosed with breast cancer. (Cancer Research UK, 2016)

Breast Cancer Risk Factors

There are many risk factors associated with breast cancer including (but not exclusively): alcohol and tobacco use, geographical location, socio-economic conditions, exposure to radiation, obesity, decreased physical activity, urbanization, sedentary lifestyle, high fat diet, frequent spontaneous miscarriages, lack of breast-feeding, hormone replacement therapy, aging, early menopause, reproductive events and incidents of ovarian cancer (Ayub, Rasool et al., 2014, Bray, Jemal et al., 2012, Bray, McCarron et al., 2004, Calderón-Garcidueñas, Ruiz-Flores et al., 2005, Collaborative Group on Hormonal Factors in Breast Cancer, 2002, Mehrgou and Akouchekian, 2016, Torre, Bray et al., 2015)

The incidence of breast cancer increases with age, with the risk doubling every 10 years until menopause (McPherson, Steel et al., 2000) The majority (81%) of breast cancers occur in women over 50 (Quante, Whittemore et al., 2012).

By far the largest risk factor is familial history, the risk is dependent on the type and number of relatives affected, age at which the relative developed breast cancer and the incidence of bilateral disease (Arver, Du et al., 2000, Pharoah, Day et al., 1997, Song, Chen et al., 2017). Depending on these factors, familial history can increase the risk by 2 or 3 fold (Antoniou, Gayther et al., 2000, Arver, Du et al., 2000, Song, Chen et al., 2017) and it is estimated that 5% to 10% of all breast cancer cases are due to these inherited genetic factors (Arver, Du et al., 2000)

The gene BRAC1 was first associated with breast cancer risk in 1990 by Hall et al. They studied breast cancer families, defined as three or more close family members having breast cancer, and discovered the genetic link. It was also found to be associated with breast-ovarian cancer, and ovarian cancer (Antoniou, Gayther et al., 2000) . BRAC1 is considered a tumour suppressor gene, and involved in DNA repair, and mutations of this gene is strongly linked to breast cancer rates (Arver, Du et al., 2000)

More recently, breast density has been associated with an increased risk of breast cancer. The measure of the relative amount of fibroglandular tissue within the breast has been established as a strong independent risk factor for breast cancer (Chen, Gulsen et al., 2015, McCormack and dos Santos Silva, 2006, Ng and Lau, 2015, Sherratt, McConnell et al., 2016) changes in breast density may be caused by hormones, age, genetic factors and body habitus (Chen, Gulsen et al., 2015, Rice, Rosner et al., 2017, Tice, Cummings et al., 2008, Vachon, van Gils et al., 2007).

Breast density can be assessed in various ways, using either visual assessment or computer based metrics (Ekpo, Hogg et al., 2015). Traditionally the breast density was visually assessed by the clinician making a visual evaluation of the mammogram, where they use all presented images to make a decision of the percentage of fibroglandular tissue within the volume of the breast.

This requires the clinician to be able to correctly assess the relative proportions of glandular and fatty tissue while accounting for variations in breast shape, radiographic texture and the presence of cancer (which increases breast density locally). The clinician is also able to take account of the variation in the technical acquisition parameters of the mammogram, for example the AEC settings. These density scores can then be presented as a percentage on a continuous scale or within discrete ranges such as the composition categories used in BI-RADS (McCormack and dos Santos Silva, 2006). Moderate agreement has been found between clinicians, but training and experience have shown to impact on the accuracy and reproducibility of these scores (Damases, Hogg et al., 2017, Martin, Helvie et al., 2006).

Introduction of digital images meant that automated measures of breast density could be assessed, with early work applying thresholds to pixel values to determine the proportion of the breast that is dense (Gilbert, Tucker et al., 2015, Martin, Helvie et al., 2006, Wang, Good et al., 2003). Later software developments estimate the volume of dense fibroglandular tissue in the whole breast rather than the density of the breast tissue within the mammogram, known as Breast Volume Density (VBD). By using both the image pixel data in combination with acquisition information from the DICOM header, algorithms are used to provide measurements of the relevant tissue volumes. Data

regarding the breast thickness and the X-ray exposure's tube potential, current, time, target and filter – in combination with knowledge of the radiation attenuation properties of different tissues – can enable a derivation of the breast composition represented by each pixel. Two main software are available for assessment of VBD, Volpara™ (Highnam, Brady et al., 2010) and Quantra (Hartman, Highnam et al., 2008)

In one large UK trial (Gilbert, Tucker et al., 2015), they found that correlation between these assessment software was high, whereas correlation between visual assessment and VBD has found to be poor, with numerous factors being described as the cause, including observers' use of a continuous scale, experience of the observer, and the fact that the observers were using processed images in comparison to the software using raw data. However, when considering risk factors a strong correlation was found with the absolute measurements of the fibroglandular tissue volume than the percentage volume from both the Quantra and Volpara software. Furthermore (Astley and Harkness, 2015) found variability in VBD between observers and the two soft wares, and found that the observers showed the strongest relationship with risk of cancer, they discussed that is was likely that observers, whilst asked to quantify density, also took into account higher-level features, such as the pattern and location of dense areas.

Accuracy of the breast thickness measurement is important within VBD assessment, as the thickness used within the algorithm is taken from the recorded value given by the machine. Any inaccuracy in this measurement could lead to inaccurate calculation of VBD, and this will be discussed within this thesis in Section 1.1.3 - Breast density calculation

Breast Screening Programmes

Health screening is a way of identifying those at risk of a condition within a healthy asymptomatic population so early diagnosis and treatment can improve client morbidity and mortality, or to allow the client to make informed decisions. There are various screening programmes undertaken worldwide, and the risks and costs of any screening program have to be taken into account when considering implementation.

Early detection and treatment of breast cancer is directly linked with better patient prognosis (Ng and Muttarak, 2003), and mammography is, at present, the best mass population screening method for diagnosing non-palpable breast cancer. Within the UK, the aim of the NHS Breast Screening Programme (NHS BSP) is to aide early detection of breast cancer through screening of a non-symptomatic population. Women in the UK are offered breast screening every 3 years, between the

ages 47-73.(Moss, Wale et al., 2015) with the exception for individuals at moderate and high risk who are eligible for enhanced surveillance using mammography and MRI (Evans, Graham et al., 2013) due to familial risk factors.

Screening programmes in other countries vary in both frequency and age range. In the US screening starts as young as 40 up to the age of 75 and is performed as frequently as every 1-2 years, depending on state. Australia have a similar program to the US, whereas most European countries have similar programmes to the UK, except Sweden where they screen every 18 months for 40-49 and every 2 year for 50-74 (International Cancer Screening Network, 2012). The age range and frequency each of the program is based on that country's demographic and risk/benefit analysis, and the variation in these approaches can lead to controversy. As with the UK those women considered to have a high risk of breast cancer are offered enhanced surveillance, for example in the US high risk clients can receive early cancer screening from the age of 25

False positive results and over diagnosis raises the sensitivity measure of the test but is an issue in breast screening (Hofvind, Ponti et al., 2012, Løberg, Lousdal et al., 2015). Women can be diagnosed and treated for disease that may not have impacted on their lives and the psychological impact of this needs to be considered (Bond, Pavey et al., 2013a, Bond, Pavey et al., 2013b, Brett, Bankhead et al., 2005). Additionally, there is also the risk from ionising radiation and the costs that need to be taken into account (Jatoi, 2015, Jørgensen and Bewley, 2015, Lauby-Secretan, Loomis et al., 2015).

In 2012 an independent review was undertaken by The Independent UK Panel on Breast Cancer Screening (Independent UK Panel on Breast Cancer Screening, 2012) to assess the risks and harms of breast cancer screening, the results were published in October of that year, and summarised that for all women attending screening, there is about 1% chance of a cancer being diagnosed and treated that would never had caused an issue for the women. For breast cancer detection they suggested that for each death prevented, there will be three over diagnosed cases. It is recognised in the report that this data is very hard to assess, and therefore limited in its accuracy and reliability. However, the report concludes that the breast screening service should still be considered as having a significant benefit and should therefore be continued.

Full Field Digital Mammography (FFDM) as a screening tool

Full field digital mammography (FFDM) is still the most effective screening tool for breast cancer diagnosis, a screening mammogram consists of two projections of each breast; one in the

craniocaudal projection known as the CC and one in the medio-lateral oblique projection known as the MLO (Hogg, Kelly et al., 2015).

For the CC the inferior portion of the breast is placed on the image receptor and the compression paddle is applied onto the superior portion of the breast; the mammography machine gantry is parallel to the floor (Figure 2) For the MLO the arm of the mammography gantry is tilted from the vertical and angled to be parallel to the pectoral muscle (Figure 3).

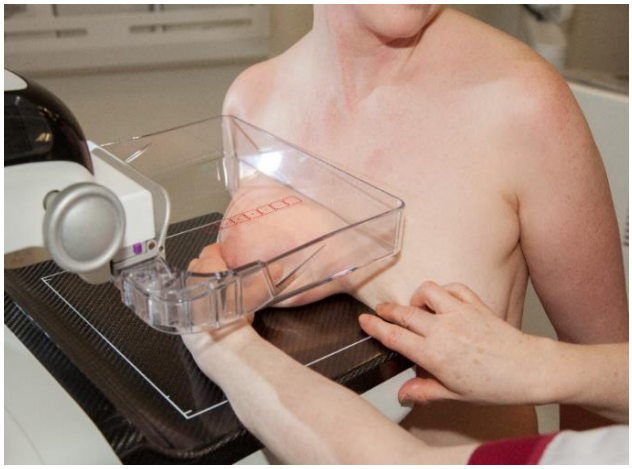


Figure 2 The craniocaudal (CC) mammogram (Hogg, Kelly et al., 2015)

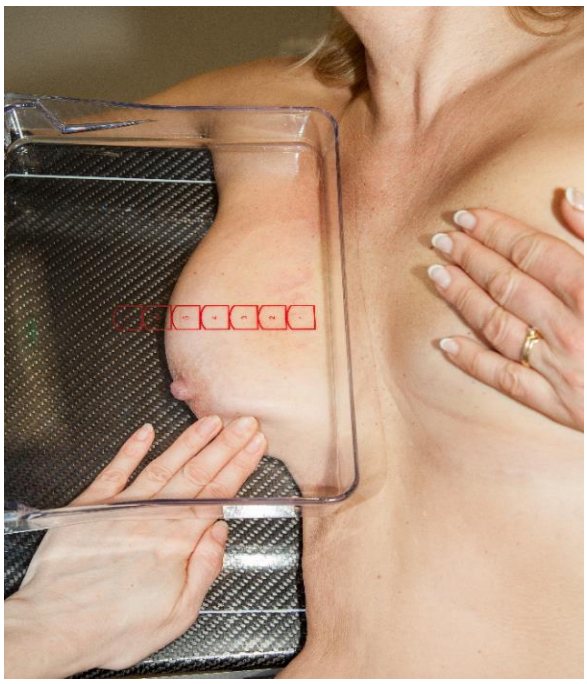


Figure 3The medio-lateral oblique (MLO) mammogram (Hogg, Kelly et al., 2015)

Compression in mammography

The compression force that is applied to the breast during mammographic imaging is one of the most important parts of the image acquisition. However, despite its importance in terms of image quality

and dose, there is a lack of supporting information for the appropriate level of compression, which inevitably gives the potential for significant variations in practice with the application of breast compression force (Mercer, Hogg et al., 2013b, Poulos, McLean et al., 2003)

Image Acquisition is heavily reliant on automatic exposure control (AEC) in clinical practice, where the kV, exposure, exposure time (mAs) and anode/filter combination are automatically selected based on the thickness of the compressed breast. Some systems work by additionally taking a short exposure of the breast to measure the resultant exposure to accurately set the AEC values (Bick and Diekmann, 2010). All these factors directly impact on the image quality and patient dose, so it can be seen that the compression force applied needs to be optimised and standardised to ensure appropriate imaging parameters are selected.

Taplin et al (Taplin, Rutter et al., 2002) demonstrated a correlation between poor image quality and cancer developing after a negative screening, with poor positioning having the largest impact on the sensitivity (66.3%). This indicates that poor image quality has an impact on detection of early stage lesions that are seen on subsequent images.

Although they found other factors such as noise, compression, sharpness, contrast, exposure and artefacts had only a moderate correlation with later cancer detection, Rauscher et al (Rauscher, Conant et al., 2013) found that compression forced used and the resultant sharpness was associated with a later stage diagnosis, and suggests that improved positioning, compression and sharpness would improve stage at diagnosis.

As there is a lack of evidence or guidelines for the application of compression force, and with the negative implications of patient discomfort, this inevitably leads to variability in clinical practice (Mercer, Hogg et al., 2013b, Poulos, McLean et al., 2003, Waade, Sanderud et al., 2017). The practitioners decide when enough compression force has been applied. Various ways of describing “enough” compression force have been suggested in the literature (Eklund, Cardenosa et al., 1994, Kopans, 2007, Long, Miller et al., 2010, Mercer, Hogg et al., 2013b, Poulos and McLean, 2004) but there are no evidence-based agreed guidelines for practitioners to identify optimal compression force levels. Previous work has demonstrated a lack of consistency in the way compression force is applied (Branderhorst, de Groot et al., 2015, de Groot, Branderhorst et al., 2015, de Groot, Broeders et al., 2013, Dustler, Andersson et al., 2012, Mercer, Hogg et al., 2013b, Poulos, McLean et al., 2003, Waade, Moshina et al., 2017) and the impact of this leads to a lack of optimisation, which is essential in a screening programme, and a requirement within national and international legislation (Great Britain. Health and Safety, 2000, ICRP, 2006)

Presented Published Work

All works presented for consideration have been published in peer review journals. The labelling of figures and texts, and the referencing style are consistent with the requirements of each journal.

Paper 1

The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units

Ingrid H. R. Hauge[a)]
Oslo and Akershus University College of Applied Sciences, Faculty of Health Sciences, Department of Radiography and Dental Technology, P. O. Box 4, St. Olavs plass, NO-0130 Oslo, Norway and Norwegian Radiation Protection Authority, P. O. Box 55, NO-1332 Østera˚s, Norway
Peter Hogg and Katy Szczepura
Directorate of Radiography, University of Salford, Salford M6 6PU, United Kingdom
Paul Connolly
Integrated Radiological Services Ltd., Unit 188 Century Building, Tower Street, Brunswick Business Park, Liverpool L3 4BJ, United Kingdom
George McGill
The Christie NHS Foundation Trust, Wilmslow Road, Manchester M20 4BX, United Kingdom
Claire Mercer
Royal Bolton Hospital NHS Foundation Trust, Minerva Road, Farnworth, Bolton BL4 0JR, United Kingdom

(Received 13 May 2011; revised 28 October 2011; accepted for publication 2 November 2011; published 20 December 2011)
 [DOI: 10.1118/1.3663579]
Key words: mammography, breast thickness, breast compression

Purpose: To establish a simple method to determine breast readout accuracy on mammography units.

Methods: A thickness measuring device (TMD) was used in conjunction with a breast phantom. This phantom had compression characteristics similar to human female breast tissue. The phantom was compressed, and the thickness was measured using TMD and mammography unit readout. Measurements were performed on a range of screen film mammography (SFM) and full-field digital mammography (FFDM) units (8 units in total; 6 different models/manufacturers) for two different sized paddles and two different compression forces (60 and 100 N).

Results: The difference between machine readout and TMD for the breast area, when applying 100 N compression force, for nonflexible paddles was largest for GE Senographe DMR+ (24 cm30 cm paddle: +14.3%). For flexible paddles the largest difference occurred for Hologic Lorad Selenia (18 cm24 cm paddle: +26.0%).

Conclusions: None of the units assessed were found to have perfect correlation between measured and readout thickness. TMD measures and thickness readouts were different for the duplicate units from two different models/manufacturers.

Introduction

Accurate breast thickness estimation is required in order to calculate the mean glandular dose (MGD).[1–3] Accuracy is also required for density measurements (which can be used for predicting breast cancer risk)[4] and for estimation of breast tissue volume.[5,6] Compression paddles may deform/ tilt during mammography and this can lead to differences between the actual and readout (displayed by the mammography machine) thickness of the compressed breast. Under realistic clinical imaging conditions (phantom-simulated), this study aimed to conduct a comparative analysis of readout versus measured thicknesses over a range of mammography units.

Previous studies have highlighted inaccuracies with thickness readouts of mammography machines; some of these studies have also proposed methods which may provide a better estimate of the compressed breast thickness.[3,7–9] Diffey *et al.*[10] found a maximum variation of 21.1 mm in the chest wall to nipple direction, while the paddle deformation in the lateral direction was found to be insignificant in comparison to the chest wall to nipple direction. Tyson *et al.*[9] described a technique for measuring breast thickness by using optical stereoscopic photogrammetry. This method had a precision of >1 mm, and a measurement accuracy of >0.2 mm. The readout thickness for a number of different mammography systems was found to vary by as much as 15 mm when compressing the same breast or phantom.[9] The value of the method developed by Tyson *et al.*[9] was its accuracy; system use however is labor intensive, being highly dependent on room lighting and also on image quality. Mawdsley *et al.*[7] developed functions that can estimate the compressed breast thickness based upon the machine readout thickness and compression force reported by the machine.

This study aimed to develop a simple, clinically adaptable and accurate method to measure the difference between the readout and measured thickness. Building on previous research there was particular interest in, the creation and documentation of the physical breast phantom characteristics, particularly in relation to in-vivo female human breast tissue. In order to investigate how the thickness readout and the thickness across the breast correlated, a breast thickness measuring device (TMD) was constructed.

Methods and Materials

The method comprised of three stages. First, a clinically realistic breast phantom and backing plate with the creation of a rigid torso was tested. Second, the TMD was designed and tested. Finally, using the TMD, the breast phantom with its backing plate was used to assess several mammography units/paddle combinations.

Design, creation, and validation of breast phantom

Three breast prostheses (small (220 cm^3), medium (360 cm^3), and large (700 cm^3), Trulife, Sheffield, United Kingdom) were assessed for their compression characteristics. Each of the breast prostheses were adhered onto a semiflexible backing plate. The backing plate was mounted onto a rigid torso (Fig. 1) in order to simulate how a real breast will behave when it is compressed. The resistance to compression incurred by the torso changed the compressibility of the phantom to better simulate a real breast.

Six rubber balloons were glued onto the flexible backing plate. The balloons gave minor mobility similar to pectoral muscle and fascia. The phantom was glued onto the balloons and covered with layers of latex. The latex was painted across the surface of the phantom and along the edges, with fewer layers across the surface than around the edges. The backing plate was mounted onto a rigid torso (CIRS, Norfolk) using two ratchet straps, one above and one below the breast phantom. Before compressing the breast phantom, a lubricant was applied to the phantom. This allowed the compression paddle to slide smoothly over the breast surface when pressure was applied.

Using the three breast phantoms, mounted as described, compression (N)/thickness (mm) graphs were generated from 40 to 100 N stepping through 10 N values. For each phantom, the compressed breast thickness data were averaged and normalized (the data were normalized to 1 for 40 N compression force). For comparison the normalized average of 29 female human datasets were acquired (Fig. 2).

The 29 female datasets were acquired on a Hologic Lorad Selenia, while the phantom data were collected from a GE Senographe 800 T. The normalized compression curve of the large prosthesis was compared with the normalized correlation curve of the real breast, and it was found that the compression characteristics correlated well, with a correlation coefficient of 0.95. On this basis the large phantom (700 cm^3) was chosen as our breast phantom.

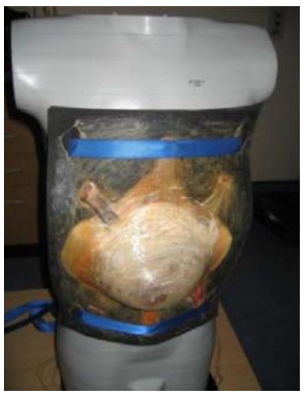


FIG. 1. Breast mounted to semiflexible background plate and rigid torso.

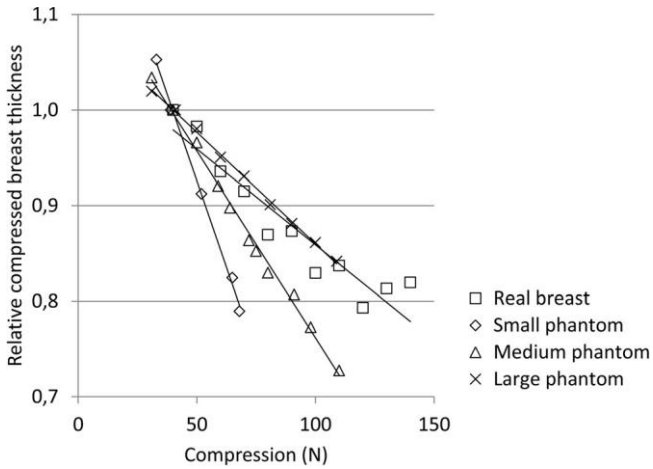


FIG. 2. Compressed breast thickness (mm) as a function of compression force (N) for real breasts and the three breast phantoms.

Compression paddle bend and distortion measuring device

The TMD was constructed of poly methyl methacrylate (PMMA) (Fig. 3). TMD dimensions (depth: 17.1 cm, width: 36.0 cm, and height: 21.8 cm) were such that they would fit the mammography machines/paddles that were to be included in the study. Wooden rods, diameter approximately 5 mm, and of different lengths (10–25 cm) were used (Fig. 3) to measure thickness. The top of the TMD had a matrix of 5 mm diameter holes drilled through it; the centers were 20 mm apart.

How the study was conducted

The measurements were performed on different mammography units from three different manufacturers [General Electric (GE Medical Systems, Buc, France), Hologic Inc. (Bedford, MA) and Siemens (Siemens Healthcare, Erlangen, Germany)]. Both screen film mammography (SFM) and fullfield digital mammography systems (FFDM) were included (Table I). This selection is representative of machines that were in clinical use at the time of the study. Two different paddle sizes, standard [approximately 18 cm24 cm (18x24)] and large [approximately 24 cm30 cm (24x30)] were used (Table I).

FIG. 3. Thickness measuring device (TMD) and rods.

TABLE I. Mammographic units included in this study.

Location	Manufacturer=Model	SFM=FFDM size	Compressed breast thickness accuracy QC: maximum difference (specified by manufacturer)	in measured and readout thickness[b]	Paddle	Flexible=Nonflexible paddle	Tilting=Nontilting
A	GE Senographe 800T	SFM	610 mm	60.4 cm	18 cm24 cm	Nonflexible	Nontilting
			610 mm		24 cm30 cm	Nonflexible	Nontilting
A	GE Senographe DMR+	SFM	610 mm	+0.5 cm	18 cm24 cm	Nonflexible	Nontilting
			610 mm		24 cm30 cm	Nonflexible	Nontilting
B	GE Senographe DMR+	SFM	610 mm	+0.5 cm	18 cm24 cm	Nonflexible	Nontilting
			610 mm		24 cm30 cm	Nonflexible	Nontilting
C	Siemens Mammomat Inspiration	FFDM	39–45 mm[a]	0.1 cm	18 cm24 cm	Nonflexible	Nontilting
					24 cm30 cm	Nonflexible	Nontilting
B	GE Senographe Essential	FFDM	610 mm	0.3 cm	19 cm23 cm[d]	Nonflexible	Nontilting
			610 mm		19 cm23 cm[d]	Flexible	Tilting
			610 mm		24 cm31 cm	Flexible	Tilting
D	Hologic Lorad Selenia	FFDM	60.5 cm	0.1 cm	18 cm24 cm	Flexible	Tilting
			60.5 cm		24 cm30 cm	Flexible	Tilting
D	Hologic Selenia Dimensions	FFDM	60.5 cm	0.1 cm	18 cm24 cm[d]	Flexible	Tilting
			60.5 cm		24 cm29 cm[d]	Flexible	Tilting
E	Hologic Lorad Selenia	FFDM	60.5 cm	0.4 cm[c]	18 cm24 cm	Flexible	Tilting
			60.5 cm		24 cm30 cm	Flexible	Tilting

[a] The thickness of a compressible phantom should be between 39 and 45 mm. The thickness of the compressible phantom (RMI 156, Gammex RMI, Middleton, WI) is 42 mm.

[b] In the UK the compressed breast thickness accuracy is measured during quality control (QC) which is conducted every six months. This consists of measuring the compressed thickness for a PMMA phantom of known thickness. Difference in compressed breast thickness=Thickness of Perspex—Readout thickness. An under- and=or underestimation is considered equally faulty.

[c] All quality control measurements were conducted with a nonflexible paddle.

[d] Even if Hologic Selenia Dimensions and GE Senographe Essential were a bit different in size than the others, they are referred to as 18 cm24 cm (18x24) and 24 cm30 cm (24x30) in the figures.

The TMD was placed on top of the table, with the long side (36.0 cm) parallel and along the edge of the chest side of the table top and centered left to right. The compression paddle was fastened such that it was located between the top and bottom plate of the TMD (Fig. 4), with the breast pros thesis resting on the bottom plate of the TMD. Two different compression forces were applied when compressing the breast prosthesis (60 and 100 N).

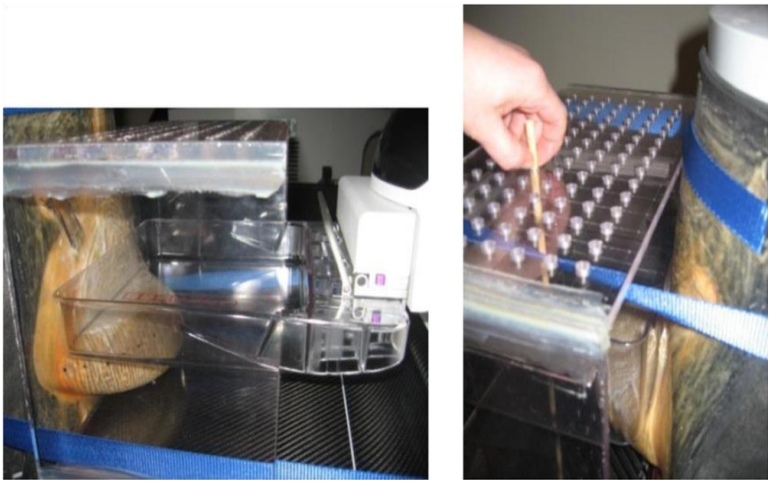


FIG. 4. How the measurements were conducted.

In order to estimate the compressed breast thickness, the distance from the top of the TMD to the top of the compression paddle was measured across the whole area (Fig. 4). The distance was measured by using a rod that was dropped into the hole at the top of the TMD. A fingernail was used to mark where the rod touched the top plate, the rod was then removed and the length of the rod from the bottom (where it touched the top of the compression paddle) up to the fingernail was measured using a ruler. This was repeated until the height of the rod for all the holes that covered the compression paddle in question had been measured. Row 1 was defined as the row parallel to the breast chest wall and closest to the breast chest wall. Column 1 was defined as the column perpendicular to the breast chest wall and out to the left side. Column 15 was then the last column on the right. A full set of thickness measurements (105) took approximately 20 min to conduct.

Mawdsley *et al.*[7] defined a reference point along the midline in the chest wall to nipple direction, 20 mm in from the chest wall side. They found that for most images the maximum height occurred at this reference point. We defined the same reference point in our study—hole in row 1, column 8 (located 2.5 cm from the breast chest wall side of the imaging table, and 18.0 cm from the short edge side).

Calculation of breast thickness

The measurements performed to find the readout and measured thickness of the phantom is illustrated in Fig. 5.

The readout thickness (d) is given by the following equation:

$$d = D - t \qquad (1)$$

where *D* is the system readout thickness including the thickness of the bottom plate. The thickness of the bottom plate (*t*) had to be subtracted from the total readout thickness (*D*) in order to obtain the readout thickness for the phantom (*d*). The measured thickness (*M*) of the object was calculated as follows:

$$M = H - t - p - 1 \qquad (2)$$

where *H* is the total height of the TMD, *p* is the thickness of the compression paddle, and *l* is the distance from the top of the compression paddle to the top of the TMD. Using a Vernier calliper, the thickness of the compression paddles (*p*) was measured to be 1.00 mm for Siemens Mammomat Inspiration and 2.75 mm for all the other paddles in this study. The area covering the compressed phantom (row 1 columns 3–13, row 2 columns 4–12, row 3 columns 6–10, and row 4 column 8) was defined as the breast area. The thickness for the area covering the compressed breast phantom was measured (breast area), and the minimum, maximum and average measured breast thickness for this area was compared to the readout thickness, and the difference between them were found, as follows-

$$\text{Percentage} = \frac{(\text{Average/min/ max measured breast area}) - \text{readout thickness}}{\text{readout thickness}}$$

A positive value implies that the measured thickness is larger than the readout thickness, which suggests the machine underestimates thickness. A negative value implies that the measured thickness is smaller than the readout thickness, which suggests the machine overestimates the thickness. An over- or underestimation is considered equally faulty, and a difference close to zero is preferred.

TMD - precision and observer variability

Prior to commencing the study, a precision and operator variability study was conducted. A wooden block (depth: 96 mm, width: 253 mm, and height: 55 mm) was placed inside the TMD device, centred in the middle and parallel to the long side of the TMD device. The thickness was measured three times by the person who would perform the thickness measurements. Average measured thickness was 55.5 mm, with a standard deviation of 0.4 mm across the whole area measured by the reader for all three measurements. The deviation in the measured thickness varied between 1 and 2 mm (only one measurement varied with 2 mm) with an average of 0.0460.12 mm (95% confidence interval). Concluding from this, this person would conduct the study with good precision. However, in the study itself 15% of the actual measurements were repeated on a blind sampling basis to

minimize random error. The average difference between the first measurement and the second measurement (blind testing) was 0.1760.07 mm (95% confidence interval). Concluding from this their precision and repeatability was more than adequate for this study.

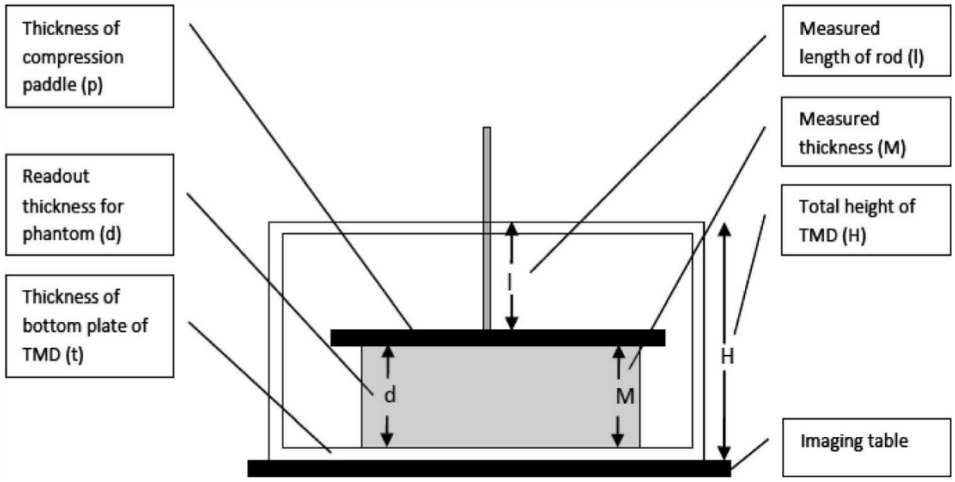


FIG. 5. Diagram to illustrate the measurements performed to calculate readout and measured thickness of the object.

Quality control: checking the readout thickness

In the United Kingdom (the location for all the mammography units in this study) the allowed difference between readout and measured thickness is 65 mm.[11] Each machine was tested every six months (Table I); all units were operating within manufacturer specification.

Quality control: checking the compression force

Accuracy of compression force is assessed on traceably calibrated scales and noted to an accuracy of 5 N every 6 months by a medical physicist and monthly by radiographers. The readout compression force is checked for 40, 80, and 120 N and also at maximum compression force (200 N). The accuracy of the readout compared to the measured compression force was 610 N (in accordance with IPEM 89 Ref. 11) for all the units.

Results

Figures 6 and 7 illustrate a 3D representation of the difference between the measured thickness and the readout thickness for a nonflexible and flexible paddle across the whole measured area. Since the primary interest is the variation across the breast area, and the average percentage difference in compressed breast thickness, the minimum percentage difference in breast thickness and the percentage difference between readout and measured thickness for the reference point are shown in Fig. 8.

Difference between measured and readout thickness across paddle area

The smallest and largest difference between the measured and readout thickness of the compressed phantom across the whole measured area of the paddle is shown in Fig. 6 for the 18x24 flexible paddle (smallest difference: 12 mm and largest difference: 19 mm) and Fig. 7 for the 18x24 nonflexible paddle (smallest difference: 3 mm and largest difference: 7 mm). The average difference between the smallest and largest measured thickness across the whole area was smaller for nonflexible paddles compared to flexible paddles (nonflexible/flexible 18x24: 5.0/16.0 mm, nonflexible/flexible 24x30: 5.3/10.0 mm). Figure 7 illustrates that the compression paddle may be uneven in the left to right direction.

The average, minimum, maximum percentage, and reference point percentage difference between measured compressed breast thickness and the readout compressed breast thickness for the breast area for the 18x24 paddle for 60 and 100 N applied compression force is shown in Fig. 8.

Figure 8 shows that there is a larger spread in the average percentage difference for the flexible than for the nonflexible compression paddle for both 60 N (range: 5.5%–6.8% (nonflexible), 4.5%–9.0% (flexible)) and 100 N (range: 8.0%–11.2% (nonflexible), 6.0%–26.0% (flexible)), and the difference is larger for 100 N than for 60 N applied compression force. For the nonflexible paddles Siemens Mammomat Inspiration (60 N: 1.0%, 100 N: 2.6%) came closest to 0% difference for the average percentage difference, and for the flexible paddle Hologic Selenia Dimensions (60 N: 1.5%) came closest to 0% difference when 60 N compression force was applied and GE Senographe Essential (100 N: 3.1%) came closest to 0% difference when 100 N compression force was applied.

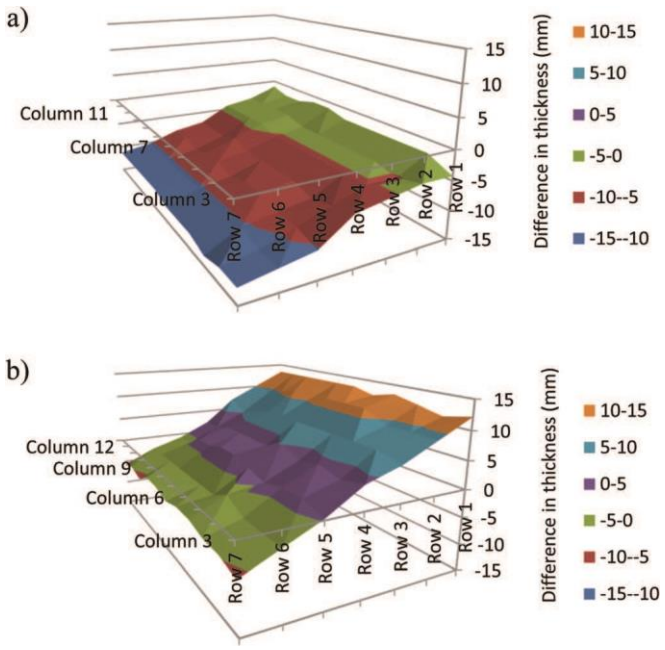


FIG. 6. Map of differences in thickness for the whole area for 18 cm x 24 cm flexible compression paddle for (a) Hologic Selenia Dimensions, which had the smallest (12 mm) difference in thickness across the whole area and (b) Hologic Lorad Selenia, which had the largest (19 mm) difference in thickness across the whole area, when applying 100 N compression force.

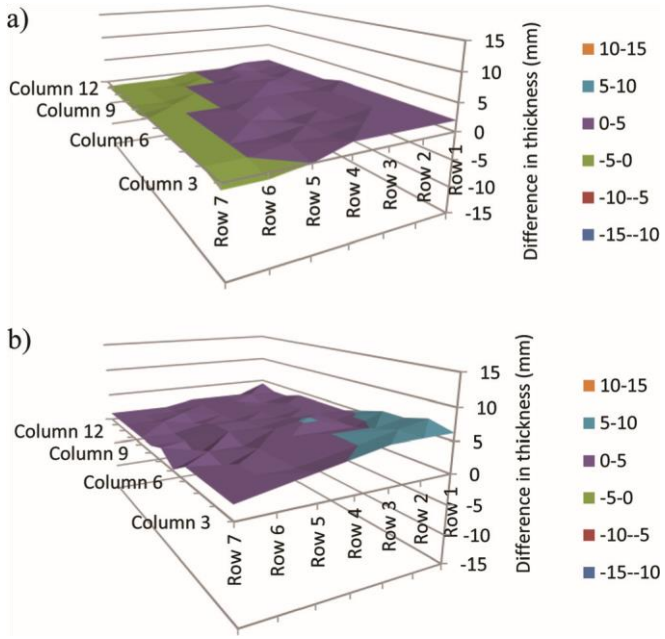


FIG. 7. Map of differences in thickness for the whole area for 18 cm24 cm nonflexible compression paddle for (a) Siemens Mammomat Inspiration, which had the smallest (3 mm) difference between measured and readout thickness across the whole area and (b) GE Senographe 800 T, which had the largest (7 mm) difference in measured and readout thickness across the whole area, when applying 100 N compression force.

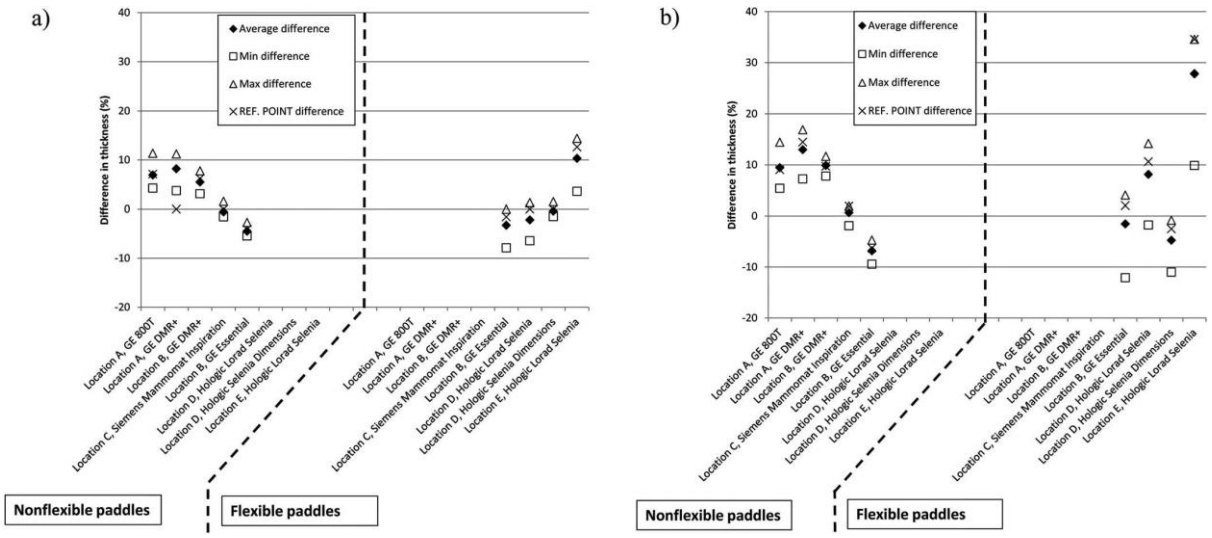


FIG. 8. The percentage difference between measured thickness and readout thickness for the breast area for 18 cm24 cm nonflexible and flexible compression paddle for (a) 60 N and (b) 100 N applied compression force.

Variation in thickness across breast area

The average, minimum, and maximum differences (measured in mm) for the compressed breast area is shown in Table II.

The difference between machine readout and measured thickness for nonflexible paddles for the breast area, applying 100 N compression force was smallest for the Siemens Mammomat Inspiration (18x24 paddle: +2.6% (p<0.01), 24x30 paddle: +0.7% (p=0.05)) and largest for GE Senographe DMR+ (18x24 paddle (location A): +11.2% (p<0.01), 24x30 paddle (location B): +14.3% (p<0.01)). For the 18x24 flexible paddle, and with an applied compression force of 100 N, the smallest difference between machine readout and measured thickness for the breast area occurred for GE Senographe Essential [3.1% (p<0.01)], and the largest for a Hologic Lorad Selenia [26.0% (p<0.01)]. For the 24x30 flexible paddle, and with an applied compression force of 100 N, the smallest difference between machine readout and measured thickness for the breast area occurred for a Hologic Lorad Selenia [3.0% (p<0.01)] and the largest difference occurred for the other Hologic Selenia Dimensions [8.9% (p<0.01)].

The average differences for both paddles, both compression forces (60 and 100 N) and all modalities in this study were +2.6% (60 N: +1.3%, 100 N: +2.8%).

In this study, two Hologic Lorad Selenia and two GE Essential DMR+ units were included. When comparing the results for the two units of equal manufacturer and model, it was found that the average difference between the readout thickness and the measured thickness for the breast area is different for the two units [GE DMR+: 11.2 vs 8.4% (18x24), 0.7 vs 14.3% (24x30), Hologic Lorad Selenia: 6.8 vs 26.0% (18x24), 3.0 vs 8.3% (24x30)].

Change in measured compressed breast thickness when increasing the compression force

When increasing the compression force from 60 to 100 N an 18% decrease in measured compressed breast thickness was observed for the breast area (18x24: 17.861.4%, 24x30: 17.765.4%) when using nonflexible paddles. When using flexible paddles a larger decrease in measured compressed breast thickness can be observed for the 18x24 paddles (18.662.6%) versus the 24x30 paddles (17.161.9%).

Reference point

The average difference for both compression forces, both paddles (nonflexible/flexible) and both paddle sizes between the measured thickness for the average breast area and the measured thickness for the reference point is 0.760.2 mm (in percentage: 1.460.5%).

TABLE II. Average, minimum and maximum difference in thickness (mm) for the breast area for the compression forces 60 and 100 N for the different mammography units included in this study.

	Compression Force 60 N				Compression force 100 N			
	Average difference mm (%)[a]	Min difference mm (%)[b]	Max difference mm (%)[c]	Ref. point difference mm (%)[d]	Average difference mm (%)[a]	Min difference mm (%)[b]	Max difference mm (%)[c]	Ref. point difference mm (%)[d]
Nonflexible paddle, 18x24								
Location A, GE 800T	4.1 (5.9)	2.3 (3.2)	7.3 (10.3)	4.3 (6.0)	4.5 (8.1)	2.3 (4.1)	7.3 (10.3)	4.3 (7.7)
Location A, GE DMR+	3.6 (6.8)	1.3 (2.3)	5.3 (9.8)	4.3 (7.9)	4.6 (11.2)	2.3 (5.4)	6.3 (15.1)	5.3 (12.7)
Location B, GE DMR+	2.8 (4.3)	1.3 (1.9)	4.3 (6.6)	3.3 (5.0)	4.3 (8.4)	3.3 (6.3)	5.3 (10.2)	4.3 (8.3)
Location B, GE Essential	2.8 (4.5)	0.8 (1.2)	5.8 (9.1)	1.8 (2.8)	1.5 (3.1)	1.3 (2.5)	14.8 (13.6)	0.3 (0.5)
Location C, Siemens Mammomat Inspiration	0.7 (1.0)	0.0 (0.0)	2.0 (3.1)	1.0 (1.6)	1.3 (2.6)	0.0 (0.0)	2.0 (3.8)	2.0 (3.8)
Nonflexible paddle, 24x30								
Location A, GE 800T	2.8 (5.0)	2.3 (4.1)	4.3 (7.7)	3.3 (5.9)	3.4 (7.7)	1.3 (2.8)	4.3 (9.6)	3.3 (7.3)
Location A, GE DMR+	3.9 (7.4)	3.3 (6.1)	5.3 (9.8)	4.3 (7.9)	0.3 (0.7)	0.8 (1.8)	1.3 (2.9)	1.3 (2.9)
Location B, GE DMR+	4.6 (9.7)	2.3 (4.7)	7.3 (15.3)	5.3 (11.1)	5.6 (14.3)	3.3 (8.2)	7.3 (18.4)	6.3 (15.7)
Location C, Siemens Mammomat Inspiration	0.1 (0.1)	1.0 (1.6)	2.0 (3.3)	0.0 (0.0)	0.3 (0.7)	1.0 (1.9)	2.0 (3.8)	1.0 (1.9)
Flexible paddle, 18x24								
Location B, GE Essential	2.8 (4.5)	0.8 (1.2)	5.8 (9.1)	1.8 (2.8)	1.5 (3.1)	1.3 (2.5)	6.8 (13.6)	0.3 (0.5)
Location D, Hologic Lorad Selenia	2.4 (3.2)	0.3 (0.3)	5.8 (7.4)	0.8 (1.0)	3.8 (6.8)	1.8 (3.1)	7.3 (12.8)	5.3 (9.3)
Location D, Hologic Selenia Dimensions	1.0 (1.5)	0.3 (0.4)	1.8 (2.6)	0.8 (1.1)	3.6 (6.0)	1.3 (2.1)	7.3 (12.3)	2.3 (3.8)
Location E, Hologic Lorad Selenia	5.0 (9.0)	1.3 (2.3)	7.3 (13.1)	6.3 (11.3)	10.5 (26.0)	3.3 (8.0)	13.3 (32.7)	13.3 (32.7)
Flexible paddle, 24x30								
Location B, GE Essential	2.9 (4.4)	1.8 (2.7)	3.8 (5.8)	2.8 (4.2)	3.8 (7.0)	2.8 (5.1)	4.8 (8.7)	2.8 (5.1)
Location D, Hologic Lorad Selenia	4.1 (4.9)	2.8 (3.3)	5.8 (6.8)	3.8 (4.4)	2.0 (3.0)	1.8 (2.6)	4.3 (6.4)	3.3 (4.9)
Location D, Hologic Selenia Dimensions	4.8 (8.9)	1.8 (2.9)	2.8 (4.5)	1.8 (2.9)	4.8 (8.9)	2.3 (4.2)	8.3 (15.3)	2.3 (4.2)
Location E, Hologic Lorad Selenia	0.2 (0.3)	1.3 (1.9)	1.8 (2.6)	1.3 (1.9)	4.5 (8.3)	1.3 (2.3)	7.3 (13.3)	6.3 (11.5)

[a] Average difference: average difference between measured and readout thickness across the area defined as the breast area.
[b] Min difference: minimum difference between measured and readout thickness across the area defined as the breast area.
[c] Max difference: maximum difference between measured and readout thickness across the area defined as the breast area.
[d] Ref. point difference: difference between measured and readout thickness for the hole defined as the reference point (row 1, column 8).

Discussion

For all machine and paddle combinations the readout breast thickness was different to; reference point thickness, average thickness, minimum thickness, or maximum thickness. This resulted in the measured thickness being over-estimated and also under-estimated. The difference was more marked at 100 N compared with 60 N, suggesting that as force increases the error in thickness readout also increases. At 100 N and 18x24 paddle, only 2 (Location B GE Essential/18x24 flexible; Location C, Siemens Mammomat Inspiration/18x24/24x30 nonflexible) out of 9 machines (22%) gave reference point and average values for the breast area that were within 65% of the readout thickness. Flexible paddles had greater departure from measured thickness when compared with nonflexible paddles.

Quality control and tolerance data supplied by manufacturers

The results for the average difference in compressed breast thickness for the breast area was compared to the maximum difference in measured thickness (for phantom of known thickness) and readout thickness from the annual quality control. Only two units (GE Senographe DMR+ (Location A) and GE Senographe Essential) of the eight units (25%) were found to have an average difference between measured and readout thickness within the maximum difference found at the annual quality control. For the Hologic Lorad Selenia at Location D the average difference was larger than the difference between measured and readout thickness from the quality control for both paddles and both compression forces. For the other units (GE Senographe 800T, GE Senographe DMR+ (Location B), Siemens Mammomat Inspiration, Hologic Selenia Dimensions and Hologic Lorad Selenia (Location E)) discrepancies were found for 18x24 and/or 24x30 paddle and/or for both compression forces (60 and 100 N). The results in this study show that the test performed annually by the medical physicist might not be adequate to reveal discrepancies between the measured and the readout thickness.

Our measurements for the compressed breast thickness were compared to the tolerance data stated in the operator manuals supplied by the different manufacturers. For GE Senographe 800T and GE Senographe DMR+ our results were within the tolerance limits of 610 mm stated in the operator manuals. Hologic Lorad Selenia user manual states that compression thickness accuracy should be 60.5 cm for thicknesses between 0.5 and 15 cm. This was found to be true for one of the Hologic Lorad Selenia units (difference in measured and readout thickness for average breast area: 3.8 mm), but not for the other unit [difference in measured and readout thickness for average breast area: 10.5 mm (18x24)], when the 18x24 paddle was used and 100 N compression force was applied. For GE Senographe Essential the difference between the measured and readout thickness for the breast

area was within the tolerance limit (610 mm). Had the tolerance limit been 65 mm, in other words the same as for Hologic Lorad Selenia/Hologic Selenia Dimensions, the results for the minimum difference between measured and readout thickness for the 18x24 paddles (nonflexible and flexible), when 100 N compression force was applied, would have also been within the limits.

To calibrate the readout thickness Siemens uses a 42 mm phantom and compresses the object using a 70 N compression force. The readout thickness should read between 39 and 45 mm. If not a recalibration is performed.

A calibration of the Hologic Lorad Selenia is performed by compressing a 5 cm thick phantom (BR-12, CIRS, Norfolk, VA). A compression force of 133.5 N is applied, and then the compression thickness is calibrated for the installed paddle/receptor combination.

For Hologic Selenia Dimensions most of the calibration is done automatically. A 2 and 8 cm thick phantom (BR-12) is compressed by applying 133.5 N compression force, and the machine will then register the thickness of the phantom. For the "FAST" paddle (the flexible paddle) the same approach is taken, but without any compression. The paddle is just lowered until it touches the phantom, and the machine is told that this is 2 or 8 cm. The fact that a rigid phantom is used for this test is probably not optimal, because a tilt will probably occur. Maybe one needs to rethink how the thickness is measured, or maybe a different approach to how the paddle is constructed needs to be addressed.

GE also has routines for the calibration of the thickness, but the calibration routines are propriety.

Reference point

The difference between readout and measured thickness for the reference point and the average breast area values are similar [0.760.2 mm (in percentage: 1.4±0.5%)], suggesting that a simplistic one-point of sample could be used for accurate estimation of average breast thickness. This approach would involve sampling only at the reference point, which would mean that the measuring time for the thickness would decrease drastically (from a maximum of 105 measurements down to one). We found that there is a large variation in the chest wall to nipple direction, and a smaller lateral variation, in accordance with Diffey *et al.*[10] A better estimate would therefore be to measure the thickness for the points/holes outlining the breast area; in this way, a better average for the compressed breast thickness could be measured.

Where Diffey *et al.*[10] found for real breasts an underestimation of thickness of as much as 21.2 mm in the chest to nipple direction, our results show a maximum underestimation of 13 mm for a Hologic Lorad Selenia mammography machine, and a maximum overestimation of 8 mm for a Hologic Selenia

Dimensions mammography machine. If one takes into consideration this under-/overestimation of thickness only (and not the fact that a change in the thickness might also have implications for the choice of target/filter combination and kV), the MGD can be estimated. For a Hologic Lorad Selenia, for instance, an underestimation of 13 mm would imply a smaller estimated MGD of 17% for a thin breast (readout thickness 35 mm) and 9% for a thick breast (readout thickness 80 mm). An underestimation of thickness will in general imply that the MGD originally estimated is too large, and thus overestimate the MGD and the risk. For a Hologic Lorad Dimensions an overestimation of 8 mm would imply a larger estimated MGD of 20% for a thin breast (readout thickness 31 mm) and 6% for a thick breast (readout thickness 79 mm). An overestimation of thickness will in general imply that the MGD originally estimated is too small, and thus underestimate the MGD and the risk.

Correction factor

Varying paddle/machine combinations give different error levels between readout thickness and measured thickness. Correction factors may be applied, in order to obtain higher accuracy clinically. The correction factor can be found by dividing the measured thickness with the readout thickness for different manufacturers/models, different paddle sizes (in this study: 18x24 and 24x30) and

Study limitations

Preservation of breast phantom integrity limited our experiment to a maximum pressure force of 100 N. We propose that a more resilient breast phantom should be used across a broader range of clinically representative force values (e.g., 60 N stepping 10 to 150 N). This would provide a better understanding on how bend and distortion may vary across the higher end of the normal clinical pressure range. In this study the effect of different breast volumes or breast densities was not considered; extending these variables might be considered, as bend and distortion may be affected by them.

A further limitation in this study is the fact that a different readout thickness was achieved every time the measurements were repeated. When compressing the phantom, different thicknesses were achieved every time; as such the results are not reproducible. Positioning error was reduced by trying to position the phantom approximately in the middle of the compression paddle (along midline), but the compressed thickness still altered.

Tyson *et al.*[9] devised a method for determining the compressed breast thickness that had a thickness determination accuracy of better than 1 mm, and a measurement accuracy of better than 0.2 mm. The method described here will lead to a larger inaccuracy than the method described by Tyson *et*

al.[9] Tyson *et al.*[9] state that a mean accuracy of better than 1 mm is required to make good estimates for the volumetric breast density. It was not possible with the device used in this study to obtain such a precision, but as for use in a busy clinically environment the TMD can be used to determine the difference in measured and readout thickness.

Clinically adaptable method

In theory this method can be applied for real breasts in a clinic to measure the real compressed breast thickness for the breast. The breast must be placed inside the TMD, in the same fashion as the phantom, compression must be applied and the compressed breast thickness must be measured. Because of the time span (20 min) for measuring the compressed breast thickness in this study, it will probably be necessary to limit the number of measurements performed to only one point (e.g., the reference point). The breast must then be recompressed (applying the same compression force) in order to obtain the actual image. This last step will probably be difficult to accomplish, since it has been shown to be difficult to obtain the same thickness applying the same compression force when compressing an object similar to a breast.

Conclusion

The difference in the readout thickness and the measured thickness varies between units for the same model and between manufacturers. Individual correction factors for breast thickness may need to be established for each dependent on paddle selection and compression force applied. Any corrections to compressed breast thickness need therefore to be performed for the unit in question, and one cannot assume that the correction in compressed breast thickness applies to all mammography machines of the same model.

Acknowledgements

The authors would like to thank Steve Curtis from the Guitar Repair Workshop (Manchester, United Kingdom) for making the TMD.

[a]Electronic mail: ingrid-helen.ryste-hauge@hioa.no

[1] D. R. Dance, "Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose," Phys. Med. Biol. 35, 1211–1219 (1990).
[2] D. R. Dance, C. L. Skinner, K. C. Young, J. R. Beckett, and C. J. Kotre, "Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol," Phys. Med. Biol. 45, 3225–3240 (2000).
[3] R. P. Highnam, J. M. Brady and B. J. Shepstone, "Estimation of compressed breast thickness during mammography," Br. J. Radiol. 71, 646–653 (1998).
[4] N. F. Boyd, H. Guo, L. J. Martin, L. Sun, J. Stone, E. Fishell, R. A. Jong, G. Hislop, A. Chiarelli, S. Minkin, and M. J. Yaffe, [3 3] "Mammographic density and the risk and detection of breast cancer," N. Engl. J. Med. 356, 227–236 (2007).
[5] J. J. Heine, K. Cao, and J. A. Thomas, "Effective radiation attenuation calibration for breast density: compression thickness influences and correction," Biomed. Eng. Online 9, 73 (2010).
[6] N. Boyd, L. Martin, A. Gunasekara, O. Melnichouk, G. Maudsley, C. Peressotti, M. Yaffe, and S. Minkin, "Mammographic density and breast cancer risk: evaluation of a novel method of measuring breast tissue volumes," Cancer Epidemiol. Biomarkers Prev. 18, 1754–1762 (2009).
[7] G. E. Mawdsley, A. H. Tyson, C. L. Peressotti, R. A. Jong, and M. J. Yaffe, Accurate estimation of compressed breast thickness in mammography," Med. Phys. 36, 577–586 (2009).

[8] A. Burch and J. Law, "A method for estimating compressed breast thickness during mammography," Br. J. Radiol. 68, 394–399 (1995).
[9] A. H. Tyson, G. E. Mawdsley, and M. J. Yaffe, "Measurement of compressed breast thickness by optical stereoscopic photogrammetry," Med.Phys. 36, 569–576 (2009).
[10] J. Diffey, A. Hufton, C. Beeston, J. Smith, T. Marchant, and S. Astley, "Quantifying breast thickness for density measurement," in *Digital Mammography, 9th International Workshop, IWDM 2008*, Tucson, AZ, edited by E. A. Krupinski (Springer-Verlag, Berlin/Heidelberg, 2008), pp. 651–658.
[11] A. C. Moore, D. R. Dance, D. S. Evans, C. P. Lawinski, E. M. Pitcher, A. Rust, and K. C. Young, *IPEM Report 89: Commissioning and Routine Testing of Mammographic X-Ray Systems* (Institute of Physics and Engineering in Medicine, 2005).

Paper 2

Tissue bulge during stereotactic core biopsy

L. Hackney[a,*], S. Williams [a], P. Hogg [b], K. Szczepura [b]

a Department of Breast Imaging, University Hospital of North Staffordshire, UK b University of Salford, Manchester, UK

Article history: Received 23 April 2013, Received in revised form 27 June 2013, Accepted 30 June 2013, Available online 26 July 2013

Keywords: Stereotactic, Breast, Biopsy, Image quality, Tissue bulge

Abstract

In full field digital mammography (FFDM) the whole breast is subjected to compression with a perspex compression paddle in order to reduce breast thickness and improve image quality. Once a mammographic abnormality has been detected using FFDM and a decision to proceed with a stereotactic (X-ray) guided core biopsy has been made, a different compression paddle is utilised. This paddle has a central aperture in order to allow access to the lesion for biopsy.

Clinical observations made during biopsy procedures have revealed that a bulge of tissue forms within the aperture. The magnitude of the bulge of tissue and BI-RAD breast density was recorded in 15 consecutive patients. Results showed an average of 18.7% (range 11.3e30%) increase in the breast thickness (over the bulge region) compared to the surrounding compressed breast.

BI-RAD breast density category 3 had on average the lowest measured thickness and the greatest percentage of tissue bulge. Overall, results confirm that for all patients there was a measurable tissue bulge that varied from 6 mm to 10 mm, representing between 10.14% and 23.08% of additional tissue not measured by the machine. In clinical practice a perceivable difference in lesion visibility was subjectively indicated between the FFDM images and the stereotactic scout biopsy image.

The suggested hypothesis from these observations is that there may be an association between the magnitude of the tissue bulge and the ability to accurately perceive certain lesions during stereotactic biopsy procedures. A phantom study is in progress to determine how lesion visibility varies with the amount of tissue bulge.

Introduction

The UK NationalHealth Service Breast Screening Program(NHSBSP) was implemented to detect breast cancer at an early stage, with a subsequent associated reduction in mortality.[1] The mammographic signs are often subtle, minimal and confined to a small range of tissue densities.[2,3] Mammographicimagequalityenhancementisdetermined by several factors and breast thickness reduction is considered to be a fundamental factor in achieving this.[4,5] In full field digital mammography (FFDM) the whole breast is subjected to compression with a Perspex compression paddle in order to reduce the breast thickness.

Once a mammographic abnormality has been detected using FFDM and a decision to proceed with a stereotactic (X-ray) guided core biopsy has been made, a different compression paddle is utilised. This paddle is smaller and has a central aperture in order to allow access to the lesion for biopsy (Fig. 1). The area of concern containing the abnormality is placed as centrally as possible within

the aperture and the tissue surrounding the lesion is compressed to immobilise the breast and reduce tissue thickness.

Clinical observations made during biopsy procedures have revealed that a bulge of tissue forms within the aperture (Fig. 1). It was also noted that some lesions become less conspicuous when using the biopsy paddle compared with images obtained using the standard mammography paddles.

A measuring instrument (Fig. 2) was developed to determine tissue bulge magnitude. As seen in Fig. 2, the instrument comprised a metal tool with a scale in millimetres and a plastic plunger. When the measuring instrument is placed on the outer lip of the biopsy compression paddle the plunger can be lowered to touch the upper surface of the breast bulge. At this point the measurement of breast bulge magnitude can be taken from the scale on the measuring instrument. Using the instrument, 15 consecutive patients had measurements taken to determine bulge magnitude.

Mammography units routinely provide breast thickness measurements which normally show the thickness of the breast under compression; for the biopsy paddle these measurements would represent the tissue immediately surrounding the area of the aperture.

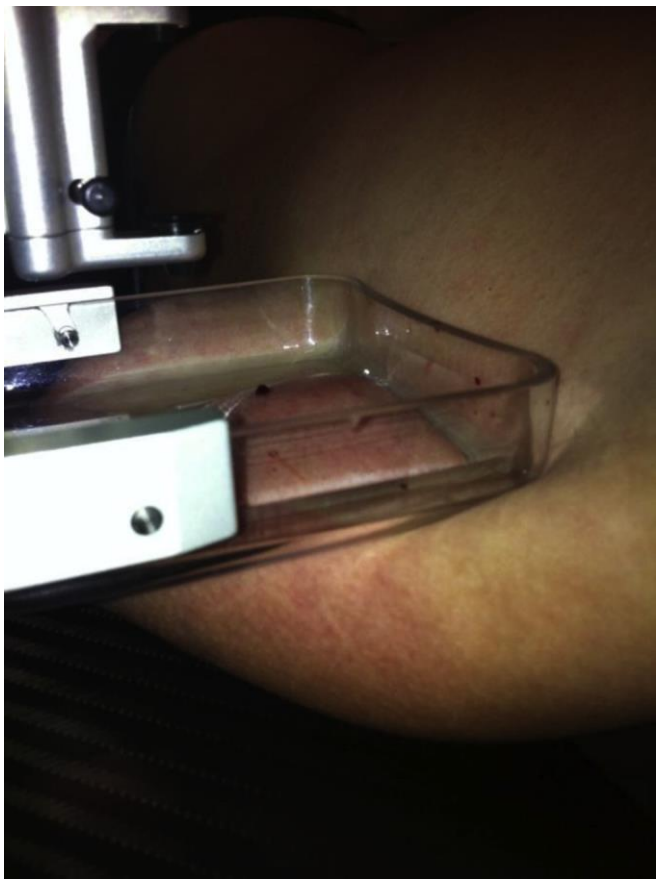


Figure 1. Typical breast bulge observed during Stereotactic needle core biopsy procedure.

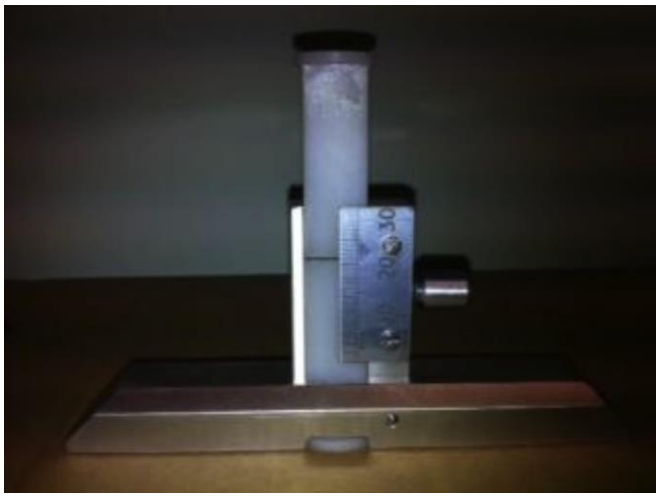


Figure 2. Instrument devised to measure the breast bulge.

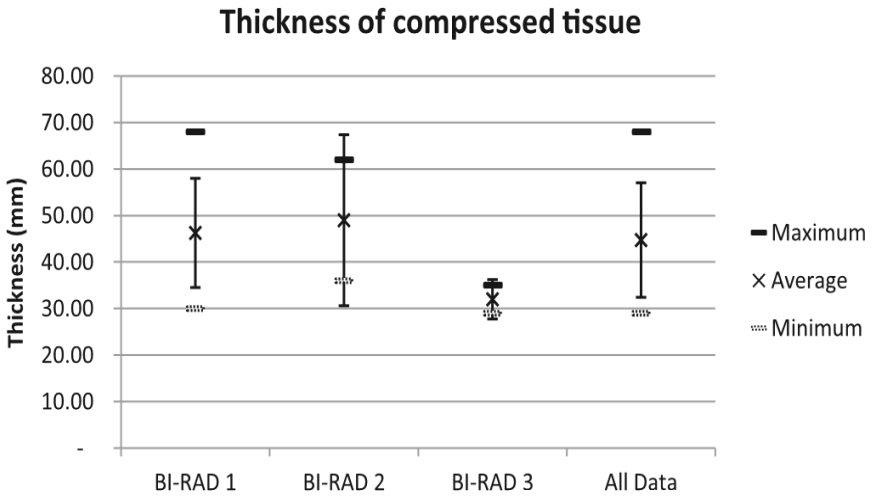
Also recorded was the BI-RAD density classification[6] as this indicates the overall percentage of glandular tissue of the breast. There is a recognised relationship between breast density and perception of abnormalities.[7]

The thickness of breast given by the mammography unit was compared to those corrected for bulge magnitude through use of the measuring instrument. The results (Table 1) of the 15 patients show

an average of 18.7% (range 11.3e30%) increase in the breast thickness (over the bulge region) compared to the surrounding compressed breast.

BI-RAD classification for breast density (Type1e4)	Machine recorded thickness (mm)	Additional depth of tissue (bulge) (mm)	Overall tissue depth (mm)	% Of additional tissue (bulge)
1	55.00	7.00	62.00	11.76
1	68.00	8.00	76.00	12.73
1	30.00	9.00	39.00	12.73
1	46.00	10.00	56.00	14.55
1	55.00	7.00	62.00	16.33
1	55.00	8.00	63.00	18.42
1	38.00	7.00	45.00	18.42
1	45.00	10.00	55.00	21.74
1	38.00	7.00	45.00	22.22
1	30.00	7.00	37.00	23.33
1	49.00	8.00	57.00	30.00
2	62.00	7.00	69.00	11.29
2	36.00	6.00	42.00	16.67
3	29.00	8.00	37.00	22.86
3	35.00	8.00	43.00	27.59

Table 1 Additional measured tissue for 15 Patients.

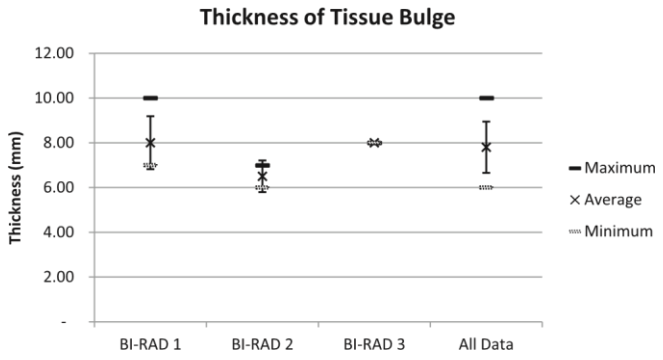


Graph 1. Demonstrating the range of thickness of compressed breast tissue relative to breast density.

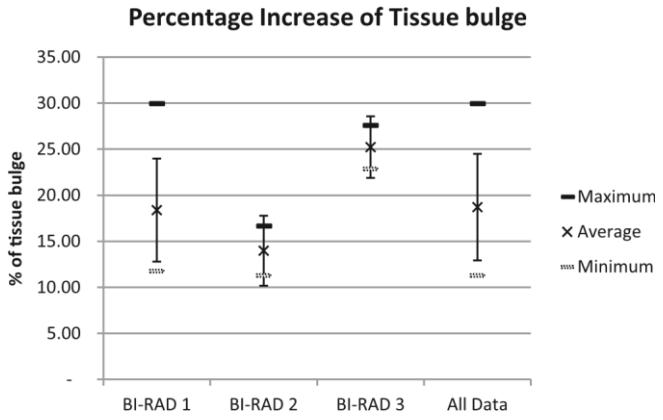
Table 1 demonstrates tissue bulge measurements from 15 patients undergoing stereotactic core biopsy. BI-RAD categories were assigned to determine percentage of glandular tissue. There were no patients classified as BI-RAD category 4, this was most likely due to the fact that this represents very dense tissue and is uncommon in the breast screening age group.

Graph 1 demonstrates that BI-RAD 3 category had on average the lowest measured breast thickness. Graph 2 demonstrates that BI-RAD 3 category had the greatest amount of additional tissue, leading to the greatest percentage of additional tissue; this is shown in Graph 3. Any difference between BI-RAD categories is difficult to establish due to the small number of patients. However Graphs 1e3

present data which suggests that for all patients there was tissue bulge, and this varied from 6 mm to 10 mm, representing between 11.29% and 30.00% of additional tissue.



Graph 2. Demonstrating the range of tissue bulge thickness relative to breast density.



Graph 3. Demonstrating the % increase of tissue bulge relative to breast density.

The additional breast thickness (bulge) would not be measured/ detected by the mammography machine. Machine detected breast thickness is essential for auto-setting of exposure factors, notably kVp. Having an inaccurate thickness could result in suboptimal exposure factors being used; this might affect image quality and lesion visibility.

Lesion visibility was compared between the FFDM screening images and the stereotactic scout biopsy images. The images were viewed under standardised reporting conditions by two Consultant Radiographers. A perceivable difference was noted in lesion visibility, with the images used for biopsy procedures showing the lesions less clearly. This has led to the suggestion that there may be an association between the tissue bulge and the clinicians ability to accurately perceive and target certain lesions during stereotactic biopsy procedures. Currently a phantom study is being conducted to determine how lesion visibility varies as bulge increases.

References

Schopper D, De Wolf C. How effective are breast cancer screening programmes by mammography? Review of the current evidence. European Journal of Cancer 2009;45(11):1916e23.
Houssami N, Given-Wilson R, Ciatto S. Early detection of breast cancer: overview of the evidence on computer-aided detection in mammography screening. Journal of Medical Imaging and Radiation Oncology 2009;53:171176.
Zanello PA, Robim AFC, de Oliveira TMG, Junior JE, Moreira de Andrade J, Monteiro CR. Breast ultrasound diagnostic performance and outcomes for mass lesions using Breast Imaging Reporting and Data System category 0 mammogram. Clinics 2011;66(3):443e8.
Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? Australasian Radiology 2003 Jun;47(2):121e6.
Poulos A, McLean D. The application of breast compression in mammography: a new perspective. Radiography May 2004;10(2):131137.
American College of Radiology. Breast imaging and reporting data system. 5th ed. Reston V A: American College of Radiology; 2004.
Pinker K, Perry N, Vinnicombe S, Shiel S, Weber M. Conspicuity of breast cancer according to histopathological type and breast density when imaged by full-field digital mammography compared with screen-film mammography. European Radiology January 2011;21(1):18e25.

Paper 3

Breast tissue bulge and lesion visibility during stereotactic biopsy phantom study

S. Williams [a,*], L. Hackney[a], P. Hogg [b], K. Szczepura [b]

a University Hospital of North Staffordshire, United Kingdom b University of Salford, United Kingdom

Article history: Received 18 March 2014, Accepted 24 April 2014 Available online 28 May 2014

Keywords: Image quality, Stereotactic core biopsy, Breast imaging, Mammography

Abstract

Background: During mammography guided stereotactic breast biopsy a bulge of tissue can form in the paddle needle biopsy aperture. This bulge has been estimated to have a height of up to 30% of the breast itself. During clinical biopsy we have noticed that lesions can appear to be less visible when tissue bulges are evident. This can make biopsy more difficult in some cases.

Objectives: This experiment investigates how lesion visibility varies with breast bulge magnitude. Method: Using a phantom to represent breast and breast bulge, lesion visibility was assessed using a two alternative forced choice methodology. To mimic clinical conditions, imaging was performed on a full field digital mammography system with the biopsy paddle attached using an automatic exposure device. Organ dose (breast) was estimated.

Results: As breast bulge increases lesion visibility decreases; organ dose increases as breast bulge magnitude increases.

Conclusion: Consideration should be given to the impact of breast bulge magnitude and lesion visibility when performing image guided biopsy.

Advances in knowledge: The authors found no similar studies and the results of this study demonstrate a potential clinical risk.

Introduction

Associated with the establishment of the National Health Breast Screening Program (NHSBSP) and continuing technological advancements in mammographic imaging systems, smaller and more

subtle breast abnormalities are being detected, however various factors confound this.[1,2] The ability to achieve an accurate histo-pathologic diagnosis is fundamental, and often this involves tissue sampling through ultrasound or mammography guided biopsy. Stereotactic (mammography) guided needle core biopsy is a well-established technique for sampling nonpalpable breast lesions and sensitivity/sensitivity rates can be very high.[3,4]

In stereotactic biopsy we have noticed perceivable differences in lesion visibility between the full field digital mammographic (FFDM) views obtained during diagnostic work-up and the small field images used during stereotactic biopsy. Lesions seen on small field images during biopsy procedures can be harder to see in the clinical room compared with the same lesions acquired using FFDM in the reporting room. This reduction in lesion visibility can make biopsy harder in some cases. Image quality is dependent on many factors; these include type, design and performance of the imaging equipment. Differences also exist between FFDM and small field images for biopsy. Display differences also exist e lower resolution screens tend to be used for biopsy, whereas higher resolution screens are used for diagnosis. A further confounding factor could be related to breast tissue bulge.[5]

In 2013 Hackney et al. described a bulge of breast tissue which can form within the aperture of the paddle during biopsy procedures.[5] An average of 18.7% (range 11.3e30%) increase in breast thickness over the bulge region occurred compared to the surrounding compressed breast. Hackney hypothesised that breast bulge might diminish lesion visibility during stereotactic biopsy procedures. Despite an intensive literature search no other publications were found on this phenomenon. Using a phantom, this paper builds on Hackney's tissue bulge work to determine whether an association does exist between bulge magnitude and lesion visibility.

Materials and methods

A phantom study was conducted to simulate clinical conditions. Phantom design was informed by an audit of compressed breast sizes.[5] Perceptual measures of lesion visibility were used to assess the effect of tissue bulge across a range of breast phantom/breast bulge thicknesses.

Breast phantom

To inform breast phantom size, breast thickness readout from an FFDM mammography machine (Siemens Inspiration system (Siemens PLC, Berkshire)) was noted on a convenience sample of 100 female clients for right Medio-Lateral Oblique (MLO), left MLO, right Cranio-Caudal (CC) and left CC (Table 1). Minimum breast thickness was 21 mm, maximum 95 mm, mean 56.6 mm, SD 13.47.

Perspex has similar scattering and attenuation characteristics to human soft tissue.[6,7] Perspex was therefore used for breast phantom and breast bulge construction. Four Aluminium disks of different thicknesses (5 mm diameter; thicknesses e 0.001, 0.002, 0.004 and 0.009 mm) were used to simulate lesions, as these imitate high density lesions.[8,9]

Using data from Table 1, the breast phantom was constructed. The phantom comprised of Perspex blocks measuring 205 mm (L) 105 mm (W) 10 mm (D). These could be stacked to represent different breast thickness. The four aluminium discs were encased between the two Perspex blocks (Fig. 1).

Data from Hackney's study[5] were used to determine phantom bulge magnitudes. Perspex, measuring 53 mm 41 mm 5 mm (the size of the biopsy paddle aperture), were placed directly over the aluminium disks located within the biopsy compression paddle aperture (Image 1).

Phantom imaging

Siemens Inspiration (Siemens PLC, Berkshire) FFDM was used in conjunction with a digital X-ray stereotactic breast biopsy system (Siemens PLC, Berkshire). Both met the quality control standards for screening and assessment.[10]

The compression plate used for the stereotactic views was a 3D biopsy compression paddle measuring 100 mm (L) 94 mm (W) with a 55 mm (W) 44 mm (L) aperture (Fig. 2); the standard compression paddle measured 245 mm (L) 175 mm (W).

Normal clinical practice was followed to replicate clinical conditions with the exposure factors being auto-selected by the equipment (Table 2).

Stereotactic scout image

As there are a number of confounding factors influencing image quality, the decision was taken to use the stereotactic scout view for all test images therefore minimising the number of variables. In the scout view the X-ray beam is perpendicular to the phantom whereas the stereotactic image pairs involve a 10 angulation of the tube head to the left and right. This change in beam angulation could result in the beam passing through a variation in bulge thickness, which in turn may affect Aluminium

disc visibility. Also, the aluminium discs may be distorted with the change in angle of the X-ray beam, again this may also influence disc visibility.

	Right MLO	Left MLO	Right CC	Left CC
Minimum	24 mm	21 mm	27 mm	25 mm
Maximum	95 mm	91 mm	91 mm	86 mm
Mean	57.22	58.21	55.42	55.69
SD	13.88	14.64	12.46	12.91

Table 1 Breast thickness for all views.

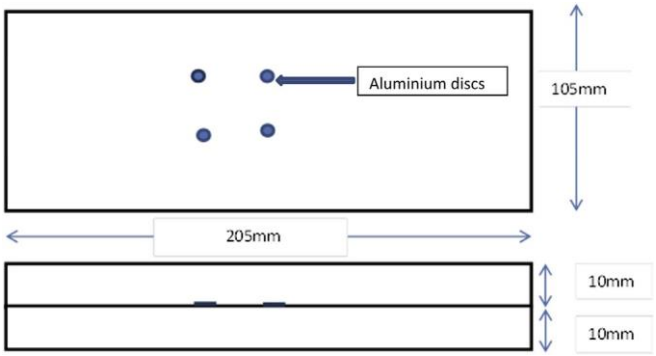


Figure 1. Top e plan view; bottom e side view. Both demonstrate the four aluminium disks encased by the two perspex blocks.

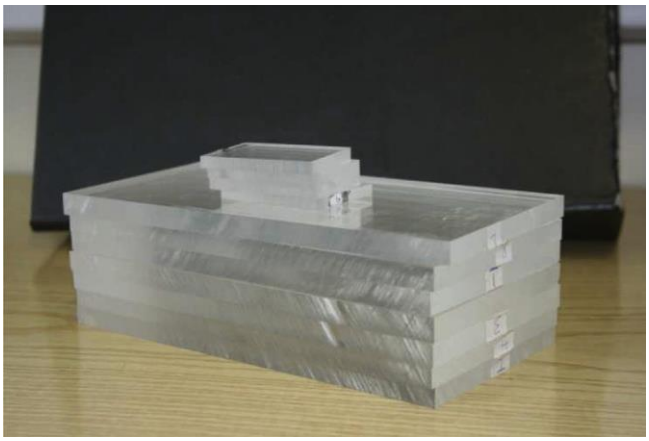


Image 1. Perspex blocks which were utilised for varying breast and tissue bulge thicknesses.

Images were acquired in the CC plane. The phantom was placed on the detector so the long edge (205 mm) was in line with the chest wall edge and central to the unit. The four aluminium discs were positioned within the biopsy aperture of the compression paddle. The phantom and the additional Perspex blocks were numbered and always orientated in the same direction for each set of images. For all images, the compression setting (Table 2) was constant (75 N).

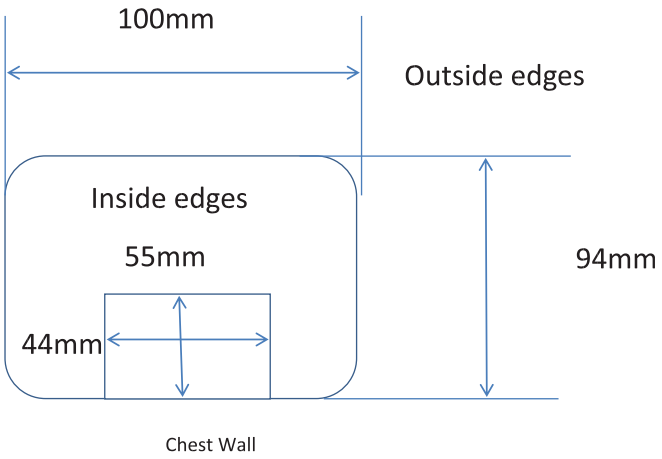


Figure 2. Representation of the biopsy compression paddle.

Image set	Phantom breast thickness (mm)	Bulge thickness (mm)	kV	mAs	Exposure time (ms)	Compression thickness (mm)	Compression force (N)	Anode/filter	Organ dose (mGy)	Ent. dose (mGy)
Set 1	40	0	27	36	572	38	78	Mo/Rh	0.89	2.56
	40	5	27	45	573	38	76	Mo/Rh	1.12	3.21
	40	10	27	59	574	38	79	Mo/Rh	1.46	4.17
	40	15	27	75	576	38	79	Mo/Rh	1.85	5.30
Set 2	50	0	28	51	579	48	74	Mo/Rh	1.19	4.24
	50	5	28	64	579	48	73	Mo/Rh	1.5	5.33
	50	10	28	82	651	48	73	Mo/Rh	1.92	6.84
	50	15	28	103	812	48	73	Mo/Rh	2.42	8.64
	50	20	28	129	1003	48	73	Mo/Rh	3.02	10.79
Set 3	60	0	29	70	581	57	78	Mo/Rh	1.62	6.82
	60	5	29	87	710	58	76	Mo/Rh	1.97	8.49
	60	10	29	110	893	57	78	Mo/Rh	2.55	10.75
	60	15	29	137	1100	57	79	Mo/Rh	3.17	13.37
	60	20	29	170	1356	57	78	Mo/Rh	3.93	16.59
	60	25	29	206	1637	57	79	Mo/Rh	4.77	20.14
Set 4	70	0	30	95	802	67	77	Mo/Rh	2.15	10.63
	70	5	30	117	977	67	77	Mo/Rh	2.65	13.09
	70	10	30	146	1210	67	77	Mo/Rh	3.31	16.34
	70	15	30	175	1441	67	76	Mo/Rh	3.97	19.61
	70	20	30	217	1782	67	77	Mo/Rh	4.94	24.38
	70	25	30	250	2044	67	76	Mo/Rh	5.68	28.06
	70	30	30	330	2687	67	77	Mo/Rh	7.51	37.10
Set 5	80	0	33	82	769	77	72	Mo/Rh	2.29	12.90
	80	5	31	149	1276	77	72	Mo/Rh	3.38	19.32
	80	10	31	182	1553	77	73	Mo/Rh	4.14	23.63
	80	15	31	223	1888	77	72	Mo/Rh	5.06	28.91
	80	20	31	269	2268	77	72	Mo/Rh	6.11	34.87
	80	25	31	318	2670	77	72	Mo/Rh	7.21	41.18
	80	30	31	394	3305	77	72	Mo/Rh	8.96	51.14
	80	35	33	269	2413	77	72	Mo/Rh	7.51	42.34
Set 6	90	0	32	212	1477	86	77	W/Rh	2.31	10.67
	90	5	32	252	1756	86	77	W/Rh	2.76	12.73
	90	10	32	306	2118	86	77	W/Rh	3.34	15.41
	90	15	32	354	2449	86	77	W/Rh	3.88	17.86
	90	20	32	437	3012	86	77	W/Rh	4.78	22.02
	90	25	32	533	3665	86	77	W/Rh	5.83	26.86
	90	30	34	349	2565	86	77	W/Rh	4.5	19.26
	90	35	34	415	3037	86	77	W/Rh	5.34	22.86
	90	40	34	510	3730	86	77	W/Rh	6.58	28.14
Set 7	100	0	35	158	1218	96	77	W/Rh	2.03	9.87
	100	5	35	185	1417	97	73	W/Rh	2.36	11.58

100	10	35	215	1642	97	73	W/Rh	2.75	13.47
100	15	35	252	1920	96	78	W/Rh	3.25	15.74
100	20	35	284	2158	97	73	W/Rh	3.63	17.81
100	25	35	344	2604	96	78	W/Rh	4.43	21.48
100	30	35	415	3127	96	77	W/Rh	5.33	25.85
100	35	35	500	3763	96	77	W/Rh	6.43	31.18
100	40	35	462	3479	96	77	W/Rh	5.94	28.80
100	45	35	457	3447	97	77	W/Rh	5.84	28.64

Table 2 Phantom characteristics and dose data.

For stereotactic biopsy settings the small collimation field was active. The grid was in the out position. The system was configured to remove the grid automatically for stereotactic images as grid cutoff would occur when acquiring the stereotactic image pair which involves 10 angulation of the X-ray beam across the detector. In the advanced settings the auto-decompression was switched off, and the Opcomp[11] was switched on. Opcomp informed the automatic exposure parameters e kV, mAs, anode/filter.

The automatic exposure control (AEC) segmentation[11] was switched on. The AEC calculates the mAs based on the pixel value received in the prepulse. The purpose is to optimise exposure parameters breast size and composition, whilst achieving optimal image contrast. A total of 7 sets of phantom images were taken (see Table 2); dose information was also recorded.

The number of image sets was determined by the phantom plus the number of additional Perspex blocks, up to a maximum of 100 mm as informed by the breast thickness study. In our study we included a much larger bulge size than measured by Hackney et al. Bulge magnitude was limited by the dimension of the Perspex pieces which only allowed 5 mm increments (Image 2).

Visual appraisal

Images were visually appraised with dimmed ambient room lighting on two 5-megapixel monitors (Siemens PLC, Berkshire); monitor quality control results fell within manufacturer specifications.

14 clinical colleagues (3 radiologists and 5 radiographers with film reading qualifications and 6 mammographers) appraised the images. These volunteers were blinded to other volunteer scores and the conditions under which the images were produced. Images were displayed in a randomised order. No alteration to the images was permitted by the volunteers. Images were viewed with single tiling format, with each image expanded to the full monitor screen. This resulted in a pixel size magnification of 2.19, a true size magnification of 4.46, original 81 mm and displayed 37 mm. A cardboard shield was applied to the monitor face to reduce distracting glare and reflections from outside the field of view.

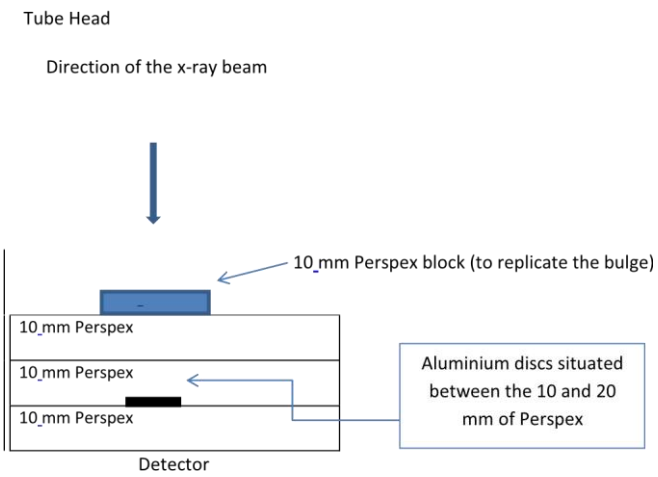


Figure 3. A diagrammatical representation of the Perspex configuration from image set 1.

Two alternative forced choice (2AFC) was used for visual grading.[12] For each image set (Table 2) a reference image was selected for comparison against the other images in its set. Volunteers performed the 2AFC task by comparing the reference image against the evaluation images. The 2AFC task used a 5 point Likert scale: score range of 1e5 in 1 = much worse than the reference image, 2 = worse than the reference image, 3 = equal to the reference image, 4 = better than the reference image, 5 = much better than the reference image. Volunteers also recorded how many aluminium discs they could visualise on each evaluation image.

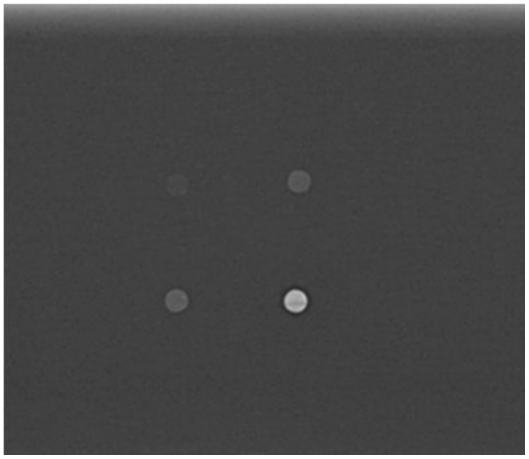


Image 2. Image from the configuration shown in Fig. 3.

Results

Data from the 14 volunteers' 2AFC visual grading tasks were combined to calculate a visibility score. The visibility score is defined as the number of lesions seen (1-4) multiplied by the score given using the 2AFC method on the 5 point Likert scale (1-5):

$$V = n.S$$

where:

V = visibility score

n = number of lesions seen in the image

S = 2AFC score given for image

The visibility score therefore ranged from 1 to 20; 20 was the highest possible score.

Graph 1 demonstrates the average visibility score for all observers, for each breast phantom thickness, with increasing phantom bulge thickness. As expected, increasing breast thickness decreases the visibility score.

Graph 2 demonstrates the average visibility score against added bulge tissue. Increasing bulge thickness causes deterioration in the visibility score.

The recorded organ dose is demonstrated in Graph 3. As automatic exposure chambers were used there are a few anomalies on the graph. These data points represent a change in kV, or target/ filter combination, which leads to a change in organ dose. This can be seen clearly at 30 mm of bulge tissue for 80 mm breast thickness, 25 mm of bulge tissue for 90 mm breast thickness, and 35 mm of bulge tissue for 100 mm breast thickness.

The average visibility score was then divided by the organ dose, to give an optimisation score, where:

$$O.S: = V/OrgD(mGy^{1})$$

where:

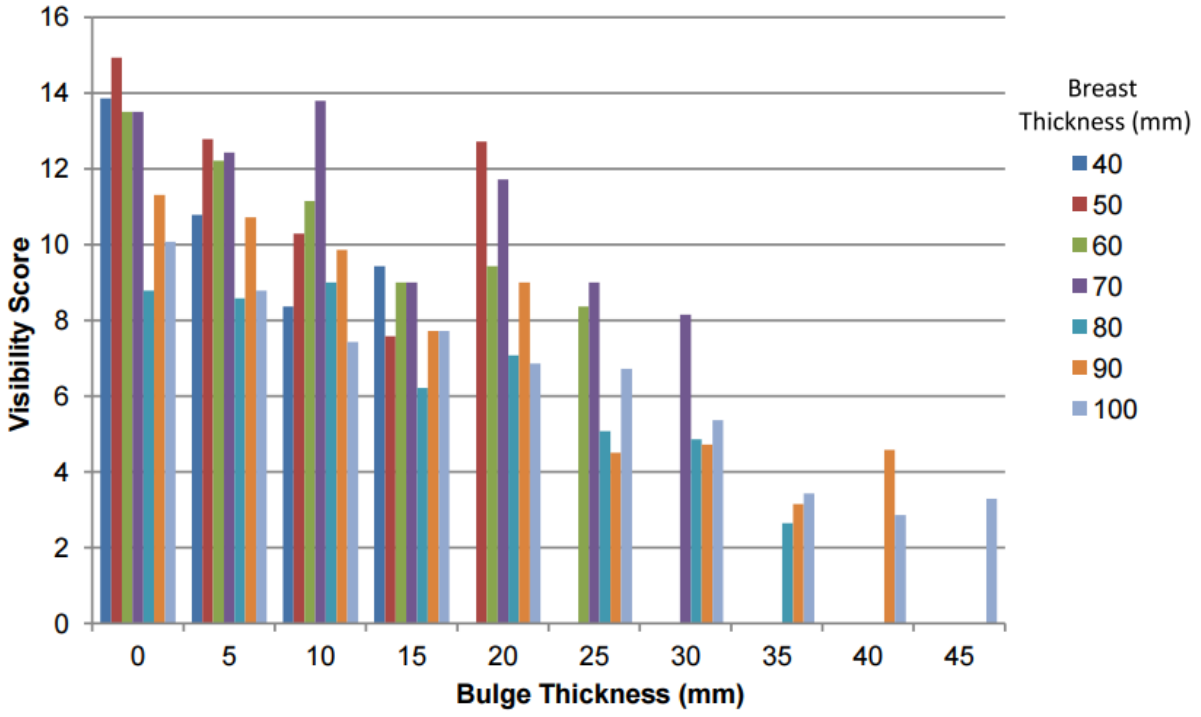
O.S. = optimisation score

V = average visibility score for all observers

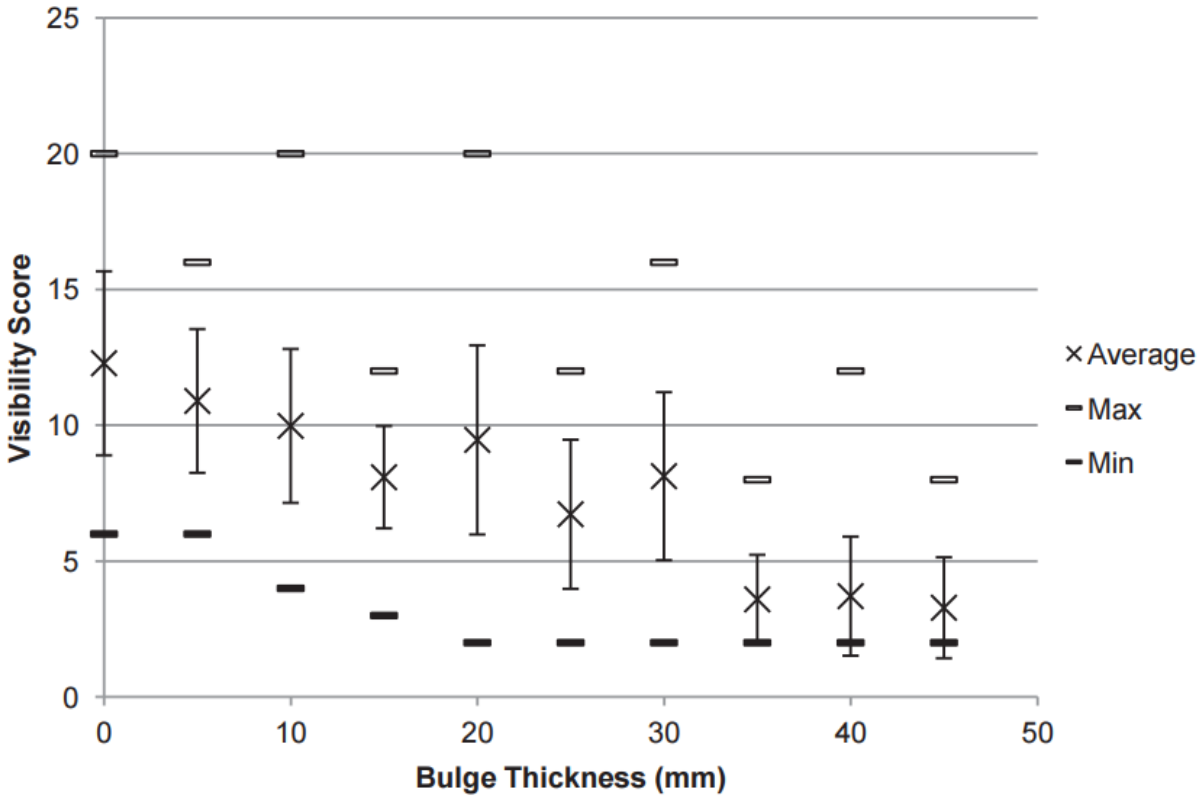
OrgD = organ dose for the image (mGy)

A higher value of O.S. indicates a better quality image, for lower organ dose; a low value of O.S. indicates a poor image quality with high organ dose.

As the breast tissue thickness increases the O.S. decreases (Graph 4). As the bulge thickness increases, the O.S. also decreases, indicating a higher organ dose with poorer image quality.



Graph 1. Visibility score compared to bulge thickness, for each phantom breast thickness.



Graph 2. Average visibility score compared to bulge thickness.

Discussion

The Aluminium disks become less visible as the breast bulge thickness increases (Graph 2). One explanation for lesions becoming less visible, as noted by Hackney et al., could be due to breast bulge.

The data can be explained by understanding how the automatic exposure system works in FFDM. In FFDM the machine given readout for breast thickness is used to auto-select kV. Usually the readout thickness is estimated from the rigid supporting system towards the rear of the machine which holds the paddle in place; this mechanism is prone to readout thickness inaccuracies.[13] If additional tissue (e.g. breast bulge) is added to the breast and that additional thickness is not accounted for, so the thickness could be underestimated and the kV set to be too low. This is demonstrated in sets 5, 6 and 7. Auto-select resulted in too low a kV and a "flash-exposure".

The kV was selected according to the breast thickness but as the bulge thickness was increased adjustment to the exposure parameters was required to penetrate the additional thickness of Perspex as shown in Table 2. This meant an increase in kV, and additionally in set 7 with 9 bulges, the "low quality" algorithm had to be selected in order to obtain an image. During the exposure the ionisation chamber allows the exposure to continue until adequate radiation has been received to permit termination. If kV is too low the ionisation chamber could increase exposure time to compensate, resulting in an increase in mAs. For phantom thicknesses of 100 mm and above 80 mm of additional Perspex to the 20 mm thick phantom, bulges greater than 40 mm resulted in images not being acquired due to termination of the exposure. This was due to the automatic selection of kV not compensating for the bulge. If the initial kV is too low then inadequate penetration could occur. This could result in lesion visibility being reduced.

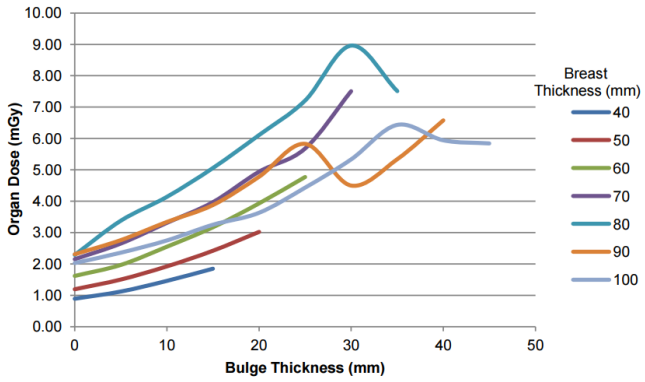
The combination of breast tissue bulge, auto-selection of kV and the potential compensation by increasing mAs through the automatic exposure device creates another problem, as demonstrated in Fig. 3. The mean glandular dose increases as bulge thickness increases. mAs has a linear relationship with organ dose whilst kV has a positive, non-linear relationship. If kV is reduced on its own this would lead to a reduction in organ dose. However the increase in mAs demanded by the automatic exposure device is likely to result in a net increase in organ dose. This phenomenon could be mitigated if the kV was correctly auto-selected in the first place. Graph 4 clarifies the situation, as the optimisation score demonstrates the combined effect of organ dose and lesion visibility due to increasing breast bulge thickness e as organ dose increases whilst lesion visibility decreases.

We believe the trends illustrated in our paper will occur when imaging human breast tissue during stereotactic procedures. Steps should be taken to minimise bulge magnitude and/or reduce the impact of bulge on lesion visibility.

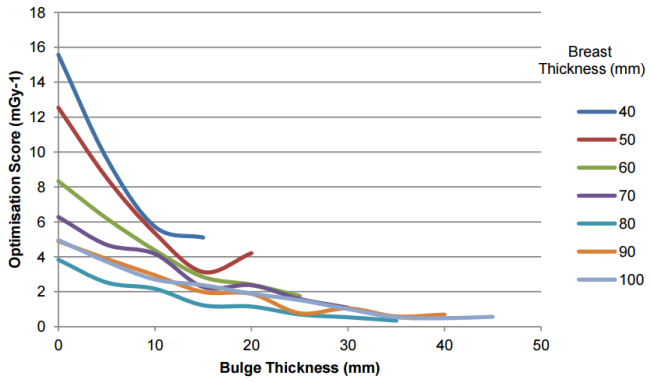
Although the vertical approach of performing stereotactic biopsy is more frequently used, there is the option of undertaking biopsies with a “lateral arm” mechanism. This method of sampling would allow full compression of the breast tissue, as there is no aperture required, but it is not without its limitations.

In the clinical setting the equipment used in our experiment has been configured to give a diagnostic image with the lowest dose. In fact, it may be appropriate to increase the dose to improve image quality, especially for thicker breasts. On the equipment used in our experiment Mo/Rh is the default target/filter setting for stereotactic images regardless of compressed breast thickness. For a thin breast lesion visibility could be improved by selection of the Mo/Mo option as demonstrated by the test images in Image 3 below.

Manual manipulation of the anode filter combination, adjusting the kV and selecting the “high quality” algorithm for subtle, low density microcalcification may optimise lesion visibility. The equipment used also locks the exposures factors for the stereotactic “scout” image and utilises these for the stereo pair images. This confounds the phenomenon as the X-ray beam is penetrating a different volume of tissue and density in an oblique projection (10) comparative to a perpendicular angle of penetration.



Graph 3. Organ dose compared to bulge thickness, for each breast thickness.



Graph 4. Optimisation index compared to bulge thickness, for each breast thickness.

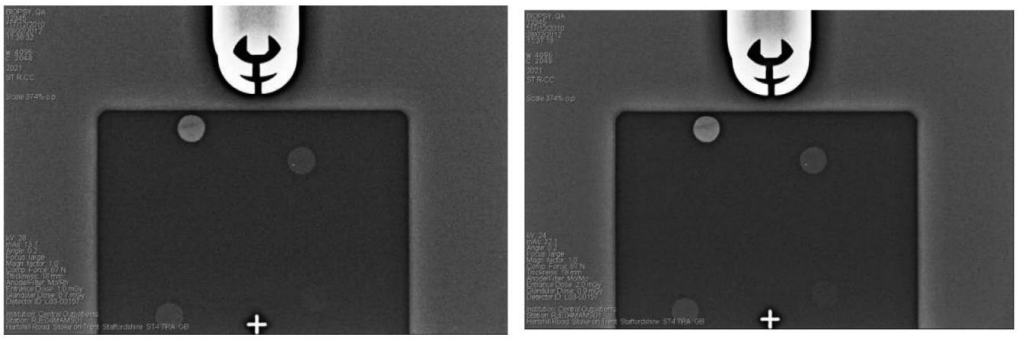


Image 3. Phantom stereo image. (Left) Auto expose –28 kV and Mo/Rh target/filter. (Right) Optimum setting – decreased kV to 24, Mo/Mo target/filter enables the 4th low density disc to be visualised.

Conclusion

For our phantom, as breast bulge size increases, lesion visibility decreases. At the same time organ dose increases. A practical way in which visibility could be improved and organ dose reduced would be to limit the formation of the bulge in the first place, this could be by improving the biopsy paddle design. Alternatively, the actual thickness of breast and bulge could be more accurately estimated. Conflict of interest None.

References

Burhenne LJW, Wood SA, D'Orsi CJ, Feig SA, Kopans DB, O'Shaughnessy KF, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. Radiology 2000;215:554e62.

Cole EB, Pisano EB, Kistner EO, Muller KE, Feig SA, et al. Diagnostic accuracy of digital mammography in patients with dense breasts who underwent problem-solving mammography: effects of image processing and lesion type. Radiology 2003 Jan;226(1):153e60.

Verkooijen HM. Diagnostic accuracy of stereotactic large-core needle biopsy for nonpalpable breast disease: results of a multicenter prospective study with 95% surgical confirmation. Int J Cancer 2002;99: 853e9.

Verkooijen HM, Peeters PHM, Buskens E, et al. Diagnostic accuracy of large-core needle biopsy for nonpalpable breast disease: a meta-analysis. Br J Cancer 2000;85:1017e21.

Hackney L, Williams S, Hogg P, Szczepura K. Tissue bulge during stereotactic core biopsy. Radiography November 2013;19(4):366e8.

Young KC, M Oduko J, Bosmans H, Nijs K, Martinez L. Optimal beam quality selection in digital mammography. Br J Radiol 2006;79:981e90.

Dance DR, Skinner CL, Young KC, Beckett JR, Kotre CJ. Additional factors for the estimation of mean glandular dose using the UK mammography dosimetry protocol. Phys Med Biol 2000;45:3225e40.

Pachoud M, Lepori D, Valley JF, Verdun FR. A new test phantom with different breast tissue compositions for image quality assessment in conventional and digital mammography. Phys Med Biol 2004;49:5267e81.

Carton AK, Bosmans H, Vandenbroucke D, Souverijns G, Van Ongeval C, Dragusin O, et al. Quantification of Al-equivalent thickness of just visible microcalcifications in full field digital mammograms. Med Phys 2004;31:2165e 76.

NHSBSP equipment report 0604. Commissioning and routine testing of full field digital mammography systems. Version 3.http://www.screening.org.uk/ breastscreen/publications/nhsbsp-equipment-report-0604.pdf; 2009.

http://www.healthcare.siemens.co.uk/mammography.

Krupinski E. The role of perception in imaging: past and future. Semin Nucl Med November 2011;41(6):392e400.

Hauge I, Hogg P, Szczepura K, Connolly P, McGill G, Mercer C. The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. Med Phys January 2012;39(1).

Paper 4

Practitioner compression force variation in mammography: A 6-year study

Claire E. Mercer[a,*], Peter Hogg [b], Katy Szczepura [b], Erika R.E. Denton [c]

aUniversity Hospital of South Manchester, The Nightingale Centre & Genesis Prevention Centre, Wythenshawe Hospital, Southmoor Road, Wythenshawe, Manchester M23 9LT, UK b University of Salford, Directorate of Radiography, Allerton Building, Frederick Road Campus, UK cUniversity of East Anglia and Norfolk & Norwich University Hospital, Norfolk & Norwich University Hospital Trust, Colney Lane, Norwich NR4 7UY, UK

Article history: Received 28 February 2013, Received in revised form 18 May 2013, Accepted 2 June 2013 Available online 24 June 2013

Keywords: Mammography, Breast, Breast compression, Compression force

Abstract

The application of breast compression in mammography may be more heavily influenced by the practitioner rather than the client. This could affect image quality and will affect client experience. This study builds on previous research to establish if mammography practitioners vary in the compression force they apply over a six-year period.

This longitudinal study assessed 3 consecutive analogue screens of 500 clients within one screening centre in the UK. Recorded data included: practitioner code, applied compression force (daN), breast thickness (mm), BI-RADS density category and breast dose. Exclusion criteria included: previous breast surgery, previous/ongoing assessment, breast implants. 344 met inclusion criteria. Data analysis: assessed variation of compression force (daN) and breast thickness (mm) over 3 sequential screens to determine whether compression force and breast thickness were affected by practitioner variations.

Compression force over the 3 screens varied significantly; variation was highly dependent upon the practitioner who performed the mammogram. Significant thickness and compression force differences over the 3 screens were noted for the same client (<0.0001). The amount of compression force applied was highly dependent upon the practitioner. Practitioners fell into one of three practitioner compression groups by their compression force mean values; high (mean 12.6 daN), intermediate (mean 8.9 daN) and low (mean 6.7 daN).

For the same client, when the same practitioner performed the 3 screens, maximum compression force variations were low and not significantly different ($p > 0.31$). When practitioners from different compression force groups performed 3 screens, maximum compression force variations were higher and significantly different ($p < 0.0001$).

The amount of compression force used is highly dependent upon practitioner rather than client. This has implications for radiation dose, patient experience and image quality consistency.

Introduction

In mammographic practice, breasts are compressed until adequate thickness reduction is induced. Various descriptors have been proposed to indicate when enough compression force has been applied.[1e5] The main aims of compression include the requirement to improve image quality[6] and the need to

minimise breast radiation dose.[7] However, within the National Health Service Breast Screening Programme (NHSBSP), there are no specific guidelines for optimal compression force levels required to achieve effective breast thickness reduction, other than a statement indicating that 'the force of the compression on the X-ray machine should not exceed 200 N'.[8]

Previous research[9] has established that practitioners vary in the amount of compression force they apply to breast tissue during mammography. This finding was independent of specific client characteristics (e.g. breast density). This research involved the cross sectional evaluation of 14 practitioners and 344 clients' compression force data on one mammography unit. Statistical analysis demonstrated a highly significant difference in mean compression used by different practitioners ($p < 0.0001$ for each BI-RADS density). Practitioners applied compression force in using low, intermediate or high compression force, with no significant difference in mean compression force within each group ($p ¼ 0.99$, $p ¼ 0.70$, $p ¼ 0.54$, respectively). It concluded that practitioners routinely apply either low, intermediate or high levels of compression force. Consequently, it was suggested that the amount of compression force applied to the breast could be highly dependent upon the practitioner.

As NHSBSP requires serial imaging to occur at regular intervals, with images reviewed to assess for subtle changes,[10] if compression force variability between practitioners exists then comparison between images over time may become more challenging. Additionally, and importantly, client experience may vary too and this may affect re-attendance rates. As such, we conducted a retrospective analysis to establish whether practitioner variations in compression force existed over time. For this analysis we identified a consecutive analogue sample from NHSBSP client data over 3 screening cycles e 6 years in total.

Materials/method

Study characteristics

The study was performed in a regional breast screening service located in the North of England (UK). Hospital audit and University ethics committees approved access to a sample of 500 clients from which data could be drawn. In order to reduce variability between mammogram machines, data was gathered from one static site, using one mammogram machine (analogue GE DMRþ mammography machine; Chalfont St. Giles, UK). The machine was operating within NHSBSP and manufacturer specifications[11,12] during the study period.

Client sample

Analogue mammogram images and associated data were gathered retrospectively. Data was gathered from clients who attended three consecutive screens. Only three screening rounds could be included as the required data for this study was unavailable prior to 2004. Data and images were therefore included from 2004, enabling 2004, 2007 and 2010 screening rounds for inclusion.

Identification of clients who were included into this study was through a consecutive convenience sampling basis. To be included each client had to have 3 consecutive screening mammograms 2004, 2007 and 2010; their first recorded mammogram experience at 2004. Each would have had the 4 standard projections acquired (left/right CC (cranialecaudal) and left/right MLO (medio-lateral oblique)). For each client the following information was recorded e size of film, breast compression force value in deca-Newtons (daN), compressed breast thickness (mm) and the name of practitioner who performed the mammogram. The latter was coded for anonymity purposes.

Mean glandular dose (MGD) estimations[12] were calculated retrospectively for specific clients. Together with this, breast density was established by one reader for each image using the 4 point BI-RADS scale (Breast Imaging Reporting and Data System[13]) e BIRADS 1 <25% dense, BI-RADS 2: 25%-50% dense, BI-RADS 3: 51%-75% dense, and BI-RADS 4 >75% dense. This reader was an experienced breast practitioner who had good BI-RADS classification scoring agreement with 3 other experienced breast clinicians (Kappa 0.83, 0.92, 0.83). [9]

Exclusion criteria

Exclusion criteria included: the inability of clients to tolerate compression force, clients who had breast pain, previous breast surgery, breast implants or cysts/abscesses, disabled clients, clients with arm/shoulder movement limitations. As the study was retrospective, some client data we would

have liked to consider was not available e for example point in menstrual cycle and whether the client had pain upon pressure application. Consequently these parameters could not be considered in our analysis. Due to exclusion criteria 156 of the 500 clients were not included; 344 clients remained. This represented 1032 'mammogram sets' over the 3 screening rounds - 4128 individual images.

Practitioners

The clients were imaged by 14 trained practitioners; these consisted of all the staff who rotated through the breast imaging department at the time of the study. They comprised of Advanced Practitioners, mammographers and Assistant Practitioners with experience ranging from 1 to 12 years. These practitioners were the same as those used in a previous study[9]; this permitted direct comparison of results between these two studies. The average number of mammograms performed per practitioner was 73 (range 10-146).

Results

For the 344 clients the following analysis was carried out.

Breast density change

Data was categorised into BI-RADS breast density distribution for each mammogram visit. Only 7% of clients (n = 24) showed a change in BI-RADS density over time. This represented a reduction of one BI-RADS density grade. These clients were not removed from the sample prior to analysis in the first instance as images were analysed separately. It was only when sequential patient images were considered together that these BI-RADS density variations were removed.

Compression force values

Regardless of BI-RADS density grade, practitioner data was first analysed for mean compression force on each mammogram projection (MLO and CC). All mammograms were assigned to the practitioner who performed the mammogram, regardless of year imaged. Figs. 1a and 1b demonstrate the mean compression force values, standard deviations and confidence intervals for each practitioner.

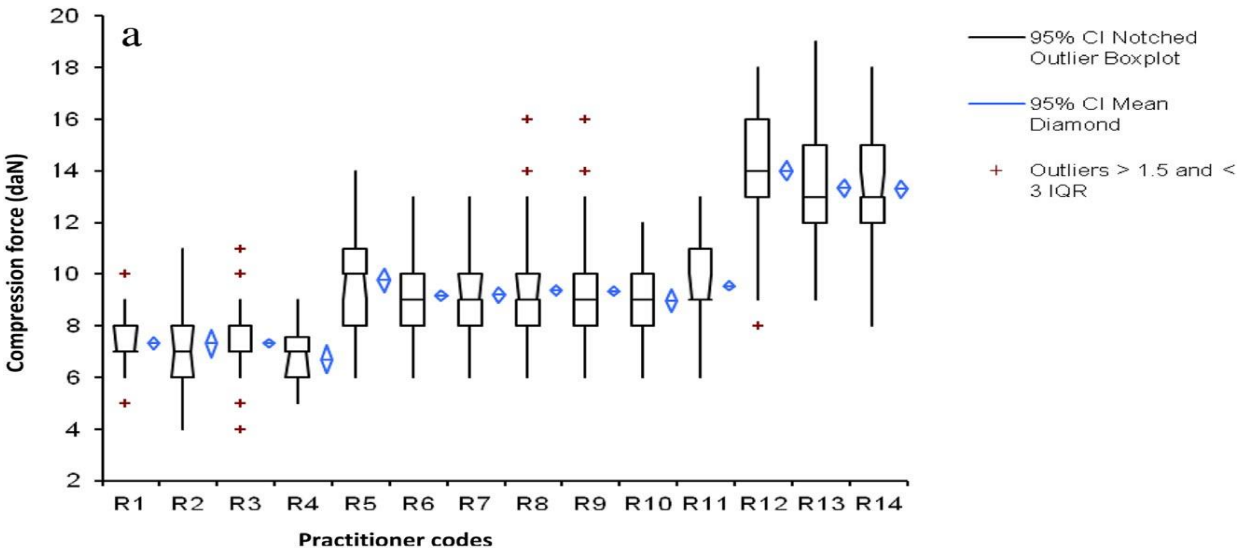


Figure 1a. Compression force variation (daN) on MLO mammography images per practitioner.

Within a previous study[9] these practitioners were placed into compression force groups because of their similar compression force means; this provided a way of classifying them. For the current study the same practitioner groupings/classifications were applied e the practitioners had similar compression force means as the previous study[9] (rank sum correlation coefficient ¼ 0.9). The coefficient of 0.9 indicates that the practitioners performed very similarly in their compression force behaviours for both client datasets. In the current study 4 practitioners fell into the low compression force group, 7 into the intermediate group and 3 into the high group. Dispersal of practitioner grade and length of experience across the three compression groups appeared to demonstrate no particular trend for the purposes of this study.

For the low compression force practitioner group: in the MLO projection, practitioners imaged with compression force mean values (regardless of BI-RADS density grade) between 7.17 and 7.4 daN and in the CC projection between 6 and 6.27 daN.

For the intermediate compression force practitioner group: in the MLO projection, practitioners imaged with compression force mean values (regardless of BI-RADS density grade) between 8.6 and 9.6 daN and in the CC projection between 7.95 and 8.71 daN.

For the high compression force practitioner group: in the MLO projection, practitioners imaged with compression force mean values (regardless of BI-RADS density grade) between 12.6 and 14 daN and in the CC projection between 11.45 and 11.7 daN.

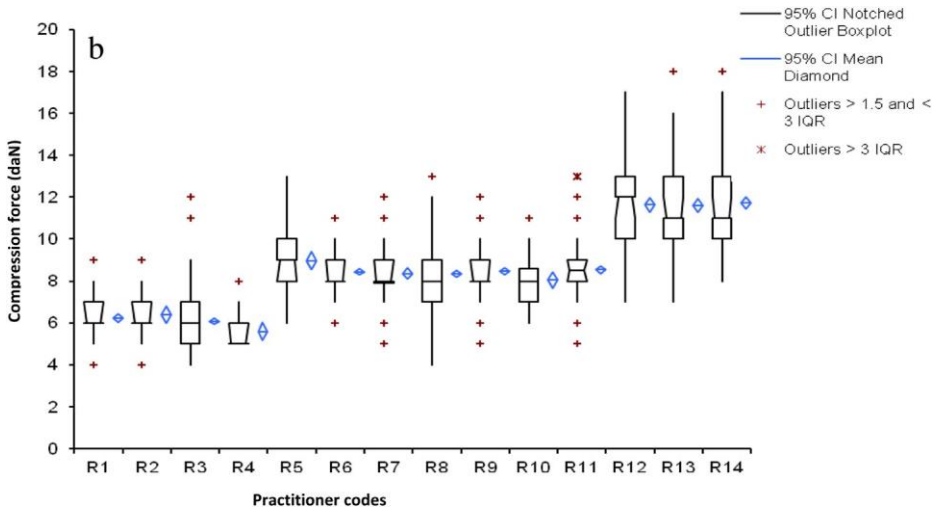


Figure 1b. Compression force variation (daN) on CC mammography images per practitioner.

There is a highly significant difference in the mean compression force values between the practitioners in the low and the intermediate group, the low and the high group and the intermediate and the high group ($p < 0.0001$); this holds true within each BIRADS density classification

Breast thickness values

Mean thickness of breast tissue for each practitioner is presented, distributed by BI-RADS density grade, in Fig. 2. There is a highly significant difference between the breast thicknesses from the intermediate practitioner group and the high practitioner group ($p < 0.0001$). There is a significant difference between the breast thicknesses from the low practitioner group and the high practitioner group ($p < 0.001$). There is no statistical difference between the breast thicknesses from the low practitioner group and the intermediate practitioner group.

Longitudinal assessment of compression force and thickness due to practitioner variation

In order to assess if there was variation of compression force and breast thickness over the three screening rounds, specifically due to practitioner variation, we applied additional inclusion/exclusion criteria. From the remaining 344 clients we assessed which clients had been imaged either: sequentially by the same practitioner for each of the 3 screens, sequentially by practitioners from the same practitioner group for each of the 3 screens, or sequentially from the practitioners from different compression force groups for each of the 3 screens. From the remaining 344 clients, 134 remained for further analysis for the exacting purposes of analysing longitudinal variation of compression force and thickness.

To achieve the assessment of compression force and thickness variations within these clients, we set a 'reference value' of zero to the client's initial mammogram. Any increase or decrease from that value was represented by a plus (an increase in compression force or thickness) or a minus (a decrease in compression force or thickness). The term 'maximum absolute compression force variation' is the maximum compression force difference displayed between each screening mammogram for the three years for clients. Similarly the term 'maximum absolute thickness variation' is the maximum breast thickness difference displayed between each screening mammogram for the three years for clients; their first mammogram experience (incident round) being assigned as their reference value.

Clients imaged sequentially by practitioners from the same practitioner compression force group

From the 134 clients, 81 were imaged by a practitioner from the same compression force group on each attendance. Of these, 6 clients had a change in BI-RADS density over time and at this stage of the analysis they were removed to minimise any variation in compression force/thickness which may be caused by density change.

Seven clients were imaged by practitioners in the low compression force group each time they attended their 3 screens. These clients experienced maximum absolute compression force variations between the sequential screens of -2 daN and +1 daN (MLO projections) and -2 daN and +1 daN (CC projections). There were no statistically significant differences in the compression force values of these clients over their three screening episodes. Maximum absolute breast thickness variations between the sequential screens of these clients were -18 mm and +9 mm (MLO projection) and -17 mm and +6 mm (CC projection). Again there were no statistically significant differences in the breast thickness values of these clients over their three screening episodes.

Sixty-eight clients were imaged by practitioners from the intermediate compression force group each time they attended. These clients experienced maximum absolute compression force variations between the sequential screens of -4 daN and +2 daN (MLO projection) and -3 daN and +2 daN (CC projection). There were no statistically significant differences in the compression force values of these clients over their three screening episodes. Maximum absolute breast thickness variations between the sequential screens of these clients were -22 mm and +10 mm (MLO projection) and -14 mm +15 mm (CC projection). Again there were no statistically significant differences in the breast thickness values of these clients over their three screening episodes.

Over the three sequential screening rounds 14 clients were imaged by the same practitioner on each attendance. These clients experienced maximum absolute compression force variations between the 3 screens of -2 daN and +2 daN (MLO projection) and -2 daN and +1 daN (CC projection). There were no statistically significant differences in the compression force values of these clients over their three screening episodes. Maximum absolute breast thickness variations between the 3 screens of these clients were +16 mm and -15 mm (MLO projection) and +17 mm and -6 mm (CC projection). Again, there were no statistically significant differences in the breast thickness values of these clients over their three screening episodes.

In summary the clients who saw the same practitioner or practitioners from the same practitioner compression force group on their three sequential screening mammograms had no significant differences in their breast thickness or their breast compression force levels.

Clients imaged sequentially by practitioners from different compression force groups

Thirty-nine clients were imaged by a practitioner from each compression force group (low, intermediate and high) during their three screens in a variety of orders. As above their first screening attendance was assigned a 'zero' and changes calculated from this figure.

These clients experienced maximum absolute compression force variations over the three sequential screens of -2 daN and +10 daN (MLO projection) and +3 daN and +14 daN (CC projection). For these 39 clients, in order to represent this change in breast compression force longitudinally over the 3 screens, the results have been displayed time independently and averaged for the two MLO and CC projections for each attendance (Fig. 3). T-tests indicate highly significant differences in compression force values ($p < 0.0001$) for CC and the MLO projections. This level of significance is the same for the low and high compression force groups, the intermediate and high compression force groups and the low and intermediate groups.

For the 39 clients, in order to represent change in breast thickness longitudinally over three sequential screens, results have been displayed time independently and include both MLO and CC projections for each attendance (Fig. 4). Absolute thickness reductions between the low and intermediate group reduced by 1 mm (MLO projection) and 1.7 mm (CC projection). Absolute thickness reductions between the intermediate and high group reduced by 5.7 mm (MLO projection) and 6 mm (CC projection). Absolute thickness reductions between the low and high group reduced by 6.2 mm (MLO projection) and 7.7 mm (CC projection). T-tests indicate highly significant differences in breast thickness reductions ($p < 0.0001$) between the low and high compression force

groups and the intermediate and high compression force groups for both projections. The differences between low and intermediate groups did not achieve the same level of significance; for the MLO projections there was no significant difference, for the CC projections it was significant ($p < 0.05$).

MGD for the 39 clients was calculated retrospectively. These are illustrated in Fig. 5. Maximum dose differences were 2.64 mGy (MLO) and 1.12 mGy (CC) when clients were imaged by a practitioner from a low practitioner group and then a high practitioner group. Some clients experienced differences in dose of 1.57 mGy (MLO) and 1 mGy (CC) when they were imaged by a practitioner from a low compression force group followed by a practitioner from an intermediate compression force group.

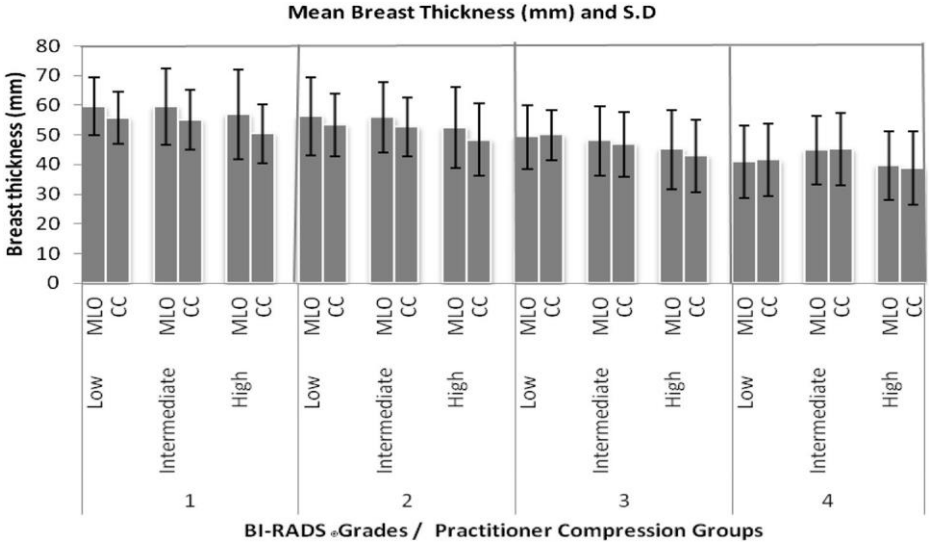


Figure 2. Mean breast thickness (mm) and SD within BI-RADS grades and practitioner group.

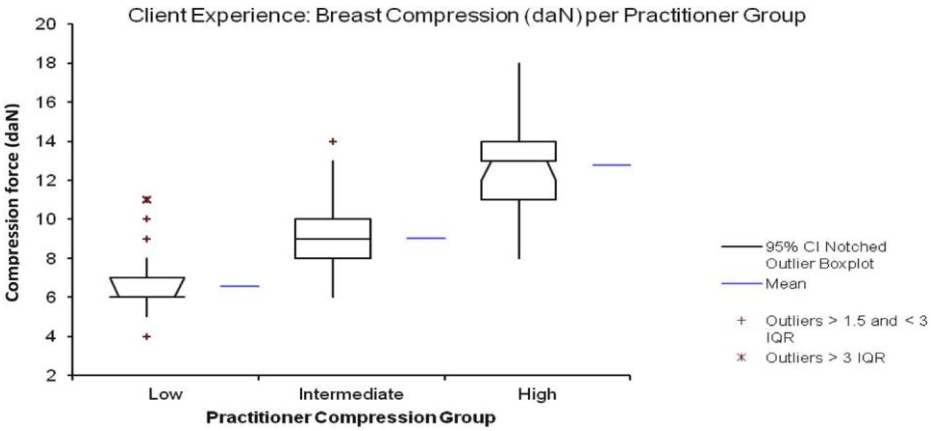


Figure 3. Breast compression force changes (daN) for clients imaged by different practitioner compression group - time independent.

Overall percentage dose differences demonstrated a mean difference of 10.2% (MLO) and 6.9% (CC) when clients were imaged by practitioners from a low practitioner group followed by practitioners from a high compression force group. These dose differences would likely represent a clinically important difference and are due to the differences in breast thickness levels on mammogram acquisition (Fig. 5).T-tests highlight significant dose differences between the low and high compression force groups in both projections ($p < 0.01$). Differences from low to intermediate groups are not significant for the MLO though significant for the CC view ($p < 0.05$). For the intermediate to high group for both projections there were no significant differences.

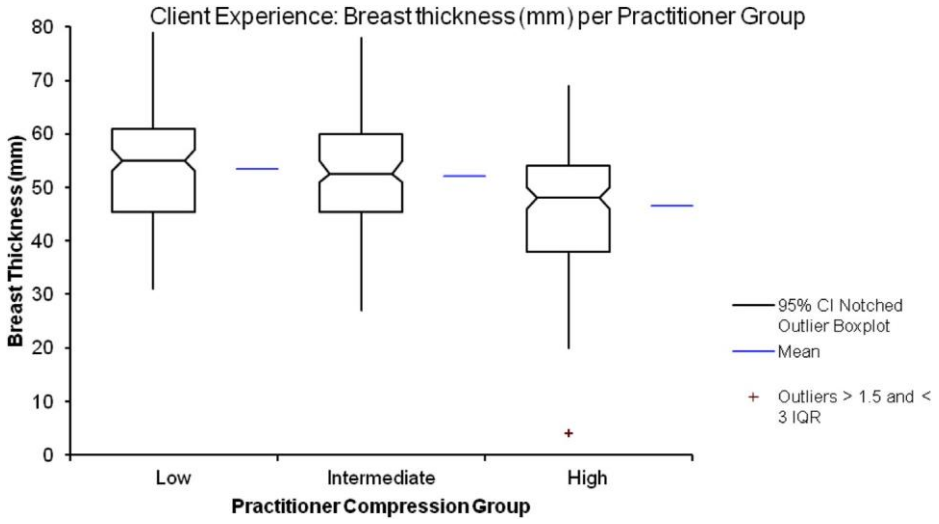


Figure 4. Breast thickness changes (mm) for clients imaged by different practitioner compression group - time independent.

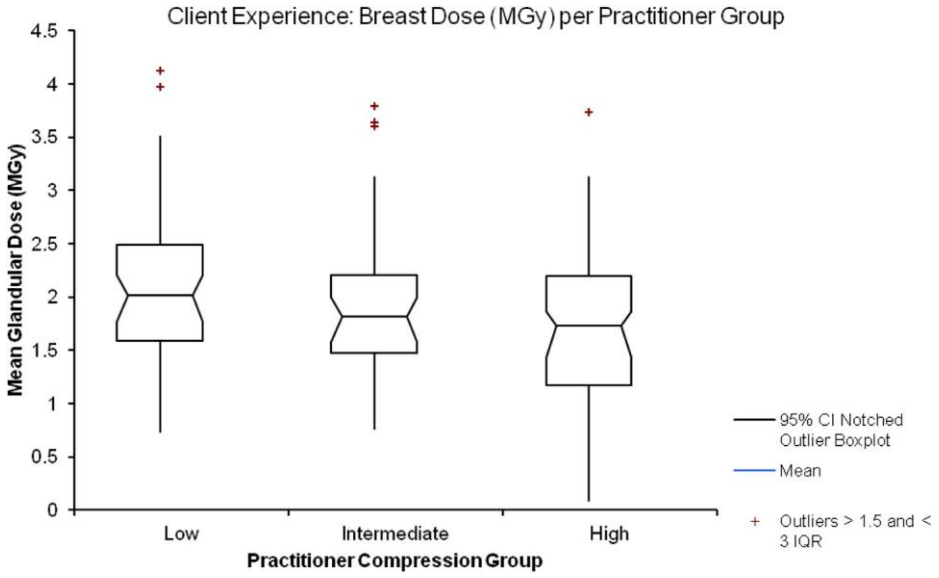


Figure 5. Breast dose changes (mGy) for clients imaged by different practitioner compression group - time independent.

In summary, clients who saw practitioners from a different compression force group on each attendance had significant differences in their compression force levels and some of their thickness

levels. Depending upon which practitioner compression force group is considered significant differences in dose have also been demonstrated.

Discussion

Implications for practice

Our study establishes that the amount of breast compression force seems highly dependent upon practitioner rather than client. This has implications for radiation dose and image quality consistency for sequential screening within the National Health Service Breast Screening Programme together within the symptomatic setting.

We have highlighted four areas for consideration. Firstly, the practitioners fall into the same compression force groups as with Mercers' previous study.[9] Secondly, from a client perspective, the compression force that is applied to client breasts during each mammogram can vary over time and this is dependent upon the practitioner who images them. Thirdly, for the clients who were imaged with a practitioner from a different group on each attendance, breast compression force values are significantly different. Breast thickness reduction was also significantly different apart from between the low and intermediate compression force groups on the MLO view. This suggests that there is significance to the application of higher compression force in the reduction of breast thickness. Finally, it has been highlighted that for certain cases, the larger thickness reductions have resulted in lower mean glandular doses (MGDs). Though T-tests show that some of these were not statistically significant in some cases, there has to be consideration of the clinical importance of this e doses should be kept as low as practical.

It appears that each practitioner is consistent over time in the amount of compression force that they apply. We have also indicated that there is a close correlation between mean compression force values from this study and Mercer's cross sectional study.[9] This suggests that individual practitioners are applying compression force consistently over time, and also within different client groups. This could mean that practitioners are applying their own tolerance levels to compression force application. We have also demonstrated that changes in BI-RADS density grades made little difference to the practitioner's behaviour in their application of compression force. This again could suggest that practitioners are applying compression force to the breast using their own tolerance levels regardless of breast type.

The relevance to clients being imaged by practitioners applying different levels of compression force may give rise to different levels of pain and discomfort experienced whilst having mammography

and this may have consequences for future attendance. Studies[14e18] have suggested varying thresholds of compression force for pain tolerances varying from 9 daN to 16 daN. As such, consistency of optimal compression force applied over time could be paramount in the maintenance of client experience. The same argument would hold true for the consistency of image quality over time.

Our data has demonstrated statistically significant variations in breast compression force and breast thickness levels when clients are imaged by different practitioners over their 3 screening rounds. Our study has also demonstrated that clients imaged by the same practitioner on each screen have less breast compression force and breast thickness variation. It is likely that these clients have had more a consistent experience.

For the third and final issues, breasts might be imaged with breast thickness reduction (rather than compression force) in mind in order to reduce radiation burden.[1,4] This will likely achieve better consistency of breast dose and image quality[19] for clients imaged serially within the NHSBSP.

Limitations

Our study has several limitations. Firstly, this study was at a single site with a relatively small group of mammographers. The study has now been extended to a multicentre study in order to assess if the results will be similar at other screening centres. Secondly, as these were retrospective important factors such as point in menstrual cycle, breast pain upon compression force and weight changes, for example, could have effect on the results of this study.

Conclusion

We have established that compression force and breast thicknesses can fluctuate for the same client when they are imaged by different practitioners. Implications from this can result in variations in mean breast glandular dose between 3 yearly screening events. The possibility exists for variations to occur in image quality and lesion visibility. Given that compression force differences can occur over time it is possible that client experience may vary too with possible implications to clients screening attendance within the future.

Conflict of interest

We can confirm that there are no financial and personal relationships with other people or organisations that could inappropriately influence (bias) our work.

References

1. Long S. The handbook of mammography. 4th ed. Edmonton: Mammography Consulting Services Ltd.; 2000.

2. Wentz G. Mammography for radiologic technologist. New York: Mc Graw Hill; 1992.
3. Kopans D. Breast imaging. 3rd ed. Lippincott Williams and Wilkins; 2007.
4. Eklund GW, Cardenosa G. The art of mammographic positioning. Breast imaging current status and future directions. Radiologic Clinics of North America 1992;30:21e53.
5. Poulos A, McLean D. The application of breast compression in mammography: a new perspective. Radiography 2004;10:131e7.
6. Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography, how much is enough? Australasian Radiology 2003;47:121e6.
7. Brnic Z, Hebrang A. Breast compression and radiation dose in two different mammographic oblique projections: 45 and 60_. European Journal of Radiology 2001;40:10e5. Elsevier Science Ireland Ltd.
8. NHSBSP 63. Quality assurance guidelines for mammography, ISBN 1 84463 028 5; April 2006. p. 42.
9. Mercer CE, Hogg P, Lawson R, Diffey J, Denton ERE. Practitioner compression force variability in mammography: a preliminary study. British Journal of Radiology 2013;86:20110596.
10. NHSBSP 59. Quality assurance guidelines for breast cancer screening radiology. 2nd ed. March 2011.
11. NHSBSP 33. Quality assurance guidelines for medical physics services. 2nd ed.,ISBN 1 84463 016 1; June 2005.
12. Report No. 89 (Revision of IPEM Report 59)The commissioning and routine testing of mammographic X-ray systems. Institute of Physics and Engineering in Medicine (IPEM); 2005.
13. D'Orsi CJ, Bassett LW, Berg WA. Mammography. In: Breast imaging reporting and data system: ACR BI-RADS_. 4th ed. Reston VA: American College of Radiology; 2003.
14. Chida K, Komatsu Y, Sai M, Nakagami A, Yamada T, Yamashita T, et al. Reduced compression mammography to reduce breast pain. Clinical Imaging2009;33:7e10.
15. Sullivan DC, Beam CA, Goodman SM, Watt DL. Measurement of force applied during mammography. Radiology 1991;181:355e7.
16. Myklebust AM, Seierstad T, Stranden E, Lerdal A. Level of satisfaction during mammography screening in relation to discomfort, service provided, level of pain and breast compression. European Journal of Radiography 2009;1:66e72.
17. Pisano ED, Gatsonis C, Hendrick E. Diagnostic performance of digital versus film mammography for breast-cancer screening. New England Journal of Medicine 2005;353:1773e83.
18. Oliveira M, Nogueira MS, Guedes E, Andrade MC, Peixoto JE, Joana GS, et al .Average glandular dose and phantom image quality in mammography. Nuclear Instruments and Methods in Physics Research 2007;580:574e7. Elsevier B.V.
19. Korf A, Herbst CP, Rae WID. The relationship between compression force, image quality and radiation dose in mammography. South African Journal of Radiology 2009;13(4):86.

Paper 5

A 6-year study of mammographic compression force: Practitioner variability within and between screening sites

Claire E. Mercer a, Katy Szczepura b, Judith Kelly c, Sara R. Millington c, Erika R.E. Denton e, Rita Borgen d, Beverley Hilton d, Peter Hogg b

A The Nightingale Centre & Genesis Prevention Centre, University Hospital of South Manchester, Wythenshawe Hospital, Manchester M23 9LT, UK, b University of Salford, UK c Countess of Chester Hospital NHS Foundation Trust, UK d East Lancashire Breast Screening Unit, Burnley General Hospital, UK e University of East Anglia and Norfolk & Norwich University Hospital, UK

Article history: Received 26 April 2014, Received in revised form 21 June 2014, Accepted 7 July 2014 Available online 29 July 2014
Keywords: Compression force, Breast compression, Compression variability

Abstract

Background: The application of compression force in mammography is more heavily influenced by the practitioner rather than the client. This can affect client experience, radiation dose and image quality. This research investigates practitioner compression force variation over a six year screening cycle in three different screening units.

Methods: Data were collected from three consecutive screening events in three breast screening sites. Recorded data included: practitioner code, applied compression force (N), breast thickness (mm), BIRADS® density category. Exclusion criteria included: previous breast surgery, previous/ongoing assessment and breast implants. 975 clients (2925 client visits, 11,700 mammogram images) met inclusion criteria across three sites. Data analysis assessed practitioner and site variation of compression force and breast thickness.

Results: Practitioners across three breast screening sites behave differently in the application of compression force. Two of the three sites demonstrate variability within themselves though they demonstrated no significant difference in mean, first and third quartile compression force and breast thickness values CC ($p > 0.5$), MLO ($p > 0.1$) between themselves. However, in the third site, where mandate dictates a minimum compression force is applied, greater consistency was demonstrated between practitioners and clients; a significant difference in mean, first and third quartile compression force and breast thickness values ($p < 0.001$) was demonstrated between this site and the other two sites.

Conclusion: Variability within these two sites and between the three sites could result in variations. Stabilisation of these variations may have a positive impact on image quality, radiation dose reduction, re-attendance levels and potentially cancer detection. The large variation in compression forces could negatively impact on client experience between the units and within a unit.

Further research is required to establish best practice guidelines for compression force within mammography.

Advances in knowledge: Practitioners vary in the compression forces they apply to clients over sequential screening attendances. Establishing practice guidance with cessation guidelines could help to minimise this problem.

Introduction

It is acknowledged that one of the most important factors in determining the success of a screening programme is screening uptake.[1,2] The causes of any non-uptake are multifactorial.

A systematic review in 2013 measured the extent of non-uptake. This review indicated clients not re-attending for screening because of breast pain from prior mammography was a significant issue.[3] Whelehan and colleagues suggested that between 47,000 and 77,000 women within England do not re-attend for breast screening in a year due to pain directly related to a previous mammogram.[3]

Pain from mammography can arise from the application of compression force.[3] It has also been identified that the position of the breast under the mammography compression paddle can directly affect the amount of pressure in different portions of the breast[4] with potential for direct association with increased breast pain.

Quality assurance standards within the National Health Service Breast Screening Programme (NHSBSP) are essential to ensure its continued effectiveness. The 2012[5] annual review of breast screening highlighted that 'ultimately decisions based around screening programmes must be evidence based' and that it should be 'a first class system ensuring excellent training for all professional staff'. It seems extraordinary that such a service has no standards or guidelines on the application of compression force other than a statement 'the force of the compression on the X-ray machine should not exceed 200 Newtons (N)[6] with various proposed descriptors such as 'taut to touch' or 'until the skin blanches'.[7-11]

This research investigates practitioner compression force variation over a six year screening cycle in three different screening units. It builds on earlier research, which was single centre. Previous research[12,13] identified practitioner variability in compression force application during mammography imaging within a single NHSBSP screening programme. The current research includes two additional regional breast screening services located in the North of England (UK).

Materials/method

Hospital (service evaluation) and University ethics committees approved access to a sample of 1500 screening events at each screening unit (a screening event is defined as one mammogram series which includes four images). In order to exclude mammography machine variability[14] as a confounding factor in terms of data quality, data was gathered from one mammogram machine at each location (GE Seno Essential, Lorad Mk4 and Siemens Mammomat 3000). The three analogue mammogram machines were operated within NHSBSP and manufacturer specifications[15,16] during the study period. The study period was for a consecutive six year period; only analogue images were included as NHSBSP screening sites had not been converted to digital technology for a six year period at the time of the study. Design characteristics of compression paddles tend to be similar between analogue and digital units, though it should be noted that recently paddles on the latter have started to introduce changes to their design.

Client sample

Data were gathered retrospectively at all three sites from clients who attended three consecutive screening events. Only three screening events could be included as the required data for this study was unavailable prior to 2004 at certain screening sites.

Identification of clients was through consecutive stratified sampling. For inclusion each client had to have three consecutive screening events, with their first recorded mammogram experience as their first event. Each would have four standard projections acquired (left/right CC (cranial-caudal) and left/right MLO (medio-lateral oblique). For each client the following information was recorded directly from the mammography image - size of film, breast compression force value in deca-Newtons (daN) or Newtons (N), compressed breast thickness (mm) and the practitioner who performed the mammogram, coded for anonymity.

Breast density was established by 5 observers in the three screening units using the 4 point BI-RADS® scale (Breast Imaging Reporting and Data System)[17] - BI-RADS® 1 < 25% dense, BI-RADS® 2: 25%-50% dense, BI-RADS® 3: 51%-75% dense, and BI-RADS® 4 > 75% dense. In order to establish inter and intra observer characteristics of the 5 observers for BI-RADS scoring, fifty film-screen mammograms were used.[18] These images comprised of left and right CC and MLO and were scored by each observer independently under the same viewing conditions; blinded to the findings of other observers. To provide data to assess intra-observer variability, mammography image sets were re-scored after an interval of at least two weeks, to minimise recall bias. Near complete intraobserver agreement

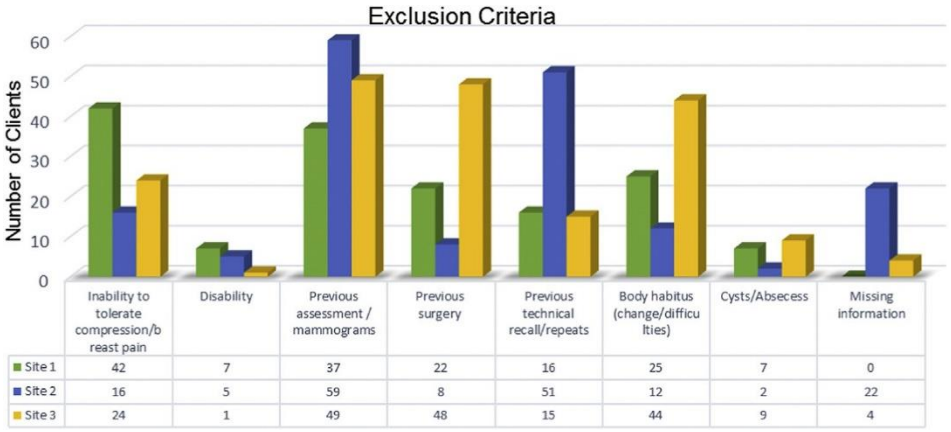
(Kappa >0.81) and strong or above interobserver variability was demonstrated (First score Fleiss kappa 0.77 second score 0.65).[18]

Exclusion criteria

Exclusion criteria were established (Fig.1). Clients with less than or more than four standard projections were also excluded. Following application of exclusion criteria the number of clients remaining for analysis at each unit were: site 1 =344,[13] site 2 = 325, site 3 = 306.

Practitioners

Practitioners at all sites consisted of staff working in the breast imaging department at the time of the study. The staff included a mixture of Advanced Practitioners, Mammographers and Assistant Practitioners, all are referred to as practitioners for the purposes of this study. Clients were imaged by similar numbers of trained practitioners at the three sites; 14 at site one, 11 at site two and 15 at site three.



	Inability to tolerate compression/breast pain	Disability	Previous assessment / mammograms	Previous surgery	Previous technical recall/repeats	Body habitus (change/difficulties)	Cysts/Absecess	Missing information
Site 1	42	7	37	22	16	25	7	0
Site 2	16	5	59	8	51	12	2	22
Site 3	24	1	49	48	15	44	9	4

Figure 1. Exclusion criteria

Recorded data

Compression force and compressed breast thickness, together with practitioner details of those who performed the imaging were noted for all images.

Results

Practitioners

Firstly, analysis of practitioner grade between sites was compared (Table 1). The range of the number of clients the practitioners imaged at each site was: site one (10e146); site two (10e155); site three (12e139). The mean number of clients imaged by all practitioners at each site was, site one: (73.7),

site two: (88.6), site three (61.2). The median number of clients imaged at each site was, site one: (73.5), site two: (100), site three:(75).

BI-RADS breast density classification

The distribution of BI-RADS density within each site was assessed for similarity between sites by documenting the number of mammograms imaged per site in a percentage for each BI-RADS breast density category (Table 2). For the purposes of statistical analysis, combination of BI-RADS® 1 and 2 (Group A) and also BIRADS® 3 and 4 (Group B) was required due to the low numbers of images in BI-RADS® group 1 with BI-RADS® group 4, having zero figures for some practitioners. Pearson Chi Squared test was used for the comparison of BI-RADS® Group A and Group B amongst sites. Pearson's X^2 156 (Group A) and 107 (Group B), ($p < 0.0001$) suggests there is a significant difference in the distribution of BIRADS® grades between different sites. The authors would like to acknowledge the updated release of the BI-RADS scale in 2013 with a change in scale for BI-RADS breast density from 1e4 to A-D; this study was completed prior to that grading release and as such is not recognised within this paper.

Whilst it is recognised that this could be considered as a study limitation, it has been established previously[12] that practitioners display the same compression behaviours across BI-RADS density classifications and do not necessarily vary their application of compression force according to breast density.

Practitioner variability

To establish practitioner variability, the mean compression values for all practitioners, at all sites, were analysed (Figs. 2 and 3). Compression force values varied across the three sites, with CC average at site one 86N, site two 84N, site three 125N. For the MLO, site one 97N, site two 88N, site three 132N. Analysis of variance (ANOVA) of mean compression force values of practitioners demonstrated a significant difference ($p < 0.0001$) between sites 'one and three', and 'two and three'. Sites 'one and two' demonstrated no significant difference (CC $p > 0.5$, MLO $p > 0.1$). These levels of significance hold true within each BI-RADS density classification.

Site	Assistant practitioners	Practitioners (radiographers)	Advanced practitioners	Total practitioners
Site one	2	10	2	14
Site two	0	9	2	11

Site three	2	8	5	15

Table 1 Practitioner grade per site.

Site	% Mammograms BI-RADS 1	% Mammograms BI-RADS 2	% Mammograms BI-RADS 3	% Mammograms BI-RADS 4
One	11	64	21	4
Two	28	28	28	16
Three	21	40	29	10

Table 2 Percentage of mammograms within each BI-RADS breast density category.

First and third quartile results at all sites were analysed (Table 3). In CC and MLO, ANOVA of first and third quartile compression force levels of practitioners demonstrated a significant difference ($p < 0.0001$) between sites 'one and three' and sites 'two and three'. Sites 'one and two' demonstrated no significant difference (first quartile $p > 0.1$, third quartile $p > 0.5$). This holds true within each BI-RADS grade. Having removed the outliers (see Figs. 2 and 3), minimum and maximum compression force values for CC views ranged as follows: Site one 47N-122N (75N), site two 42Ne-114N (72N), site three 103N-158N (55N). For MLO: site one 65N-136N (71N), site two 48N-139N (91N), site three 103N-163N (60N).

Percentage changes in breast compression force

Analysing the mean percentage change between minimum and maximum compression force values per client, from their three screening mammograms, establishes one aspect of variability from a client perspective.

The mean percentage change between minimum and maximum compression force was calculated for each BI-RADS grade for both CC and MLO (Fig. 4). Average values of mean percentage change for each site for the MLO: site one 55%, site two 66%, site three 27% and the CC: site one 57%, site two 60% and site three 26%.

ANOVA was performed on percentage changes. For MLO, sites 'one and three' and 'two and three' demonstrated a significant difference ($p < 0.0001$) and this holds true within each BI-RADS grade. Sites one and two demonstrated no significant difference ($p > 0.2$), this holds true for each BI-RADS grade. No significant difference was demonstrated between sites 'one and two' ($p > 0.5$). It can be concluded that site three displays low client variability over the three screens.

Breast thickness

Compressed breast thickness ranges at all sites were compared by mean, first and third quartile values for CC and MLO.

Mean compressed breast thickness values at all sites were analysed (Table 4). Over the three screens, in both the CC and MLO, ANOVA of mean compressed breast thickness values of practitioners demonstrated a significant difference ($p < 0.0001$) between 'site one and three' and site 'two and three'. Site one and two demonstrated no significant difference in mean CC values of thickness ($p > 0.5$). This holds true within each BI-RADS grade. Practitioners at site three applied higher compression values and this would explain why the breast thicknesses at this site are smallest.

First and third quartile compressed breast thickness values at all sites were analysed (Table 5). For both the CC and MLO, ANOVA demonstrated significant differences ($p < 0.0001$) in first and third quartile breast compressed thickness values between sites 'one and three' and sites 'two and three'. Site 'one and two' demonstrated no significant difference in values of thickness ($p > 0.5$). This holds true within each BI-RADS grade.

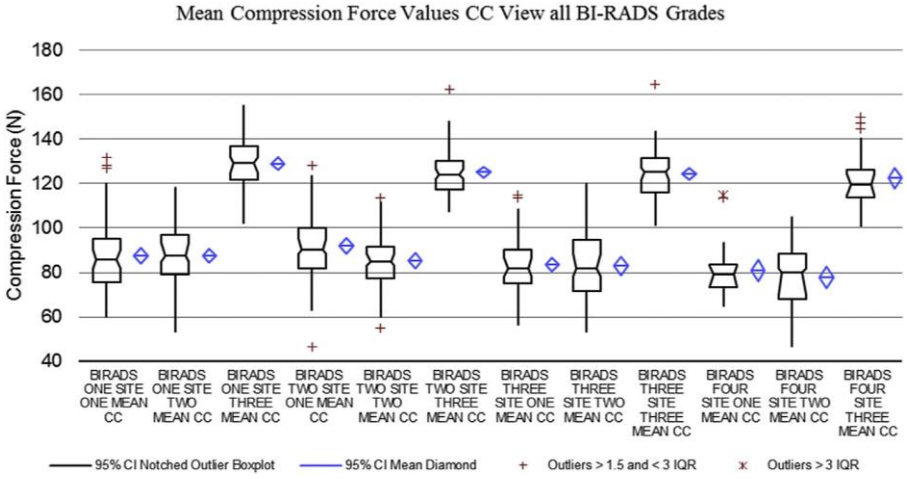


Figure 2. Mean compression force values CC view.

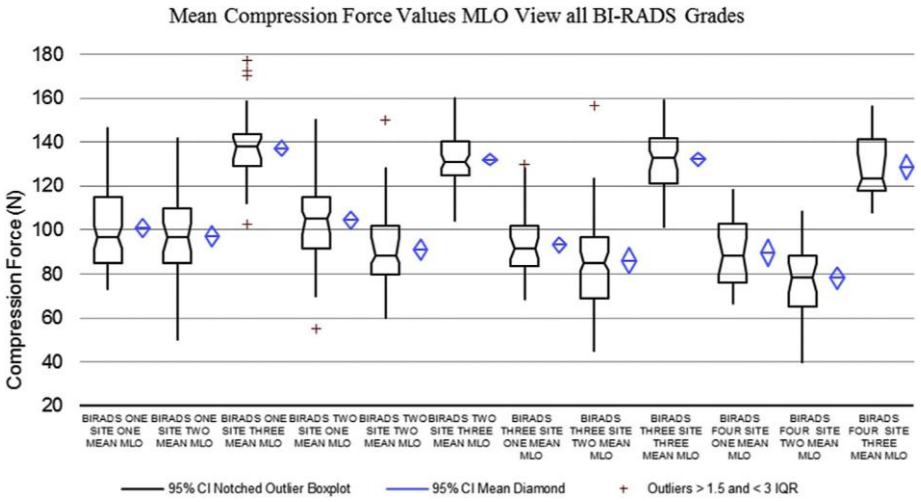


Figure 3. Mean compression force values MLO view.

Discussion

Compression force variability

This research has demonstrated that the amount of breast compression force applied by practitioners is not consistent within and between three NHSBSP screening sites.

For site one, within each of the three subgroups variability is low'.13,14 At site two practitioners apply compression force across a wide range of values and they do not fall into subgroups. Overall, practitioners from site one and site two apply compression forces within the same mean values, first and third quartiles and there is no statistical difference between them. Sites one and two permitted their practitioners to define their own compression force values, within NHSBSP maximum tolerance levels. Whilst there is no statistical difference between sites one and two, a client attending either or both of these sites would potentially be subject to large variations in compression force on subsequent visits. However, on average, for sites one and two, a client would have a lower level of compression force applied compared with site three. However for site three a client would likely have a higher though more consistent level of compression forced applied over time.

Site three had a protocol in place which mandates that a minimum level of 100N compression force is used. Some sites within NHSBSP have protocols similar to this. Therefore, the lack of a consistent approach within NHSBSP exposes clients to variation in compression force if they moved between sites. It might be worthwhile speculating that higher compression force values could be associated

with reduced client experience and pain and reduced re-attendance. Equally variability could also cause this problem too - perhaps even at lower levels of compression force.

It is also worth noting that no data exists to illustrate that image quality is better when compression forces of 100N or higher are used, as in site three; rather anecdote dictates that higher compression forces are likely to result in better image quality. A pilot study19 identified no differences in image quality with higher compression forces, however the image quality scoring mechanism may not be sensitive enough to identify subtle changes in image quality.

A noted limitation of this study is that the three sites studied are located in the same geographical region and therefore practitioners could have been trained similarly, thereby reflecting a local variability problem. However, in 2013 Murphy and colleagues,20 from a UK-wide analysis of compression force behaviours, identifiethat practitioners vary in their approach to the application of compression force. This current study is therefore likely to reflect behaviour nationally.

	FIRST QUARTILE				THIRD QUARTILE			
SITE	MLO compression (N)	S.D	CC compression (N)	S.D	MLO compression (N)	S.D	CC compression (N)	S.D
SITE ONE	84.85	21.63	75.5	17.07	106.1	26.07	92.7	22.87
SITE TWO	73.13	11.73	71.27	11.57	104.3	15.5	95.87	12.42
SITE THREE	118.21	12.75	111.99	10.09	144.34	14.65	135.41	15.25

Table 3 First and third quartile compression forces all sites

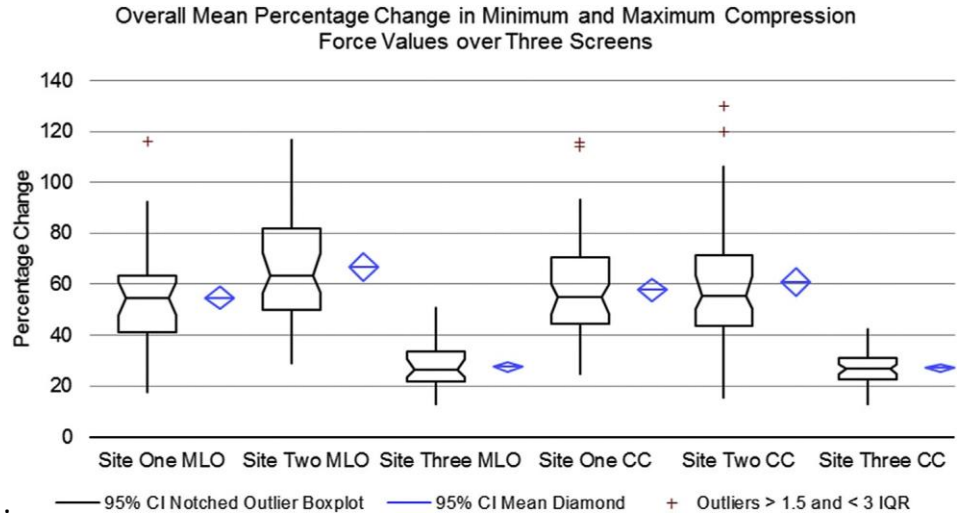


Figure 4. Overall mean percentage change in minimum and maximum compression force values over three screens.

Breast thickness variability

The inconsistency in compression force application across the three sites has a direct association with an inconsistency of compressed breast thickness values. Site one and two have similar means, first and third quartile compressed thickness values with no statistical difference ($p > 0.5$). Site three has significant differences in compressed breast thickness levels to the other two sites ($p < 0.001$); this has obvious direct implications for radiation dose and may have an impact on image quality e especially when sequential imaging comparison is considered. On this basis site three might be considered superior for consistency and dose minimisation.

Site	MLO thickness (mm)	S.D	CC thickness (mm)	S.D
Site one	53.8	13.7	50.9	11.3
Site two	57.9	12.2	56.8	10.9
Site three	47.1	12.7	43.5	10.5

Table 4 Mean breast thickness value (mm): comparison all sites.

National standards

From this and prior research13,14 there is a need for the NHSBSP to consider the introduction of national guidance on compression force levels. Hogg and colleagues21 in 2013 highlighted minimum and cessation compression force levels for one mammography machine. They suggested that cessation should be considered based upon rate of change of compression force and thickness reduction, rather than by compression force alone.

Taking a different perspective, a recent study by de Groot and colleagues22 questioned if standardisation by compression force was meaningful and they suggested a focus towards pressure. They explained that clients with small breasts would experience more pressure than clients with large breasts with the same applied compression force. They suggested standardisation based upon pressure and this shows promise.

Possible impact on client experience

The findings of this research have possible implications for clients. These will be discussed in turn.

Radiation risk

With respect to radiation risk there remain uncertainties about absolute cancer risk from low dose mammography screening. A recent report states that the risk of radiation induced cancer is approximately 1 in 20,000 per screening visit.23 This equates to 154 cancers detected for every one

induced and 80 lives saved for every life lost to radiation induced cancers.23 Benefit thus exceeds risk. This research demonstrated that site three had lower breast thickness levels than the other two sites overall within the six year screening cycle ($p < 0.001$). Reducing breast thickness has potentially quantifiable reductions in radiation risks to clients within the screening programme.

			First Quartile			Third Quartile		
Site	MLO Thickness (mm)	S.D.	CC Thickness (mm)	S.D.	MLO Thickness (mm)	S.D.	CC Thickness (mm)	S.D.
Site one	44.55	3.43	43.56	2.86	63.6	3.80	59.73	2.54
Site two	49.78	2.94	50.46	3.02	65.36	3.08	62.61	2.58
Site three	38.23	3.70	36.32	2.66	56.52	2.85	50.74	2.91

Table 5 First and third quartile compressed breast thickness value (mm): comparison all sites.

Image comparison

Direct comparison between images on successive screens is vital to ensure accurate visualisation of subtle changes within the breast. Direct comparison is not only essential within the same screening site but across the whole NHSBSP as clients can attend different sites. Our research has demonstrated compression force and breast thickness differences exist between and within sites, and the latter could influence image quality. If differences in quality exist for the same client then this could confound comparison of images on successive screens.

Re-attendance

Pain and non-re-attendance are related. Having a standardised approach to compression force levels within a specified range might improve client experience by offering them a consistent expectation and experience. Further research is needed into client pain and levels of applied compression force.

Conclusion

Our research demonstrates that practitioners across three breast screening sites behave differently in the application of compression force when undertaking mammography. Two of the three sites demonstrate variability. Variability within these two sites and between the three sites could result in variations in image quality, radiation dose together with client experience which in turn could influence re-attendance. When mandate dictates a minimum compression force standard this results in greater consistency between practitioners and clients. This may have a positive impact on image quality, radiation dose reduction and potentially cancer detection.

Conflict of interest statement/role of funding

The authors have no conflict of interest to declare. Part funding for the study was received by the Countess of Chester Breast Unit and Burnley General Breast Unit to cover associated staff costs whilst

data collection ensued for this study. This funding was from the North West Regional Quality Assurance Reference Centre. The sponsors have had no role in study design, collection of data, analysis or interpretation of data, writing of the manuscript of decision to submit the manuscript.

References

Marmot M, Altman DG, Cameron DA, Dewer JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380(9855):1778e86.
Weller DP, Campbell C. Uptake in cancer screening programmes: a priority incancer control. Br J Cancer 2009 Dec 3;101(Suppl. 2):S55e9. http://dx.doi.org/ 10.1038/sj.bjc.6605391.
Whelehan P, Evans A, Wells M, Macgillivray S. The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. Breast 2013 Aug;22(4):389e94.
Smith H, Hogg P, Maxwell A, Mercer C, Szczepura K. An analysis of the compressed breast area and image receptor/compression paddle pressure balance in different mammographic projections. UKRC; 2013 [Publication].
NHS Breast Screening Programme. Annual review published 2012. ISBN 978-184463-093-6; 2012.
NHSBSP 33. Quality assurance guidelines for medical physics services. 2nd ed. June 2005. ISBN 1 84463 016 1.
Long S. The handbook of mammography. 4th ed. Edmonton: Mammography Consulting Services Ltd; 2000.
Wentz G. Mammography for radiologic technologist. New York: Mc Graw Hill; 1992.
Kopans D. Breast imaging. 3rd ed. Lippincott: Williams and Wilkins; 2007.
Eklund GW, Cardenosa G. The art of mammographic positioning. Breast imaging: current status and future directions. Radiol Clin North Am 1992;30: 21e53.
Poulos A, McLean D. The application of breast compression in mammography: a new perspective. Radiography 2004;10:131e7.
Mercer CE, Hogg P, Lawson R, Diffey J, Denton ERE. Practitioner compression force variability in mammography: a preliminary study. Br J Radiol Feb 2013;86:20110596.
Mercer CE, Hogg P, Szczepura K, Denton ERE. Practitioner compression force variation in mammography: a 6-year study. Radiography 2013;19:200e6.
Hauge I, Hogg P, Szczepura K, McGill GG, Connolly P, Mercer CE. The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. Med Phys 2012;39:1.
Report No. 89 (Revision of IPEM Report 59). The commissioning and routine testing of mammographic X-ray systems. Institute of physics and engineering in medicine (IPEM); 2005.
NHSBSP 63. Quality assurance guidelines for mammography, 42; April 2006. ISBN 1 84463 028 5.
D'Orsi CJ, Bassett LW, Berg WA. Mammography. Breast imaging reporting and data system: ACR BI-RADS®. 4th ed. Reston VA: American College of Radiology; 2003.
Mercer C, Hogg P, Kelly J, Borgen R, Enion D, Hilton B, et al. A mammography image set for research purpose utilizing BI-RADS density classification. Radiologic Technology 2014 July/August;85:609e13.
Mercer C, Hogg P, Cassidy S, Denton ERE. Does an increase in compression force really improve visual image quality in mammography? An initial investigation, 2013. Radiography 2013;19:363e5.
Murphy FJ, Nightingale JM, Mackay SM, Robinson L, Seddon D, Hogg P.
Compression behaviours e an exploration of the beliefs and values influencing the application of breast compression during screening mammography. UKRC; 2013.
Hogg P, Taylor M, Szczupera K, Mercer C, Denton E. Pressure and breast thickness in mammographydan exploratory calibration study. Br J Radiol 2013;86. 20120222.
De Groot JE, Broeders MJM, Branderhorst W, den Heeten GJ, Grimbergen CA. A novel approach to mammographic breast compression: improved standardization and reduced discomfort by controlling pressure instead of force. Med Phys August 2013;40(8).
The Royal College of Radiologists. Guidance on screening and symptomatic breast imaging. 3rd ed. London: The Royal College of Radiologists; 2013, June. http:// www.rcr.ac.uk/docs/radiology/pdf/BFCR(13)5_breast.pdf.

Paper 6

A method to measure paddle and detector pressures and footprints in mammography

Peter Hogg[a)] and Katy Szczepura
University of Salford, Salford, Greater Manchester M6 6PU, United Kingdom
Alison Darlington
Pennine Acute Hospital NHS Trust, Manchester M8 5RB, United Kingdom
Anthony Maxwell
Royal Bolton Hospital NHS Foundation Trust, Bolton BL4 0JR, United Kingdom

(Received 18 October 2012; revised 18 January 2013; accepted for publication 6 February 2013; published 18 March 2013)
Key words: mammography, breast compression, breast pressure balancing

Purpose: Compression is necessary in mammography to improve image quality and reduce radiation burden. Maximizing the amount of breast in contact with the image receptor (IR) is important. To achieve this, for the craniocaudal projection, there is no consensus within the literature regarding how the IR should be positioned relative to the inframammary fold (IMF). No information exists within the literature to describe how pressure balancing between IR and paddle, and IR breast footprint, might be optimized. This paper describes a novel method for measuring the respective pressures applied to the breast from the IR and the paddle and a method to simultaneously measure the breast footprints on the IR and the paddle.

Methods: Using a deformable breast phantom and electronic pressure-sensitive mat, area and pressure readings were gathered from two mammography machines and four paddles at 60, 80, and 100 N with the IR positioned at −2, −1, 0, +1, and +2 cm relative to the IMF (60 combinations in total).

Results: Paddle and IR footprints were calculated along with a uniformity index (UI). For all four paddle/machine/pressure combinations the greatest IR footprint was achieved at IMF +2 cm. The UI indicates that the best pressure/footprint balance is achieved at IMF +1 cm.

Conclusions: The authors' method appears to be suited to measuring breast footprints and pressures on IR and paddle and a human female study is planned.

I. INTRODUCTION

Compression of the breast during mammography is necessary in order to obtain acceptable image quality.[1–3] The resulting reduction in breast thickness improves image contrast, reduces the overlap of normal breast structures, and improves image sharpness by reducing the tissue to image receptor (IR) distance. It has the added benefit of reducing the radiation dose to the breast.[4] While the technique for the application of compression has been well described in the literature,[3,5,6] little has been written about how to maximize the volume of breast tissue imaged (i.e., the footprint of the breast upon the detector) or how to optimize the relative pressures exerted on the breast from the compression paddle and IR. For the craniocaudal projection there is no consensus within the literature regarding how the IR should be positioned relative to the inframammary fold (IMF) and the effect this may have on image quality and breast footprint.[3,7] Furthermore, it would seem reasonable for the contributions to breast compression from the paddle and from the IR to be evenly balanced so as to minimize client discomfort. No information exists within the literature to describe how pressure balancing and IR breast footprint might be optimized.

In this paper we describe a novel method for measuring the respective pressures applied to the breast from the IR and the paddle and a method to simultaneously measure the breast footprints on the IR and the paddle. Using our methods this paper describes a preliminary phantom study to investigate the effect of changing the positions of the IR relative to the IMF on the balance of compression pressure and on breast footprint. Our methodology is not intended to be used during mammographic acquisitions as the pressure sensing device will produce unacceptable image artefacts.

II. METHOD

No method is known to have been published previously to describe how to measure the pressures exerted from the paddle and IR on the breast or the footprint that the breast defines onto the paddle and IR. Knowing that practical and ethical problems would arise in using humans to develop a method, we choose to use a physical representation of the human female breast.

II.A. Deformable breast phantom

To validate our method a deformable breast phantom was created and mounted in a semiflexible fashion to a rigid structure which simulated the human thorax. The phantom was the same as that used by Hauge.[8] The phantom had compression characteristics similar to a human female breast,

with a correlation coefficient of 0.95 (compared with human female breast compression data gathered from the same mammography machine and same paddle).

II.B. Pressure and footprint measuring system

An electronic pressure-sensitive mat [Xsensor (Ref. 9)] developed for measuring pressure distribution on patients for bed sore prediction[10] and for tyre-road contact analysis in the motor vehicle industry was used. Xsensor offers a range of pressure monitoring products with suitable temporal and spatial resolution characteristics. The pressure mat used is designed for measuring patients in seated/lying positions. The pressure values applied to the breast during compression are greater than the pressures exerted in seated/lying positions; therefore, the compression values used were restricted to the lower range that might be applied clinically.

The flexible design of the Xsensor pressure recording mat, together with the associated acquisition module and PC software, provides an easy to use pressure recording system. The pressure mat consists of an array of pressure sensors embedded within a flexible material. The array comprises 1296 sensing points with an overall mat size of 63.5 × 63.5 cm^2 and an active sensing area of 45.72 × 45.72 cm^2, with spatial resolution of 1.27 cm. Manufacturer's specifications state an accuracy of ±10%.[9] Prior to conducting the study, the XSensor was calibrated in accordance with the manufacturer's instructions to provide values in mmHg and to ensure all pixels were functioning. Data from the pressure mat could be represented as a 2D or 3D image on a color scale and as numeric (pressure) data exported to a spreadsheet.

II.C. Pressure readings, footprint, and IR position

The Xsensor pressure mat was wrapped around the phantom so that pressure readings at the phantom-IR and phantompaddle interfaces could be taken simultaneously. The pressure readings were taken with the phantom compressed in two different mammography machines—a Hologic Selenia and a Hologic Selenia Dimensions. Each had two flexible paddles— an 18 × 24 cm^2 paddle and a 24 × 29 cm^2 paddle, plus a 24 × 29^2 cm nonflexible paddle which was common to both machines. Three compression forces were applied (as measured by the calibrated readout on each machine)-– 60, 80, and 100 N. The values of 60, 80, and 100 N were chosen as compression values as values greater than 100 N resulted in most of the pixel values being greater than the detectable range (>256 mmHg), rendering analysis unfeasible. Less than 60 N is an unrealistic compression value clinically, and so was not considered within this project. Through clinical experience, 100 N is representative of a mid-range compression value. Five vertical IR positions were used: −2, −1, 0, +1, and +2 cm,

measured relative to the IMF. In total, pressure data were recorded at 60 IMF position and pressure value combinations. A maximum pressure of 100 N was used because this represented the maximum limit for the XSensor in our experiment.

Pressure readings were acquired at a framing rate of 1/s; the frames were later averaged.

After the application of pressure had ceased, a few minutes were allowed for the paddle and the phantom to stabilize. This stabilization period was necessary as it was noted that the mammography compression values displayed on the mammography machines dropped for several seconds after compression was initially applied.

II.D. Data cleaning and data analysis

Prior to analysis the spreadsheet data were cleaned. This involved deleting artifactual datapoints which were created by pressure which was not attributable to the phantom. An example of an artifactual datapoint generated by a minor crease in the pressure mat is shown in Fig. 1.

Consideration was also made for pressure values greater than the detectable range (>256 mmHg). Table I demonstrates the number of pixels >256 mmHg in terms of percentage of all pixels within the pressure mat (total area) and the percentage of pixels that received a reading (active area).

These values did not affect the area of pressure measured. However, these values need to be taken into account when considering the maximum pressure applied.

Each datapoint represented a sensor reading on the pressure mat, and each sensor had an area of 1.6129 cm^2 (the pixel area as stated by the manufacturer[9]). Therefore, the area of phantom that was compressed by the IR and the area of phantom that was compressed by the paddle (the “footprints”; area in cm^2) could be calculated by simply multiplying the number of pixels that had a reading >0 by 1.6129 cm^2.

For each of the IR positions, paddle and detector footprints (cm^2) and pressures (mmHg) were recorded. From this average pressure applied by the IR, average pressure applied by the paddle (in mmHg), average pressure per unit area applied by the IR, and average pressure per unit area applied by the paddle (mmHg/cm^2) were calculated.

A uniformity index (U.I.) was derived; this considered the distribution of average pressure per unit area applied by the IR and by the paddle. The equation used was as follows:

$$\text{Uniformity Index} = (A - B)/(A + B),$$

where A is the average pressure per unit area applied by the paddle (mmHg/cm^2) and B is the average pressure per unit area applied by the IR (mmHg/cm^2).

The U.I. value has the following implications. If U.I. = 0, there is equal pressure per unit area from the IR and the paddle (equal distribution).

TABLE I. Percentage of pixels greater than detectable range.

	Percentage of total area			Percentage of active area		
Compression (N)	Average	SD	Maximum	Average	SD	Maximum
60	0.03	0.07	0.23	0.42	0.96	3.26
80	0.23	0.20	0.62	3.12	2.78	10.13
100	0.47	0.27	0.93	5.70	3.57	13.79

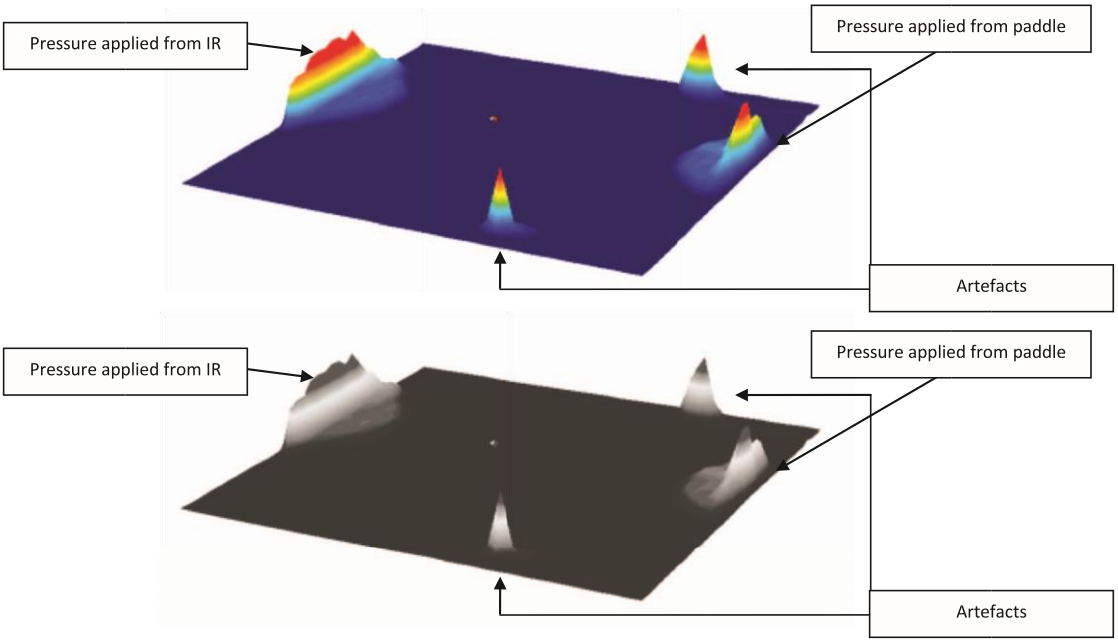


FIG. 1. Xsensor pressure map data using a flexible paddle at 80 N with receptor 2 cm pushed up. This demonstrates artifactual datapoints generated by a minor crease in the pressure mat

If 0 < U.I. < 1, there is greater pressure per unit area from the paddle on the top of the phantom, with 1 = all pressure per unit area is applied by the paddle.

If –1 < U.I. < 0, there is greater pressure per unit are[1]from the IR on the underside of the phantom, with −1 = all pressure per unit area is applied by the IR.

III. RESULTS

[1] Note: That is, B + 2, 18 × 24 cm^2 nonflexible = image receptor placed 2 cm above the inframammary fold using the Hologic Selenia Dimensions, 18 × 24 cm^2 nonflexible paddle.

Data recorded by the XSensor were displayed visually using a color or grey scale (Fig. 1) and as a numerical array. After removing artifacts the numerical arrays were analyzed within a spreadsheet. Figure 2 (18 × 24 cm^2 flexible paddle) demonstrates the area of the phantom in contact with the paddle and the IR. As can be seen, the area of phantom in contact with the IR is lowest at IMF −2 cm and highest at IMF +2 cm. The area of phantom in contact with the paddle increases gradually from the −2 cm position to the +2 cm position. This trend is preserved for the four paddle types on the two machines for all three applied compression levels. Table II demonstrates the key to Figs. 2 and 3, in terms of machine, IR position in relation to the IMF, and the paddle size and type.

Figure 3 (18 × 24 cm^2 flexible paddle) demonstrates the pressure exerted onto the phantom from the paddle and from the IR. As can be seen, the pressure exerted from the paddle onto the phantom is highest at IMF −2 cm and lowest at IMF +2 cm. Similarly, the pressure exerted onto the phantom from the IR is lowest at IMF −2 cm and highest at IMF +2 cm. This trend is preserved for the four paddle types on the two machines for all three applied compression levels.

Table III illustrates the combined uniformity index for the four paddles on the two machines for all three compression levels. An assumption is made that a uniformity index of zero is desirable, indicating an equal balance of pressure and area between the top and bottom of the phantom. The absolute values of these have been used to calculate the average, standard deviation, maximum, and minimum values of each position, irrespective of unit type, paddle size/type, and pressure applied. Absolute values were used so that the negative values do not skew the data.

Figure 4 demonstrates graphically the absolute uniformity index for each IR position. The uniformity index is closest to zero for positions above the IMF, these positions also have smaller SD values, demonstrating greater consistency in the technique. Figure 4 also demonstrates that there is not only a poorer distribution of pressure but also greater variability of the uniformity index when the IR is positioned below the IMF.

TABLE II. Key to Figs. 2 and 3.

Machine type	
A =	Hologic Selenia
B =	Hologic Selenia Dimensions
IR Position	
−2 =	2 cm below IMF
−1 =	1 cm below IMF
0 =	on IMF
+1 =	1 cm above IMF
+2 =	2 cm above IMF

Paddle size and flexible/nonflexible as stated.

TABLE III. Uniformity index.

Unit[a], paddle position[b], paddle size, paddle type	60 N	80 N	100 N	Ave[c]	SD[c]	Max[c]	Min[c]
A + 2, 18 × 24, Flexible	−0.17	−0.20	−0.17	0.12	0.06	0.20	0.03
A + 2, 24 × 30, Flexible	0.09	0.11	0.06				
B + 2 24 × 30, Flexible	0.08	0.11	0.08				
B + 2, 24 × 30, Nonflexible	−0.20	−0.10	−0.03				
A + 1, 18 × 24, Flexible	0.01	0.03	0.04	0.03	0.02	0.06	0.01
A + 1, 24 × 30, Flexible	−0.01	0.06	0.06				
B + 1, 24 × 30, Flexible	0.03	0.01	0.04				
B + 1, 24 × 30, Nonflexible	−0.02	0.01	0.02				
A0, 18 × 24, Flexible	0.32	0.33	0.33	0.19	0.10	0.33	0.08
A0, 24 × 30, Flexible	0.22	0.25	0.23				
B0, 24 × 30, Flexible	−0.10	0.09	0.08				
B0, 24 × 30, Nonflexible	0.08	0.09	0.19				
A − 1, 18 × 24, Flexible	0.38	0.38	0.35	0.37	0.11	0.50	0.18
A − 1, 24 × 30, Flexible	0.46	0.45	0.39				
B − 1, 24 × 30, Flexible	0.50	0.45	0.44				
B − 1, 24 × 30, Nonflexible	0.21	0.21	0.18				
A − 2, 18 × 24, Flexible	0.57	0.53	0.47	0.52	0.13	0.90	0.39
A − 2, 24 × 30, Flexible	0.58	0.51	0.40				
B − 2, 24 × 30, Flexible	0.90	0.53	0.39				
B − 2, 24 × 30, Nonflexible	0.49	0.45	0.44				

[a]A—Hologic Selenia, B—Hologic Selenia Dimensions.
[b]+2—image receptor 2 cm above IMF, +1-–image receptor 1 cm above IMF, 0-–image receptor on IMF, −1-–image receptor 1 cm below IMF, −2-–image receptor 2 cm below IMF. [c]Average, standard deviation, maximum, and minimum values based on absolute values of uniformity index.

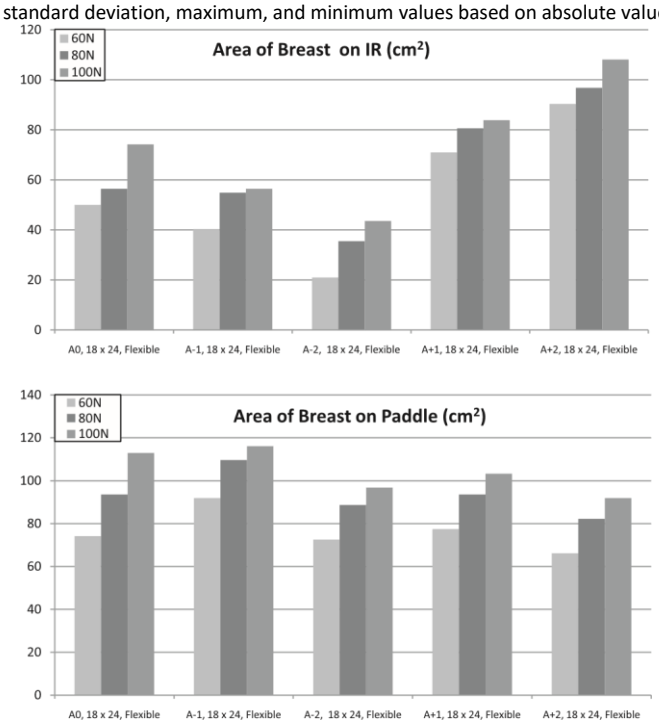


FIG. 2. Area of the phantom in contact with the paddle and the IR.

IV. DISCUSSION

Overall the method was easy to conduct, though the Xsensor pressure mat had to be wrapped carefully around the phantom to avoid creases and therefore artifacts. Data generated by the Xsensor were suited to the calculation of footprints and pressure, however, the resolution was coarse (sensor represented = 1.6129 cm^2). A pressure recording system with a finer resolution would be expected to give more accurate footprints. The main limitation of this study was that the phantom is unlikely to behave similar to the live human female breast. A further limitation is that pressures above 100 N (which are not infrequently used in clinical practice) could not be applied due to the limitations of the pressure mat. On this basis we have been granted ethical permission to conduct a prospective human female study.

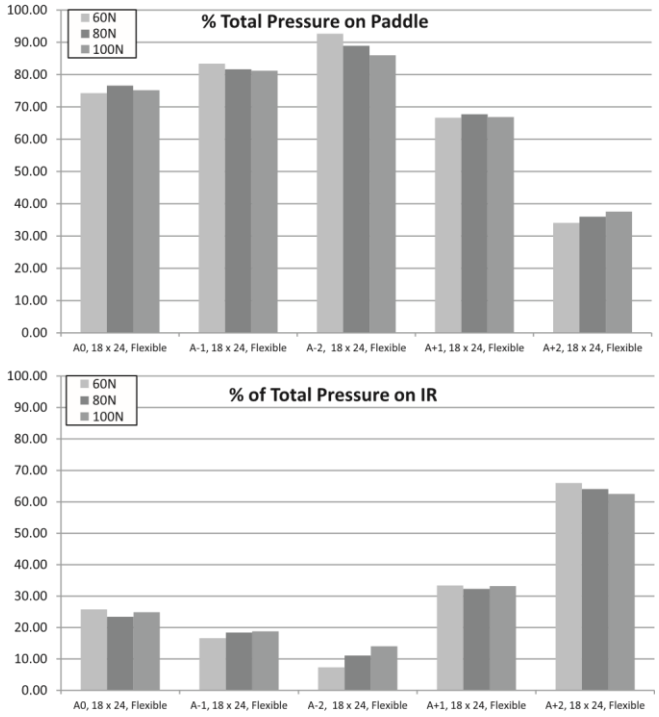


FIG. 3. Pressure exerted onto the phantom from the paddle and from the IR.

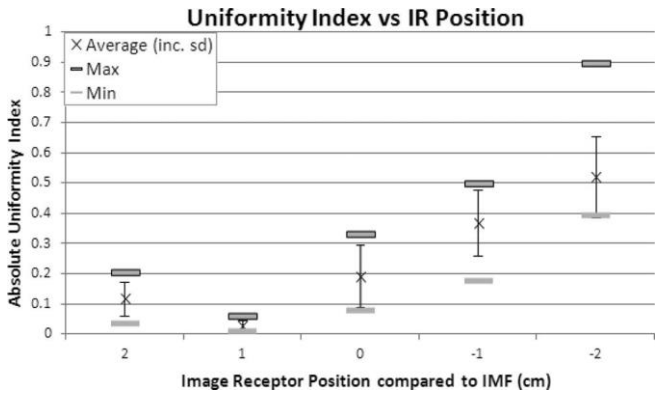


FIG. 4. Absolute uniformity index for each IR position.

Mammography literature suggests that the IR could be positioned at the level of the IMF or raised slightly from it, but no data have been published to demonstrate the effect of this on breast footprint or the balance of pressure exerted on the breast from IR and paddle. For our breast phantom the data demonstrate clearly that as the IR is elevated in relation to the IMF, then the breast footprint on the IR increases to a maximum at +2 cm from the IMF. The balance of pressure exerted onto the breast from the paddle and the IR is also related clearly to the IMF position. For instance, as the IR position decreases to −2 cm in relation to IMF, the pressure exerted from the paddle increases and the pressure from the IR decreases. A balanced pressure from paddle and IR for all paddle/machine combinations is with the IR elevated 1–2 cm from the IMF. Considering the uniformity index (Fig. 4) with the IR elevated +1 cm from the IMF, the best balance of pressure and area is given. On the basis of the phantom findings the data suggest that elevating the IR 1–2 cm will increase breast IR footprint and potentially improve image quality. Similarly, this elevation may more evenly distribute the compression of the breast from above and below which may improve comfort for the person.

V. CONCLUSIONS AND RECOMMENDATIONS

We have created a method for measuring the pressure applied to the breast during mammographic compression by the compression paddle and the IR and for measuring the area of the breast in contact with the paddle and IR. This method has been used to assess the effect of different relative positions of the IR and IMF on these factors during simulated mammography of a female breast phantom. Positioning the IR 1-2 cm above the IMF in the craniocaudal projection results in a better balance of compression on the breast. At the same time this increases the footprint on the IR. A study in human female volunteers using our method is in progress in order to more closely simulate clinical practice. The purpose of this study is to verify whether the data generated from the female breast phantom produces similar results on human female breasts.

References

[1]G. W. Eklund and G. Cardenosa, "The art of mammographic positioning," Radiol. Clin. North Am. 30(1), 21–53 (1992).
[2]A. K. Tucker and Y. Y. Ng, *Textbook of Mammography*, 2nd ed. (Churchill Livingstone, London, 2001).
[3]D. B. Kopans, *Breast Imaging*, 3rd ed. (Lippincott/Williams and Wilkins, Philadelphia, 2007).
[4]G. Barnes, "Mammography equipment: Compression, scatter control, and automatic exposure control," in *Syllabus: A Categorical Course in Physics*, edited by A. Haus and M. Yaffe (RSNA Publications, Oak Brook, 1993).
[5]National Health Service Breast Screening Programme (N.H.S.B.S.P.), "Quality assurance guidelines for mammography, including radiographic control," Publication No. 63 (NHS Cancer Screening Programmes,
Sheffield, 2006).
[6]A. Poulos and M. Rickard, "Compression in mammography and the perception of discomfort," Australas. Radiol. 41(3), 247–252 (1997).
[7]L. Lee, V. Strickland, R. Wilson, and E. Roebuck, *Fundamentals of Mammography* (WB Saunders, London, 1995).
[8]I. Hauge, P. Hogg, K. Szcepura, P. Connolly, G. McGill, and C. Mercer, "The readout thickness versus the measured thickness for a range of screen film mammography and full field digital mammography units," Med. Phys. 39(1), 263–271 (2012).
[9]XSensor, http://www.xsensor.com/pressure-imaging/pressure-imaging (2012).
[10]E. Call and L. Baker, "How does bed frame design influence tissue interface pressure? A comparison of four different technologies designed for longterm or home care," J. Tissue Viability 17(1), 22–29 (2008).

Paper 7

Does elevating image receptor increase breast receptor footprint and improve pressure balance?

H. Smith*, K. Szczepura, C. Mercer, A. Maxwell, P. Hogg

University Hospitals of Morecambe Bay NHS Foundation Trust, Breast Care Unit, Ashton Road, Lancaster, United Kingdom

Article history: Received 1 December 2014, Received in revised form 28 January 2015, Accepted 3 February 2015 Available online 24 March 2015

Keywords: Breast, Cancer, Mammography, Technique, Radiography

Abstract

There is no consensus in the literature regarding the image receptor (IR) position for the cradio-caudal projection in mammography. Some literature indicates the IR should be positioned to the infra mammary fold (IMF); other literature suggests the IR be raised 2 cm relative to the IMF. Using 16 female volunteers (32 breasts) and a pressure sensitive mat we investigated breast footprint and pressure balance with IR at IMF and IR 2 cm above the IMF. Breast area on IR and paddle and interface pressure between IR/breast and paddle/breast were recorded. A uniformity index (UI) gave a measure of pressure balance between IR/ breast and paddle/breast. IR breast footprint increases significantly by 13.81 cm^2 ($p < 0.02$) when IR is raised by 2 cm. UI reduces from 0.4 to 0.00 ($p = 0.04$) when positioned at IMF +2 cm demonstrating an improved pressure balance. Practitioners should consider raising the IR by 2 cm relative to the IMF in clinical practice. Further work is suggested to investigate the effects of practitioner variability and breast asymmetry.

Introduction

Breast compression during mammography is necessary to produce an image of diagnostic quality.[1,2] Effective compression spreads out overlapping tissues to enable better visualisation of breast structures; compression also reduces breast thickness, which minimises radiation to the breast.[3] Good radiographic technique ensures that the maximum amount of breast tissue is imaged adequately so as to optimise lesion visualisation.

There is no consensus in the literature regarding the position for the image receptor (IR) during the exposure for the cranio caudal (CC) projection. Some authors suggest the IR be located at the infra mammary fold (IMF), whilst others indicate it can be elevated slightly from the IMF.[4,5] The intention of elevating the IR relative to the IMF is to increase the amount of breast tissue (the 'breast

footprint') on the IR. This action would bring the object (breast) closer to the IR and potentially enhance image quality by reducing geometric unsharpness. Additionally, elevating IR relative to IMF might improve the balance of pressure on the breast from above and from below, which could result in the procedure being less uncomfortable, as noted by Hogg et al., in 2013.[6] Despite IMF elevation being proposed within the literature no human study has been performed to determine whether the breast footprint increases or pressure balance improves when the IR is elevated.

In 2013 Hogg et al.[6] conducted a phantom study to validate a proposed method to determine the effect of changing the relative positions of the IMF and IR on the breast footprint and pressure balance. The study demonstrated that as IMF is elevated the footprint of the breast phantom increases and a better balance of pressure can be achieved. The paper concluded by suggesting that a human study should be conducted to establish whether the phantom findings hold true in human females. In this paper we used the method described by Hogg et al.[6] on a cohort of 16 human females (32 breasts). Our aim was to evaluate breast footprint and pressure balance with IR at IMF and IR 2 cm above the IMF.

Method

The study was approved as service evaluation by University Hospitals of Morecambe Bay NHS Foundation Trust UHMB, UK; ethical approval was granted by the University of Salford, UK. All women aged 47 to 66 employed by University Hospitals of Morecambe Bay NHS Foundation Trust UHMB were invited to take part in the study. After applying exclusion criteria, 16 participants were selected. Exclusion criteria were: previous breast surgery; pacemaker; current breast symptoms; local skin conditions; currently under investigation for possible breast cancer; breast bra size less than 5 (C cup equivalent). A GE Senographe Essential full field digital mammography (FFDM) with a 24 30 cm fixed compression paddle was used. No x-ray images were taken during this study as the participants were hospital employees and not screening clients or patients.

Using an Xsensor pressure mapping device, which comprises of an array of pressure sensors with a resolution of 1.6129 cm (Fig. 1), the breast phantom method described by Hogg et al. was adopted to collect human data.[7] The Xsensor is a pressure mapping tool in the form of a flexible mat which records, in real time, the pressure (in mmHg) between two contacting surfaces. In this case the interface pressure was recorded between the IR and under-surface of the breast and between the compression paddle and the upper surface of the breast. Pressure readings were taken with the Xsensor pressure mat wrapped around participant breasts (Fig. 1).

For each participant, for left and right breasts, the pressure between the breast and compression paddle and between the breast and the image receptor, and breast footprint on IR with IR at IMF and IR at 2 cm above IMF were recorded. For all but one participant a compression force of 80 N was applied to the breasts.

Two experienced female Health and Care Professions Council registered radiographers qualified in mammography carried out the breast compressions.[8] To simulate clinical conditions each radiographer was instructed to use their normal technique for 'IR at IMF'. IMF +2 cm was achieved by elevating the IR by 2 cm whilst repositioning the compression paddle. For consistency, one radiographer performed the left breast compressions; the other radiographer performed the right breast compressions. In order to minimise the potential for artefacts in the pressure map data one radiographer performed the participant positioning and compression, the other ensured that there were no creases in the pressure mat.

Each participant received four separate breast compressions, two for each breast. A drop in compression force values displayed on the mammography unit was observed for several seconds after compression was initially applied; this phenomenon has been noted previously by Hauge et al.[9] and Ma et al.[10] It was therefore necessary to adjust the compression force until a steady reading of 80 N was maintained. Once the pressure was stable Xsensor pressure data was recorded for 5 s.

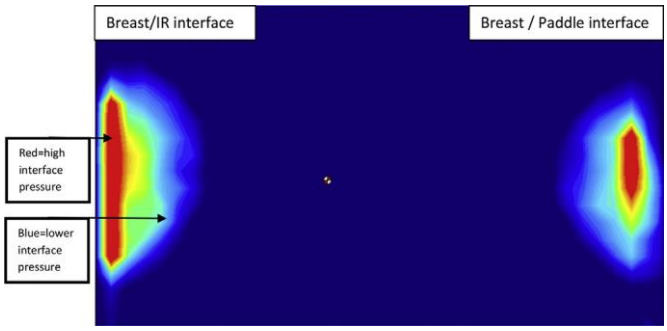


Figure 2. Pressure balance: IR at IMF +2 cm.

Data for the 16 participants was transferred from the Xsensor acquisition module to a password protected laptop computer. Pressure mat data was visually displayed as 2D images, where blue signifies low and red signifies high pressure readings (Fig. 2). Data was also recorded as matrices of pixel values in mmHg to allow analysis of the data, which was performed using Excel.

Prior to analysis the numeric data within Excel was cleaned. This involved deleting artefactual data points not attributable to pressure on the breast. These data points sat outside the breast area and were created by folds in the Xsensor pressure mat.

For each position ('at IMF' and 'IMF + 2 cm') the following measurements were made: paddle and detector footprints (cm^2), percentage of area on IR and average pressure on the paddle and detector (mmHg).

Using these values the uniformity index was calculated where:

$$\text{Uniformity Index (UI)} = (A - B)/(A + B)$$

where:

A = average pressure per unit area applied by the paddle ($mmHg/cm^2$)

B = average pressure per unit area applied by the detector ($mmHg/cm^2$)

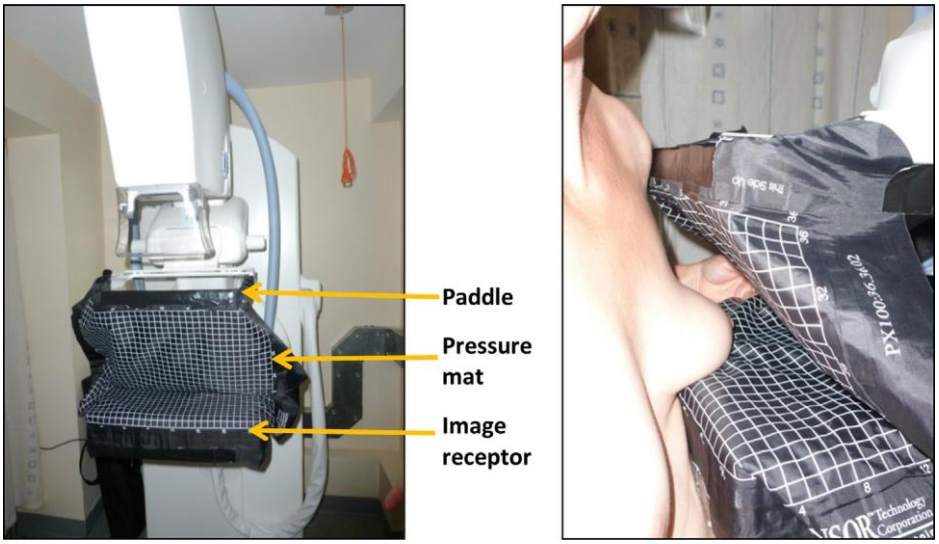


Figure 1. GE Senograph Essential with Xsensor Pressure Mat. Left e Xsensor mat in position on the mammography IR and paddle; Right e Xsensor mat about to be wrapped around participant breast.

The UI value has the following implications. If UI = 0, there is equal pressure per unit area from the IR and the paddle (equal distribution); if 0 < UI > 1, there is greater pressure per unit area from the paddle on the top of the breast, with 1 = all pressure per unit area is applied by the paddle; if 1<UI > 0, there is greater pressure per unit area from the IR on the underside of the breast, with 1 = all pressure per unit area is applied by the IR.

The difference between the area, percentage area, average pressure and UI were calculated between the two positions, and comparisons were made between the radiographers (for right and left breast).

Results

One participant was excluded from the final data analysis due to intolerance of the procedure, resulting in 15 participants (30 breasts, 60 compression readings) being available for analysis. Fig. 2 illustrates an image of the pressure distribution at the 'breast/paddle' and 'breast/IR' interfaces, with IR at +2 cm; Fig. 3 illustrates an image of the pressure distribution with the IR at IMF.

The difference in breast footprint (cm^2) between IR at IMF and IR +2 cm is demonstrated in Graph 1 and Table 1. It is clear that for both left and right breasts, there is a significant increase in IR breast footprint when IR is raised by 2 cm; on average this increase is 13.81 cm^2 ($p < 0.02$). No significant difference was found for paddle breast footprint when raising the IR by 2 cm, with an average decrease in area of 1.06 cm^2 ($p > 0.26$). Graph 2 and Table 2 illustrate these differences in terms of percentage increase in area. For left and right breasts, IR breast footprint percentage area increases significantly by 13.81% ($p < 0.02$) when IR is raised by 2 cm. By contrast there is no significant change in percentage area on the paddle when the IR is raised by 2 cm 0.81% ($p = 0.51$).

Graph 3 and Table 3 illustrate the percentage difference in pressure between IR at IMF and IR +2 cm. As can be seen there are significant differences, however the changes are small in comparison to the increase in footprint. On average the pressure decreased by 0.04% ($p < 0.05$) when IR was raised to +2.

Graph 4 and Table 4 show the uniformity index for IR at IMF and IR +2 cm. As shown, the UI was closer to zero when the IR was positioned at +2 cm. On average there was a significant difference between the UI, which was 0.04 for IR at IMF and 0.00 for IR +2 cm ($p = 0.04$).

All graphs have the following legend:

LCC IR = left breast, Image receptor

RCC IR = right breast, Image receptor

Ave IR = average of left and right breast, Image receptor

LCC P = left breast, paddle

RCC P = right breast, paddle

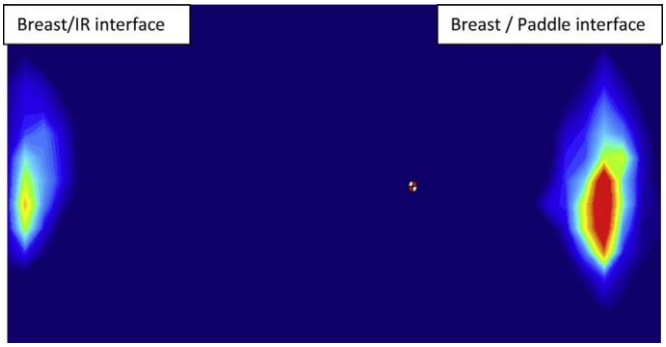
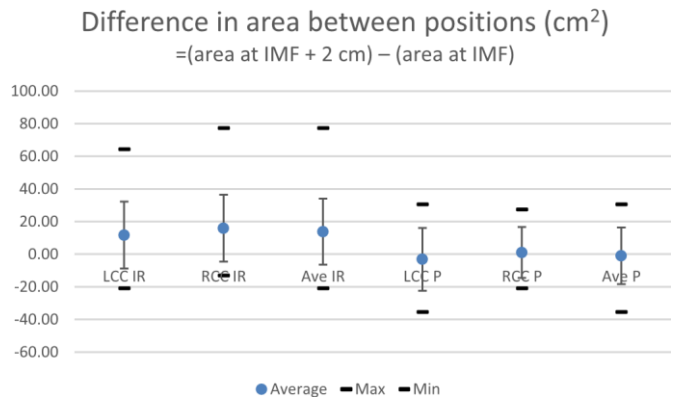


Figure 3. Pressure imbalance: IR at IMF.

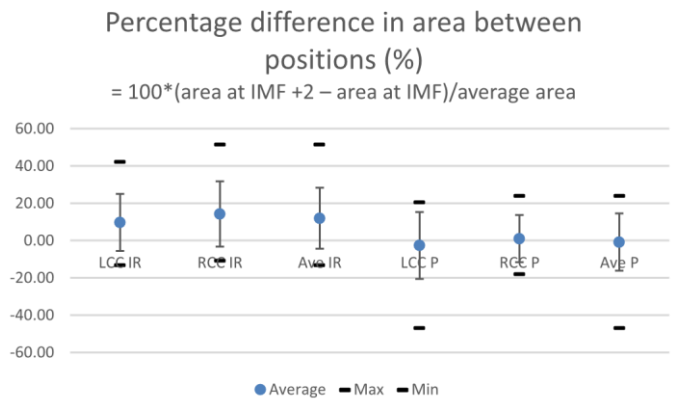


Graph 1. Difference in area between positions (cm^2) = (area at IMF + 2 cm) (area at IMF).

Table 1 Difference in area between positions (cm^2) = (area at IMF + 2 cm) (area at IMF).

	LCC IR	RCC IR	Ave IR	LCC P	RCC P	Ave P
Average	11.69	15.93	13.81	3.12	1.01	1.06
Max	64.52	77.42	77.42	30.65	27.42	30.65
Min	20.97	12.90	20.97	35.48	20.97	35.48
SD	20.51	20.48	20.27	19.23	15.58	17.35
P value	0.04	0.01	0.02	0.53	0.80	0.73

p Values in bold are significant.

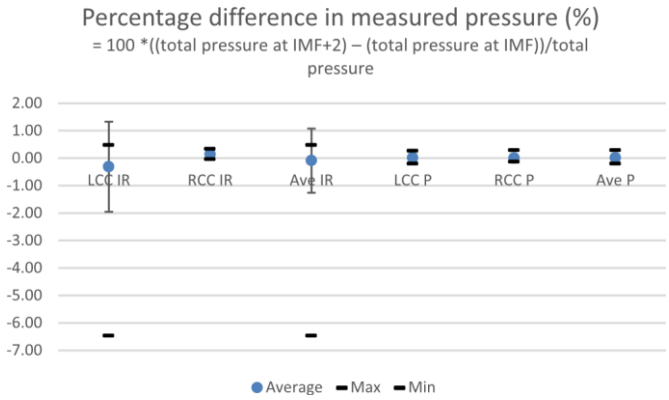


Graph 2. Percentage difference in area between positions (%) = 100*(area at IMF +2 - area at IMF)/average area.

Table 2 Percentage difference in area between positions (%) = 100*(area at IMF +2 area at IMF)/average area. p Values in bold are significant.

	LCC IR	RCC IR	Ave IR	LCC P	RCC P	Ave P
Average	9.68	14.25	11.96	2.62	1.01	0.81
Max	42.11	51.61	51.61	20.54	24.11	24.11
Min	13.33	10.67	13.33	46.81	17.91	46.81
SD	15.33	17.46	16.33	17.92	12.68	15.38
P value	0.04	0.01	0.02	0.53	0.80	0.73

p Values in bold are significant.



Graph 3. Percentage difference in pressure between positions (%) = 100*((total pressure at IMF+2) (total pressure at IMF))/total pressure.

Table 3 Percentage difference in pressure between positions (%) = 100*((total pressure at IMF+2)-(total pressure at IMF)/total pressure.

	LCC IR	RCC IR	Ave IR	LCC P	RCC P	Ave P
Average	0.31	0.14	0.09	0.01	0.01	0.01
Max	0.49	0.34	0.49	0.27	0.30	0.30
Min	6.45	0.04	6.45	0.20	0.14	0.20
SD	1.64	0.12	1.17	0.12	0.10	0.11
P value	0.09	0.01	0.02	0.52	0.53	0.36

p Values in bold are significant.

Ave P = average of left and right breast, paddle

Total = average between left and right breast, paddle and image receptor

At IMF is the baseline

Discussion

Breast footprint increases significantly when the IR is raised by 2 cm from the IMF. There are significant pressure differences between IR at IMF and IR at IMF +2 cm, however these changes are small in comparison to the increase in footprint. UI at IR +2 cm is close to zero, compared with IR at IMF, suggesting a better balance when the IR is raised by 2 cm.

Overall, raising the IR by 2 cm appears to be justified. One data point has a large pressure difference for the LCC IR (Graph 3); this could have been due to an error during data acquisition. Because the data could only be analysed after the data had been collected and because of constraints relating to access to the pressure recording instrumentation we were not able to repeat this measurement. Consequently it cannot be said with any certainty why this anomaly exists

Table 4 Uniformity index. p Values in bold are significant.

LCC at IMF		LCC +2 RCC at IMF		RCC +2 LCC and RCC at IMF		LCC and RCC at +2
Average	0.04	0.02	0.05	0.01	0.04	0.00
Max	0.20	0.16	0.26	0.12	0.26	0.16
Min	0.15	0.10	0.18	0.12	0.18	0.12
SD	0.09	0.08	0.13	0.08	0.10	0.08
P value	0.43		0.07		0.04	

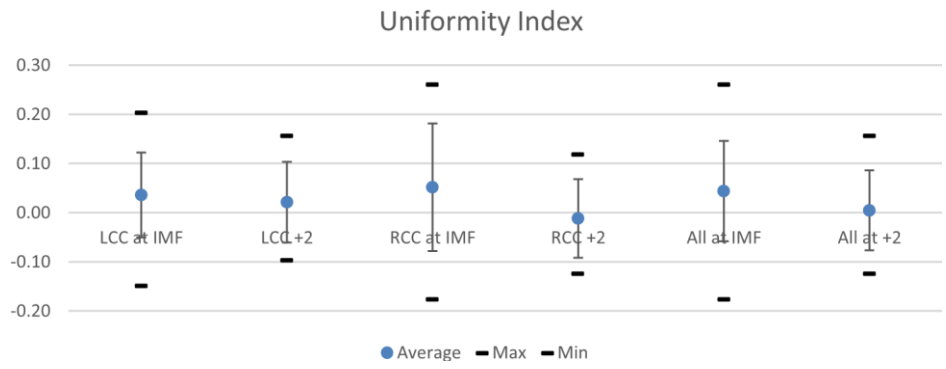
p Values in bold are significant.

Differences existed between left and right breasts. These differences could be due to asymmetry between the left and right breasts, or differences in radiographer technique. Female breasts are rarely the same shape or volume and variation is common,[11-13] which might help explain our findings. As part of our study we could not assess breast volume or shape because bra size is not a reliable indicator of breast size.[14-17] Further work should be considered to examine the potential effects that asymmetry (shape and volume) might have on pressure balance, UI and IR footprint.

Differences in compression forces used in mammography have been reported within and between practitioners[18-20]; these differences are likely to be explained by underlying differences in technique.[21] Such technique differences could extend to where practitioners position the IR, relative to the IMF. In this respect, if technique differences did exist between the two radiographers in our study then this might explain why UI and pressure were different between left and right breasts and this could represent a limitation to our work. Conversely, if practitioner differences are the explanation for UI and pressure differences, between left and right breasts, then this could add external validity to our work by reflecting the practitioner variability within clinical practice. In any event, for left and right breasts, and therefore for both practitioners, breast footprint on the IR increased when the IR was elevated by 2 cm from the IMF.

Conclusion

The data suggests that raising the IR by 2 cm relative to IMF increases the breast footprint on the IR, gives a better pressure balance between breast/IR and breast/paddle and gives a uniformity index close to zero. On this basis practitioners should consider raising the IR by 2 cm relative to the IMF in the clinical practice.



Graph 4. Uniformity index.

Further work is suggested to investigate the effects of practitioner variability and breast asymmetry for breast footprint on IR, pressure balance between IR/breast and paddle/breast and UI.

Conflict of interest statement None.

References

1. Tabar L, Duffy SW, Vitak B, Chen H, Prevost TC. The natural history of breast carcinoma: what have we learned from screening? Cancer 1999;86:449e62.
2. Eklund GW. Mammographic compression: science or art? Radiology 1991;181: 339e41.
3. nhs.uk [homepage on the internet]. Sheffield, UK: NHS Cancer Screening Programmes; 2006 [accessed 16.09.14]. Available from: www.cancerscreening. nhs.uk/breastscreen/publications/nhsbsp61.pdf.
4. Lee L, Strickland V, Wilson R, Roebuck E. Fundamentals of mammography. 2nd ed. London: Churchill Livingstone. Ltd; 2003. p. 31e46; a Daniel Kopans B. Breast imaging. Philadelphia: Lippincott; 1989. p. 43e50; b NHSBSP 63. Quality assurance guidelines for mammography including radiographic quality control. Sheffield: NHS Cancer Screening Programmes; 2006.
5. Huynh PT, Jarolimek AM, Daye S. The false-negative mammogram. Radiographics 1998 SepeOct;18(5):1137e54. quiz 1243-4, Department of Radiology, University of Texas Health Science Center, Houston 77030, USA.
6. Hogg P, Szczepura K, Darlington A, Maxwell A. A method to measure paddleand detector pressures and footprints in mammography. Med Phys 2013;40(4): 041907. http://dx.doi.org/10.1118/1.4792720.
7. http://www.xsensor.com/ [accessed 16.09.14].
8. http://www.hcpc-uk.org.uk/ [accessed 16.09.14].
9. Hauge I, Hogg P, Szczepura K, McGill GG, Connolly P, Mercer CE. The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. Med Phys 2012;39:1.
10. Ma WK, Brettle D, Howard D, Kelly J, Millington S, Hogg P. Extra patientmovement during mammographic imaging: an experimental study. Br J Radiol 2014. http://dx.doi.org/10.1259/bjr.20140241.
11. Manning JT, Scutt D, Whitehouse GH, Leinster SJ. Breast asymmetry and phenotypic quality in women. Evol Hum Behav July 1997;18(4):223e36.
12. Manning JT, Scutt D, Whitehouse GH, Leinster SJ, Walton JM. Asymmetry and the menstrual cycle in women. Ethol Sociobiol 1996;17(2):129e43.
13. Denoel C, Ismael Aguirre MF, Bianco G, Mahaudens PH, Vanwijck R, Garson S, et al. Idiopathic scoliosis and breast asymmetry. J Plastic Reconstr Aesthetic Surg October 2009;62(10):1303e8.
14. Bowles K, Steele JR. Sports brassieres: is there a need for better education? PreOlympic congress. Brisbane, Australia: Sports Medicine Australia; 2000. p. 36.
15. Greenbaum AR, Heslop T, Morris J, Dunn KW. An investigation of the suitability of bra fit in women referred for reduction mammaplasty. Br J Plast Surg 2003;56(3):230e6.
16. McGhee DE, Steele JR. How do respiratory state and measurement method affect bra size calculations? Br J Sports Med 2006;40:970e4.
17. Pechter EA. A new method for determining bra size and predicting postaugmentation breast size. Plast Reconstr Surg 1998;102(4):1259e65.
18. Mercer CE, Hogg P, Lawson R, Diffey J, Denton ERE. Practitioner compression force variability in mammography: a preliminary study. Br J Radiol Feb 2013;86:20110596.
19. Mercer CE, Hogg P, Szczepura K, Denton ERE. Practitioner compression force variation in mammography: a 6-year study. Radiography 2013;19:200e6.

20. Claire E. Mercer, Katy Szczepura, Judith Kelly, Sara R. Millington, Erika R.E.Denton, Rita Borgen, Beverley Hilton, Peter Hogg, A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, http://dx.doi.org/10.1016/j.radi.2014.07.004.
21. Fred Murphy, Julie Nightingale, Peter Hogg, Leslie Robinson, Doreen Seddon,Stuart Mackay, Compression force behaviours: An exploration of the beliefs and values influencing the application of breast compression during screening mammography, http://dx.doi.org/10.1016/j.radi.2014.05.009.

Paper 8

Pressure and breast thickness in mammography—an exploratory calibration study

[1] P HOGG, FCR, [2] M TAYLOR, BSc (Hons), MSc, [1] K SZCZEPURA, BSc (Hons), MSc, [3] C MERCER, BSc (Hons), MSc and [4] E DENTON, FRCR, FRCP

[1]School of Health Sciences, University of Salford, Salford, UK, [2] Radiology Department, North Manchester General Hospital, Manchester, UK, [3]The Nightingale Centre & Genesis Prevention Centre, Wythenshawe Hospital, Manchester, UK, and [4]Radiology Department, Norfolk and Norwich University Hospital, Norwich, UK

Received 24 April 2012, Revised 3 September 2012, Accepted 1 October 2012 DOI: 10.1259/bjr.20120222

Abstract

Objective: To perform a calibration study to provide data to help improve consistency in the pressure that is applied during mammography.

Methods: Automatic readouts of breast thickness accuracy vary between mammography machines; therefore, one machine was selected for calibration. 250 randomly selected patients were invited to participate; 235 agreed, and 940 compression data sets were recorded (breast thickness, breast density and pressure). Pressure (measured in decanewtons) was increased from 5daN through 1-daN intervals until the practitioner felt that the pressure was appropriate for imaging; at each pressure increment, breast thickness was recorded.

Results: Graphs were generated and equations derived; second-order polynomial trend lines were applied using the method of least squares. No difference existed between breast densities, but a difference did exist between ''small'' (15629cm) and ''medium/large'' (18624/24630cm) paddles. Accordingly, data were combined. Graphs show changes in thickness from 5-daN pressure for craniocaudal and mediolateral oblique views for the small and medium/large paddles combined. Graphs were colour coded into three segments indicating high, intermediate and low gradients [≤ -2 (light grey); -1.99 to -1 (mid-grey); and ≥-0.99 (dark grey)]. We propose that 13daN could be an appropriate termination pressure on this mammography machine.

Conclusion: Using patient compression data we have calibrated a mammography machine to determine its breast compression characteristics. This calibration data could be used to guide practice to minimise pressure variations between practitioners, thereby improving patient experience and reducing potential variation in image quality.

Advances in knowledge: For the first time, pressure–thickness graphs are now available to help guide mammographers in the application of pressure.

In 2008, within the UK, breast cancer was the second most diagnosed cancer in females. Internationally, it accounted for nearly 11% of female cancer deaths [1]. For breast cancer detection, mammography plays an important role in screening symptomatic populations and rigorous quality assurance procedures are applied accordingly [2, 3]. There is a particular emphasis on equipment performance [4] and image reader ability to identify abnormalities [5]. By contrast, surprisingly little quality assurance emphasis is placed on the clinical image acquisition phase—especially the optimisation of pressure to reduce breast thickness.

Pressure is considered necessary to reduce breast thickness and for many years this reduction has been associated with image quality enhancement and radiation dose limitation [6]. Within the UK, there is no specific protocol for thickness reduction, but it is generally accepted that pressure should be applied slowly and gently to ensure that the breast is held firmly in place and the skin is taut to touch or that blanching occurs [3, 7, 8]. The National Health Service Breast Screening Programme (NHSBSP) suggests that pressure should not exceed 20 daN. Limited literature exists about the application of pressure. However, Sullivan et al [9] demonstrated a relationship between pressure and thickness, and a maximum value of 16 daN was suggested. By contrast, Chida et al [10] used a standard compression force of 12 daN; if patients experienced pain a reduced force of 9 daN was suggested. Documented variation of opinion therefore exists.

Practitioner subjectivity associated with pressure application has been a concern for many years [11], and in 2004 Poulos and McLean [12] predicted that lack of attention to this could lead to large variations. In 2011, Mercer et al [13] concluded, from a cross-sectional clinical study of 500 females and 14 practitioners (radiographers and assistant practitioners), that large variations existed, and 3 categories of ‘‘compressor’’ were identified by their mean compression values: low—7.4 daN [standard deviation (SD) 1.5]; medium—8.8 daN (SD 1.5); and high—11.1 daN (SD 2.1). Importantly, Mercer et al concluded that the variation is highly dependent upon the practitioner. The study by Mercer et al raises concerns about the consistency of care, radiation dose and image quality, and suggests that more objective criteria for the application of pressure in mammography are required.

On reviewing the literature it is clear that little is published on the optimisation of pressure in mammography; for instance, almost no empirical data are available to describe how the in vivo female breast behaves when pressure is applied to it. This may partly explain why the NHSBSP guidance is lacking in detail and also why this aspect of practice is not adequately quality assured.

In this exploratory study we present a method and data to describe the relationship between pressure and female breast thickness. Because mammography machine and paddle combinations

have readout thickness inaccuracies [14, 15], we have verified the relationship only for one machine by using a sample from its ‘‘typical’’ clinical population. It is worth remembering that Hauge et al [14] used a deformable breast phantom to determine how readout thickness varied from actual thickness; the experiment was conducted under clinically realistic conditions, which incurred bend and distortion across the paddle surface. These are not accounted for in standard medical physics quality control tests. With this in mind, it might be that, for the same pressure, thickness values will be different between mammography machines and different paddles. Similarly, there may be patient differences too, particularly between screening and symptomatic caseloads. Calibrating a mammography unit based on its local caseload would therefore seem an important first step.

Our study follows a similar design to work conducted by Hoflehner et al [16] and Poulos and McLean [12]. For one mammography machine, we outline a method to determine breast compression characteristics which include typical end points for pressure cessation and critical stages within the compression cycle. We conclude by proposing that our approach could be used to establish local pressure standards on which practice might be based and assessed.

Methods and materials

The mammography machine (Hologic^{TM} Selenia; Hologic UK Ltd, West Sussex, UK, full field digital) served only a symptomatic female patient population, from which a sample of 250 patients was drawn. Three paddle sizes were used for imaging [1—small (15629 cm), 2—medium (18624 cm) and 3—large (24630 cm)]. Routine medical physics quality assurance tests performed on the machine indicated it to be operating within expected manufacturer specifications. Owing to refusals (7) and exclusions (8), only 235 patients participated. Reasons for exclusion included breast implants and incomplete sets of pressure/thickness data. To minimise bias, computer-generated randomisation tables were used to select the patients. To meet ethics approval requirements, informed consent was established prior to commencement. Ethics approval was granted by North Manchester General Hospital, Manchester, UK, and the University of Salford Ethics Committee, Salford, UK; the hospital in which the study was conducted considered the work to be ‘‘service evaluation’’, and approval was granted accordingly. As part of the normal mammogram imaging routine, 940 compression sets were acquired, of which 470 were craniocaudal (CC) and 470 were mediolateral oblique (MLO), with left and right described as l and r, respectively.

Five practitioners who held recognised mammography qualifications conducted the mammograms. Prior to the study, to minimise practitioner technique and data recording variability, a 2-week training review was conducted. To help the practitioners, the same assistant was present in the room

for all mammograms to record the pressure and breast thickness data. For the study, all practitioners followed the same technical and positioning procedures; these were in line with published techniques [7]. For rCC, rMLO, lCC and lMLO, automatic machine readouts for breast thicknesses were recorded along with the applied pressures (measured in decanewtons). For the most part, this recording procedure commenced at 5 daN and increased through 1-daN increments until the practitioner had reached the termination pressure and thickness for the patient's mammogram. Factors affecting termination of pressure included patient tolerance and the practitioner deciding that enough had been applied. These factors meant that the lower pressures had more data and the higher pressures had less data. Overall, per patient, the pressure and thickness recording process added to examination time by approximately 2–3 min. Breast density scoring was performed by two experienced observers using the Breast Imaging Reporting and Data System (BI-RADS) classification [17]. Their agreement was high (79%), and to resolve differences in opinion a third experienced observer arbitrated so that agreement was reached in 100% of the cases. Additional data collected on each patient included age and menstrual status.

Results

Each practitioner collected data on different numbers of patients (40%, 13%, 25%, 17% and 5%). Of the patients, 96% were attending the one-stop diagnostic clinic; for 58%, it was their first mammogram attendance. There was a fairly even distribution across the menstrual cycle [1–7 days (16%); 8–14 days (11%); 15–21 days (11%); 21– 28 days (9%); 28+ days (12%); and unknown (1%)], with almost half of the patients being post menopause (40%). Age distribution demonstrates that there was close similarity to the previous 3 years' clients (Pearson's correlation indicates: 2008/study, r=0.926601; 2009/ study, r=0.923102; 2010/study, r=0.944200); BI-RADS density distribution indicates that BI-RADS 4 was undersampled (2%) but BI-RADS 1, 2 and 3 were fairly well represented (20%, 59% and 19%, respectively). Paddles were used with the following frequencies: small, n=19 (8%); medium, n=96 (41%); and large, n=120 (51%). Prior to generating graphs of pressure and breast thickness the data were examined for quality. As noted earlier, it was observed that less sampling was performed at higher compression values. Because of this, to minimise error, for each pressure value, data were excluded that did not have adequate sample size. The cut-off sample size was √N, where N was the maximum number of patients acquired within the chosen group. As the pressure increased, the number of patients able to be sampled decreased, owing to either imaging requirements or patient tolerance. This meant that, as pressure increased, sample numbers decreased. A cut-off sample number was required and this was chosen to be the square

root of N (√N), where N was the number of patients at the initial pressure, as this is the standard error value within a sample (assuming a normal distribution). Sample numbers lower than this value would mean that the sample was below the standard error, leading to high standard deviations. This meant that for all samples a value of 14 daN was the cut-off pressure value.

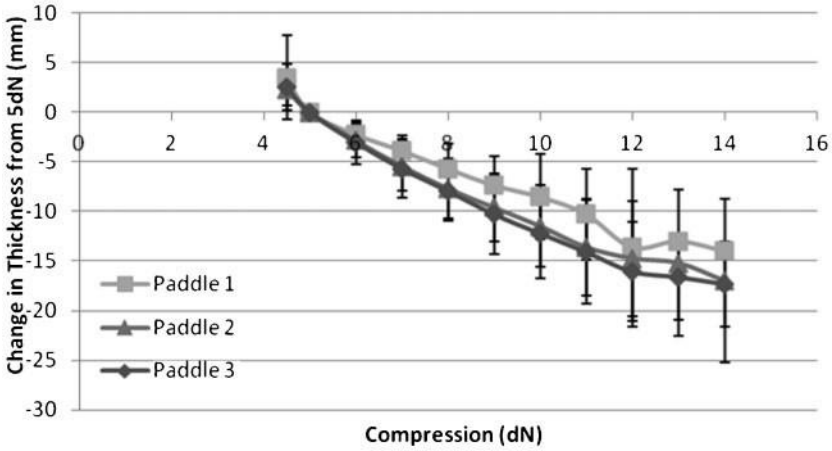


Figure 1 Paddle comparison—craniocaudal view. Paddle 1, small (15x29 cm); Paddle 2, medium(18x24 cm) Paddle 3 large, (24x30cm).

The initial thickness of breast tissue inevitably varied, depending on the patient size; therefore, the change in thickness (measured in millimetres) was evaluated to observe the effect the compression had on the deformation of the tissue. Using graphs, the data are therefore described as the absolute change in breast tissue thickness measured from the thickness at 5 daN in millimetres. Knowing that paddles may have different compression characteristics, data from the three paddles were presented in graphical form (Figures 1 and 2). As can be seen for MLO and CC, Paddles 2 and 3 (medium and large) describe similar characteristics while Paddle 1 (small) is different. Graphs were generated for the BIRADS categories (Figures 3 and 4). It is worth noting that no graph is presented for BI-RADS 4, as only four sets of patient data were available. Because the scatter plot of these four and all of BI-RADS 3 had similar distributions, we included the four into the BI-RADS 3 group to increase sample size.

In Figures 3 and 4, divergences in the graphs can be seen at around 11 daN. These divergences could be explained by the reduced sampling at the higher pressure values; this is illustrated in Figure 5a,b. For MLO and CC, little difference is noted until 11 daN; consequently, accepting that the divergence beyond this point is due to sampling error, all BI-RADS for the small paddle (Figures 6 and 7) and all BI-RADS for the medium and large paddles (Figures 8 and 9) were combined, and composite graphs were created. Error bars demonstrate the standard deviation of the data. Second-order polynomial trend lines were applied to the data using the method of least squares. These gave good correlation

($r^2 > 0.98$) for all data sets. Extrapolation of the data demonstrates the point at which further compression force no longer decreases breast tissue thickness (zero gradients).

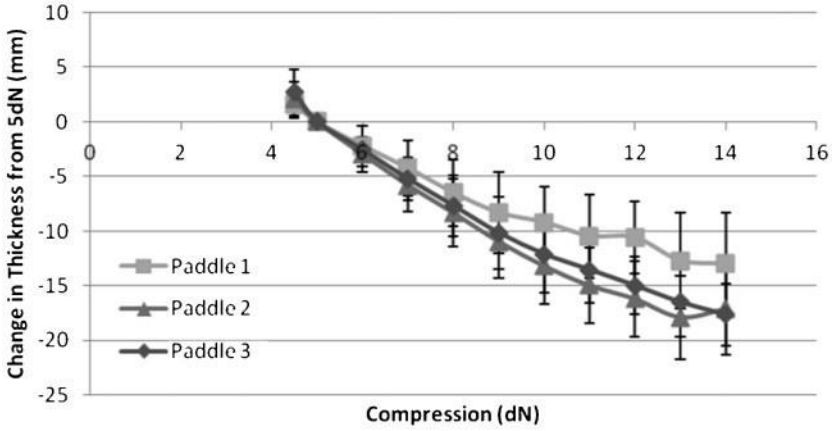


Figure 2 Paddle comparison—mediolateral oblique view. Paddle 1, small (15x29cm); Paddle 2, medium (18x24 cm); Paddle 3, large (24x30cm)

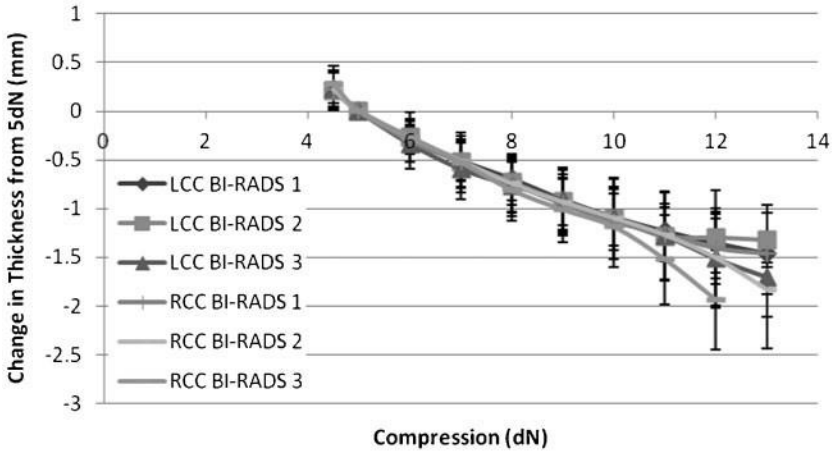


Figure 3 Breast Imaging Reporting and Data System (BI-RADS) comparison—craniocaudal (CC) view.

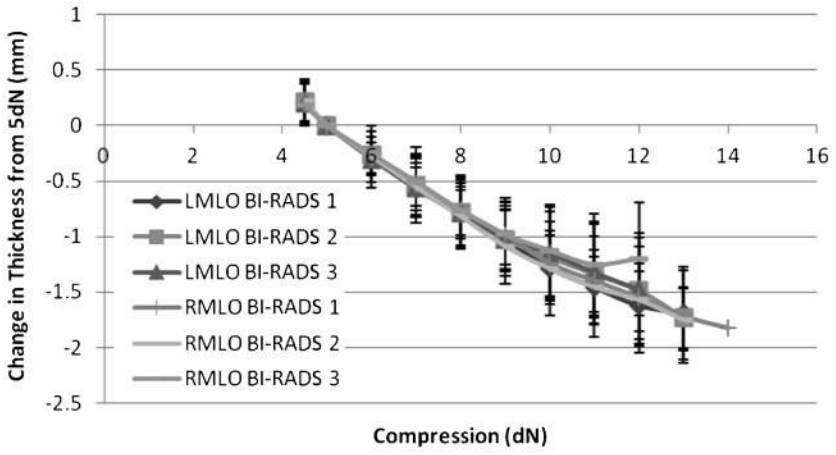


Figure 4 Breast Imaging Reporting and Data System (BI-RADS) comparison—mediolateral oblique (MLO) view. LMLO, left MLO; RMLO, right MLO

Maximum compression forces derived from the composite graphs are: small paddle, CC 18.4 daN, MLO 15.9 daN; medium and large paddles, CC 16.9 daN, MLO 17.3 daN.

Using the applied polynomial trendlines, the equations were differentiated to enable calculation of the gradient at various points. The gradient demonstrated the amount of change of thickness of

tissue, per unit of pressure applied. A higher gradient means a greater reduction in tissue thickness per unit of pressure applied. On this basis, we have colour coded the graphs into three gradient segments: ≤ - 2 (light grey); -1.99 to -1 (mid-grey); and ≥0.99 (dark grey). The use of this gradient calculation and the colour coding is described in the discussion section below.

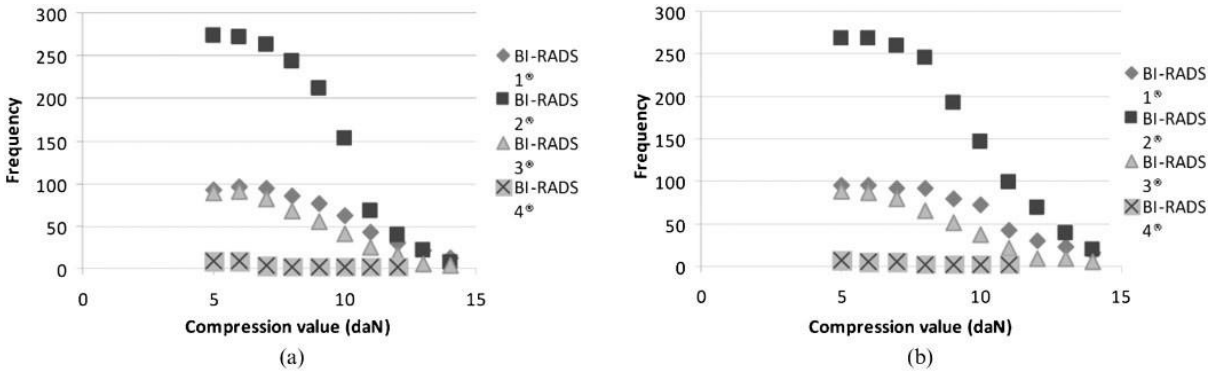


Figure 5 (a) Craniocaudal compressions; (b) mediolateral oblique compressions. BI-RADS, Breast Imaging Reporting and Data System.

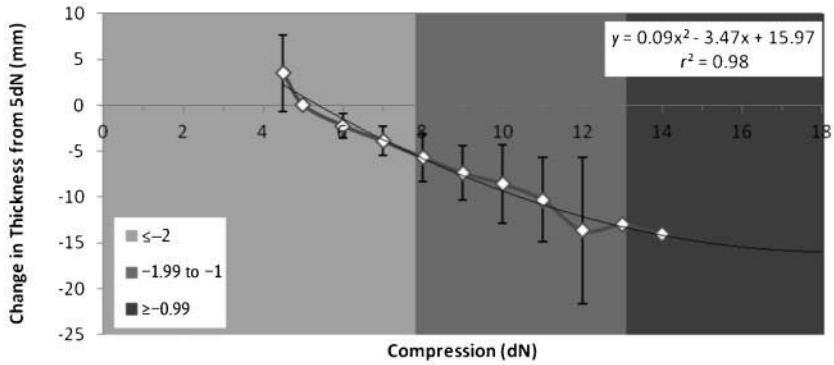


Figure 6 Small paddle—average craniocaudal.

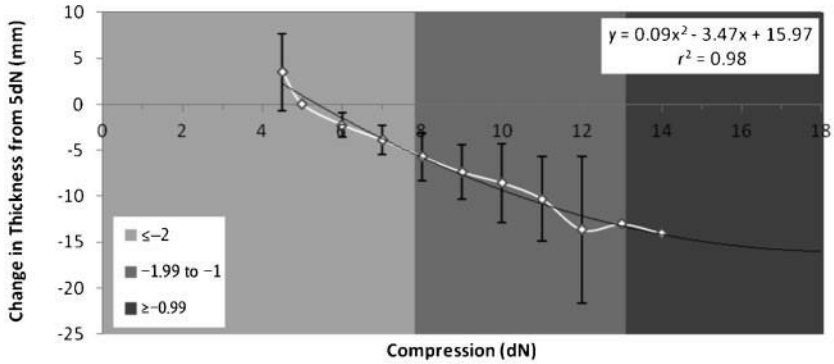


Figure 7 Small paddle—average mediolateral oblique

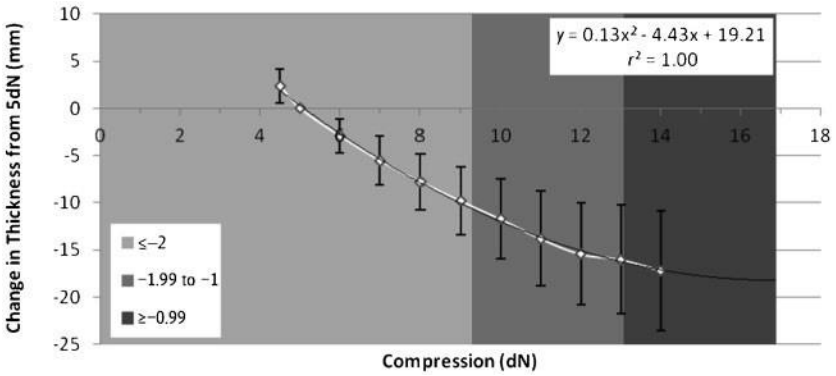


Figure8 Medium and large paddles—average craniocaudal.

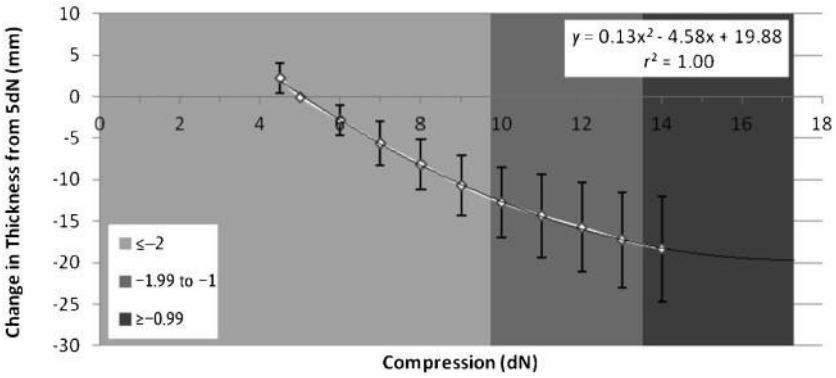


Figure 9 Medium and large paddles—average mediolateral oblique

Discussion

This study was carried out in a symptomatic unit where a larger proportion of younger females are imaged than in a screening setting; 63% of patients imaged were under the age of 50 years. While this may represent a study limitation, it does reflect the clinical norm for this machine's usage in symptomatic practice. Given that the intention was to propose a pressure calibration for the mammography machine using its own patient population, ''oversampling'' of BI-RADS 1–3 would seem to be appropriate, because BI-RADS 4 is likely to be associated with a much younger age.

Surprisingly, on reviewing Figures 3 and 4, there were almost no differences between the BI-RADS densities up to 11 daN (with some divergence beyond this, as explained earlier). This minimal difference may be because of the limited precision for the thickness measurements, suggesting that minor compressibility differences may exist but the machine cannot differentiate them. By contrast, differences did exist between the small and the medium/large paddles (Figures 1 and 2). Patient and paddle factors are likely to account for this. Firstly, the small paddle is used exclusively on small breasts and for these breasts there tends to be less mobility with a much smaller compression capability range. Secondly, the small paddle is non-tilting, unlike the medium and large paddles, which do tilt. Hauge et al [14] noted that larger thickness readout errors are associated with tilting paddles, so the differences could partly be owing to precision. Overall, the lack of difference between BI-RADS scores is helpful because it means that for this machine all BI-RADS scores can be combined

for the small and medium/large paddles, allowing for a simpler process of calibration because only two composite CC and two composite MLO graphs would be required. Applying the data to the clinical setting would also be simplified.

Figures 6–9 demonstrate that SDs tend to increase with increasing pressures. This was explained earlier in relation to the reduced sampling for the higher-pressure values. Should this study be repeated, consideration should be given to how more data might be recorded for higher-pressure values, with due regards to patient comfort and tolerance. However, for all four graphs (Figures 6–9), extrapolation suggests that the NHSBSP maximum of 20 daN was not reached. This indicates that the machine's maximum average pressure falls within the NHSBSP recommendation; on the other hand, it might suggest that for this mammography machine a lower maximum absolute value could be proposed (e.g. 19 daN for small and 18 daN for medium/large paddles). The colour-coded graphs (Figures 6–9) demonstrate areas of different gradients as described within the method. The gradient describes the amount of reduction in tissue thickness per unit of pressure, i.e. the rate of change of tissue thickness. In all cases the light-grey zone depicts a high rate of change, with average gradients of -2.0 and higher. The mid-grey zone depicts a medium rate of change, with average gradients varying from -1.99 to -1.0. Finally, the dark-grey zone depicts a low rate of change, with average gradients varying from 0 to -0.99. On comparison with the light-grey zone, once the dark grey zone has been entered the amount of breast thickness reduction is relatively small compared with the pressure required to effect that change. By contrast, in the light grey zone there is a very high level of thickness reduction achieved for relatively small amounts of applied pressure. As the dark-grey zone is entered, resistance increases rapidly and the potential for pain and discomfort is also likely to increase quickly per applied decanewton. The thickness reduction in the dark-grey zone is low compared with the pressure required to effect that change; therefore, the benefit of applying additional pressure from the point of entering that zone ought to be questioned. On this basis, we propose that the practitioner enter the midgrey zone and then attempt to reach but not necessarily enter the dark-grey zone before ceasing the application of pressure. Consideration for terminating compression for this machine would, therefore, on average, begin approaching 13daN.

Practitioner latitude for the application of pressure would still be expected for patients who experience pain/discomfort and further research is required to assist the practitioners in using graphs of this type. At first presentation for mammography, the graphs could be used to help guide initial pressure and thickness values; for subsequent visits previous thicknesses and pressures should be noted but attention should still be paid to the graphs. It may be valuable to overlay a measure of

pain/discomfort on Figures 6–9 and further research is proposed on this basis. It is also important to recognise that the selection of the critical gradients which differentiate the three shaded grey zones was arbitrary; it is likely that they will be redefined based on experience.

Conclusion

The lack of detail in national guidelines and published literature for the application of pressure in mammography can allow for variation to occur between and within practitioners. This variation may have consequences for mammographic image quality, radiation dose and patient experience.

Using female breast compression data for one mammography machine, we have proposed a method which may help minimise practitioner variability. Our method acknowledges that mammography machines have inherent differences and because of these each machine may require calibration. Additionally, we have acknowledged that different machines will serve different populations and those populations might also affect the calibration. We anticipate that our method and calibration data could be used to inform local practice and also serve as an audit standard. Consequently, we believe that our approach provides evidence for breast compression limits specific to the machine and its population and is therefore likely to have value within other mammography imaging centres. Finally, we would like to propose that our approach may be worth replicating on other mammography machines and paddles, because the resultant data could be used to help improve consistency in the application of pressure.

References

1. cancerresearchuk.org [homepage on the internet]. London,UK: Cancer Research UK; 2010 [accessed 3 February 2012]. Breast cancer UK: Incidence Statistics. Available from: http://info.cancerresearchuk.org/cancerstats/types/ breast/incidence/uk-breast-cancer-incidence-statistics
2. NHSBSP. Consolidated guidance on standards for the NHSbreast screening programme. Publication no. 60. Sheffield, UK: NHS Cancer Screening Programmes; 2005.
3. NHSBSP. Quality assurance guidelines for mammographyincluding radiographic quality control. Publication no. 63. Sheffield, UK: NHS Cancer Screening Programmes; 2006.
4. Moore AC, Dance DR, Evans DS, Lawinski CP, Pitcher EM,Rust A, et al. Commissioning and routine testing of mammographic X-ray systems. IPEM Report 89. York, UK: Institute of Physics and Engineering in Medicine; 2005.
5. lboro.ac.uk [homepage on the internet]. Loughborough,UK: Loughborough University [accessed 3 February 2012]. Available from: http://performs.lboro.ac.uk/index.htm
6. Barnes G. Mammography equipment: compression, scattercontrol, and automatic exposure control. In: Haus A, Yaffe M, eds. Syllabus: a categorical course in physics. Oak Brook, IL: RSNA Publications; 1993.
7. Kopans D. Mammography equipment and basic physics. In: Kopans D, ed. Breast imaging. 3rd edn. Philadelphia, PA: Lippincott Williams and Wilkins; 2007. p. 255.
8. Poulos A, Rickard M. Compression in mammography andthe perception of discomfort. Australas Radiol 1997;41: 247–52.
9. Sullivan DC, Beam CA, Goodman SM, Watt DL. Measurement of force applied during mammography.
Radiology 1991;181:355–7.
10. Chida K, Komatsu Y, Sai M, Kakagami A, Yamada T,Yamashita T, et al. Reduced compression mammography to reduce breast pain. Clin Imaging 2009;33:7–10.
11. Keefe FJ, Hauck ER, Egert J, Rimer B, Kornguth P. Mammography pain and discomfort: a cognitive-behavioural perspective. Pain 1994;56:247–60.
12. Poulos A, McLean D. The application of breast compressionin mammography: a new perspective. Radiography 2004;10: 131–7.
13. Mercer CE, Hogg P, Lawson R, Diffey J. Radiographervariability of breast compression in mammography. Br J Radiol 2013; in press.
14. Hauge IH, Hogg P, Szczepura K, Connolly P, McGill G,Mercer C. The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. Med Phys 2012;39: 263–71.

15. Mawdsley GE, Tyson AH, Peressotti CL, Jong RA, Yaffe MJ.Accurate estimation of compressed breast thickness in mammography. Med Phys 2009;36:577–86.
16. Hoflehner H, Pierer G, Rehak P. ‘‘Mammacompliance’’: anobjective technique for measuring capsular fibrosis. Plast Reconstr Surg 1993;92:1078–84.
17. acr.org [homepage on the internet]. Reston, VA: AmericanCollege of Radiography; 2003 [accessed 19 November 2012]. Available from: http://www.acr.org/Quality-Safety/Resources/BIRADS/Mammography

Paper 9

Validated novel software to measure the conspicuity index of lesions in DICOM images

K. R. Szczepura, D. J. Manning
University of Salford, Manchester, UK

Proc. SPIE 9787, Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment, 978703 (24 March 2016); doi: 10.1117/12.2216211

Abstract

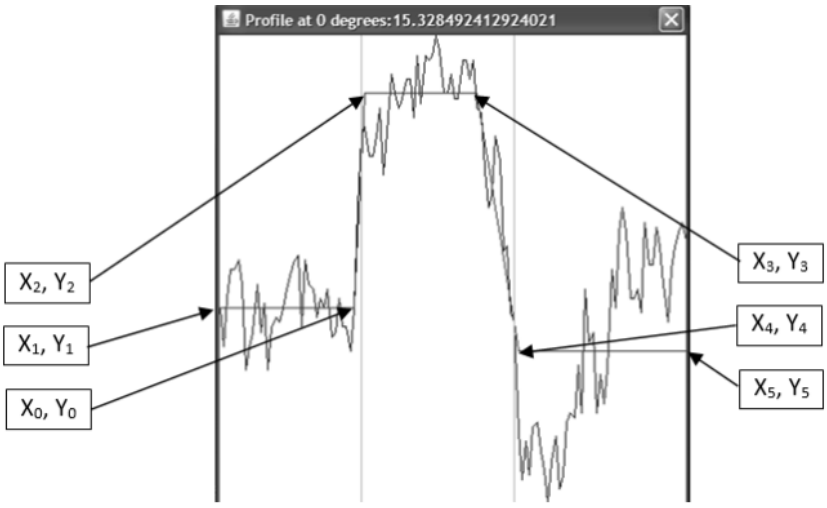
Description of purpose

A novel software programme and associated Excel spreadsheet has been developed to provide an objective measure of the expected visual detectability of focal abnormalities within DICOM images.

Methodology

ROIs are drawn around the abnormality, the software then fits the lesion using a least squares method to recognise the edges of the lesion based on the full width half maximum. 180 line profiles are then plotted around the lesion, giving 360 edge profiles.

The co-ordinates show in Figure 1 are captured, as well the standard deviation of the pixel values within the background and lesion (representing anatomical noise and lesion noise respectively).



An Excel spreadsheet has been developed to allow variables to be calculated, including SNR and CNR. A conspicuity index has also been developed:

$$\chi = \frac{d \tan[\theta - 1]\Delta GL}{\sqrt{\sigma_s^2 + \sigma_n^2}}$$

Where:

χ = conspicuity index

d = lesion dimension ($X_3 - X_2$)

θ = maximum slope angle of all the line profiles

For each side of the line profile:

$$\theta = \tan^{-1}\left[\frac{Y_1 - Y_0}{X_1 - X_0}\right]$$

$$\theta = \tan^{-1}\left[\frac{Y_1 - Y_0}{X_1 - X_0}\right]$$

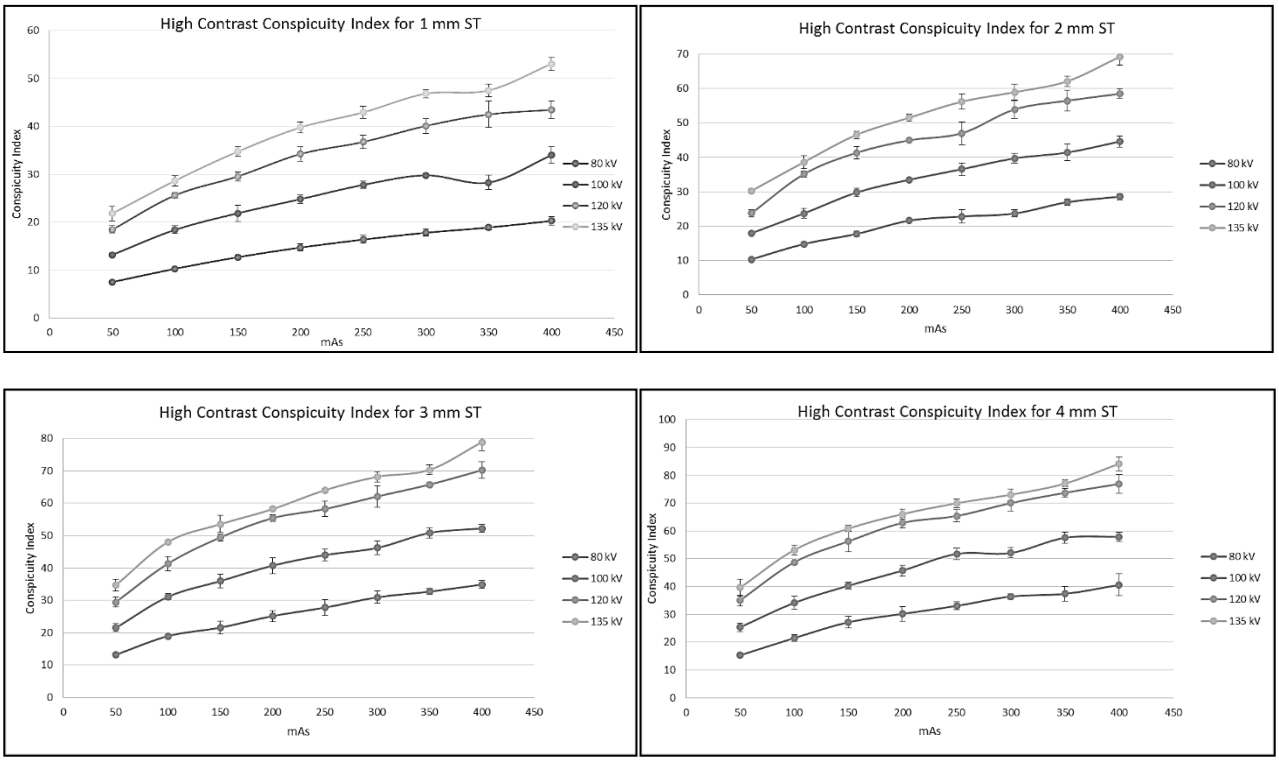
$\Delta GL = Y_2 - \frac{(Y_1 + Y_4)}{2}$ = the difference in average grey levels between the lesion and the background

$\sigma_s = sdTop$

$\sigma_n = \frac{sd1 + sd2}{2}$

Results

The software has been validated using the GAMMEX ACR CT accreditation phantom, varying mA, kVp and slice thickness (ST) and the results have been found to give a linear response:



Conclusion

A novel software programme has been validated to allow calculation of many physical properties of lesions. Additionally, a new measure of conspicuity index has been developed for focal lesions.

The analysis could be further developed to incorporate reader decision-analysis data and eye-tracking data allowing correlations between physical and perception measures to be made beyond

basic CNR calculations. It could also be used as a tool to distinguish between perceptual and cognitive error.

Further refinements could lead to measures of the detectability of more diffuse disease features.

Keywords: objective, conspicuity, index, image quality

INTRODUCTION

It is well known that the performance of people reporting medical images is not perfect, both false positive and false negative decisions can be made, which can impact on the patient pathway. Errors such as these can be put down to two main factors, poor lesion detectability due to the imaging process, and cognitive or perceptual errors.

The quality of an image can be determined in numerous ways. Medical imaging systems are maintained and quality assured using structured phantoms, and measures such as signal to noise ratio (SNR) contrast to noise ratio (CNR), spatial resolution in terms of modulation transfer function (MTF) [1, 2] are used to ensure the system is still functioning within expected limits. It needs to be recognised however, that these measures, although essential for maintaining and assuring standards of equipment, do not represent the complexities of a diagnostic image[1] and their outcome measure is not directly related to any specific radiological task. As well as the natural tissue variations within patients, quality control measures are taken under strict protocols, positioning and acquisition parameters, which do not represent the acquisition of an image of a patient. The term lesion conspicuity is an attempt to represent the visibility of a lesion taking into consideration the structure of the lesion, and the tissues surrounding it as well as taking into account the diagnostic potential of the image for a well-defined radiological task.

Conspicuity

The term 'conspicuity' was first used in 1974 by Revesz et al[3], and defined as the lesion contrast divided by the surrounding complexity, where the lesion contrast is the density change across the lesion border, and the surround complexity is the rate of fluctuation of the density around the lesion border. This led to a definition of structural, or anatomical, noise which takes into account the surrounding structures and artefacts. However is was recognised within this seminal piece of work that this is a simple equation and does not take into account other features that could affect the conspicuity of a lesion.

Factors affecting conspicuity

An important consideration to take into account is that a diagnostic radiograph is not a simple shadow of the anatomy, it is a complex summation of a polyenergtic beam of X-rays that have interacted in a 3D object within multiple layers. Therefore, as it is a summation of several layers, subtle lesions may be obscured by overlying anatomical structures[4]. Samei et al[5] inserted simulated lesions at different positions on a chest radiographs and found significant variation in lesion detectibility as a function of position, finding that local anatomic variations surrounding and overlying subtle lesions on a chest radiograph that are created by the projection of anatomic features in the thorax, such as ribs and pulmonary vessels, can greatly influence the detection of lesions.

Even in 3D imaging such as CT, adjacent anatomy can obscure lesions. Li et al[6] undertook a review of missed cancers during CT screening using low dose helical CT in a general population, and found that all 32 missed lesions were found to be in the interpulmonary region, and that of the 23 missed lesions due to detection (rather than interpretation) error, 19 of those were due to the fact that the lesions appearance was similar to that of normal structures. When measuring signal to noise or contrast to noise ratio, usually the noise is measured from a background that doesn't include complex surrounding structures, therefore these complexities that affect the detectability of a lesion are not recognised with these standard measures.

Noise is defined as unwanted information on the image[1], and there are two types that need to be taken into account when considering the conspicuity of the lesion: the structural noise and the radiographic noise. Structural noise, as discussed above, is task dependant. However radiographic noise is not dependant on the subject being imaged, but is stochastic in nature and dependent on many factors, such as the exposure factors used and the capability of the detector[1]. Radiographic noise is measured using SNR, and provides information about the system capability, but does not give the full information about the noise within a clinical image.

The size of a focal lesion is an important factor in determining its conspicuity. Many studies have found that lesions that are smaller than 1cm are missed when viewing 2D images[7], and 3mm is deemed the threshold size limit for detecting a lesion[4]. In mammography Birdwell et al[8] found that 81% of missed lesions in their study were found to be less than 20mm even when retrospectively assessed, and Michealson et al[9] found that the median size of lesions to be detectable by mammographic screening was 7mm, with only 40% of lesions being seen at 5mm. As previously discussed, Li et al[6] found that in CT 11 of the 32 missed lesions could not be identified as they were barely discernible due to their size (<2mm). It is recognised, however, that even the smallest image

feature can be detected if its contrast against the background is high enough, and this principle is used to good effect in the contrast/detail test objects in quality assurance programmes [33].

Contrast is an essential measure of the diagnostic capability of a system. CNR and SNR are standard measures used to represent this, in clinical assessment of images as well as in quality control. They are both size independent if they are used in isolation as quantitative measures of image quality. They represent the difference in signal amplitude between the lesion or test feature and the background [2, 10]. Both SNR and CNR are used as an objective measure of the quality of an image. Measures are taken of the signal of the lesion, and the signal of the surrounding areas and this data is used to calculate either the SNR or the CNR. Contrast agents such as barium sulphate in fluoroscopy, iodine in CT etc. can be added to improve visibility of the structures of interest. Improved contrast leads to greater detectability of a lesion, and so is an important measure of the image quality capabilities of a system.

The sharpness refers to the ability of the system to represent distinct anatomical features within the object being imaged[1], therefore the sharpness of the border of a lesion impacts on its visibility. There are mathematical techniques to measure the sharpness of a system. Point and line spread functions and their Fourier transform, the MTF all measure the resolving capability[1] or spatial resolution of a system, and take into account issues such as focal spot size, the polyenergetic beam, and any magnification that can cause blurring of edges. Blurring can also occur due to voluntary or involuntary patient motion. Blurring causes reduced visibility of details, image un-sharpness and reduced spatial resolution. The 3 mm minimum size of lesion visible, as discussed above is only applicable if the edges of the structure are parallel to the X-ray beam, if the margins are bevelled (either due to blur, or anatomical causes) then this influences the visible threshold size[4]. Edge sharpness has a powerful influence on the probability that an image feature will be detected. This is because loss of spatial resolution reduces the effectiveness of the Mach band phenomenon in the visual system which so enhances narrow gradients between adjacent regions of different grey-level [11].

In summary, the factors that have been reported to affect the conspicuity of a lesion have been found to be the structural noise within and surrounding a lesion, the size of the lesion, the contrast and the sharpness of the edges.

Measuring Conspicuity

Likert scales are often used in observer studies in attempts to quantify conspicuity [12-18] , when comparing modalities, varying acquisition factors, or using different contrast agents. The likert scales used vary greatly within these studies, in terms of number of points on the scale, and the terms used for each of the points, and the scale is always task dependant. However, the use of likert scales are open to bias[19], because both perceptual and cognitive errors may be included within the task, therefore, the use of likert scales is not an objective measure of conspicuity.

Conspicuity has also been assessed using the two alternate forced choice method, 2AFC are those where the participant is asked to choose between two images, one is a static reference image the other varies depending on the research question being asked. Due to the way the studies are designed, these techniques are less susceptible to bias than likert scales and present low variability. However large sets of images are needed and it can be time consuming both in terms of the individual observer, and the number of observers required in the study, making it a costly technique. It needs to be noted however that numerous questions can be asked of the observer, not just the conspicuity, during 2AFC, adding to the complexity depth of the data obtained. An issue to consider with 2AFC is that it provides ordinal data, where the images are sorted in order, and the difference between the ranks are assumed to be equal, therefore the findings are dependent on the dataset used within the study.

Various studies have used region of interest data to measure conspicuity[20, 21] comparing the lesion to the surrounding tissue, and although this is an objective measure that indicates the conspicuity of the lesion it does not take into account the other factors that affect the conspicuity, such as lesion size and sharpness as discussed above.

Manning et al[22] first proposed a method of combining the factors that describe conspicuity into a single equation. This equation took into account those factors that impact on conspicuity of a lesion based on the saliency of the image feature to the visual system, Line profiles were plotted across a lesion, extending into the immediate surrounding background and the following equation was developed.

$$\chi = \frac{d \tan[\theta - 1]\Delta GL}{\sqrt{\sigma_s^2 + \sigma_n^2}}$$

Equation 1 calculation of conspicuity index[22]

Where:

χ = conspicuity index

d = maximum lesion dimension

θ = maximum edge angle

ΔGL = mean contrast (difference in grey level)

σ_s = mean noise within the lesion

σ_n= mean background noise

However, there were limitations to this work. The analysis was performed on chest films that were converted to digital format, rather than digital images and only four line profiles were plotted. This was unable to capture the features of the entire lesion, gave only limited information on the distribution of background noise and introduced losses in the digitisation process, It was also dependent on the operators' choice of placement of the lines. A further limitation of the work was the decision to use the maximum angle from the four profiles rather than the mean value of the edge angles. This gave little insight into the radial extent of the maximum edge angle and could therefore overestimate its influence on the salience of the whole lesion. Nevertheless it showed there was value in making such measurement because it could help discriminate detection and decision errors, provide a perceptual guide to the technology of CAD and image enhancement processes and it extended the range of parameters available in eye-tracking experiments. The present work builds on this original set of ideas and provides a more refined and precise means of quantifying the perceptual salience of focal image features.

Perceptual Measures

Kundel et al [23] defined three lesion detection error classifications; search or scanning errors, recognition errors, and decision making errors. Scanning errors are due to the failure of the observer to fixate on the lesion, recognition errors are when the lesion is fixated on yet failing to detect it as a lesion, and decision making errors are incorrect interpretation of the observed lesion[23], by far the largest proportion of errors have found to be in decision making[23]. Other perceptual errors come from satisfaction of search (SOS) error, where the observers' attention is diverted from the lesion by a more conspicuous finding [24-26].

Perception based image quality assessment is far more useful for determining diagnostic capabilities than objective measures, as it is directly related to the way the operator looks at the images, where the quality is directly related to the ability to detect and recognise the lesion[27]. However human

observer studies such as ROC-AUC[28] or 2AFC[29, 30] studies can be time consuming, and potentially expensive, and can also yield inconclusive findings, and so additional objective measures are useful to support human observer studies[27].

When Revesz et al[3] measured the conspicuity of lung nodules, and film readers were asked to determine the presence and location of any lesions, they found that the calculated conspicuity correlated with the probability of detection within a limited range of conspicuity values. Samei et al[5] measured the contrast-diameter product based on local anatomical noise, and found a correlation with detectability of the lesions with this measure.

However, attempting to correlate objective measures with human observer studies has not always proved successful, Manning et al[22] found poor correlation between the measure of conspicuity index and missed lesions in chest radiography and indicated that decision errors were more common than those of detection. However it is recognised that the approach to calculating conspicuity was limited, as only four profiles and also the lack of correlation involved other types of observer error as originally noted by Kundel et al[22]. Mello-Thomas et al[31] also found during an eye tracking study, that unreported lesions often received adequate visual attention and other eye tracking studies have found that the conspicuity of a lesion, or the amount of time it is observed, is not the only reason they are not reported[22, 32-34]. In short, although errors in radiology are not confined to readers missing features because they are poorly demonstrated it is valuable to inform efforts to improve the radiological task with measures of image quality that use functional data on visual performance. Conspcuity indices make a contribution to this endeavour.

Methodology

Software design

A novel software programme was developed to enable analysis of focal lesions within DICOM images. Using a JAVA based programme, an operator is enabled to draw a region of interest (ROI) around a lesion. Any number of ROIs within the image can be drawn.

The software then plots 180 line profiles around the ROI, giving data 360° around the lesion.

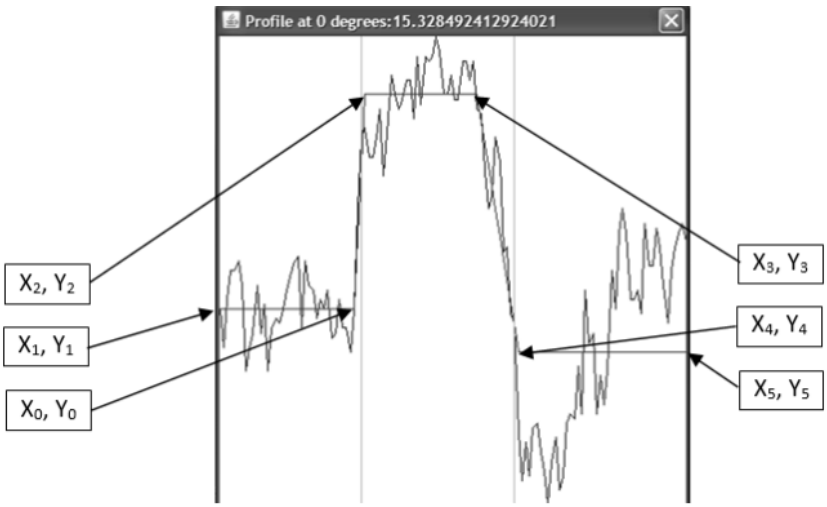


Figure 1 line profiles from ROI

Figure 1 gives an example of a single line profile across an ROI. The least squared method is used to fit the profile, this improves objectivity if the ROI is not drawn exactly around the lesion.

The following data is then exported as a TSV file for use in Excel; the profile values (in grey scale values) of $X_{0,1,2,3,4\&5,}$ $Y_{0,1,2,3,4\&5,}$ the standard deviations of the Y values between X_0 & X_1 (sd_1) and X_4 & X_5 (sd_2) are used to represent the (structural) background noise, and the standard deviation of the Y values between X_2 & X_3 (*sdTop*) are used to represent the noise within the lesion.

Data analysis

An Excel spreadsheet has been developed to import the TSV files and then to calculate the following:

$$\bar{d} = \overline{(X3 - X2)}$$ = average lesion dimension

Equation 2

$$\bar{\theta} = \overline{\tan^{-1}\left[\frac{Y_2 - Y_1}{X_2 - X_1}\right], \tan^{-1}\left[\frac{Y_4 - Y_3}{X_4 - X_3}\right]}$$ = average slope angle of all the line profiles

Equation 3

$$\overline{\Delta GL} = \overline{Y_2 - \frac{(Y_1 + Y_4)}{2}}$$

$=$ the average difference in grey levels between the lesion and the background

Equation 4

$$\overline{\sigma_s} = sdTop$$

Equation 5

$$\overline{\sigma_n} = \frac{sd1 + sd2}{2}$$

Equation 6

This data was then used to calculate the Conspicuity Index (χ):

$$\chi = \frac{\bar{d}\tan[\bar{\theta} - 1]\overline{\Delta GL}}{\sqrt{\overline{\sigma_s^2} + \overline{\sigma_n^2}}}$$

Equation 7

A distinction between Equation 1 and Equation 7 is that instead of using the maximum slope angle and lesion dimension has been used, this is due to the sensitivity of the data, and the maximum values having the possibility of representing noise instead of real values. However, this can be amended to calculate any values as required.

CNR and SNR can also be calculated using the following equations

$$CNR = \frac{\overline{Y_2 - \left(\frac{Y_1 + Y_4}{2}\right)}}{Y_2}$$

Equation 8

$$SNR = \frac{\overline{Y_2 - \left(\frac{Y_1 + Y_4}{2}\right)}}{\sqrt{\overline{sdTop^2} + \left(\frac{\overline{sd1} + \overline{sd2}}{2}\right)^2}}$$

Equation 9

Software reliability

The software was tested for reliability by using the Gammex ACR CT accreditation phantom. A Toshiba Aquillion 16 CT scanner was used, and images were acquired using the following parameters: kVp 80, 100, 120, 135; 50-400mAs, in 50 mAs increments; 1, 2, 3, 4 mm slice thickness (acquired and reconstructed) all using singe slices through Section 4 of the phantom, an example image is shown in Figure 2 below. ROIs were drawn and the conspicuity index was calculated. This was repeated four times to assure reproducibility of the conspicuity index calculation.

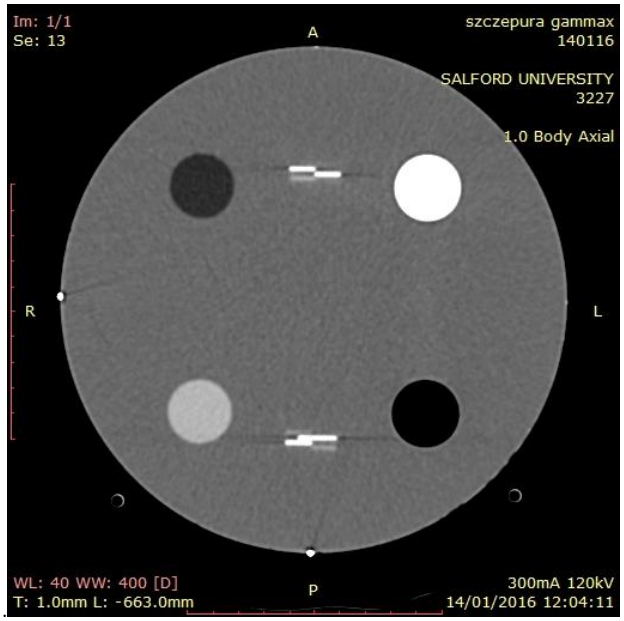


Figure 2 - Gammex ACR CT Phantom

Phantom Data

The software was also tested using a realistic anthropomorphic chest phantom “Lungman” [35], with 4 lesions inserted Figure 3, Figure 4 and Table 1 below.

Figure 3 - Lungman phantom

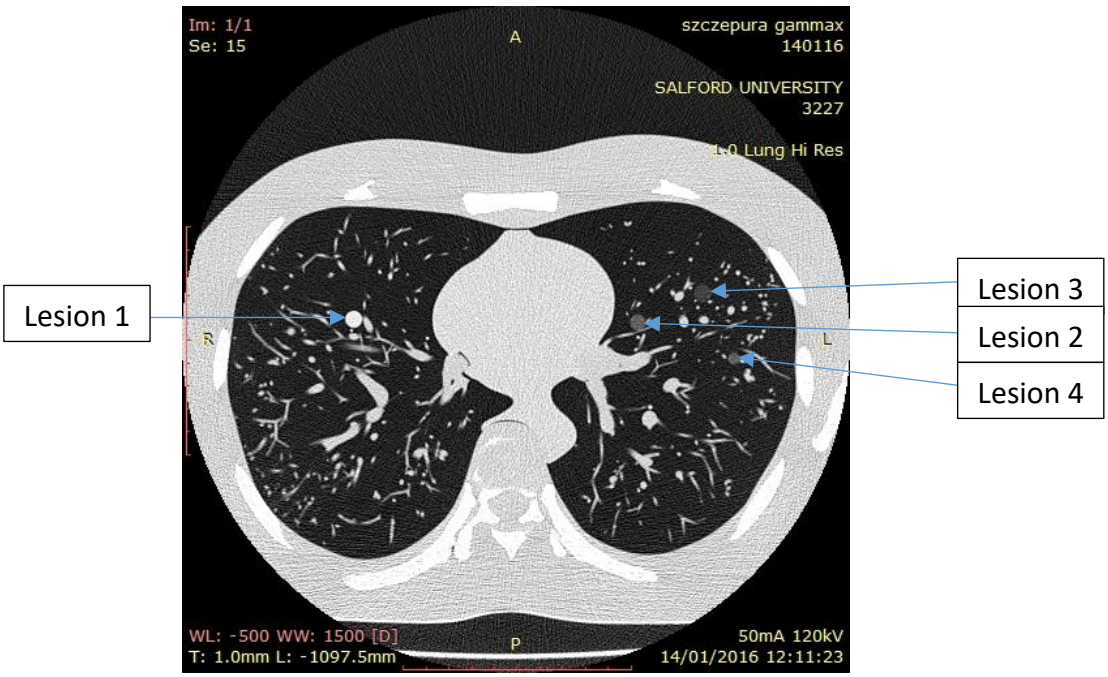


Figure 4 - CT image of Lungman phantom with lesions

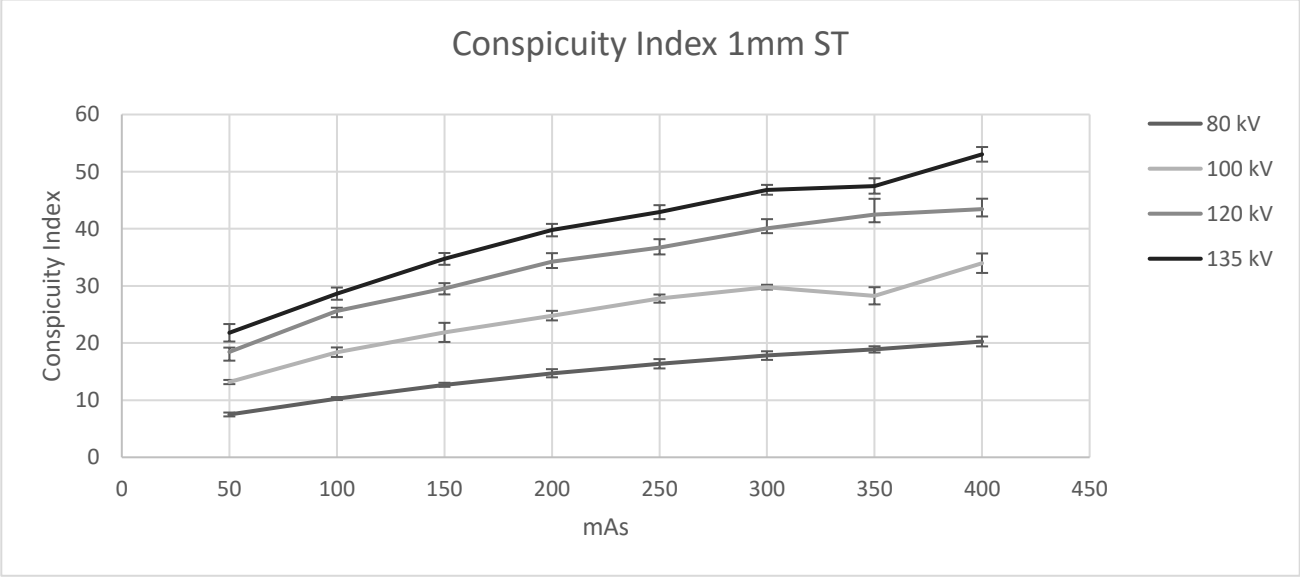
Lesion Number	Size (mm)	Hounsfield Units
1	8	+100
2	8	-630
3	8	-800
4	5	-800

Table 1 - Lesion inserted into Lungman phantom

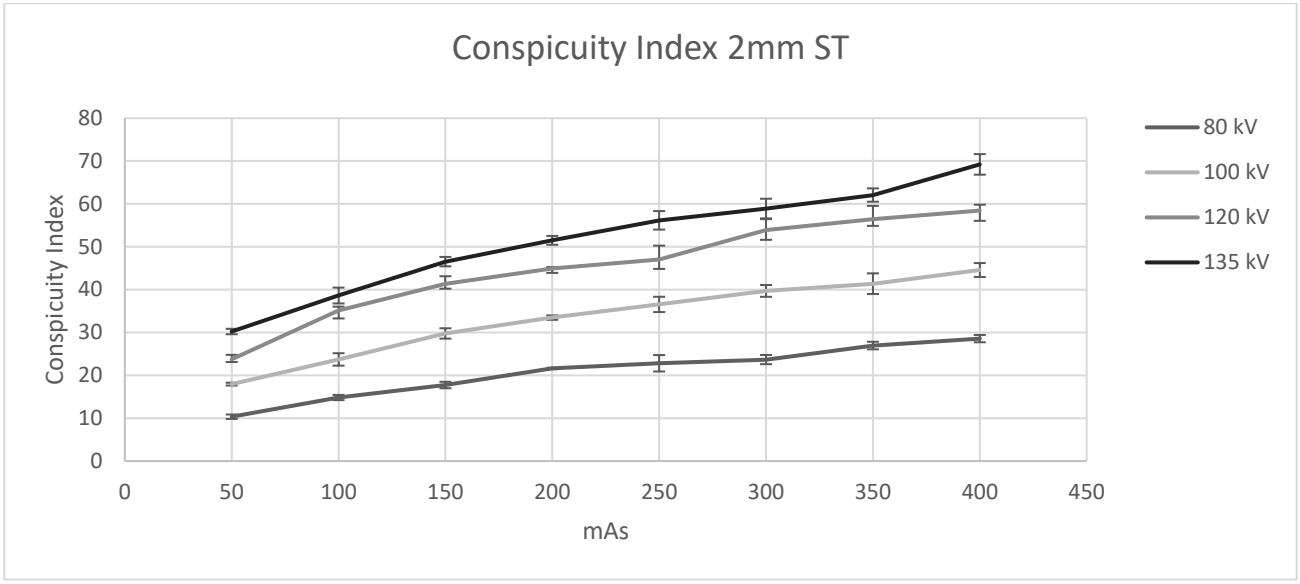
The phantom was imaged using the same CT scanner on a 1mm High Resolution Chest CT protocol, with varying mAs of between 25-600 mAs.

Results

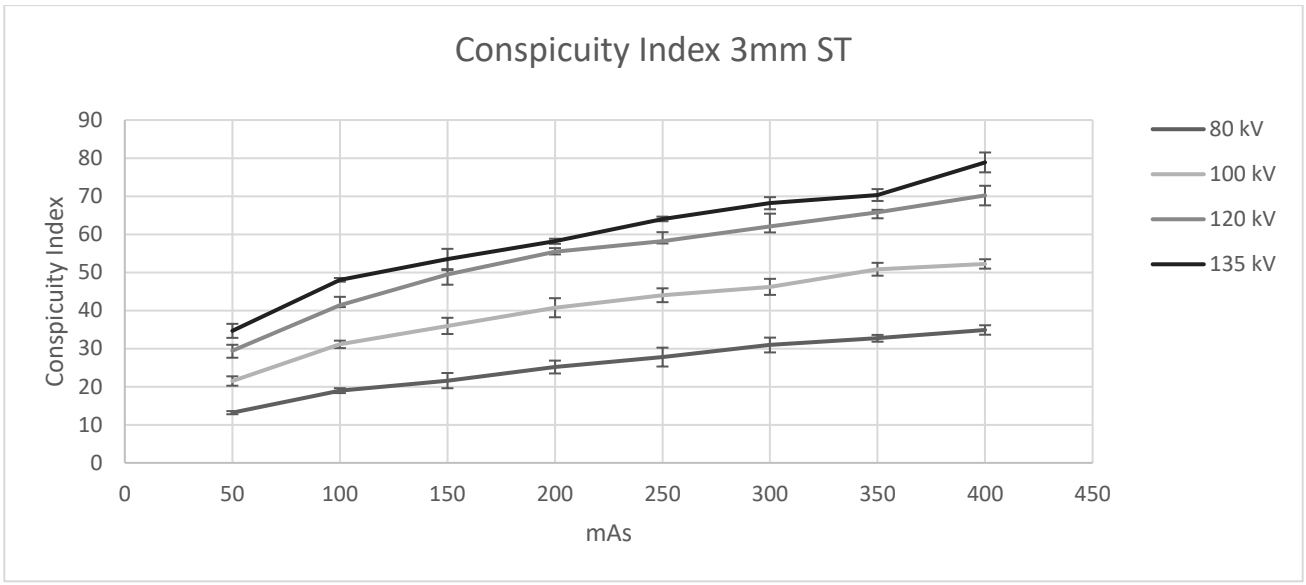
Software reliability



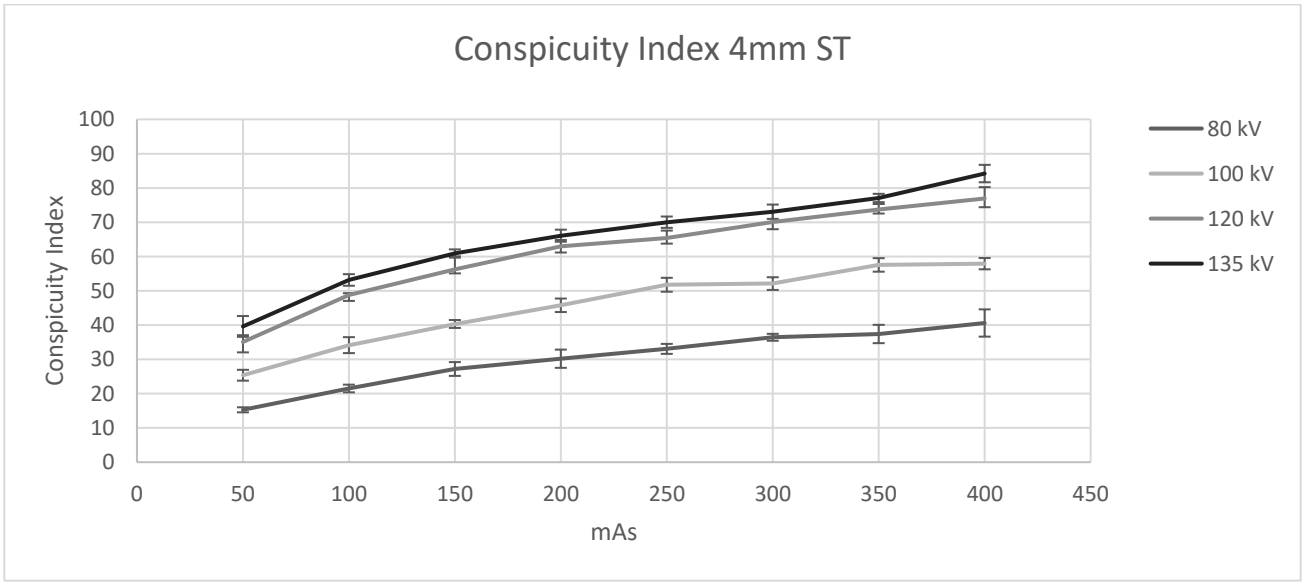
Graph 1 Conspicuity Index - 1mm Slice Thickness



Graph 2 Conspicuity Index - 2mm Slice Thickness

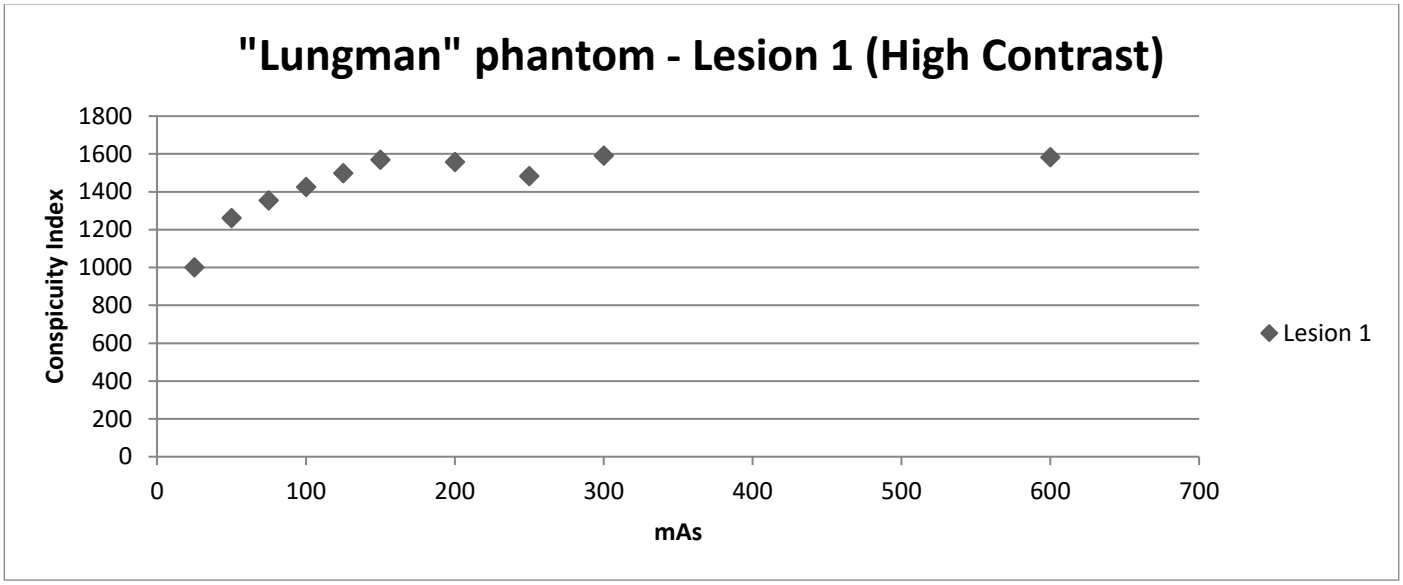


Graph 3 Conspicuity Index - 3mm Slice Thickness

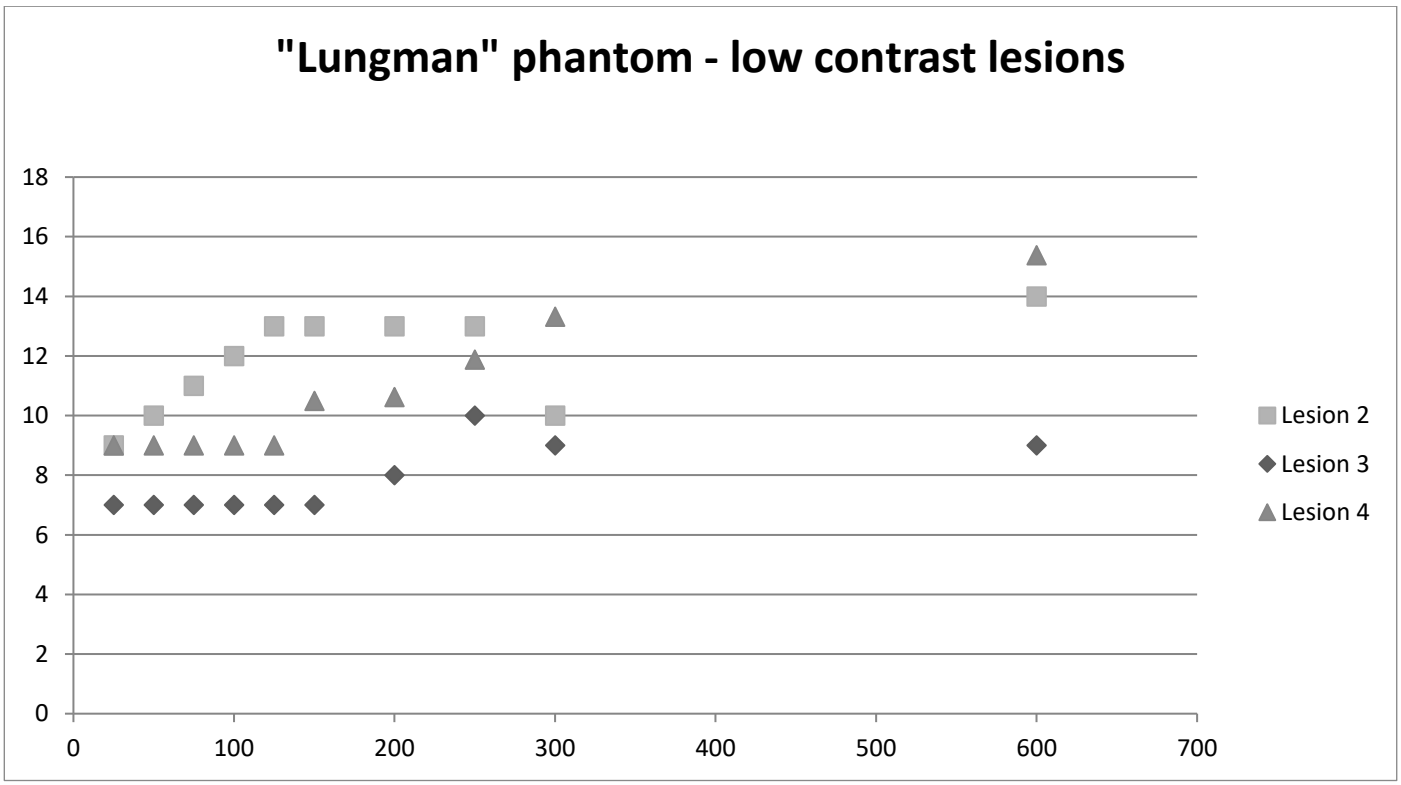


Graph 4 Conspicuity Index - 4mm Slice Thickness

Phantom Data



Graph 5 - "Lungman" phantom lesion 1



Graph 6 - "Lungman" conspicuity index- low contrast lesions

Discussion

All reliability results (Graph 1, Graph 2, Graph 3, Graph 4) were linear with mAs and behaved as expected with varying kVp and slice thickness. This demonstrates that the software is reliable in terms of representing the conspicuity within a DICOM image. Also, the ROIs were drawn 4 times, and as can be seen from these graphs, the standard deviations are very low, demonstrating reproducibility.

The more clinically representative phantom data using the “Lungman” phantom also demonstrates the reliability of the software, as even in a more complex phantom the results were still as expected with the changing parameters. Graph 5 demonstrates that the conspicuity index increases with mAs up to a point where it plateaus. This demonstrates the lack of improvement in conspicuity with increased mAs, and demonstrates the capability of the software to provide optimisation opportunities for clinical imaging. Graph 6 demonstrates that with low contrast lesions, the conspicuity is low, as expected, the lesions with a greater difference to background (lesion 3) shows the highest conspicuity index, with lesion 4 being the lowest value, as this is the smallest lesion (5mm)

The software has been used on real clinical images, and been shown to function with DICOM images from various modalities and manufacturers, however it has not yet been used to assess the diagnostic performance of real clinical images. Due to the radiation risk, the reliability of the software cannot be established using real clinical images as variables would have to be repeatedly changed on the same participant to demonstrate the behaviour of the conspicuity index, and this would require ethical approval.

Now the software and calculations have been found to be reliable, future work with this software is to compare to perceptional analysis, such as JAFROC, 2AFC and eye tracking.

Although correlations of these distinct datasets may be possible in simple controlled images, such as physics phantoms, there will be the added complications of decision errors, and so this software has the potential to distinguish between perceptual and cognitive errors.

Currently this research has focused on focal lesions, however it is recognised that there are many other lesion types, such as linear or diffuse, future developments of the software may include analysis of these types of lesions.

Conclusion

A novel software has been developed that enables the calculation of the conspicuity index of focal lesions in DICOM images.

The software and associated Excel spreadsheet have been tested for reliability and has been shown to provide reliable reproducible results.

The analysis could be further developed to incorporate JAFROC and eye-tracking data allowing correlations between physical and perception measures to be made beyond basic CNR calculations, it could also be used as a tool to distinguish between perceptual and cognitive error.

Further refinements could lead to measures of the detectability of more diffuse disease features.

References

[1] E. Samei, "Performance of Digital Radiographic Detectors: Quantification and Assessment Methods." 33-37.
[2] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt *et al.*, [The essential physics of medical imaging] Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia, Pa. ; London(2012).
[3] G. Revesz, H. L. Kundel, and M. A. Graber, "The influence of structured noise on the detection of radiologic abnormalities," Invest Radiol, 9(6), 479-86 (1974).
[4] E. E. Coche, B. Ghaye, J. Mey *et al.*, [Comparative Interpretation of CT and Standard Radiography of the Chest [electronic resource]] Springer Berlin Heidelberg, Berlin, Heidelberg(2011).
[5] E. Samei, M. J. Flynn, E. Peterson *et al.*, "Subtle lung nodules: influence of local anatomic variations on detection," Radiology, 228(1), 76-84 (2003).
[6] F. Li, S. Sone, H. Abe *et al.*, "Lung cancers missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings," Radiology, 225(3), 673-83 (2002).
[7] U. Plöckinger, "Diagnosis and Treatment of Gastrinomas in Multiple Endocrine Neoplasia Type 1 (MEN-1)," Cancers (Basel), 4(1), 39-54 (2012).
[8] R. L. Birdwell, D. M. Ikeda, K. F. O'Shaughnessy *et al.*, "Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection," Radiology, 219(1), 192-202 (2001).
[9] J. Michaelson, S. Satija, R. Moore *et al.*, "Estimates of the Sizes at Which Breast Cancers Become Detectable on Mammographic and Clinical Grounds," Journal of Women's Imaging, 5(1), 3-10 (2003).
[10] S. C. Bushong, [Radiologic science for technologists : physics, biology, and protection] Elsevier Mosby, St. Louis, Mo. ; [London](2008).
[11] A. Burgess, [Spatial vision research without noise] Cambridge University Press, Cambridge, 3 (2010).
[12] V. F. van Ravesteijn, T. N. Boellaard, M. P. van der Paardt *et al.*, "Electronic Cleansing for 24-H Limited Bowel Preparation CT Colonography Using Principal Curvature Flow," IEEE Trans Biomed Eng, 60(11), 3036-45 (2013).
[13] K. W. Yeom, S. J. Holdsworth, A. T. Van *et al.*, "Comparison of readout-segmented echo-planar imaging (EPI) and single-shot EPI in clinical application of diffusion-weighted imaging of the pediatric brain," AJR Am J Roentgenol, 200(5), W437-43 (2013).
[14] L. I. Peltonen, A. A. Aarnisalo, M. K. Kortesniemi *et al.*, "Limited cone-beam computed tomography imaging of the middle ear: a comparison with multislice helical computed tomography," Acta Radiol, 48(2), 207-12 (2007).
[15] M. P. Botelho, R. Agrawal, F. D. Gonzalez-Guindalini *et al.*, "Effect of radiation dose and iterative reconstruction on lung lesion conspicuity at MDCT: does one size fit all?," Eur J Radiol, 82(11), e726-33 (2013).
[16] R. T. Gupta, C. M. Iseman, J. R. Leyendecker *et al.*, "Diagnosis of focal nodular hyperplasia with MRI: multicenter retrospective study comparing gadobenate dimeglumine to gadoxetate disodium," AJR Am J Roentgenol, 199(1), 35-43 (2012).
[17] M. L. Zuley, K. M. Willison, E. Bonaccio *et al.*, "Full-field digital mammography on LCD versus CRT monitors," AJR Am J Roentgenol, 187(6), 1492-8 (2006).
[18] D. Marin, K. R. Choudhury, R. T. Gupta *et al.*, "Clinical impact of an adaptive statistical iterative reconstruction algorithm for detection of hypervascular liver tumours using a low tube voltage, high tube current MDCT technique," Eur Radiol, 23(12), 3325-35 (2013).
[19] I. B. Weiner, and W. E. Craighead, [The Corsini encyclopedia of psychology] John Wiley, Hoboken, N.J.(2010).
[20] O. R. Brook, S. Gourtsoyianni, A. Brook *et al.*, "Split-Bolus Spectral Multidetector CT of the Pancreas: Assessment of Radiation Dose and Tumor Conspicuity," Radiology, (2013).
[21] W. Chang, J. M. Lee, K. Lee *et al.*, "Assessment of a model-based, iterative reconstruction algorithm (MBIR) regarding image quality and dose reduction in liver computed tomography," Invest Radiol, 48(8), 598-606 (2013).
[22] D. J. Manning, S. C. Ethell, and T. Donovan, "Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph," Br J Radiol, 77(915), 231-5 (2004).
[23] H. L. Kundel, C. F. Nodine, and D. Carmody, "Visual scanning, pattern recognition and decision-making in pulmonary nodule detection," Invest Radiol, 13(3), 175-81 (1978).
[24] K. S. Berbaum, E. A. Franken, D. D. Dorfman *et al.*, "Role of faulty decision making in the satisfaction of search effect in chest radiography," Acad Radiol, 7(12), 1098-106 (2000).
[25] K. S. Berbaum, E. A. Brandser, E. A. Franken *et al.*, "Gaze dwell times on acute trauma injuries missed because of satisfaction of search," Acad Radiol, 8(4), 304-14 (2001).
[26] K. S. Berbaum, D. D. Dorfman, E. A. Franken *et al.*, "Proper ROC analysis and joint ROC analysis of the satisfaction of search effect in chest radiology," Acad Radiol, 7(11), 945-58 (2000).
[27] G. B, L. Q, P. L *et al.*, "Objectively measuring signal detectability, contrast, blur and noise in medical images using channelized joint observers." 8673.
[28] J. D. Thompson, D. J. Manning, and P. Hogg, "The value of observer performance studies in dose optimization: a focus on free-response receiver operating characteristic methods," J Nucl Med Technol, 41(2), 57-64 (2013).
[29] A. E. Burgess, "Comparison of receiver operating characteristic and forced choice observer performance measurement methods," Med Phys, 22(5), 643-55 (1995).
[30] S. Richard, and J. H. Siewerdsen, "Comparison of model and human observer performance for detection and discrimination tasks using dual-energy x-ray images," Med Phys, 35(11), 5043-53 (2008).
[31] C. Mello-Thoms, L. Hardesty, J. Sumkin *et al.*, "Effects of lesion conspicuity on visual search in mammogram reading," Acad Radiol, 12(7), 830-40 (2005).
[32] C. Mello-Thoms, "The problem of image interpretation in mammography: effects of lesion conspicuity on the visual search strategy of radiologists," Br J Radiol, 79 Spec No 2, S111-6 (2006).

[33] D. S. Brettle, E. Berry, and M. A. Smith, “The effect of experience on detectability in local area anatomical noise,” Br J Radiol, 80(951), 186-93 (2007).
[34] M. S. Chesters, “Human visual perception and ROC methodology in medical imaging,” Phys Med Biol, 37(7), 1433-76 (1992).
[35] Kyotokagaku, [Multipurpose Chest Phantom N1 “Lungman”].

Paper 10

Effects of Increased Compression with an Ultrasound Transducer on the Conspicuity of Breast Lesions in a Phantom

Katy Szczepura, Tahreem Faqir, David Manning
University of Salford, Manchester, UK

Proc. SPIE 10136, Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment, 101360T (24 April 2017); doi: 10.1117/12.2254263

Abstract

Ultrasound imaging of the breast is highly operator dependant. The amount of pressure applied with the transducer has a direct impact on the lesion visibility in breast ultrasound.

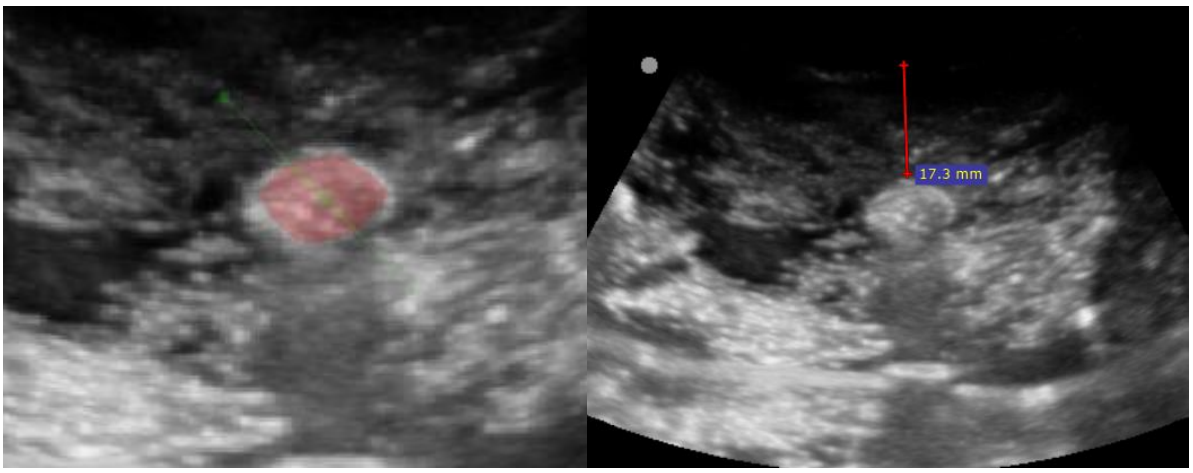
The conspicuity index is a quantitative measure of lesion visibility, taking into account more parameters than standard measures that impact on lesion detection [1].

This study assessed the conspicuity of lesions within a breast phantom using increased transducer compression in breast ultrasound.

Methods

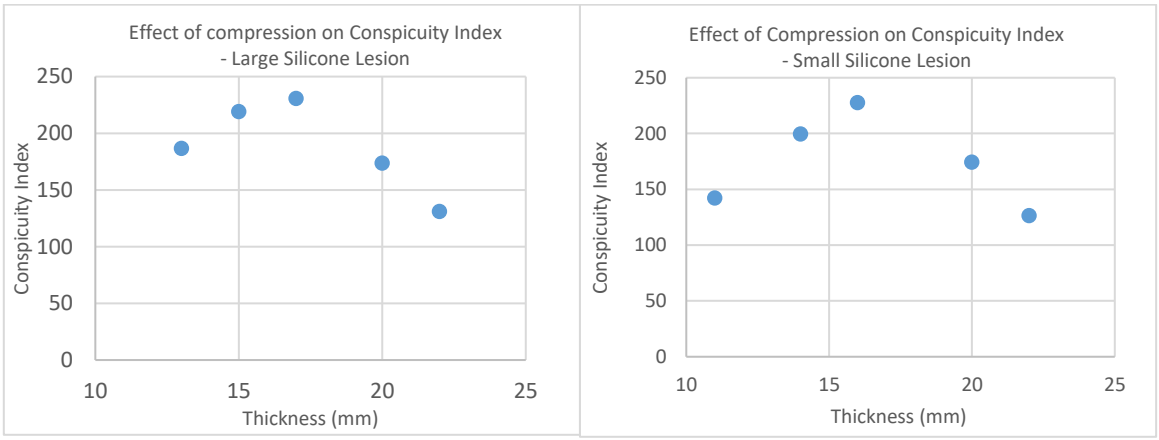
A phantom was constructed of gelatine to represent adipose tissue, steel wool for glandular/blood vessels and silicone spheres to represent lesions, this meant that the lesions were also compressible, but less than the surrounding tissue.

The phantom was imaged under increasing transducer compression.



The conspicuity index was measured using the Conspicuity Index Software. The distance between the transducer surface and lesion surface was measured as an indication of increased compression.

Results



When moderate compression (17mm) was applied, the conspicuity index increased resulting in better visualisation of the silicone lesions. However, with increased compression the conspicuity index decreased.

New work to be presented

The conspicuity index has never been demonstrated in ultrasound imaging before. This is preliminary phantom work to demonstrate the impact of increased transducer compression on quantitative lesion visibility assessment.

Conclusion

The compression applied should be moderate for optimum visualisation, as excessive pressure decreases conspicuity. However, further work needs to be conducted in order to consider other factors, such as density of the breast and lesion location, for a better understanding of the effect of compression on the visualisation of the lesion. A human study is planned.

Introduction

Ultrasound scanning is often used in conjunction with mammograms and are performed when additional information is required. Ultrasound is used to explore factors, such as the source of breast symptoms, check for breast lumps and investigate abnormal results from a mammogram. In addition, it is frequently a choice of imaging for young women with dense breast tissue, as mammograms do not demonstrate sufficient detail [2] Although, ultrasound is used in addition to a mammogram, it can be the primary choice for detecting breast lesions. This often occurs when a mammogram is not an option, due to the ionising radiation, or simply is not available [2]. Ultrasound has many characteristics, which make it an extremely useful modality. These characteristics include: simplicity (it does not require any special preparation), immediacy of the image, rapidity of examination, availability, ability to study pathologic and normal masses in real time, cost and that it does not carry any known side effects. Ultrasound is usually not uncomfortable, however discomfort can be experienced when the transducer is compressed against the area of examination [3].

Ultrasound is able to detect palpable, as well as non-palpable lesions and early carcinomas. In addition, it is often used for the staging of the lesions. However, the accuracy of the ultrasound examinations is highly dependent on the examiner's experience and the equipment [4]. Ultrasound is commonly stated as being "operator dependent" in that the examination is dependent on the dedication and experience of the operator. A suboptimal ultrasound scan can be interpreted in many different ways, which makes ultrasound a less objective technique in comparison to mammograms [3].

Reflecting structures at severe angles to the beam of the ultrasound can cause excessive shadow artefacts and reduce image quality due to the reduction of penetration [5]. During the scanning of a breast lesion, angulation and compression of the transducer is a common technique to sharpen the edges of the lesion [6]. The degree of compression applied can vary and is often dependent on the location of the lesion, the pendulousness and size of the breast.

The compression applied in combination with the patient's positioning are key elements in the ultrasound of the breast and play a vital role in discovering and the visualisation of a lesion. The patient is scanned in a contralateral posterior oblique or a supine position and they are asked to place their arms behind their head [2], this thins the tissue being scanned and forces the normal breast tissue into a parallel plane in relation to the transducer and perpendicular to the beam. This improves the conspicuity of lesions and the overall image quality by allowing greater beam penetration, as compression can decrease refraction and scatter from neighbouring structures [7].

Increased pressure can be applied to lesions that are small, have fibrous tissue lying superficially and are situated deeply near the chest wall. Occasionally, structures that lie superficial to the lesion, e.g. Cooper's ligament, can create acoustic shadows and hinder the evaluation of deeper structures. With the use of moderate compression this can be eliminated [7].

However, extensive compression with the use of a transducer may not always improve image quality. Vigorous compression can result in the attenuation or even the elimination of vascular signals when using Doppler, predominantly in benign lesions [3]. Furthermore, scanning in a perpendicular plane to the near-field tissue in combination with heavy compression, has the ability to push side-lobe artefacts and near-field reverberation echoes deeper into the breast. This can interfere with the evaluation of superficial lesions and consequently have a detrimental effect on the diagnosis and patient's pathway. In such cases, the visualisation can be improved by not applying any compression or with the use of angulation. This allows the beam to penetrate obliquely through the lesion, rather than perpendicularly [7].

This is supported by Carson et al, who proposes that deeper structures have reduced acoustic path lengths when considerable local compression is applied [5]. It is demonstrated by Fargier-Voiron *et al,* that excessive pressure, applied by the transducer, has the ability to impact visualisation and localisation of lesions. They further recommend, to achieve better accuracy, the transducer pressure should be kept as low as possible. This, in addition, will allow sufficient contact for optimum image quality [8].

Furthermore, it has been proposed by Dobler *et al,* a probe displacement of 1-2 cm was essential in order to obtain good image quality. The study established that excessive compression can result in a shift of anatomical structures [9] consequently, making it difficult to localise lesions. This study, in combination with multiple other studies, demonstrated a maximum displacement of approximately 1.7cm when excessive pressure was applied [10-12] Although, these studies were conducted on the displacement of prostates, the same principle can be applied to the ultrasound scanning of breasts.

Methodology

Phantom Design

In order to investigate new imaging techniques, it is vital that the phantom used has the properties to mimic human tissue. Although, commercial ones are an easy option, they are not very easily adaptable to specific applications. In-house, custom made options allow an individual to tailor their phantom to their needs. There are several studies that promote the use of gelatine to create a

phantom [13]. According to Farrer et al, gelatine has properties in the same order of magnitude as that of soft tissue and is easy to manufacture [14]. Therefore, in order to simulate breasts, a gelatine phantom was made. A study conducted by Sultan et al suggested using 70g (6 sachets) of Dr Oetker's gelatine and 500ml of water, which resulted in the gelatine-to-water ratio of approximately 1:7 [15]. However, when this ratio was used, the density was too firm and did not resemble human tissue. After some trial-and-error, the gelatine-to-water ratio used was 1:31, using 18g (1.5 sachets) of Dr Oetker's gelatine and 570ml of water.

The 570ml of hot water was poured into a jug and the 18g of gelatine was sprinkled into the water very slowly, whilst the mixture was being stirred, allowing the gelatine to dissolve completely.

The lesions were made from, silicone spheres, two lesions, measuring approximately 1cm and 1.5cm, were made. In order to simulate glandular tissue, fine steel wool was used [16]. The steel wool was placed in a small plastic tub, which measured 15cm x 10cm x 7cm. The lesions were placed within the steel wool towards the opposite corners of the tub and the mixture of water and gelatine was poured into the plastic tub. It was essential to ensure the mixture covered the entire steel wool to prevent the formation of air bubbles (Figure 5).

Figure 5 phantom construction

The tub was then refrigerated until 2 hours before it was required for scanning. The total depth of the phantom was 4.5cm.

Image Acquisition

The images were acquired using a SonoSite M-Turbo Ultrasound scanner with a C60X transducer. The transducer was used to apply pressure to the phantom with the aid of "Aquasonic 100 Ultrasound Gel". Each of the lesions was scanned with 5 different compressions. The pressure applied varied from minimal compression (compression 1), just enough to obtain sufficient image

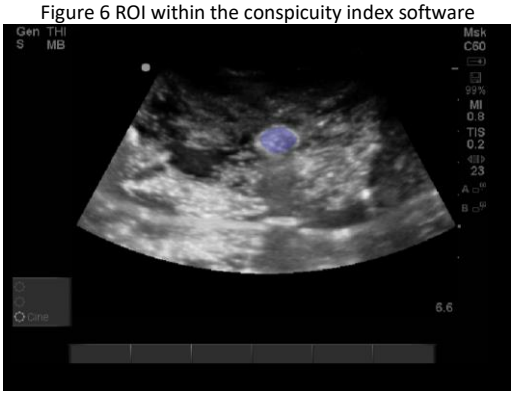
quality, to extreme compression (compression 5). The aim throughout was to maintain equal changes in pressure. During each compression, the screen was frozen and the image was exported as DICOM images. Variables, such as windowing or the field-of-view were kept constant, this was to ensure that the change of visualisation was only caused by the change in compression.

Conspicuity index analysis

Conspicuity software (ref) was used to draw a region of interest (ROI) around the abnormality (

Figure 6) avoiding artefacts.

Figure 6 ROI within the conspicuity index software



The programme then plotted 180 line profiles around the region drawn, which resulted in 360 edge profiles.

A fit was then applied for every line profile (Figure 7). The blue line demonstrated the line profile, the red line represented the plotted fit and the green line demonstrated the ROI drawn by the operator. This programme then acquired the pixel values of the lesion and the background. It calculated the standard deviation for these values, which represented the lesion and anatomical noise.

Figure 7 Line profile

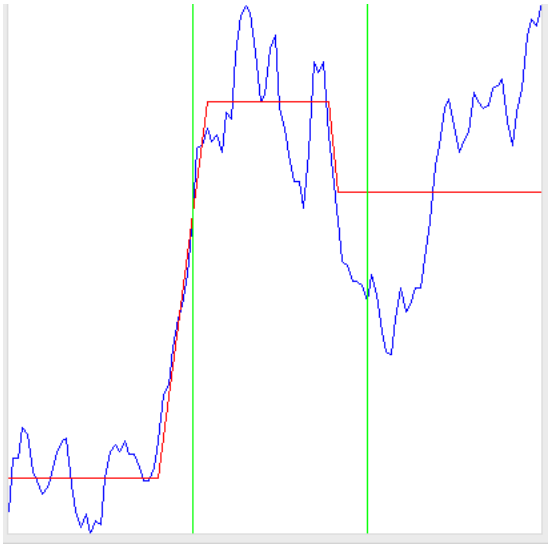
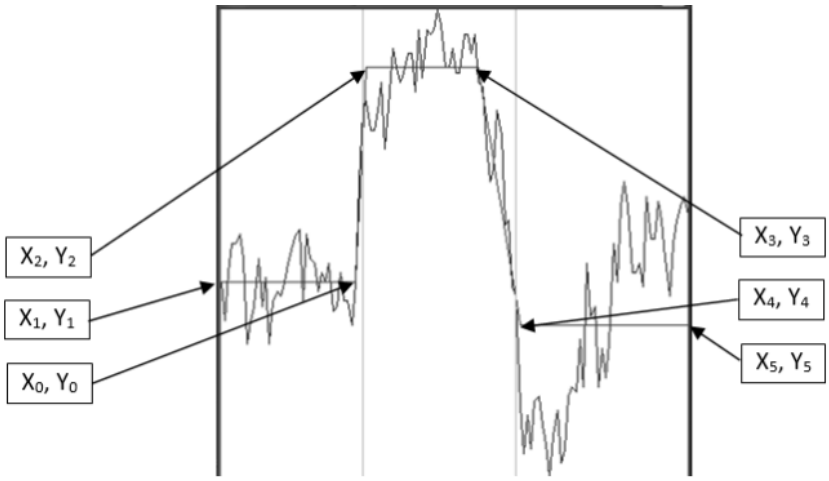


Figure 8 Example single line profile across the region of interest for a lesion



An Excel spreadsheet has been developed to import the TSV files and then to calculate the following:

$$\bar{d} = \overline{(X3 - X2)}$$ = average lesion dimension

Equation 10

$$\bar{\theta} = \overline{\tan^{-1}\left[\frac{Y_2 - Y_1}{X_2 - X_1}\right], \tan^{-1}\left[\frac{Y_4 - Y_3}{X_4 - X_3}\right]}$$ = average slope angle of all the line profiles

Equation 11

$$\overline{\Delta GL} = \overline{Y_2 - \frac{(Y_1 + Y_4)}{2}}$$

$=$ the average difference in grey levels between the lesion and the background

Equation 12

$$\overline{\sigma_s} = sdTop$$

Equation 13

$$\overline{\sigma_n} = \frac{sd1 + sd2}{2}$$

Equation 14

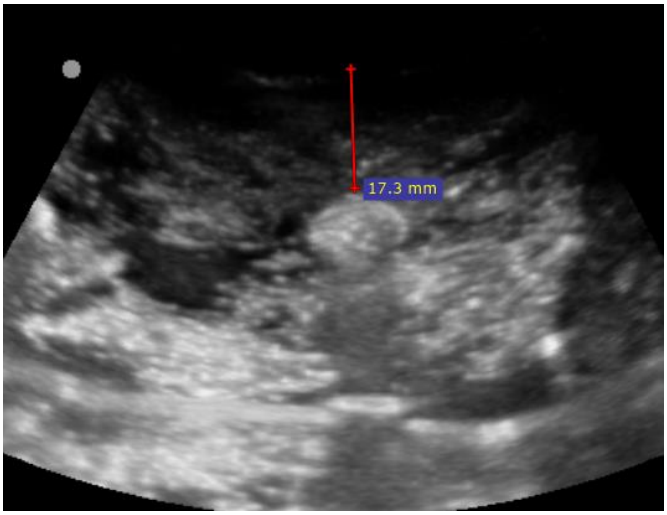
This data was then used to calculate the Conspicuity Index (χ):

$$\chi = \frac{\bar{d}\tan[\bar{\theta} - 1]\overline{\Delta GL}}{\sqrt{\overline{\sigma_s^2} + \overline{\sigma_n^2}}}$$

The ROI was drawn around each lesion three times, to work out the standard deviation and measure reproducibility. From the three conspicuity index numbers, an average was calculated.

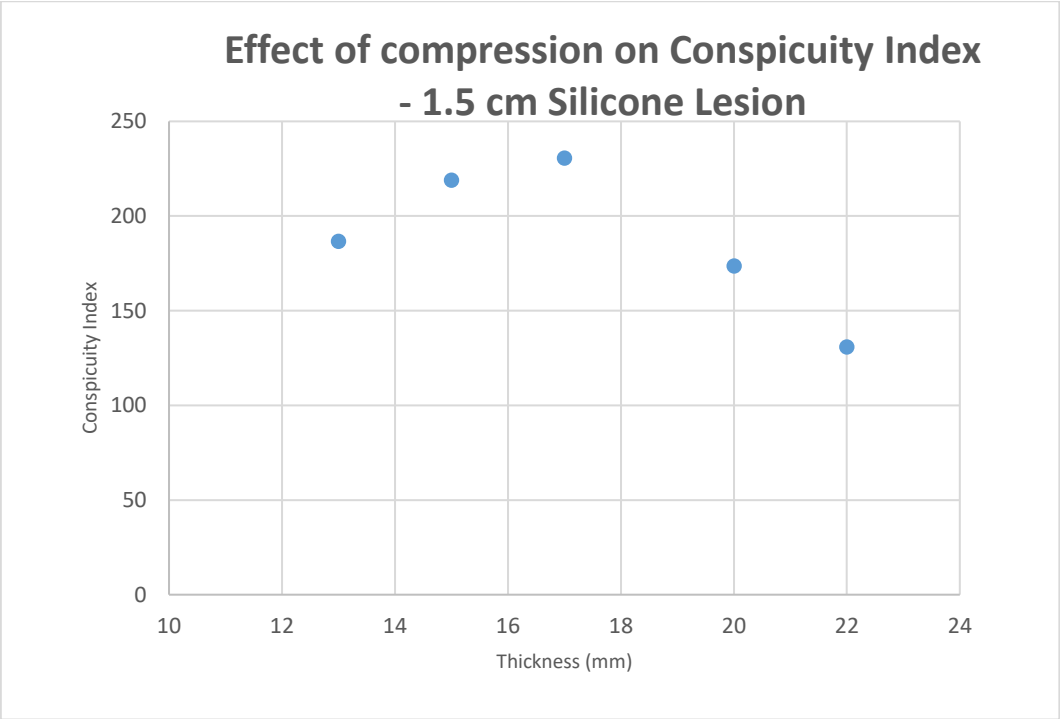
After this data was collected, the original ultrasound images were used to calculate the pixels and mm of the top of the ultrasound image to the top of the lesion (Figure 9). This was used to work out the change in thickness above the lesion in mm to indicate the level of compression.

Figure 9 measuring thickness from transducer to surface of lesion

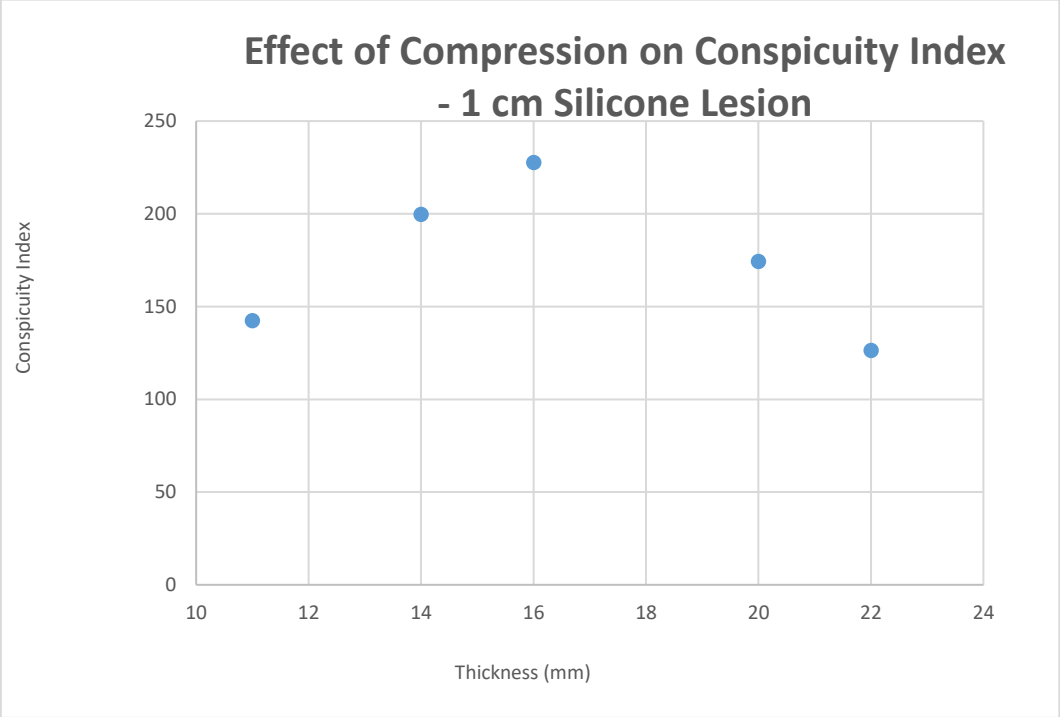


Results

Graph 7 Effect of compression on C.I. for 1.5 cm lesion



Graph 8 Effect of compression on C.I. for 1 cm lesion



Discussion

Overall, the results show when compression is applied, the conspicuity index of the lesion can improve. The compression applied should be between minimal and moderate for optimum visualisation, as excessive pressure decreases the conspicuity index. Graph 7Graph 8 establishes that the optimum compression for the silicone lesions was at a thickness reduction of 17mm, when applying moderate compression, regardless of size. The conspicuity index was at its highest at approximately 230 for both the lesions.

It is important to recognise that in addition to compression, the image quality and the visualisation of a lesion is impacted by multiple other factors in ultrasound imaging. These elements include the patient's breast size and density, the equipment used, and the amount and distribution of gel [17].

Conclusion

The compression applied should vary from minimal to moderate, as excessive pressure can result in loss of conspicuity and hinder detectability. Consequently, this can result in false negatives and impact the patient's diagnosis and pathway. However, further research is required to take other factors that can affect the visualisation into consideration.

A patient study is planned to see if this preliminary work can be replicated in humans.

Refernces

[1] K. R. Szczepura, and D. J. Manning, "Validated novel software to measure the conspicuity index of lesions in DICOM images," 9787, 978703-978703-15 (2016).

[2] A.-M. M. Dixon, [Breast ultrasound : how, why and when] Churchill Livingstone Elsevier, Edinburgh(2008).

[3] O. Catalano, A. Nunziata, and A. Siani, [Fundamentals in Oncologic Ultrasound: Sonographic Imaging and Intervention in the Cancer Patient] Springer Science & Business Media, (2009).

[4] H. Madjar, H. A. Ladner, W. Sauerbrei *et al.*, "Preoperative staging of breast cancer by palpation, mammography and high-resolution ultrasound," Ultrasound in Obstetrics & Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology, 3(3), 185-190 (1993).

[5] P. L. Carson, B. Wang, G. L. LeCarpentier *et al.*, "Local compression in automated breast ultrasound in the mammographic geometry." 1787-1790.

[6] N. Yang, "Breast ultrasound | Radiology Reference Article | Radiopaedia.org."

[7] A. T. Stavros, [Breast Ultrasound] Lippincott Williams & Wilkins, (2004).

[8] M. Fargier-Voiron, B. Presles, P. Pommier *et al.*, "Impact of probe pressure variability on prostate localization for ultrasound-based image-guided radiotherapy," Radiother Oncol, 111(1), 132-7 (2014).

[9] B. Dobler, S. Mai, C. Ross *et al.*, "Evaluation of possible prostate displacement induced by pressure applied during transabdominal ultrasound image acquisition," Strahlentherapie Und Onkologie: Organ Der Deutschen Rontgengesellschaft ... [et Al], 182(4), 240-246 (2006).

[10]

[11] J. P. McGahan, J. Ryu, and M. Fogata, "Ultrasound probe pressure as a source of error in prostate localization for external beam radiotherapy," International Journal of Radiation Oncology, Biology, Physics, 60(3), 788-793 (2004).

[12] C. F. Serago, S. J. Chungbin, S. J. Buskirk *et al.*, "Initial experience with ultrasound localization for positioning prostate cancer patients for external beam radiotherapy," International Journal of Radiation Oncology, Biology, Physics, 53(5), 1130-1138 (2002).

[13] J. Sutcliffe, R. L. Hardman, N. C. Dornbluth *et al.*, "A novel technique for teaching challenging ultrasound-guided breast procedures to radiology residents," Journal of Ultrasound in Medicine: Official Journal of the American Institute of Ultrasound in Medicine, 32(10), 1845-1854 (2013).

[14] A. I. Farrer, H. Odéen, J. de Bever *et al.*, "Characterization and evaluation of tissue-mimicking gelatin phantoms for use with MRgFUS," Journal of Therapeutic Ultrasound, 3, (2015).

[15] S. F. Sultan, G. Iohom, and G. Shorten, "A Novel Phantom for Teaching and Learning Ultrasound-guided Needle Manipulation," Journal of Medical Ultrasound, 21(3), 152-155 (2013).

[16] Y. Li, "Mammographic Density Assessment: Inter-Reader Variability And Novel Phantom Quantification," (2015).

[17] P. Chatelain, A. Krupa, and N. Navab, "Optimization of ultrasound image quality via visual servoing." 5997-6002.

Critical Review

1. Objective 1 – Error in breast thickness measurements

The compressed breast thickness measurement in FFDM is used in multiple ways. Primarily it is used to optimise the image via the AECs, additionally it is used in the calculation of mean glandular dose (MGD), and more recently is used to calculate the volumetric breast density (VBD), the latter of which can be used for cancer risk predictions (as discussed in Section 0). Accurate and reproducible measures of breast thickness are essential in mammographic imaging.

Errors in thickness measurement can arise as the paddle deforms under the compression force, and also some paddles are designed to tilt in an attempt to reduce patient discomfort. The thickness given by the machine does not take bend and distortion into account and so the given thickness may be inaccurate which has implications for image optimisation, dose calculation and risk stratification.

Previous work has demonstrated that there was a difference between the thickness reported by the machine and the physical thickness of the compressed breast. Tyson et al (Tyson, Mawdsley et al., 2009) used an optical stereoscopic photogrammetry device to measure the compressed breast thickness. They found a maximum difference between the breast thickness and the machine given reading was up to 15 mm. They followed this work by attempting to calibrate the machines based on the differences found using a variety of phantoms (Mawdsley, Tyson et al., 2009). An issue with the work was the use of the optical system. Although it had a high precision, with a reported 1 mm precision and measurement accuracy of less than 0.2 mm, the system was not a practical resource due to the equipment set-up. It needed highly specialised equipment, was very susceptible to room lighting and dependant on the image quality.

Paper 1 The readout thickness versus the measured thickness for a range of screen film mammography (SFM) and full field digital mammography (FFDM) units

The difference in thickness measurements was important to establish between machines, within the same machine make/model and across centres, if this error could be demonstrated in machines used within the NHSBSP this could have implications for both image quality and dose across the programme. It was therefore evident that there was a need to establish an easily adaptable system that could easily be used in various clinical environments to demonstrate the difference between measured and read-out thickness (Hauge, Hogg et al., 2012). In Paper 1 the first stage of the research was to develop the measurement technique, then to establish the extent of thickness measurement variability.

To develop the measurement technique a thickness measuring device (TMD) was created (Figure 4). Additionally, it was recognised that standard solid phantoms would potentially not demonstrate the issue of deformity as their structure, lack of variability and deformability does not accurately represent the human breast, therefore additionally a realistic breast phantom attached to a rigid torso was developed (Figure 5)

Figure 4 Thickness Measurement Device (Hauge, Hogg et al., 2012)

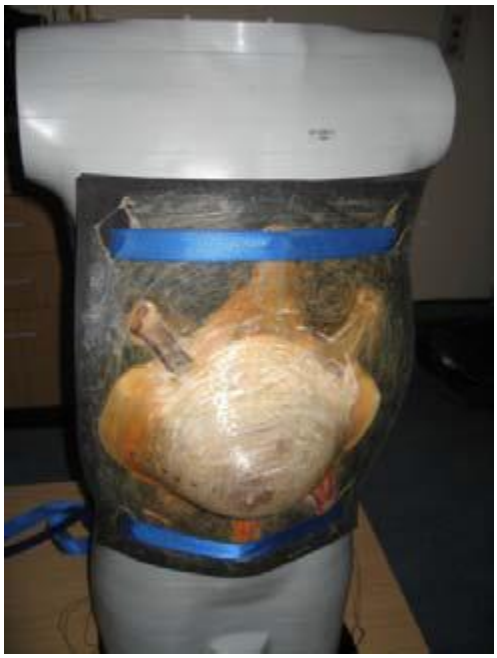


Figure 5 Breast phantom (Hauge, Hogg et al., 2012)

Preliminary work was undertaken to establish that the phantom represented human breast tissue; this was undertaken by comparing normalised thickness vs compression force curves for 3 prostheses to 29 human volunteers, the chosen prosthesis had a Pearson's correlation coefficient of 0.95 showing a very high positive correlation to the real breast data. Additionally, to ensure validity

of the work, the measurement device was repeatedly tested using 2 researchers to ensure reproducibility, reliability and consistency between users.

Once the validity, reliability and reproducibility of the equipment was established, the main part of the study was undertaken. Six different makes/model of machines were assessed (both FFDM and SFM), with 2 models repeated, giving a total of 8 machines. Two different paddles, and two compression forces were used on each machine (60 N and 100 N).

All machines had passed the required QC tests to ensure readout thickness accuracy (Perry, Broeders et al., 2008), the manufacturer stated accepted difference between measured thickness and readout thickness varied between machines as reported in Paper 1, and is between ±5 mm and ±10 mm, it is important to recognise that these QC tests are performed on solid phantoms. However, when using the deformable phantom, the largest difference was found to be 19mm, outside the manufacturers accepted range, indicating that QC using solid phantoms cannot reveal the extent of this issue.

The work found that the largest difference was with flexible, tilting paddles, and this difference increased with increasing compression force. It was also found that there were differences between the same make/model, indicating that the same correction factors cannot be assumed for a single make and model as suggested by (Mawdsley, Tyson et al., 2009).

1.1. Contributions of the work to the literature

Paper 1 - The readout thickness versus the measured thickness for a range of screen film mammography (SFM) and full field digital mammography (FFDM) units

1.1.1. Flexible paddles vs fixed paddles

Paper 1 was cited by Broeders et al (Broeders, Ten Voorde et al., 2015) when looking at implications for the increased deviation from readout thickness in flexible vs fixed paddles. They found that flexible paddles offered no pain reduction as previously assumed, and, although they did find a larger breast area and lower radiation doe when using flexible paddles, the flexible paddles removed fibroglandular tissue from the image area and reduced contrast in the clinically relevant retroglandular area at chest wall side, therefore they recommended that fixed paddles should be used for both CC and MLO views.

1.1.2. Novel Breast Phantom

The phantom developed within Paper 1 has subsequently been used in several studies where a deformable breast phantom is essential rather than a solid phantom. As shown in this paper, the

variation between the readout thickness and measured thickness were not demonstrated during the quality control testing of the machine when using solid phantoms. This indicates that solid phantoms cannot be used to demonstrate any issue with distortion or bend of the paddle, therefore for any studies of this kind a deformable phantom is essential.

These studies that used this novel phantom include (Hogg, Szczepura et al., 2013, Ma, Brettle et al., 2014, Ma, Hogg et al., 2015, Smith, Szczepura et al., 2015) and these will be discussed in more detail later in Objective 3 - Optimising image receptor position for pressure balance and breast footprint

1.1.3. Breast density calculation

Paper 1 has been cited in the following journal articles discussing breast density calculation accuracy; (Alonzo-Proulx, Jong et al., 2012, Alonzo-Proulx, Mawdsley et al., 2015, Geeraert, Klausz et al., 2014, Wang, Kato et al., 2017). As discussed earlier the measure of the relative amount of fibro glandular tissue within the breast has been established as a strong independent risk factor for breast cancer (Chen, Gulsen et al., 2015, McCormack and dos Santos Silva, 2006, Ng and Lau, 2015, Sherratt, McConnell et al., 2016).

The contribution of Paper 1 in terms of demonstrating thickness variability within and between machines was vital for these studies as software developed for the automatic calculation of breast density such as Quantra™ (Hartman, Highnam et al., 2008) and Volpara™(Highnam, Brady et al., 2010) uses machine given thickness within the algorithm to calculate the breast density, and therefore the accuracy, reproducibility and reliability of this value is essential.

Studies have repeatedly shown significant associations with breast cancer risk for both qualitative and quantitative breast density measures and a potential to improve cancer risk assessment models (Brentnall, Harkness et al., 2015, Tice, Cummings et al., 2008). In several US states it has become a legal requirement to inform the woman of their breast density as part of the screening process (Ray, Price et al., 2015) and substantial research continues to be devoted to accurate measurement of this key biomarker and to its incorporation into risk prediction models.

At the onset of the increase in evidence for the relationship between calculated volumetric breast density and cancer risk, Kopans et al (2008) discussed that there was a potential methodological flaw by not taking into account exposure values, half-value layer information, and knowledge of the compressed thickness of the breast. They discussed that this could lead to an inaccurate calculation of the breast volume (Kopans, 2008) software developments therefore now include these parameters into the algorithms ™ (Hartman, Highnam et al., 2008, Highnam, Brady et al., 2010)

As the thickness measurement is taken from the machine given value with corrections based on Mawdsley et al (2009), there is still the issue of thickness inaccuracy between the same make and model, which could lead to implications for the risk stratification of the patient. If the patient is considered high risk, they would be included into a pathway that would include additional screening attendances, leading to increased radiation risk and additional unnecessary costs. If, however, the patient was inappropriately classified as low risk, there is a risk of poor early cancer detection due to reduced screening.

It has been found that supplemental screening of women with dense breasts leads to the identification of more breast cancers (mostly invasive), but may also be associated with higher recall rates and additional biopsies (Melnikow, Fenton et al., 2016). Currently there have been no published studies of important longer-term clinical outcomes of supplemental screening.

Work by Waade et al (Waade, Highnam et al., 2016) demonstrated up to 38.5% error in breast density calculation when using Volpara when the actual thickness of the breast tissue varies from the recorded breast thickness. The recorded breast tissue was artificially increased in 1 mm increments for a ±15% change in thickness and the Volpara Density Grade (VDG) was measured. Only 11.5% of images were found to be within the QC accepted error indicating that thickness inaccuracies impacted upon the calculated VDG and this has implications for patient care pathways.

This conflicts with a large cohort study undertaken at the Mayo clinic (Olson, Sellers et al., 2012) where they found a strong association mammographic density and breast cancer risk that was not confounded by mammogram acquisition technique. However, the methodological approaches were very different. For the Mayo study breast density was assessed via the visual BI-RAD categories not automated software, and the images were film-screen, the large variability in breast size and density could have masked the impact on accuracy by not directly comparing the change in compression with a change in density as this study has done.

1.1.4. Mean Glandular Dose Calculation

The impact of thickness inaccuracy has been used to look into uncertainties in mean glandular dose calculations for women in the Norwegian Breast Screening programme (Hauge and Olerud, 2013). This study looked at the variables that impacted on MGD accuracy, and although they found that inaccuracies in air KERMA had the largest effect, an inaccuracy of breast thickness by ±10 mm leads to a ± 10 % uncertainty in the MGD calculation.

Paper 1 was cited in this work to demonstrate the variability in breast thickness accuracy, and therefore the reliability of reported MGD.

1.1.5. Limitations of the work

A limitation of this work was the limited compression force that could be applied; higher compression forces than 100 N are used clinically. However, the deformable nature of the phantom meant that to maintain phantom integrity higher forces than this could not be used. Additionally, this work was not performed on real human tissue. Although work was performed to establish the similarity between the phantom and human tissue, the use of a single phantom cannot represent the variability found in real tissue.

The measurement device was not as sensitive as Tyson et al's optical device (Tyson, Mawdsley et al., 2009), although the Thickness Measuring Device was more adaptable and it was easier to perform the measurements, it did not have the 1 mm accuracy that (Mawdsley, Tyson et al., 2009) suggest is required. More recent advancements in technology may improve the measurement sensitivity and ease of future work, depth sensor and time of flight cameras are freely available and highly sensitive and have great potential for future work (Díaz, Oliver et al., 2016, Tadano, Pediredla et al., 2015).

1.2. Papers 2 and 3

Paper 2 - Tissue bulge during stereotactic core biopsy* and *Paper 3 - Breast tissue bulge and lesion visibility during stereotactic biopsy–A phantom study

Inaccuracy of thickness measurements could have additional implications during FFDM guided biopsy. Some types of biopsy paddles have a hole within the paddle to allow needle access for the biopsy as shown in Figure 6 below. The hole is positioned above the lesion identified in the standard FFDM mammogram.

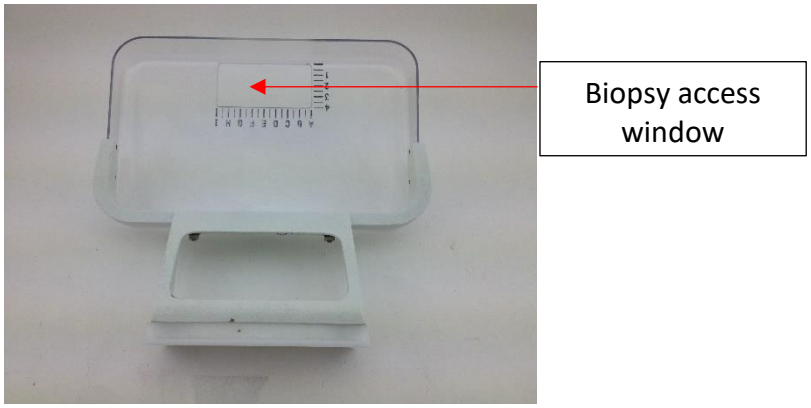


Figure 6 biopsy paddle

Discussions with clinical colleagues revealed that on occasion breast lesions seemed less visible on the biopsy images than the previous standard FFDM screening images. Observation of the clinical environment raised the question of whether the hole in the paddle could be a potential cause of the reduced visibility, as when compression was applied tissue would bulge through the hole. This led to the hypothesis that this extra tissue was not being taken into account in the measured thickness, meaning that the AECs could not function optimally.

An experiment was undertaken to test this hypothesis to both measure the extent of the tissue bulge, and the impact on lesion visibility and dose in Paper 2 and Paper 3. (Hackney, Williams et al., 2013, Williams, Hackney et al., 2014). The first stage of the work was to establish the extent of the bulge, so an audit was undertaken where a measurement device was created to measure the tissue bulge above the paddle (Hackney, Williams et al., 2013) (Paper 2).

Measurements were made in 15 consecutives clients. Results showed an average of 18.7% increase in tissue thickness, (range 11.3-30%) compared to the compressed breast thickness.

Following on from this a phantom was constructed based on traditional QC phantoms, using Perspex and aluminium discs (Carton, Bosmans et al., 2004, Pachoud, Lepori et al., 2004). Additional Perspex thicknesses were added based on the measurements made in Paper 2 (Hackney, Williams et al., 2013). Images were acquired using the AECs with combinations of breast/bulge thickness.

Relative visual grading was used with 14 observers, the machine given organ dose was also recorded. A novel “visibility score” was developed, as well as a novel “optimisation score” which based on the figure of merit methodology (Borg, Badr et al., 2012) but applying it to a visual grading scenario.

It was found that the visibility score decreased with increasing bulge thickness, and the optimisation index decreased. This showed that as the breast bulge increases the dose increases and the image quality decreases. This is due to the fact that the measured thickness is used to set the target/filter combination along with the kV, which would not be optimised for the actual breast tissue. Therefore, for the detectors to receive the required amount of photons to create the image, the mAs would have been increased with increasing tissue bulge to compensate, increasing the dose. However, as the kVp was not optimised, the image quality was reduced. This could impact on the accuracy of the biopsy. Dillon et al (2005) found that mammography guided biopsy produced more false negatives than ultrasound or clinical (palpated) guided biopsy (13%, 1.7%, and 8.9% respectively). Although they did not link this to lesion visibility during the biopsy, it was found that in a majority of the false

negatives suspicious lesions had been found on the previous standard FFDM images, implying that the correct area had not been biopsied.

1.2.1. Contribution of the work to the literature and clinical practice

The optimisation score has been used in later publications to discuss observer image quality optimisation assessment (Hogg and Lança, 2015, Reis, Ndlovu et al., 2015) but apart from this, the work has had limited citations. However, the results from this work have had direct clinical impact, because instead of allowing the AECs to manage the imaging parameters, the practitioners are now selecting the appropriate parameters based on the visual assessment of the actual breast thickness (private correspondence Appendix II)

1.2.2. Limitations of the work

There were limitations to paper 3 as it used a phantom rather than clinical images, however the benefit of using the phantom was having a controlled, known input, as well as not incurring any patient radiation risks from repeated imaging.

1.2.3. Objective 1 - conclusion

Overall papers 1-3 demonstrate that there is a difference between the thickness read by the machine and the actual thickness of the tissue. In the first instance this will have an impact on optimisation as the incorrect exposure factors will be set, impact on dose calculations, as well as having an impact on more recent developments such as breast density calculations, which potentially affects patient pathways, incurring additional patient concern and screening costs.

2. Objective 2 - Practitioner variability in applied compression force

Paper 4 - Practitioner compression force variation in mammography: A 6-year study ***and*** ***Paper 5 - A 6-year study of mammographic compression force: Practitioner variability within and between screening sites***

Previous work by Mercer et al (2013) ((Mercer, Hogg et al., 2013b) identified three groups of practitioners compression force behaviours: those who used low, intermediate and high compression across the BI-RADS density grades, they found a wide variation in compression for any given breast volume, with trends of higher compression demonstrated for increasing breast volumes.

This paper was the first to explore practitioners' application of compression force during FFDM, and further studies were conducted to assess this over time, considering sequential client screening attendances (Papers 4 and 5). Paper 4 (Mercer, Hogg et al., 2013c) was the first single centre longitudinal study where three sequential screening visits were assessed for 344 clients. Exclusion

criteria were applied so that there was a consistency of the client between visits. This would imply that the applied compression force would be similar for the three visits if a consistent and standardised approach was being taken.

Paper 4 correlated with the findings of the Paper 3, that practitioners fell into three groups; those who used low, intermediate and high compression. When looking across individual clients, those that coincidently had the same practitioner would have a similar compression force applied across the three visits but there was a significant difference in compression force for those who had different practitioners for each visit, up to a difference of 14 daN (CC view).

Paper 4 implied that the applied force was dependant on the practitioner, rather than the client. A limitation of the work was that it was undertaken in a single centre. Even so, variability of practice was found within this single centre; therefore, a multicentre project was undertaken to determine whether this variability existed in other clinical centres or if it was an issue due to local training, equipment and protocols.

Paper 5 followed the same methodology as Paper 4, but was undertaken in three centres. This was the first multi-centre study to consider compression force behaviour. Similar “low, medium and high” categories were found in two of the three centres. However, in one of the centres there was more consistency between practitioners and it was discussed that this was due to the local clinical protocol that mandated a minimum compression of 100N.

There is variability across the UK with some sites mandating a minimum compression force of 100N. However, this is not consistent across sites, and variability in approach can mean a lack of consistency in client experience, image quality and radiation dose between sites and within the practitioner variability demonstrated above. Furthermore, this variance can also occur within sites.

One complexity in data gathering for both the papers was the need to analyse the data for both sequential visits for the clients and the practitioner behaviour, as well and compression force and thickness. In Paper 4 the use of an Excel spreadsheet led to complexities in data analysis that would have been impossible to manage for the multicentre trial. Therefore, a novel Access database was developed. The benefit of Access compared to Excel is that it is a true multivariate database (whereas Excel is two-dimensional worksheet based) that can be interrogated via Queries for multiple criteria. However, Access does not enable a high level of data analysis, and so the Access database was designed to enable reports to be easily exported to Excel for analysis, based on user-selected criteria (practitioner, client, force, thickness etc).

One database was sent to each centre included in the trial, two of the three centres used the database easily and successfully, data integrity was ensured by using password protected access levels. One centre struggled with using the database, this was unresolvable due to the protection at the centre not allowing macros to be run on the hospital computers (a common protection method to reduce risk from computer viruses) therefore the data had to be inputted off site. This needs to be taken into account if further use of complex Databases are required for acquiring data in hospitals.

2.1. Contribution of the work to the literature

2.1.1. Compression force behaviours

Papers 4 and 5 demonstrated a variability in the way practitioners were applying compression force and highlighted the absence of a national recommended standard or protocol in what is a national screening procedure. However it did not explain the reason for the variability.

To try and understand the reasons for this variability, a qualitative research project was undertaken by qualitative researchers within the same research group, based on the findings from Papers 4 and 5 a grant was secured to undertake a qualitative study to explore the reasons for this variability.

Murphy et al (2015) cited papers 4 & 5 as part of the development of their research question to understand the reasons behind the varying compression force behaviours amongst practitioners. They undertook a study across six different NHSBSP screening centres across the UK, where focus group interviews were employed, including 41 practitioners. A phenomenological approach was taken, which concentrates on the study of consciousness and the objects of direct experience. The findings of this study correlated with the variability demonstrated quantitatively in papers 4 & 5, where the results demonstrated a wide variation in the application of compression force. They found that compression force was applied in many different ways due to individual practitioner experiences and behaviour based on the clinical experiences developed throughout training and subsequently as practitioners.

It is discussed in this paper that the culture and the practice of the units themselves influenced beliefs and attitudes of practitioners in compression force application, this agrees with the findings in paper 5 and can explain the differences between the 3 centres, where higher compression values were applied across all clients/practitioners for one of the centres.

The strongest recommendation to emerge from this study was the need for peer observation to enable practitioners to observe and compare their own compression force practice to that of their colleagues.

A follow-on study was conducted (Nightingale, Murphy et al., 2015) that utilised the same data set from Murphy et al (2015). However a different qualitative methodology was taken where thematic analysis was performed to extract data related to the practitioners' mammography compression problem-solving strategies. The work developed a series of categories, themes and sub-themes. Emerging themes were then peer-validated by two other researchers and developed into a model of practice.

They developed a model that included seven stages which contributed to the practitioners' application of compression force (first impressions; explanations and consent; handling the breast and positioning; applying compression force; final adjustments; feedback). They demonstrated that compression force is not a single stage decision, but in their model, it is considered a seven-stage continuum, where multiple inputs and feedback is used in the decision-making process.

These papers demonstrate the complexity of applying compression, that there are influences from practitioners' own experience, as well as cultural/centre influences that may establish a "baseline" of compression behaviour. However, the actual applied compression force is not decided upon prior to the imaging, and client verbal and non-verbal feedback, as well as visual and tactile feedback during the actual compression leads to variability during the actual event.

2.1.2. Standardisation of pressure rather than force

Papers 4 & 5 have been cited in papers that discuss the development of a new paddle design that attempts to standardise pressure rather than force. The pressure based paddle uses the applied force, and detects the breast area, to indicate pressure rather than force to the practitioner (de Groot, Broeders et al., 2013). This paddle uses a light indicator to inform the practitioner when a pre-selected pressure has been achieved; the pre-selected pressure is set at 10 kPa, which is assumed to be the best value based on venous and diastolic blood pressure. However, to date there is no empirical evidence to justify 10 kPa in terms of image quality and dose. There is evidence that using 10 kPa does not affect the visibility, contrast or sharpness of stable lesions (de Groot, Hopman et al., 2017) however there has been no assessment of lesion visibility at other pressures, however the system allows for the adjustment of pressure if required.

de Groot et al discuss that in the Dutch hospital where the study was conducted, there is a standard force of 14 daN used for all FFDM. This means that clients with smaller breasts would receive greater pressure that those with larger breasts due to the difference in area in contact with the paddle. Women with smaller breasts have been found to experience the most pain during compression (de Groot, Broeders et al., 2013) and a pressure of 10 kPa was suggested to standardise and reduce client pain.

Papers 4 & 5 were cited by these articles to demonstrate that, although the local protocol in the Dutch hospital where the study was carried out uses a standard force, this is not the same for all mammography screening programmes. The work in Papers 4 & 5 demonstrated a variability that was not client based, they were therefore suggesting that the paddle could be used to improve standardisation across clients by taking into account the breast area.

It is important to recognise however, that although this new technology may standardise pressure, positioning will still impact on the effectiveness of this compression, as will be discussed in the following section.

2.1.3. Limitations of the work

For both of these papers one shortcoming of the work is that it was performed on analogue images, this was due to the retrospective analysis over 6 years, which was before FFDM was implemented for screening. However, studies have shown that there is very little change in applied compression force when changing from film-screen to FFDM (Hendrick, Pisano et al., 2010), and so the findings from this study can be assumed to transfer.

Similar, more recent studies that have used digital rather than analogue images have been able to include larger cohort sizes. This is due to being able to use data mining tools to extract data (Waade, Moshina et al., 2017) which reduces the labour intensive process with analogue images.

3. Objective 3 - Optimising image receptor position for pressure balance and breast footprint

Paper 6 - A method to measure paddle and detector pressures and footprints in mammography and ***Paper 7 - Does elevating image receptor increase breast receptor footprint and improve pressure balance?***

As well as applied compression force, positioning is equally important for optimisation of the imaging process. For CC views little has been written about how to maximise the volume of breast tissue imaged (i.e. the area of the breast upon the image receptor (IR)) or how to optimise the relative

pressures exerted on the breast from the paddle and IR. There is also no consensus within the literature regarding how the image receptor should be positioned relative to the inframammary fold (IMF) (Kopans, 2007) (Lee, Stickland et al., 2003) and no information on the effect this may have on image quality and breast area on the image receptor.

During discussion between members of the research team and clinical colleagues, it became apparent that this lack of evidence base was causing a lack of consistency in the position of the image receptor relative to the IMF, and that this was potentially impacting on the image quality and the ability to compress the breast sufficiently. Two studies were undertaken to establish impact of varying the height of the image receptor in relation to the IMF and the breast area on the image receptor ("footprint") and the pressure exerted from the IR and the paddle when varying position of the IR in relation to the IMF and changing compression force.

The first study, Paper 6 (Hogg, Szczepura et al., 2013)) used a pressure matt and the deformable phantom developed in previous work (Hauge, Hogg et al., 2012). Combinations of compression force (60 N, 80 N & 100 N), IR distance from the IMF (-2 cm, -1 cm, 0, +1 cm and +2 cm) and 3 paddles (2 flexible, 1 non-flexible) were assessed. In total this gave 60 data combinations. Area of breast on the IR and the paddle was calculated based on the pixel size of the pressure mat, along with average pressure.

Within this work a novel Uniformity Index (U.I.) was derived; this considered the distribution of average pressure per unit area applied by the IR and by the paddle. The U.I. value has the following implications. If U.I. = 0, there is equal pressure per unit area from the IR and the paddle (equal distribution). If $0<U.I.<1$, there is greater pressure per unit area from the paddle, $-1<U.I.<0$, there is greater pressure per unit area from the IR. It was found that the largest footprint was obtained at +2 cm, and the best U.I was found at +1 cm for the CC view.

As this was a phantom study, there were limitations to the work, such as a lack of representation of the variability of human breast tissue, and the true representation of the tissue behaviour. However, this work was a methodological piece of work, where a novel method and analysis technique was used to access IR positioning and the impact on pressure and footprints to provide a proof of concept.

A follow -on study was conducted in 16 human participants to confirm the phenomena observed in the phantom study (Smith, Szczepura et al., 2015)). Using the findings from paper 6 two IR positions were selected (0 cm and + 2cm from IMF). 16 participants underwent 4 CC views at 80 N, giving 64

datasets in total. No imaging was conducted, and the same equipment was used as in the previous study (Hogg, Szczepura et al., 2013)

The human study agreed with the phantom study that by placing the image receptor at 2cm above the IMF, the breast area was increased, and the balance between the pressure exerted above and below the breast was improved.

3.1. Contribution of the work to the literature

Paper 6 - A method to measure paddle and detector pressures and footprints in mammography and ***Paper 7 - Does elevating image receptor increase breast receptor footprint and improve pressure balance?***

3.1.1. Implications of positioning for dose assessment

The papers have been cited in follow- on work that looked at compression forces used in the Norwegian Breast Cancer Screening programme, and a follow- on paper that considered the implication of this on radiation dose (Waade, Moshina et al., 2017). They found a large variation in applied compression forces, with 40% below the recommended range in the Norwegian BSP (Waade, Moshina et al., 2017). They also found that there was a positive correlation (R^2=0.8) between dose and compression force, however this is contradictory to common perception that the use of compression decreases radiation dose (Markey, 2014), as a positive correlation would imply that an increase in compression force increases radiation dose. It was discussed that as a large data set was used (17,951 clients) there could be a variety of reasons for this finding. Varying equipment was cited as a confounding variable, which agrees with Hauge et al where variability in air KERMA led to the largest uncertainty in MGD (Hauge and Olerud, 2013), however it was also considered that technique and positioning as discussed in this work (Hogg, Szczepura et al., 2013, Smith, Szczepura et al., 2015), leads to a variability in the applied compression force.

3.1.2. Force balancing in mammography

Branderhorst et al (2016) have looked into the balance of compression force, they refer to Papers 6 & 7 to demonstrate the issue of poor positioning of the IR and how this leads to imbalance of force, causing the breast tissue to be pushed up or down relative to the body during compression, which could lead to discomfort for the client due to stretching of the skin and other tissues. Many studies (Eklund, 1991, Engelman, Cizik et al., 2006, Keefe, Hauck et al., 1994, Sharp, Michielutte et al., 2003) have found that skin stretching was reported by clients as causing pain during FFDM. Rather than using a pressure mat Branderhorst et al (2016) measured the force exerted by the image receptor

onto a silicone phantom by placing weighing scales on the image receptor and applying compression force. They demonstrated that if the IR is set too low, this can cause an imbalance of the forces within the breast tissue, they suggested moving the IR during compression to compensate for this imbalance.

Dustler et al., 2012 used a force sensing resistor to observe pressure distributions of breast tissue in 103 women undergoing FFDM. They placed the pressure mat under the breast to observe the pattern of pressure exerted on the breast. They made composite images of the mammograms and the pressure mat images to see if there was greater pressure on denser areas of the breast. They found regions of high compression force in the juxtathoracic region, extending into the armpit, 42% of participants were found to have little pressure exerted onto the actual breast tissue, and that only 37% of participants had the pressure distributed mainly across the breast tissue. This indicates the importance on positioning when applying compression force in FFDM, as if the force is being applied to the juxtathoracic region this could potentially cause pain without reducing breast thickness. Popli et al (2014) (Popli, Teotia et al., 2014) found that positioning is the most important factor affecting the resultant mammography image, with optimal positioning maximising the amount of breast tissue seen on image. (O'Leary and Al Maskari, 2013) also found that women experience greater pain in the MLO view, citing the axilla and sternum as the greatest areas of discomfort.

Therefore, appropriate positioning needs to be taken into account when considering the appropriate compression force to apply, as poor positioning leads to less pressure being exerted on the breast, poor image quality and increased pain.

3.2. Limitations of the work

It was found in this work that there was a statistically significant difference between the left and right breast, this could be explained by physical variation within the patient, but the method led to one practitioner consistently performing the compression on the right breast, with another performing the left breast. Practitioner variability exists when applying compression force (Mercer, Hogg et al., 2013c, Mercer, Szczepura et al., 2015), so this could explain the observed difference. This was obviously an error in the methodology that was overlooked. The data collection was undertaken by two clinical colleagues because the data needed to be acquired with the appropriate clinical techniques and knowledge. Also, due to the sensitive nature of the clinical environment it was not appropriate for a large research team to be involved in the data acquisition. This may have impacted on the approach that was taken, led to the lack of internal validity within the work and so impacted on the results. This did, however, highlight the issue of the variability in applied compression force,

even though the practitioners were collecting data at the same time and observing each other's practice as was found in the previous work (Mercer, Hogg et al., 2013b, Mercer, Hogg et al., 2013c, Mercer, Szczepura et al., 2015).

4. Blurred digital mammography images

During the data collection for Papers 1, 4 and 5 it was found that once the compression had been applied there was a brief period of time where the machine demonstrated a drop in compression force values (Hauge, Hogg et al., 2012, Hogg, Szczepura et al., 2013), these findings were reported as a letter to the editor of Radiography journal (Hogg, Szczepura et al., 2012). This observation was compensated for by allowing a "settling period" until the compression force stabilised at 80 N. However, this work has been taken forward to look at the impact of this change on FFDM.

A study was conducted to look at the extent of paddle motion across multiple FFDM machines (Ma, Brettle et al., 2014), four mammography machines were evaluated, with 2 flexible and 2 fixed paddles for each machine, giving 16 datasets. Paddle movement assessed using two linear potentiometers placed on the paddle corners closest to the chest wall. The deformable breast phantom developed in Paper 1 was used to simulate breast tissue. For each paddle, the movement in millimetres was recorded every 0.5 s and change in compression force (N) was recorded every 1 s, continuous data was recorded for 40 s with the phantom in an initially compressed state at 80 N. Additionally clinical audit data was collected for 28 females on one mammography machine. They found that movement followed a bi-exponential curve, with motion occurring for approximately 40 seconds after compression had been applied, with the greatest movement occurring in the first 10 seconds. They also found a linear relationship between paddle movement and change in compression force.

The bi-exponential behaviour agrees with Zyganitidis et al (2007) where they were using the "spring" like behaviour of soft tissue to simulate phantoms, this methodology has since used in numerous work for breast tissue simulations for breast phantom generation (Bakic, Zhang et al., 2011, Bhatti and Sridhar-Keralapura, 2012, Bliznakova, Bliznakov et al., 2012, Bliznakova, Suryanarayanan et al., 2010, Hsu, Palmeri et al., 2013, Malliori, Bliznakova et al., 2014, Sechopoulos, Bliznakova et al., 2012) . However, this approach has not previously been applied in physical phantoms and demonstrates the application of this theory to temporal changes in breast thickness and compression force over short time frames.

This study was followed by another study (Ma, McEntee et al., 2016) where 12 mammography machines from three manufacturers with 22 flexible and 20 fixed paddles were evaluated, the findings correlated with the previous work in that the greatest movement occurred during the first 10 seconds after maximum compression had been applied (de Groot, Broeders et al., 2015) suggests that this is the “clamping” phase, used for immobilisation, which occurs after the “deformation” where tissue is reduced for image quality and dose optimisation.

Although the motion identified in both these studies (Ma, Brettle et al., 2014) (Ma, McEntee et al., 2016) was small (less than 1 mm) it was not established whether this could be perceived and whether practitioners could detect sub-millimetre motion. Work following this by Ma et al (Ma, Hogg et al., 2015) imaged ball bearings inserted onto the surface of a deformable phantom. 10 images were acquired with a 26 second time interval using a single unit, with fixed and flexible 24 x 30 cm paddles. The extent of blur was measured by assessing the change in ball bearing diameter. They found that there was a significant change in diameter with time, with 60% of motion occurring within the first 10 seconds, the greatest level of blur was found to be within the nipple region.

To establish what impact this motion had on perception, a further study was conducted using simulated image blurring, where steps from 0.1-1.5 mm of simulated motion was mathematical applied using 3 types of image masks. They used 2 observers and 25 images with 15 levels of blur, giving 1200 in total. The presence of blurring was assessed using the NHS BSP guidance through determining whether breast anatomical structures had distinct/sharp edges. They found that for the simulation that most accurately represents physical motion (termed “soft-edge” in the paper), that there was a 4% probability of detecting 0.1 mm blur, increasing to over 100% detection at greater than 0.7 mm. Although these were simulated images, it did indicate that sub-millimetre blurring can be perceived by practitioners.

This work indicates that the motion that occurs during the first stages after the clamping stages could incur detectable levels of blur within the resulting images. As has been found in this work, the greatest movement is consistently within the first few seconds after cessation of compression, therefore to reduce movement unsharpness due to paddle motion, imaging should not occur during this time.

Recent work by Abdullah et al (2017) has shown that mathematically simulated motion can be detected for simulated blur as small as 0.7mm. They found a statistically significant difference in lesion detection using a weighted jackknife alternative free response receiver operating characteristic analysis. A suggested solution to this has been proposed by Ma et al (2017) using a

closed loop system. However, is must be noted that both these papers are based on simulation and no clinical data has been acquired to date.

It is important to point out that the closed loop system suggested by Ma et al (2017) will not resolve movement unsharpness entirely. Voluntary and involuntary patient motion during imaging could still occur, and has been found to be the biggest artefact in FFDM (Choi, Kim et al., 2014). However, the immobilisation from the clamping stage, short exposure times, and clear patient instruction should reduce this (Bontrager and Lampignano, 2014), and is an issue in all medical imaging.

5. Objective 4 - Establish extent of breast thickness reduction with applied compression force

Paper 8 Pressure and breast thickness in mammography-an exploratory calibration study

On reviewing the literature, it is clear that there was little evidence for the optimised compression force to use in FFDM, and little exists in terms of explaining tissue behaviour under compression.

Poulos et al (Poulos, McLean et al., 2003) undertook a study where a standard image was acquired at the optimal compression force as determined by the practitioner. An additional CC image was acquired using a decreased compression force of between 10 – 30 Newtons. They found that the decreased compression force did not reduce the breast thickness in 24% of clients. They concluded if reducing compression force did not reduce thickness then potentially increased compression force is not appropriate, especially when taking into consideration patient comfort.

A follow up study demonstrated in 26 clients that the point at which the breast became taut varied between clients, which leads to variability in the application of compression force (Poulos and McLean, 2004). They also found that 74% of clients did not receive sufficient compression force for the acquired image, however no images were needed to be repeated within this study. The pressure mat study discussed previously (Dustler, Andersson et al., 2012), also found that a decrease in compression force from maximum applied force did not reduce the breast thickness significantly, which leaves the question: how much compression force is needed to minimise breast thickness?

To answer this question a study was designed using a convenience sampling method where 250 sequential symptomatic clients were invited to participate within the study, and 235 agreed. Machine thickness was measured for increasing values of applied compression force, increasing in 5 daN increments from 5 daN until the practitioner considered appropriate compression for imaging had been applied. The work found that increasing compression force caused a decrease in thickness

that followed a polynomial trend. It was suggested that 13 daN was the optimised compression force for the equipment that was used in the experiment (Hologic Selenia FFDM).

The work suggested that there were zones of compression force, where initially large thickness reduction was achieved, whereas at higher compression force the thickness reduction plateaued indicating no reduction in thickness with increasing compression force. This agreed with the initial work undertaken by Poulos et al (2003).

It was suggested that these zones could be used to indicate to practitioners when compression force was no longer going to optimise the imaging process, and potentially lead to increased patient discomfort.

5.1. Limitations and shortcomings within the work

Paper 8 - Pressure and breast thickness in mammography-an exploratory calibration study

This work was criticised for the terminology used within the work, a letter to the editor (Grimbergen and den Heeten, 2013) discussed the inaccurate terms used, on reply it was recognised that clinical colloquialisms had been used, instead of the appropriate scientific terms, which highlighted a disparity between clinical practice and the correct scientific expressions (Hogg, Taylor et al., 2013). However, there was a positive outcome of this initial interaction where due to the strong links between the projects, we entered into a confidentiality clause with them to share early research data.

Following from this correspondence, it was highlighted that the approach taken by Grimbergen, de Heeten et al was a far more appropriate for standardising the use of compression in FFDM. Rather than using equipment that was currently used, they had developed a new paddle which detected the area and the compression force, and combined this information to calculate pressure. (Branderhorst, de Groot et al., 2015, de Groot, Branderhorst et al., 2015, de Groot, Broeders et al., 2013, de Groot, Broeders et al., 2014, de Groot, Broeders et al., 2015). Although this technology aims to standardise the applied compression, it can only achieve this if the positioning is accurate as found in (Dustler, Andersson et al., 2012).

This work did however demonstrate that increased compression does not necessarily decrease breast thickness at higher forces. This agrees with the findings of (Holland, Sechopoulos et al., 2016, Mercer, Hogg et al., 2013a, Saunders and Samei, 2008) that high levels of compression do not improve image quality, and that moderate compression should be applied to improve lesion detection (Holland, Sechopoulos et al., 2016).

Developments emerging from these publications

6. Compression force and lesion detection

Paper 9 Validated novel software to measure the conspicuity index of lesions in DICOM images & Paper 10 Effects of Increased Compression with an Ultrasound Transducer on the Conspicuity of Breast Lesions in a Phantom

Recent work has focussed on developing a novel software that aims to measure lesion visibility; this has been termed Conspicuity Index (CI). CI can be used for focal lesions in any DICOM image, and therefore can easily be applied to mammography imaging. The intention is for further work to be undertaken that quantifies the impact of compression force on the CI of focal lesions. CI shows early promise of being able to inform the relationship between compression force (or pressure) and lesion detection.

As well as affecting patient dose and patient comfort, compression force is applied to improve lesion detection. The rationale for this is discussed in the introduction.

However, there is no supporting evidence or guidance for the appropriate compression force to use in terms of lesion visibility. Consequently, little is known about the relationship between the amount of breast compression and its effect on breast cancer detectability.

A recent study by (Holland, Sechopoulos et al., 2017) demonstrated that too much compression can decrease sensitivity, whilst too little reduces specificity. They used Volpara software to calculate the breast area and used this to calculate the applied pressure in 132,776 examinations of 57,179 women. They then separated the patients into 5 groups based on the quintiles of applied pressure across the group. They compared the calculated applied pressure to many variables, but most importantly to the 12-month sensitivity (based on the interval cancers found within 12 months of screening) and the specificity. They found that the sensitivity decreased with increasing applied pressure, where applying a higher than needed compression actually had a stronger negative effect on lesion visibility than applying insufficient compression. Although this finding was unexpected, previous reports have demonstrated a reduced lesion visibility with “spot-film” imaging, which is traditionally performed at a higher compression force (Brenner, 2001). More expected was the finding that the specificity reduced with too little compression. Clearly there is a need to investigate the relationship between compression and detectability and for appropriate compression to be applied to each individual patient.

This work used the 12-month sensitivity to define the lesion detectability. However, lesion detection is complex, and the reason a lesion was missed may additionally be due to observer decision making, rather than just the impact of decreased lesion visibility due to over compression.

6.1. Quantifying lesion detectability.

One determinant of lesion visibility is to measure its conspicuity. Conspicuity first used in 1974 by Revesz et al(Revesz, Kundel et al., 1974), and defined as the lesion contrast divided by the surrounding complexity, where the lesion contrast is the density change across the lesion border, and the surround complexity is the rate of fluctuation of the density around the lesion border. This led to a definition of structural, or anatomical, noise that considers the surrounding structures and artefacts. However, it was recognised within this seminal piece of work that this is a simple equation and does not take into account other psychophysical metrics that could affect the conspicuity or 'salience' of a lesion.

A mammogram is not a simple shadow of the internal breast anatomy, it is a complex summation of a polyenergetic beam of X-rays that have interacted in a 3D object within multiple layers. Therefore, as it is a summation of several layers, subtle lesions may be obscured by overlying anatomical structures (Coche, Ghaye et al., 2011).

One common way of predicting lesion detectability is to calculate signal to noise or contrast to noise ratios. However, when measuring these ratios, usually the noise is measured from a background that doesn't include complex surrounding structures, therefore these overlying structural complexities that affect the detectability of a lesion are not recognised with these standard measures.

Noise is defined as unwanted information on the image (Samei, 2003), and there are two types that need to be taken into account when considering the conspicuity of the lesion: the structural noise and the radiographic noise. Structural noise is task dependant. However radiographic noise is not dependent on the subject being imaged, but is stochastic in nature and dependent on many factors, such as the exposure factors used and the capability of the detector(Samei, 2003). Radiographic noise is measured using SNR, and provides information about the system capability, but does not give the full information about the noise within a clinical image.

The size of a focal lesion is an important factor in determining its conspicuity. Many studies have found that lesions smaller than 1cm are often missed when viewing 2D images (Plöckinger, 2012), and 3mm is deemed the threshold size limit for detecting a lesion (Coche, Ghaye et al., 2011). In mammography Birdwell et al (Birdwell, Ikeda et al., 2001) found that 81% of missed lesions in their

study were less than 20mm even when retrospectively assessed, and Michealson et al (Michaelson, Satija et al., 2003) found that the median size of lesions detectable by digital mammographic screening was 7.5mm, with only 40% of lesions being seen at 5mm. It is recognised, however, that even the smallest image feature can be detected if its contrast against the background is high enough, for example breast calcifications, and this principle is used to good effect in the contrast/detail test objects in quality assurance programmes (Brettle, Berry et al., 2007).

Contrast is an essential measure of the diagnostic capability of a system. CNR and SNR are standard measures used to represent this, in clinical assessment of images as well as in quality control. They are both size independent if they are used in isolation as quantitative measures of image quality. They represent the difference in signal amplitude between the lesion or test feature and the background (Bushberg, Seibert et al., 2012, Bushong, 2008).

The sharpness refers to the ability of the system to represent distinct anatomical features within the object being imaged (Samei, 2003), therefore the sharpness of the border of a lesion impacts on its visibility. There are mathematical techniques to measure the sharpness of a system. Point and line spread functions and their Fourier transform, the modulation transfer function (MTF) all measure the resolving capability (Samei, 2003) or spatial resolution of a system, and take into account issues such as focal spot size, the polyenergetic beam, and any magnification that can cause blurring of edges. Blurring can also occur due to voluntary or involuntary patient motion. Blurring causes reduced visibility of details, image un-sharpness and reduced spatial resolution. The 3 mm minimum size of lesion visible, discussed above is only applicable if the edges of the structure are parallel to the X-ray beam, if the margins are bevelled (either due to blur, or anatomical causes) then this influences the visible threshold size (Coche, Ghaye et al., 2011). Edge sharpness has a powerful influence on the probability that an image feature will be detected. This is because loss of spatial resolution reduces the effectiveness of the Mach band phenomenon in the visual system which so enhances narrow gradients between adjacent regions of different grey-level (Burgess, 2010).

In summary, the factors reported to affect the conspicuity of a lesion have been found to be the structural noise within and surrounding a lesion, the size of the lesion, the contrast and the sharpness of the edges.

Paper 9 (Szczepura and Manning, 2016) describes the development of a novel software that can be used to calculate the conspicuity of focal lesions from measurements of their image features and background. The software is used in conjunction with an Excel spreadsheet to calculate the

Conspicuity Index (C.I.), the development of the software and the algorithm to calculate C.I. can be found in Paper 9.

In Paper 10 (Szczepura, Faqir et al., 2017), this novel software was utilised to demonstrate the impact of increased compression of an ultrasound probe of lesion visibility within a breast phantom. The study showed that as compression increased the C.I. increased, peaking at 17mm of tissue reduction, after this point the C.I decreased implying that over compression with a transducer decreases lesion visibility in ultrasound.

This finding could be translated into mammography, agreeing with the findings of increased compression reducing sensitivity (Holland, Sechopoulos et al., 2017), and a decrease in lesion visibility with spot film imaging (Brenner, 2001). It is as yet, unclear why lesion conspicuity decreases with compression but for certain lesions under high compression it might be caused by a loss of contrast because of their composition. It could be that more compressible lesions may become less conspicuous with high compression because the cancer tissue may itself spread out and lose contrast with the surrounding tissues. Additionally the vascularisation of lesions needs to be taken into account, since invasive cancers are often highly vascularised, high levels of compression may lead to a reduction in blood flow (Busch, Choe et al., 2014, Carp, Kauffman et al., 2006, Carp, Selb et al., 2008) which would lead to a decrease in image contrast.

6.2. Limitations of the work

The clear limitations of Paper 10 were the fact that it was performed in a phantom, with limited variables. Further work needs to be conducted in order to consider other factors, such as transducer frequency, density of the breast and lesion location, for a better understanding of the effect of compression on the visualisation of breast lesions in ultrasound. Additionally this may impact on new technology such as Automatic Breast Ultrasound (ABUS), where only enough compression stabilise the breast is recommended (Berg and Mendelson, 2014, Carp, Kauffman et al., 2006, Shin, Kim et al., 2015), due to the operator dependency of both hand held and automatic breast ultrasound, advice on the compression to apply could be important for improved lesion visibility.

Although the software described in Paper 9 has been demonstrated to be reliable and reproducible, it has yet to be validated in terms of human perception. Kundel et al (Kundel, Nodine et al., 1978) defined three lesion detection error classifications: search or scanning errors, recognition errors, and decision making errors. Scanning errors are due to the failure of the observer to fixate on the lesion, recognition errors are when the lesion is fixated on yet failing to detect it as a lesion, and decision

making errors are incorrect interpretation of the observed lesion(Kundel, Nodine et al., 1978). By far the largest proportion of errors have found to be in decision making (Kundel, Nodine et al., 1978) but it is clear that detection is the first requirement in the process and this is strongly influenced (even dependent on) the salience of the feature. Other perceptual errors come from satisfaction of search (SOS) error, where the observers' attention is diverted from the lesion by a more conspicuous finding (Berbaum, Brandser et al., 2001, Berbaum, Dorfman et al., 2000, Berbaum, Franken et al., 2000). It is also accepted that the algorithm can only apply to discrete focal lesions and image features. It can make no contribution to estimating the perceptual salience of diffuse patterns in medical images, such as calcifications in mammograms.

Attempting to correlate objective measures with human observer studies has not always proved successful, Manning et al (Manning, Ethell et al., 2004) found poor correlation between the measure of conspicuity index and missed lesions in chest radiography and indicated that decision errors were more common than those of detection. However, their approach to calculating conspicuity was limited, as only four profiles and also the lack of correlation involved other types of observer error as originally noted by Kundel et al [22]. Mello-Thomas et al (Mello-Thoms, Hardesty et al., 2005) also found during an eye tracking study, that unreported lesions often received adequate visual attention. Other eye tracking studies have found that the conspicuity of a lesion, or the amount of time it is observed, is not the only reason they are not reported (Brettle, Berry et al., 2007, Chesters, 1992, Manning, Ethell et al., 2004, Mello-Thoms, 2006). In short, although errors in radiology are not confined to readers missing features because they are poorly demonstrated it is valuable to inform efforts to improve the radiological task with measures of image quality that use functional data on visual performance. Clearly one aspect of improving lesion visibility in mammography is through the optimisation of the applied compression force.

6.3. Future of CI work

CI has the potential to contribute to the distinction between detection and decision errors, and for image optimisation, but only if there is clear evidence that the CI is a valid measure of human performance; it is noted that the C.I. software is currently limited in this validation. To remedy this, future work has been planned to address this; correlating the C.I. calculation and the results of a human observer study will enable a distinction to be made between detection (C.I.) and decision (observer) errors. Furthermore, eye tracking can be utilised to evidence that the observers have fixated a focal lesion and differentiate between perception and cognitive errors (i.e. whether the

observer is making a detection, decision or recognition error). The findings can then be used to determine the threshold value of C.I. for focal lesions in screening mammograms.

This threshold value can be used for multiple advantages in cancer imaging. The threshold will signify the detectability of lesions, which is an improved metric compared to CNR and SNR, this will enable optimisation of the imaging process based on lesion detectability, rather than system capability, which is currently not achieved in design an optimisation of imaging processes. It could be used as a quality assurance metric to assess observer performance or used within research projects to enhance and develop current imaging techniques.

It also has potential for training, as a system of assessing observer performance with comparison to C.I. will allow feedback to the observer in terms of their performance, which may enhance the training and development of lesion detection in mammography screening.

Conclusion

The papers submitted for this thesis have shown that there are important considerations to be made during the positioning and compression of the breast during FFDM. Increasing compression force has been shown to not decrease breast thickness once the deformation stage has been passed. This implies no improvement in image quality nor dose reduction. It is essential to recognise that the applied compression force can reduce within the first moments after the clamping phase commences, and that this paddle motion has the potential to cause perceivable blurring.

The way that compression force is applied has been shown to vary between practitioners, meaning that clients might not receive the same experience for subsequent visits. Additionally, this can impact on the image quality consistency and accuracy of breast density measurements over time.

These issues resulting from technical, client and practitioner factors compound to indicate that there is a need to define the imaging technique more clearly. Optimal positioning, compression and time before imaging have all been shown to impact on the optimisation process, and although technological improvements have attempted to overcome these issues and standardise the application of compression, it is important to recognise the issues with positioning and movement will still need to be overcome, and that guidelines need to include considerations of these.

As screening images are acquired at regular intervals over a period time, change in the anatomical appearance of the breast can indicate clinical changes (Peyton, 2016) and a change in breast density

has been shown to predict the risk of cancer in the contralateral breast in a recent case matched study (Sandberg, Li et al., 2013). Prior images can also be used to monitor disease progression, or assess treatment effectiveness (Bick and Diekmann, 2010). If the image acquisition is not standardised then there is a potential that the similarity between successive images is not sufficient to allow true assessment of change (Bick and Diekmann, 2010).

Variations in compression force have been shown to impact on lesion detection (Holland, Sechopoulos et al., 2016) and breast density calculation (Waade, Highnam et al., 2016). Standardisation and consistency of the compression technique would allow greater similarity between successive visits. The complexity of breast tissue size, shape, morphology, and elasticity, indicates that the same compression force would not be applied to each patient (Holland, Sechopoulos et al., 2016), therefore standardisation rather than generalisability is essential to enable patient centred optimisation.

It can be argued that it is not important for compression force to be the same for all clients, images need to be optimised as much as practicable, taking into account many variables, such as client breast size and density, and patient comfort and tolerance (Nightingale, Murphy et al., 2015). However, if sequential intra-patient images are similar in terms of acquisition they are potentially more useful for accessing change, then this could reduce false positive rates and improve the screening process statistics (Peyton, 2016).

Guidance on the optimisation of compression force needs to take into account all of these variables including consideration of the successive nature of the screening programme and the impact on any future risk stratification to ensure a standardisation for both the client and the benefits of the screening programme.

List of Submitted Papers – Full References, Abstracts, Personal Contributions and Co Author Confirmations

Submitted paper citations (01/10/2018)

Paper	Number of Citations (via Google Scholar)
1	29
2	1
3	4
4	24
5	22
6	5
7	7
8	15
9	9
10	0

Paper 1 – statement of authorship

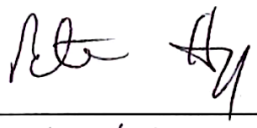
Authors	Hauge, I. H., Hogg, P., Szczepura, K., Connolly, P., McGill, G., & Mercer, C
Title	The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units.
Publication Details	Med Phys. 2012 Jan;39(1):263-71. doi: 10.1118/1.3663579.
Abstract	PURPOSE: To establish a simple method to determine breast readout accuracy on mammography units. METHODS: A thickness measuring device (TMD) was used in conjunction with a breast phantom. This phantom had compression characteristics similar to human female breast tissue. The phantom was compressed, and the thickness was measured using TMD and mammography unit readout. Measurements were performed on a range of screen film mammography (SFM) and full-field digital mammography (FFDM) units (8 units in total; 6 different models/manufacturers) for two different sized paddles and two different compression forces (60 and 100 N). RESULTS: The difference between machine readout and TMD for the breast area, when applying 100 N compression force, for nonflexible paddles was largest for GE Senographe DMR+ (24 cm × 30 cm paddle: +14.3%). For flexible paddles the largest difference occurred for Hologic Lorad Selenia (18 cm × 24 cm paddle: +26.0%). CONCLUSIONS: None of the units assessed were found to have perfect correlation between measured and readout thickness. TMD measures and thickness readouts were different for the duplicate units from two different models/manufacturers.
Candidate Contribution	Within this work I attended all research meetings to discuss the development of the method, designed and managed the manufacturing of the measuring device, and assisted with the analysis. Paper contribution included Fig 5, Fig 6 & Fig 7. Additionally assisted in editing the paper after first author had written the main body of the work.

Co-Author Confirmation

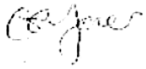
By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	**Ingrid Hauge**
Comments	
Signed	Ingrid Helen Hauge
Date	**23.02.2018**

Name	Peter Hogg
Comments	I agree with the above
Signed	
Date	11/10/18

Name	Paul Connolly
Comments	No Comments
Signed	
Date	13th March 2018

Name	Claire Mercer
Comments	Confirm contribution as above. Research meetings attended. Katy Szczepura designed and managed the manufacturing of the measuring device and assisted with the analysis. Paper contribution included Fig 5, Fig 6 & Fig 7. Draft copies of this article was sent to all co-authors for their input.
Signed	
Date	26/02/2018

Paper 2 – statement of authorship

Authors	Hackney, L., Williams, S., Hogg, P., & Szczepura, K.
Title	Tissue bulge during stereotactic core biopsy
Publication Details	Radiography. 2013 Nov; 19(4):366-368. doi: https://doi.org/10.1016/j.radi.2013.06.007
Abstract	In full field digital mammography (FFDM) the whole breast is subjected to compression with a Perspex compression paddle in order to reduce breast thickness and improve image quality. Once a mammographic abnormality has been detected using FFDM and a decision to proceed with a stereotactic (X-ray) guided core biopsy has been made, a different compression paddle is utilised. This paddle has a central aperture in order to allow access to the lesion for biopsy. Clinical observations made during biopsy procedures have revealed that a bulge of tissue forms within the aperture. The magnitude of the bulge of tissue and BI-RAD breast density was recorded in 15 consecutive patients. Results showed an average of 18.7% (range 11.3–30%) increase in the breast thickness (over the bulge region) compared to the surrounding compressed breast. BI-RAD breast density category 3 had on average the lowest measured thickness and the greatest percentage of tissue bulge. Overall, results confirm that for all patients there was a measurable tissue bulge that varied from 6 mm to 10 mm, representing between 10.14% and 23.08% of additional tissue not measured by the machine. In clinical practice a perceivable difference in lesion visibility was subjectively indicated between the FFDM images and the stereotactic scout biopsy image. The suggested hypothesis from these observations is that there may be an association between the magnitude of the tissue bulge and the ability to accurately perceive certain lesions during stereotactic biopsy procedures. A phantom study is in progress to determine how lesion visibility varies with the amount of tissue bulge.
Candidate Contribution	Within this work I attended all research meetings to discuss the development of the method, and performed all analysis and results display. Paper contribution included Table 1, Graph 1, Graph 2, & Graph 3. Writing the results and discussion section

Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Lisa Hackney
Comments	I can confirm that Katy attended all meetings relating to this research project and undertook the data analysis and resultant research findings.
Signed	
Date	27.02.18

Name	Susan Williams
Comments	During stereotactic core biopsy procedures we noticed that because of the paddle (3D biopsy compression) we used a bulge formed in the aperture and lesions to be targeted often appeared to be much less conspicuous. We discussed our observations with Peter and Katy and this paper was the result which showed support for our observation. Katy attended all of the meetings relating to this research study and she undertook the data analysis of the resultant findings and produced the graphs included in the final paper. She also contributed to all parts of the final write up
Signed	Susan Williams
Date	27/2/18

Name	Peter Hogg
Comments	I agree with the above
Signed	Peter Hogg
Date	11/10/18

Paper 3 – statement of authorship

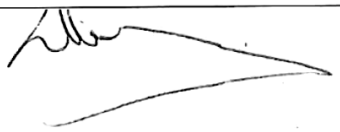
Authors	Williams, S., Hackney, L., Hogg, P., & Szczepura, K.
Title	Breast tissue bulge and lesion visibility during stereotactic biopsy–A phantom study
Publication Details	Radiography, Aug 2014, 20(3), 271-276. doi:http://dx.doi.org/10.1016/j.radi.2014.04.006
Abstract	Background: During mammography guided stereotactic breast biopsy a bulge of tissue can form in the paddle needle biopsy aperture. This bulge has been estimated to have a height of up to 30% of the breast itself. During clinical biopsy we have noticed that lesions can appear to be less visible when tissue bulges are evident. This can make biopsy more difficult in some cases. Objectives: This experiment investigates how lesion visibility varies with breast bulge magnitude. Method: Using a phantom to represent breast and breast bulge, lesion visibility was assessed using a two alternative forced choice methodology. To mimic clinical conditions, imaging was performed on a full field digital mammography system with the biopsy paddle attached using an automatic exposure device. Organ dose (breast) was estimated. Results: As breast bulge increases lesion visibility decreases; organ dose increases as breast bulge magnitude increases. Conclusion: Consideration should be given to the impact of breast bulge magnitude and lesion visibility when performing image guided biopsy. Advances in knowledge: The authors found no similar studies and the results of this study demonstrate a potential clinical risk.
Candidate Contribution	Within this work I attended all research meetings, and performed all analysis as well as developing the novel "optimisation score" algorithm. For the paper I wrote the Methods and Materials, Results and Discussion sections, contributed to all other sections, and drew figures 1- 3 and graphs 1 – 4

Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Susan Williams
Comments	Katy contributed towards the content of the final paper and played an integral part in the development and write up of the final paper. She attended all the meeting for this project and performed the data analysis
Signed	Susan Williams
Date	27/2/18

Name	Lisa Hackney
Comments	Katy undertook the data analysis and developed the optimisation score algorithm. Contributed to all sections of the final paper.
Signed	
Date	27/2/18

Name	Peter Hogg
Comments	I agree with the above
Signed	
Date	11/10/18

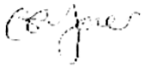
Paper 4 - statement of authorship

Authors	Mercer, C. E., Hogg, P., Szczepura, K., & Denton, E.
Title	Practitioner compression force variation in mammography: A 6-year study
Publication Details	Radiography, Aug 2013, 19(3), 200-206. doi:http://dx.doi.org/10.1016/j.radi.2013.06.001
Abstract	The application of breast compression in mammography may be more heavily influenced by the practitioner rather than the client. This could affect image quality and will affect client experience. This study builds on previous research to establish if mammography practitioners vary in the compression force they apply over a six-year period. This longitudinal study assessed 3 consecutive analogue screens of 500 clients within one screening centre in the UK. Recorded data included: practitioner code, applied compression force (daN), breast thickness (mm), BI-RADS® density category and breast dose. Exclusion criteria included: previous breast surgery, previous/ongoing assessment, breast implants. 344 met inclusion criteria. Data analysis: assessed variation of compression force (daN) and breast thickness (mm) over 3 sequential screens to determine whether compression force and breast thickness were affected by practitioner variations. Compression force over the 3 screens varied significantly; variation was highly dependent upon the practitioner who performed the mammogram. Significant thickness and compression force differences over the 3 screens were noted for the same client (<0.0001). The amount of compression force applied was highly dependent upon the practitioner. Practitioners fell into one of three practitioner compression groups by their compression force mean values; high (mean 12.6 daN), intermediate (mean 8.9 daN) and low (mean 6.7 daN). For the same client, when the same practitioner performed the 3 screens, maximum compression force variations were low and not significantly different ($p > 0.31$). When practitioners from different compression force groups performed 3 screens, maximum compression force variations were higher and significantly different ($p < 0.0001$). The amount of compression force used is highly dependent upon practitioner rather than client. This has implications for radiation dose, patient experience and image quality consistency.
Candidate Contribution	My contribution was to create the way data was recorded and categorised ready for analysis. Within the analysis offered significant input to how it was done and interpreted. Contributed to the ethical application, the article writing and editing.

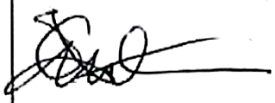
Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Claire Mercer
Comments	Confirm contribution as above. The design for data capture, data analysis and categorisation ready for analysis was completed by Katy Szczepura. The whole study was pivotal and enabled around this design. Support was also gained with data analysis and interpretation of results. Draft copies of this article was sent to all co-authors for their input.
Signed	
Date	26/02/2018

Name	Peter Hogg
Comments	I agree with the above
Signed	
Date	11/10/18

Name	Erika Denton
Comments	Confirm my contribution to this work in manuscript preparation and review of the data.
Signed	
Date	28/02/2018.

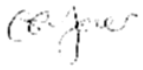
Paper 5 – statement of authorship

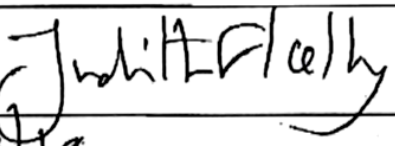
Authors	Mercer, C. E., Szczepura, K., Kelly, J., Millington, S., Denton, E., Borgen, R., Hogg, P
Title	A 6-year study of mammographic compression force: Practitioner variability within and between screening sites.
Publication Details	Radiography, Nov 2013 21(1), 68-73. doi:http://dx.doi.org/10.1016/j.radi.2014.07.004
Abstract	Background: The application of compression force in mammography is more heavily influenced by the practitioner rather than the client. This can affect client experience, radiation dose and image quality. This research investigates practitioner compression force variation over a six year screening cycle in three different screening units. Methods: Data were collected from three consecutive screening events in three breast screening sites. Recorded data included: practitioner code, applied compression force (N), breast thickness (mm), BI-RADS® density category. Exclusion criteria included: previous breast surgery, previous/ongoing assessment and breast implants. 975 clients (2925 client visits, 11,700 mammogram images) met inclusion criteria across three sites. Data analysis assessed practitioner and site variation of compression force and breast thickness. Results: Practitioners across three breast screening sites behave differently in the application of compression force. Two of the three sites demonstrate variability within themselves though they demonstrated no significant difference in mean, first and third quartile compression force and breast thickness values CC ($p > 0.5$), MLO ($p > 0.1$) between themselves. However, in the third site, where mandate dictates a minimum compression force is applied, greater consistency was demonstrated between practitioners and clients; a significant difference in mean, first and third quartile compression force and breast thickness values ($p < 0.001$) was demonstrated between this site and the other two sites. Conclusion: Variability within these two sites and between the three sites could result in variations. Stabilisation of these variations may have a positive impact on image quality, radiation dose reduction, re-attendance levels and potentially cancer detection. The large variation in compression forces could negatively impact on client experience between the units and within a unit. Further research is required to establish best practice guidelines for compression force within mammography. Advances in knowledge: Practitioners vary in the compression forces they apply to clients over sequential screening attendances. Establishing practice guidance with cessation guidelines could help to minimise this problem.
Candidate Contribution	My contribution was to create the way data was recorded and categorised ready for analysis. Within the analysis offered significant input to how it was done and interpreted. Contributed to the ethical application, the article editing and responding to reviewer comments

Co-Author Confirmation

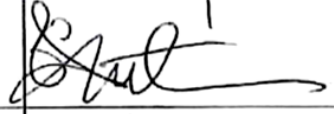
By signing the Statement of Authorship, each author certifies that

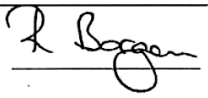
- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

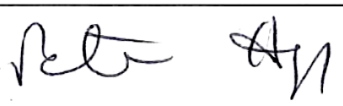
Name	Claire Mercer
Comments	Confirm Katy's contribution as above: The design for data capture, data analysis and categorisation ready for analysis was completed by Katy Szczepura. The whole study was pivotal and enabled around this design. Support was also gained with analysis and interpretation of results. Draft copies of this article was sent to all co-authors for their input.
Signed	
Date	26/02/2018

Name	**Judith Kelly**
Comments	Confirm contribution
Signed	
Date	12/03/18

Name	**Sara Millington**
Comments	Confirm contribution
Signed	SR Millington
Date	23/2/18

Name	Erika Denton
Comments	I confirm my contribution to this study - data analysis and writing up
Signed	
Date	28/2/2018.

Name	Rita Borgen
Comments	I wish to confirm the contribution to the above submission.
Signed	
Date	32.2 18

Name	Peter Hogg
Comments	I agree with the above
Signed	
Date	11/12/18

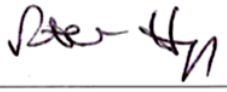
Paper 6 – statement of authorship

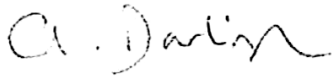
Authors	Hogg, P., Szczepura, K., Darlington, A., & Maxwell, A.
Title	A method to measure paddle and detector pressures and footprints in mammography
Publication Details	Radiography, Nov 2013 21(1), 68-73. doi:http://dx.doi.org/10.1016/j.radi.2014.07.004
Abstract	PURPOSE: Compression is necessary in mammography to improve image quality and reduce radiation burden. Maximizing the amount of breast in contact with the image receptor (IR) is important. To achieve this, for the craniocaudal projection, there is no consensus within the literature regarding how the IR should be positioned relative to the inframammary fold (IMF). No information exists within the literature to describe how pressure balancing between IR and paddle, and IR breast footprint, might be optimized. This paper describes a novel method for measuring the respective pressures applied to the breast from the IR and the paddle and a method to simultaneously measure the breast footprints on the IR and the paddle. METHODS: Using a deformable breast phantom and electronic pressure-sensitive mat, area and pressure readings were gathered from two mammography machines and four paddles at 60, 80, and 100 N with the IR positioned at -2, -1, 0, +1, and +2 cm relative to the IMF (60 combinations in total). RESULTS: Paddle and IR footprints were calculated along with a uniformity index (UI). For all four paddle/machine/pressure combinations the greatest IR footprint was achieved at IMF +2 cm. The UI indicates that the best pressure/footprint balance is achieved at IMF +1 cm. CONCLUSIONS: The authors' method appears to be suited to measuring breast footprints and pressures on IR and paddle and a human female study is planned.
Candidate Contribution	Attended all research meetings, wrote the data collection and analysis method to ensure reproducibility and reliability. Wrote the Method and Results sections, contributed to the introduction and discussion sections, figures 1-4, tables 1-4

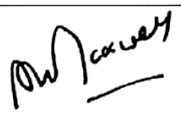
Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Peter Hogg
Comments	I agree with the above
Signed	
Date	11/10/18

Name	Alison Darlington
Comments	
Signed	
Date	05.03.18

Name	Anthony Maxwell
Comments	
Signed	
Date	23/02/2018

Paper 7 – statement of authorship

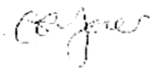
Authors	Smith, H., Szczepura, K., Mercer, C., Maxwell, A., & Hogg, P
Title	Does elevating image receptor increase breast receptor footprint and improve pressure balance?
Publication Details	Radiography, Nov 2015, 21(4), 359–363. doi:http://dx.doi.org/10.1016/j.radi.2015.02.001
Abstract	There is no consensus in the literature regarding the image receptor (IR) position for the craniocaudal projection in mammography. Some literature indicates the IR should be positioned to the infra mammary fold (IMF); other literature suggests the IR be raised 2 cm relative to the IMF. Using 16 female volunteers (32 breasts) and a pressure sensitive mat we investigated breast footprint and pressure balance with IR at IMF and IR 2 cm above the IMF. Breast area on IR and paddle and interface pressure between IR/breast and paddle/breast were recorded. A uniformity index (UI) gave a measure of pressure balance between IR/breast and paddle/breast. IR breast footprint increases significantly by 13.81 cm2 ($p < 0.02$) when IR is raised by 2 cm. UI reduces from 0.4 to 0.00 ($p = 0.04$) when positioned at IMF +2 cm demonstrating an improved pressure balance. Practitioners should consider raising the IR by 2 cm relative to the IMF in clinical practice. Further work is suggested to investigate the effects of practitioner variability and breast asymmetry.
Candidate Contribution	Attended all research meetings, developed the method to ensure reproducibility and reliability, and developed the form of analysis (uniformity index) Large contribution to the writing of the method and results section, contributed to all other sections in the paper to a lesser extent, graphs 1-4, tables 1-3.

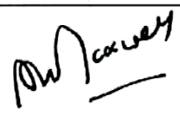
Co-Author Confirmation

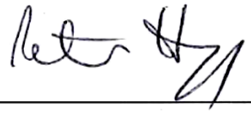
By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Helen Smith
Comments	Confirming Katy Szczepura's valuable contribution to the above publication
Signed	
Date	23/2/18

Name	Claire Mercer
Comments	Confirm contribution as above. Attended all research meetings, developed the method to ensure reproducibility and reliability, and developed the form of analysis (uniformity index) Contributed to paper with other co-authors, with the method and wrote the results section including graphs 1-4, tables 1-3.
Signed	
Date	26/02/2018

Name	Anthony Maxwell
Comments	
Signed	
Date	23/02/2018

Name	Peter Hogg
Comments	I agree with the above
Signed	11/10/18 
Date	

Paper 8 - Statement of Authorship

Authors	Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E
Title	Pressure and breast thickness in mammography-an exploratory calibration study
Publication Details	Br J Radiol. Jan 2013;86(1021):20120222. doi: 10.1259/bjr.20120222.
Abstract	OBJECTIVE: To perform a calibration study to provide data to help improve consistency in the pressure that is applied during mammography. METHODS: Automatic readouts of breast thickness accuracy vary between mammography machines; therefore, one machine was selected for calibration. 250 randomly selected clients were invited to participate; 235 agreed, and 940 compression data sets were recorded (breast thickness, breast density and pressure). Pressure (measured in decanewtons) was increased from 5 daN through 1-daN intervals until the practitioner felt that the pressure was appropriate for imaging; at each pressure increment, breast thickness was recorded. RESULTS: Graphs were generated and equations derived; second-order polynomial trend lines were applied using the method of least squares. No difference existed between breast densities, but a difference did exist between "small" (15×29 cm) and "medium/large" (18×24/24×30 cm) paddles. Accordingly, data were combined. Graphs show changes in thickness from 5-daN pressure for craniocaudal and mediolateral oblique views for the small and medium/large paddles combined. Graphs were colour coded into three segments indicating high, intermediate and low gradients [≤-2 (light grey); -1.99 to -1 (mid-grey); and ≥-0.99 (dark grey)]. We propose that 13 daN could be an appropriate termination pressure on this mammography machine. CONCLUSION: Using patient compression data we have calibrated a mammography machine to determine its breast compression characteristics. This calibration data could be used to guide practice to minimise pressure variations between practitioners, thereby improving patient experience and reducing potential variation in image quality. ADVANCES IN KNOWLEDGE: For the first time, pressure-thickness graphs are now available to help guide mammographers in the application of pressure.
Candidate Contribution	Attended all meetings, performed the analysis and contributed to writing and editing of all sections. Responded along, with first author, to letters to the editor about the work. All figures within the paper.

Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	Peter Hogg
Comments	I agree with the above
Signed	Peter Hogg
Date	11/01/18

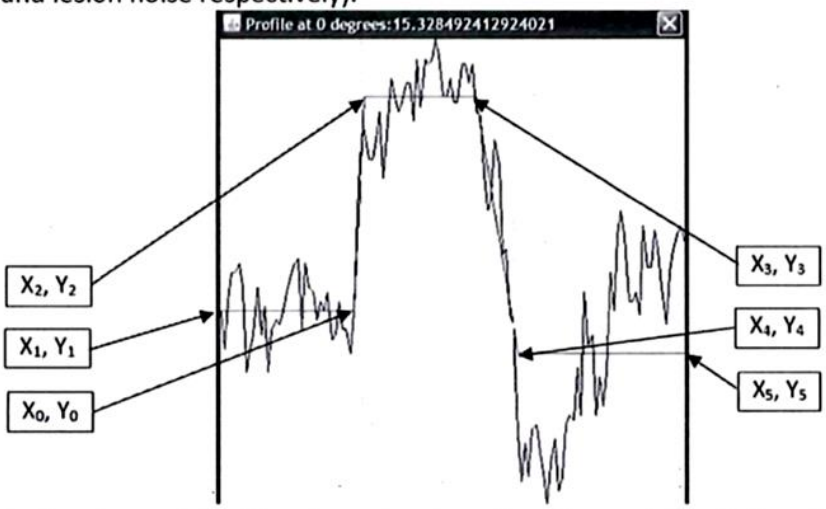
Name	Melanie Taylor
Comments	
Signed	M. Taylor
Date	28.2.18

Name	Claire Mercer
Comments	Confirm Katy's contribution as above: Attended meetings, coordinated discussion around design for data capture, performed the analysis and produced all figures within the paper. Contributed to writing and editing of sections with other co-authors. Responded to letters to the editor about the paper.
Signed	
Date	26/02/2018

Name	Erika Denton
Comments	I confirm my contribution as an author to Planning, data analytics and manuscript preparation.
Signed	
Date	28/02/2018.

Paper 9 - Statement of Authorship

Authors	Szczepura, KR, Manning, DJ
Title	Validated novel software to measure the conspicuity index of lesions in DICOM images
Publication Details	Proceedings Volume 9787, Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment; 978703 (2016); doi: 10.1117/12.2216211
Abstract	Description of purpose A novel software programme and associated Excel spreadsheet has been developed to provide an objective measure of the expected visual detectability of focal abnormalities within DICOM images. Methodology ROIs are drawn around the abnormality, the software then fits the lesion using a least squares method to recognise the edges of the lesion based on the full width half maximum. 180 line profiles are then plotted around the lesion, giving 360 edge profiles. The co-ordinates show in Figure 1 are captured, as well the standard deviation of the pixel values within the background and lesion (representing anatomical noise and lesion noise respectively).



An Excel spreadsheet has been developed to allow variables to be calculated, including SNR and CNR. A conspicuity index has also been developed:

$$\chi = \frac{d \tan[\theta - 1]\Delta GL}{\sqrt{\sigma_s^2 + \sigma_n^2}}$$

Where:

χ = conspicuity index

d = lesion dimension ($X_3 - X_2$)

θ = maximum slope angle of all the line profiles

For each side of the line profile:

$$\theta = \tan^{-1}\left[\frac{Y_1 - Y_0}{X_1 - X_0}\right]$$

$$\theta = \tan^{-1}\left[\frac{Y_1 - Y_0}{X_1 - X_0}\right]$$

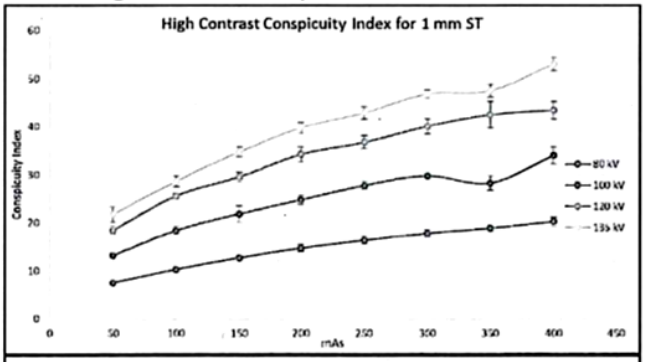
$\Delta GL = Y_2 - \frac{(Y_1 + Y_4)}{2}$ = the difference in average grey levels between the lesion and the background

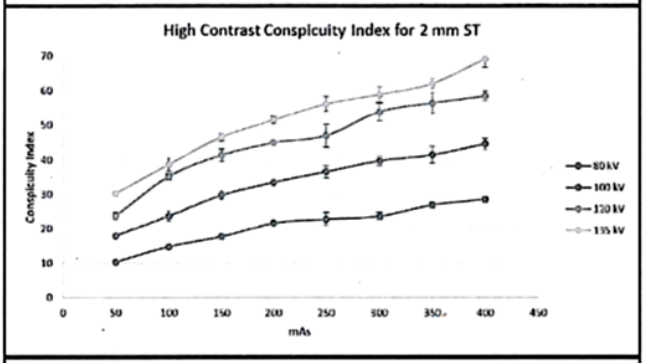
$\sigma_s = sdTop$

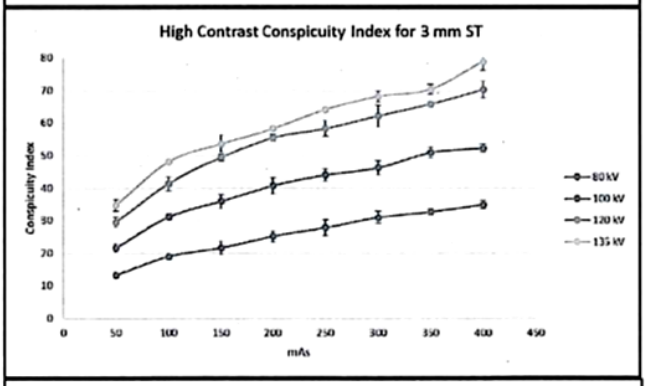
$\sigma_n = \frac{sd1 + sd2}{2}$

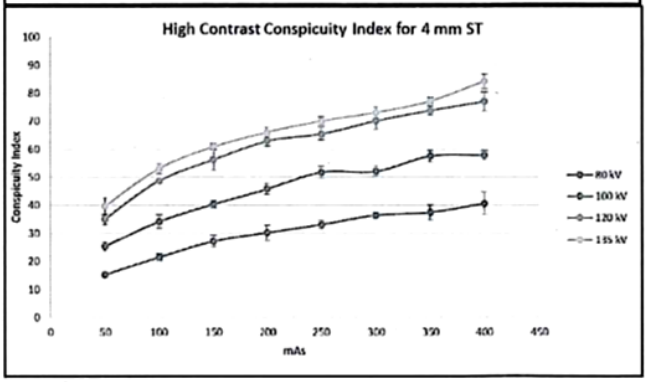
Results

The software has been validated using the GAMMEX ACR CT accreditation phantom, varying mA, kVp and slice thickness (ST) and the results have been found to give a linear response:









Conclusion

A novel software programme has been validated to allow calculation of many physical properties of lesions. Additionally, a new measure of conspicuity index has been developed for focal lesions.

The analysis could be further developed to incorporate reader decision-analysis data and eye-tracking data allowing correlations between physical and perception measures to be made beyond basic CNR calculations. It could also be used as a tool to distinguish between perceptual and cognitive error.

	Further refinements could lead to measures of the detectability of more diffuse disease features.
Candidate Contribution	Designed software, collected and analysed all data, wrote all sections, all graphs and figures Presented the paper at SPIE Medical Imaging 2016

Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	David Manning
Comments	I confirm the statement of authorship above.
Signed	
Date	23.02.2018

Paper 10 - Statement of Authorship

Authors	Szczepura, KR, Faqir, T, Manning, DJ
Title	Effects of increased compression with an ultrasound transducer on the conspicuity of breast lesions in a phantom
Publication Details	Proceedings Volume 10136, Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment; 101360T (2017); doi: 10.1117/12.2254263
Abstract	Ultrasound imaging of the breast is highly operator dependent. The amount of pressure applied with the transducer has a direct impact on the lesion visibility in breast ultrasound. The conspicuity index is a quantitative measure of lesion visibility, taking into account more parameters than standard measures that impact on lesion detection. This study assessed the conspicuity of lesions within a breast phantom using increased transducer compression in breast ultrasound. Methods A phantom was constructed of gelatine to represent adipose tissue, steel wool for glandular/blood vessels and silicone spheres to represent lesions, this meant that the lesions were also compressible, but less than the surrounding tissue. The phantom was imaged under increasing transducer compression. The conspicuity index was measured using the Conspicuity Index Software. The distance between the transducer surface and lesion surface was measured as an indication of increased compression. Results When moderate compression (17mm) was applied, the conspicuity index increased resulting in better visualisation of the silicone lesions. However, with increased compression the conspicuity index decreased. New work to be presented The conspicuity index has never been demonstrated in ultrasound imaging before. This is preliminary phantom work to demonstrate the impact of increased transducer compression on quantitative lesion visibility assessment. Conclusion The compression applied should be considered for optimum visualisation, as excessive pressure decreases conspicuity. However, further work needs to be conducted in order to consider other factors, such as density of the breast and lesion location, for a better understanding of the effect of compression on the visualisation of the lesion. A human study is planned.
Candidate Contribution	Designed method, supervised collection and analysis of all data, wrote all sections, all graphs and figures Presented the paper at SPIE Medical Imaging 2017

Co-Author Confirmation

By signing the Statement of Authorship, each author certifies that

- the candidate's stated contribution to the publication is accurate (as detailed above)
- permission is granted for the candidate in include the publication in the thesis

Name	David Manning
Comments	I confirm the authorship indicated above
Signed	
Date	23.02.2018

Journals - Aims, Scope and Impact Factors

Radiography

Radiography is an international, English language, peer-reviewed journal of radiographic imaging and radiation therapy.

Radiography is the official professional journal of the Society and College of Radiographers and is published quarterly by Elsevier Ltd.

Radiography aims to publish the highest quality clinical scientific and educational material, on all aspects of radiographic imaging (to include diagnostic radiography, computed tomography, nuclear medicine, sonography and magnetic resonance imaging) and all aspects of radiation therapy (to include patient care, dosimetry, treatment planning, verification, treatment delivery and oncology).

Radiography includes original research, novel review articles, technical analyses, evaluations and case studies. In addition it provides a forum for the exchange of information and views on all matters related to the profession of radiography.

Radiography promotes excellence in the profession of radiography by its commitment to the publication of original research, encouragement and dissemination of best clinical, scientific and educational practice and support for education.

Medical Physics

Medical Physics is the scientific journal of the American Association of Physicists in Medicine and is an official science journal of the Canadian Organization of Medical Physicists, the Canadian College of Physicists in Medicine, and the International Organization for Medical Physics (IOMP). It publishes research concerned with the application of physics and mathematics to the solution of problems in medicine and human biology. Manuscripts covering theoretical or experimental approaches are published.

Impact Factor 3.208

British Journal of Radiology

BJR is the international research journal of the British Institute of Radiology and is the oldest scientific journal in the field of radiology and related sciences.

Dating back to 1896, BJR's history is radiology's history, and the journal has featured some landmark papers such as the first description of Computed Tomography "Computerized transverse axial tomography" by Godfrey Hounsfield in 1973 [1]. A valuable historical resource, the complete BJR archive has been digitized from 1896.

BJR is an international, multi-disciplinary journal covering the clinical and technical aspects of medical imaging, radiotherapy, oncology, medical physics, radiobiology and the underpinning sciences. BJR is essential reading for radiologists, medical physicists, radiotherapists, radiographers and radiobiologists.

Impact Factor 1.533

Proceedings of SPIE Medical Imaging

Proceedings of SPIE offer access to the latest innovations in research and technology in the broad field of optics and are among the most cited references in patent literature. Presenting and publishing research in *Proceedings of SPIE* has multiple professional and business benefits but above all it offers an opportunity for first disclosure of research findings both at the stage of completion and as work in progress. *The Value of Proceedings* white paper provides a comprehensive discussion. *Proceedings of SPIE* are uniquely valuable as they are used by researchers worldwide, and lead in patent cites among optics and photonics collections. They frequently rank among the 50 most used serials, out of 50,000 analyzed by Ex Libris. If it is ready for journal publication, a proceedings paper can be submitted with minimal or no revision for consideration in the relevant SPIE journal.

SPIE partners with relevant scientific databases to enable researchers to find *Proceedings of SPIE*. They are indexed in Web of Science, Scopus, Ei Compendex, Inspec, Google Scholar, Astrophysical Data System (ADS), CrossRef, and other scholarly indexes, and are widely accessible to leading research organizations, conference attendees, and individual researchers.

Reference List

Alonzo-Proulx, O., R. A. Jong and M. J. Yaffe (2012). Volumetric breast density characteristics as determined from digital mammograms. *Phys Med Biol* 57(22) 7443-7457.

Alonzo-Proulx, O., et al. (2015). Reliability of automated breast density measurements. *Radiology* 275(2) 366-376.

Antoniou, A. C., et al. (2000). Risk models for familial ovarian and breast cancer. *Genet Epidemiol* 18(2) 173-190.

Arver, B., et al. (2000). Hereditary breast cancer: a review. *Semin Cancer Biol* 10(4) 271-288.

Astley, S. and E. Harkness. (2015). *Automated measurements for personalized breast cancer screening* [*online*]. Available at: http://spie.org/newsroom/6198-automated-measurements-for-personalized-breast-cancer-screening#B12017].

Ayub, S. G., et al. (2014). Mutational analysis of the BRCA2 gene in breast carcinoma patients of Kashmiri descent. *Mol Med Rep* 9(2) 749-753.

Bakic, P. R., C. Zhang and A. D. Maidment (2011). Development and characterization of an anthropomorphic breast software phantom based upon region-growing algorithm. *Med Phys* 38(6) 3165-3176.

Berbaum, K. S., et al. (2001). Gaze dwell times on acute trauma injuries missed because of satisfaction of search. *Acad Radiol* 8(4) 304-314.

Berbaum, K. S., et al. (2000). Proper ROC analysis and joint ROC analysis of the satisfaction of search effect in chest radiology. *Acad Radiol* 7(11) 945-958.

Berbaum, K. S., et al. (2000). Role of faulty decision making in the satisfaction of search effect in chest radiography. *Acad Radiol* 7(12) 1098-1106.

Berg, W. A. and E. B. Mendelson (2014). Technologist-performed handheld screening breast US imaging: how is it performed and what are the outcomes to date? *Radiology* 272(1) 12-27.

Bhatti, S. N. and M. Sridhar-Keralapura (2012). A novel breast software phantom for biomechanical modeling of elastography. *Med Phys* 39(4) 1748-1768.

Bick, U. and F. Diekmann (2010). *Digital mammography*: Springer Verlag.

Birdwell, R. L., et al. (2001). Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology* 219(1) 192-202.

Bliznakova, K., Z. Bliznakov and I. Buliev (2012). Comparison of algorithms for out-of-plane artifacts removal in digital tomosynthesis reconstructions. *Comput Methods Programs Biomed* 107(1) 75-83.

Bliznakova, K., et al. (2010). Evaluation of an improved algorithm for producing realistic 3D breast software phantoms: application for mammography. *Med Phys* 37(11) 5604-5617.

Bond, M., et al. (2013a). Systematic review of the psychological consequences of false-positive screening mammograms. *Health Technol Assess* 17(13) 1-170, v-vi.

Bond, M., et al. (2013b). Psychological consequences of false-positive screening mammograms in the UK. *Evid Based Med* 18(2) 54-61.

Bontrager, K. L. and J. P. Lampignano (2014). *Textbook of radiographic positioning and related anatomy*. 8th ed. / Kenneth L. Bontrager, John P. Lampignamo. ed. St. Louis, Mo. ; [London]: Elsevier Mosby.

Borg, M., I. Badr and G. J. Royle (2012). The use of a figure-of-merit (FOM) for optimisation in digital mammography: a literature review. *Radiat Prot Dosimetry* 151(1) 81-88.

Branderhorst, W., et al. (2015). Mammographic compression--a need for mechanical standardization. *Eur J Radiol* 84(4) 596-602.

Branderhorst, W., et al. (2016). Force balancing in mammographic compression. *Med Phys* 43(1) 518.

Bray, F., et al. (2012). Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. *Lancet Oncol* 13(8) 790-801.

Bray, F., P. McCarron and D. M. Parkin (2004). The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Res* 6(6) 229-239.

Brenner, R. J. (2001). Asymmetric densities of the breast: strategies for imaging evaluation. *Semin Roentgenol* 36(3) 201-216.

Brentnall, A. R., et al. (2015). Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Res* 17(1) 147.

Brett, J., et al. (2005). The psychological impact of mammographic screening. A systematic review. *Psychooncology* 14(11) 917-938.

Brettle, D. S., E. Berry and M. A. Smith (2007). The effect of experience on detectability in local area anatomical noise. *Br J Radiol* 80(951) 186-193.

Broeders, M. J., et al. (2015). Comparison of a flexible versus a rigid breast compression paddle: pain experience, projected breast area, radiation dose and technical image quality. *Eur Radiol* 25(3) 821-829.

Burgess, A. (2010). Spatial vision research without noise. *In* E. Samei and E. A. Krupinski eds. *The handbook of medical image perception and techniques*. Cambridge, Cambridge University Press. 23.

Busch, D. R., et al. (2014). Blood flow reduction in breast tissue due to mammographic compression. *Acad Radiol* 21(2) 151-161.

Bushberg, J. T., et al. (2012). *The essential physics of medical imaging*. 3rd ed., International ed. ed. Philadelphia, Pa. ; London: Wolters Kluwer/Lippincott Williams & Wilkins.

Bushong, S. C. (2008). *Radiologic science for technologists : physics, biology, and protection*. 9th ed. ed. St. Louis, Mo. ; [London]: Elsevier Mosby.

Calderón-Garcidueñas, A. L., et al. (2005). Clinical follow up of mexican women with early onset of breast cancer and mutations in the BRCA1 and BRCA2 genes. *Salud Publica Mex* 47(2) 110-115.

Cancer Research UK. (2016). *Breast Cancer Statistics* [*online*]. Cancer Research UK: Cancer Research UK. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero.

Carp, S. A., et al. (2006). Compression-induced changes in the physiological state of the breast as observed through frequency domain photon migration measurements. *J Biomed Opt* 11(6) 064016.

Carp, S. A., et al. (2008). Dynamic functional and mechanical response of breast tissue to compression. *Opt Express* 16(20) 16064-16078.

Carton, A. K., et al. (2004). Quantification of Al-equivalent thickness of just visible microcalcifications in full field digital mammograms. *Med Phys* 31(7) 2165-2176.

Chen, J. H., G. Gulsen and M. Y. Su (2015). Imaging Breast Density: Established and Emerging Modalities. *Transl Oncol* 8(6) 435-445.

Chesters, M. S. (1992). Human visual perception and ROC methodology in medical imaging. *Phys Med Biol* 37(7) 1433-1476.

Choi, J. J., et al. (2014). Mammographic artifacts on full-field digital mammography. *J Digit Imaging* 27(2) 231-236.

Coche, E. E., et al. (2011). *Comparative Interpretation of CT and Standard Radiography of the Chest [electronic resource]*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Coleman, M. P., et al. (2008). Cancer survival in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol* 9(8) 730-756.

Collaborative Group on Hormonal Factors in Breast Cancer (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet* 360(9328) 187-195.

Damases, C. N., P. Hogg and M. F. McEntee (2017). Inter country analysis of breast density classification using visual grading. *Br J Radiol* 20170064.

de Groot, J. E., et al. (2015). Towards personalized compression in mammography: a comparison study between pressure- and force-standardization. *Eur J Radiol* 84(3) 384-391.

de Groot, J. E., et al. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. *Med Phys* 40(8) 081901.

de Groot, J. E., et al. (2014). Mammographic compression after breast conserving therapy: controlling pressure instead of force. *Med Phys* 41(2) 023501.

de Groot, J. E., et al. (2015). Pain-preventing strategies in mammography: an observational study of simultaneously recorded pain and breast mechanics throughout the entire breast compression cycle. *BMC Womens Health* 15 26.

de Groot, J. E., et al. (2017). Pressure-standardised mammography does not affect visibility, contrast and sharpness of stable lesions. *Eur J Radiol* 86 289-295.

Dillon, M. F., et al. (2005). The accuracy of ultrasound, stereotactic, and clinical core biopsies in the diagnosis of breast cancer, with an analysis of false-negative cases. *Ann Surg* 242(5) 701-707.

Dustler, M., et al. (2012). Breast compression in mammography: pressure distribution patterns. *Acta Radiol* 53(9) 973-980.

Díaz, O., et al. (2016). Feasibility of Depth Sensors to Study Breast Deformation During Mammography Procedures. International Workshop on Digital Mammography, Springer, Cham.

Eklund, G. (1991). Mammographic compression: science or art? *Radiology* 181(2) 339-341.

Eklund, G. W., G. Cardenosa and W. Parsons (1994). Assessing adequacy of mammographic image quality. *Radiology* 190(2) 297-307.

Ekpo, E. U., et al. (2015). Breast composition: Measurement and clinical use. *Radiography* 21(4) 324-333.

Engelman, K. K., A. M. Cizik and E. F. Ellerbeck (2006). Women's Satisfaction with Their Mammography Experience: Results of a Qualitative Study. *Women & Health* 42(4) 17-35.

Evans, D. G., et al. (2013). Familial breast cancer: summary of updated NICE guidance. *BMJ* 346 f3829.

Geeraert, N., et al. (2014). Comparison of volumetric breast density estimations from mammography and thorax CT. *Phys Med Biol* 59(15) 4391-4409.

Gilbert, F. J., et al. (2015). The TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme--a multicentre retrospective reading study comparing the diagnostic performance of digital breast tomosynthesis and digital mammography with digital mammography alone. *Health Technol Assess* 19(4) i-xxv, 1-136.

Great Britain. Health and Safety, E. (2000). *The Ionising Radiation (Medical Exposure) Regulations 2000 [electronic resource]*. [London]: [Stationery Office].

Grimbergen, C. A. and G. J. den Heeten (2013). Pressure and breast thickness in mammography--what about physics? *Br J Radiol* 86(1027) 20130208.

Hackney, L., et al. (2013). Tissue bulge during stereotactic core biopsy. *Radiography* 19(4) 366-368.

Hall, J. M., et al. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250(4988) 1684-1689.

Hartman, K., et al. (2008). Volumetric Assessment of Breast Tissue Composition from FFDM Images. *In* E. A. Krupinski ed. *Digital mammography : 9th international workshop, IWDM 2008 Tucson, AZ, USA, July 20-23, 2008 : proceedings*. Berlin, Springer.

Hauge, I. H., et al. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Med Phys* 39(1) 263-271.

Hauge, I. H. and H. M. Olerud (2013). Uncertainties involved in the estimation of mean glandular dose for women in the Norwegian Breast Cancer Screening Program (NBCSP). *Radiat Prot Dosimetry* 155(1) 81-87.

Hendrick, R. E., et al. (2010). Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the American College of Radiology Imaging Network digital mammographic imaging screening trial. *AJR Am J Roentgenol* 194(2) 362-369.

Highnam, R., et al. (2010). Robust Breast Composition Measurement - VolparaTM. International Workshop on Digital Mammography, Springer, Berlin, Heidelberg.

Hofvind, S., et al. (2012). False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. *J Med Screen* 19 Suppl 1 57-66.

Hogg, P., J. Kelly and C. Mercer (2015). *Digital Mammography: A Holistic Approach*: Springer International Publishing.

Hogg, P. and L. Lança (2015). *Radiation dose and image quality optimisation in medical imaging*. University of Salford: Open Source.

Hogg, P., et al. (2013). A method to measure paddle and detector pressures and footprints in mammography. *Med Phys* 40(4) 041907.

Hogg, P., et al. (2012). Blurred digital mammography images. *Radiography* 18(1) 55-56.

Hogg, P., et al. (2013). Pressure and breast thickness in mammography-what about physics? Author reply. *British Journal of Radiology* 86(1027).

Holland, K., et al. (2016). Performance of Breast Cancer Screening Depends on Mammographic Compression. International Workshop on Digital Mammography, Springer, Cham.

Holland, K., et al. (2017). Influence of breast compression pressure on the performance of population-based mammography screening. *Breast Cancer Res* 19(1) 126.

Hsu, C. M., et al. (2013). Generation of a suite of 3D computer-generated breast phantoms from a limited set of human subject data. *Med Phys* 40(4) 043703.

ICRP (2006). The Optimisation of Radiological Protection - Broadening the Process. *ICRP: ICRP Publication 101b* Ann. ICRP 36 (3).

Independent UK Panel on Breast Cancer Screening (2012). The benefits and harms of breast cancer screening: an independent review. *Lancet* 380(9855) 1778-1786.

International Cancer Screening Network. (2012). *Breast Cancer Organization of Screening Programs* [*online*]. Available at: https://healthcaredelivery.cancer.gov/icsn/breast/screening.html.

Jatoi, I. (2015). Breast-Cancer Screening--Viewpoint of the IARC Working Group. *N Engl J Med* 373(15) 1478-1479.

Jørgensen, K. J. and S. Bewley (2015). Breast-Cancer Screening--Viewpoint of the IARC Working Group. *N Engl J Med* 373(15) 1478.

Keefe, F. J., et al. (1994). Mammography pain and discomfort: a cognitive-behavioral perspective. *Pain* 56(3) 247-260.

Kopans, D. B. (2007). *Breast imaging*. 3rd ed. ed. Philadelphia, Pa. ; London: Lippincott Williams & Wilkins.

Kopans, D. B. (2008). Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk. *Radiology* 246(2) 348-353.

Kundel, H. L., C. F. Nodine and D. Carmody (1978). Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 13(3) 175-181.

Lauby-Secretan, B., D. Loomis and K. Straif (2015). Breast-Cancer Screening--Viewpoint of the IARC Working Group. *N Engl J Med* 373(15) 1479.

Lee, L., V. Stickland and R. Wilson (2003). *Fundamentals of mammography*. 2nd ed. ed. Edinburgh: Churchill Livingstone.

Long, S. M., et al. (2010). *Handbook of Mammography*: Mammography Consulting Services.

Løberg, M., et al. (2015). Benefits and harms of mammography screening. *Breast Cancer Res* 17 63.

Ma, W. K., et al. (2014). Extra patient movement during mammographic imaging: an experimental study. *Br J Radiol* 87(1044) 20140241.

Ma, W. K., et al. (2015). A method to investigate image blurring due to mammography machine compression paddle movement. *Radiography* 21(1) 36-41.

Ma, W. K., et al. (2016). Analysis of motion during the breast clamping phase of mammography. *Br J Radiol* 89(1059) 20150715.

Malliori, A., et al. (2014). Breast tomosynthesis with monochromatic beams: a feasibility study using Monte Carlo simulations. *Phys Med Biol* 59(16) 4681-4696.

Manning, D. J., S. C. Ethell and T. Donovan (2004). Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol* 77(915) 231-235.

Markey, M. (2014). *Physics of Mammographic Imaging*.

Martin, K. E., et al. (2006). Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. *Radiology* 240(3) 656-665.

Mawdsley, G. E., et al. (2009). Accurate estimation of compressed breast thickness in mammography. *Med Phys* 36(2) 577-586.

McCormack, V. A. and I. dos Santos Silva (2006). Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 15(6) 1159-1169.

McPherson, K., C. M. Steel and J. M. Dixon (2000). ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ* 321(7261) 624-628.

Mehrgou, A. and M. Akouchekian (2016). The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. *Med J Islam Repub Iran* 30 369.

Mello-Thoms, C. (2006). The problem of image interpretation in mammography: effects of lesion conspicuity on the visual search strategy of radiologists. *Br J Radiol* 79 Spec No 2 S111-116.

Mello-Thoms, C., et al. (2005). Effects of lesion conspicuity on visual search in mammogram reading. *Acad Radiol* 12(7) 830-840.

Melnikow, J., et al. (2016). Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. *Ann Intern Med* 164(4) 268-278.

Mercer, C. E., et al. (2013a). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography* 19(4) 363-365.

Mercer, C. E., et al. (2013b). Practitioner compression force variability in mammography: a preliminary study. *Br J Radiol* 86(1022) 20110596.

Mercer, C. E., et al. (2013c). Practitioner compression force variation in mammography: A 6-year study. *Radiography* 19(3) 200-206.

Mercer, C. E., et al. (2015). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography* 21(1) 68-73.

Michaelson, J., et al. (2003). Estimates of the Sizes at Which Breast Cancers Become Detectable on Mammographic and Clinical Grounds. *Journal of Women's Imaging* 5(1) 3-10.

Moss, S. M., et al. (2015). Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. *Lancet Oncol* 16(9) 1123-1132.

National Quality Assurance Coordinating Group for Radiology (2006). Quality Assurance Guidelines for Mammography. NHS Cancer Screening Programmes.

Ng, K. and M. Muttarak (2003). Advances in mammography have improved early detection of breast cancer. *JOURNAL-HONG KONG COLLEGE OF RADIOLOGISTS* 6 126-131.

Ng, K. H. and S. Lau (2015). Vision 20/20: Mammographic breast density and its clinical applications. *Med Phys* 42(12) 7059-7077.

Nightingale, J. M., et al. (2015). Breast compression – An exploration of problem solving and decision-making in mammography. *Radiography* 21(4) 364-369.

O'Leary, D. and Z. Al Maskari (2013). PB.15: Pain in mammography: where and why does it arise? *Breast Cancer Research* 15(1) P15.

Olson, J. E., et al. (2012). The influence of mammogram acquisition on the mammographic density and breast cancer association in the mayo mammography health study cohort. *Breast Cancer Research : BCR* 14(6) R147.

Pachoud, M., et al. (2004). A new test phantom with different breast tissue compositions for image quality assessment in conventional and digital mammography. *Phys Med Biol* 49(23) 5267-5281.

Perry, N., et al. (2008). European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. *Ann Oncol* 19(4) 614-622.

Peyton, K. (2016). Prior Mammogram Access Important as Screening Protocols in Flux | Diagnostic Imaging.

Pharoah, P. D., et al. (1997). Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer* 71(5) 800-809.

Pisano, E. D., et al. (2005). Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 353(17) 1773-1783.

Plöckinger, U. (2012). Diagnosis and Treatment of Gastrinomas in Multiple Endocrine Neoplasia Type 1 (MEN-1). *Cancers (Basel)* 4(1) 39-54.

Popli, M. B., et al. (2014). Breast Positioning during Mammography: Mistakes to be Avoided. *Breast Cancer (Auckl)* 8 119-124.

Poulos, A. and D. McLean (2004). The application of breast compression in mammography: a new perspective. *Radiography* 10(2) 131-137.

Poulos, A., et al. (2003). Breast compression in mammography: how much is enough? *Australas Radiol* 47(2) 121-126.

Quante, A. S., et al. (2012). Breast cancer risk assessment across the risk continuum: genetic and nongenetic risk factors contributing to differential model performance. *Breast Cancer Res* 14(6) R144.

Rauscher, G. H., et al. (2013). Mammogram image quality as a potential contributor to disparities in breast cancer stage at diagnosis: an observational study. *BMC Cancer* 13 208.

Ray, K. M., E. R. Price and B. N. Joe (2015). Breast density legislation: mandatory disclosure to patients, alternative screening, billing, reimbursement. *AJR Am J Roentgenol* 204(2) 257-260.

Reis, C., et al. (2015). Optimisation of paediatrics computed radiography for full spine curvature measurements using a phantom: a pilot study.

Revesz, G., H. L. Kundel and M. A. Graber (1974). The influence of structured noise on the detection of radiologic abnormalities. *Invest Radiol* 9(6) 479-486.

Rice, M. S., B. A. Rosner and R. M. Tamimi (2017). Percent mammographic density prediction: development of a model in the nurses' health studies. *Cancer Causes Control*.

Samei, E. (2003). Performance of Digital Radiographic Detectors: Quantification and Assessment Methods. Advances in Digital Radiography: RSNA Categorical Course in Diagnostic Radiology Physics.

Sandberg, M. E., et al. (2013). Change of mammographic density predicts the risk of contralateral breast cancer--a case-control study. *Breast Cancer Res* 15(4) R57.

Saunders, R. S. and E. Samei (2008). The effect of breast compression on mass conspicuity in digital mammography. *Med Phys* 35(10) 4464-4473.

Sechopoulos, I., et al. (2012). Characterization of the homogeneous tissue mixture approximation in breast imaging dosimetry. *Med Phys* 39(8) 5050-5059.

Sharp, P. C., et al. (2003). Reported pain following mammography screening. *Arch Intern Med* 163(7) 833-836.

Sherratt, M. J., J. C. McConnell and C. H. Streuli (2016). Raised mammographic density: causative mechanisms and biological consequences. *Breast Cancer Res* 18(1) 45.

Shin, H. J., H. H. Kim and J. H. Cha (2015). Current status of automated breast ultrasonography. *Ultrasonography* 34(3) 165-172.

Smith, H., et al. (2015). Does elevating image receptor increase breast receptor footprint and improve pressure balance? *Radiography* 21(4) 359–363.

Song, J. L., et al. (2017). The association between prognosis of breast cancer and first-degree family history of breast or ovarian cancer: a systematic review and meta-analysis. *Fam Cancer*.

Szczepura, K., T. Faqir and D. Manning (2017). Effects of increased compression with an ultrasound transducer on the conspicuity of breast lesions in a phantom. SPIE Medical Imaging, SPIE.

Szczepura, K. R. and D. J. Manning (2016). Validated novel software to measure the conspicuity index of lesions in DICOM images.

Tadano, R., A. K. Pediredla and A. Veeraraghavan (2015). Depth Selective Camera: A Direct, On-Chip, Programmable Technique for Depth Selectivity in Photography, IEEE Computer Society.

Taplin, S. H., et al. (2002). Screening mammography: clinical image quality and the risk of interval breast cancer. *AJR Am J Roentgenol* 178(4) 797-803.

Tice, J. A., et al. (2008). Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 148(5) 337-347.

Torre, L. A., et al. (2015). Global cancer statistics, 2012. *CA Cancer J Clin* 65(2) 87-108.

Tyson, A. H., G. E. Mawdsley and M. J. Yaffe (2009). Measurement of compressed breast thickness by optical stereoscopic photogrammetry. *Med Phys* 36(2) 569-576.

Vachon, C. M., et al. (2007). Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Res* 9(6) 217.

Waade, G. G., et al. (2016). Impact of errors in recorded compressed breast thickness measurements on volumetric density classification using volpara v1.5.0 software. *Med Phys* 43(6) 2870.

Waade, G. G., et al. (2017). Compression forces used in the Norwegian Breast Cancer Screening Program. *The British Journal of Radiology* 90(1071) 20160770.

Waade, G. G., A. Sanderud and S. Hofvind (2017). Compression force and radiation dose in the Norwegian Breast Cancer Screening Program. *European Journal of Radiology* 88 41-46.

Wang, J., et al. (2017). Automatic Estimation of Volumetric Breast Density Using Artificial Neural Network-Based Calibration of Full-Field Digital Mammography: Feasibility on Japanese Women With and Without Breast Cancer. *J Digit Imaging* 30(2) 215-227.

Wang, X. H., et al. (2003). Automated assessment of the composition of breast tissue revealed on tissue-thickness-corrected mammography. *AJR Am J Roentgenol* 180(1) 257-262.

Williams, S., et al. (2014). Breast tissue bulge and lesion visibility during stereotactic biopsy–A phantom study. *Radiography* 20(3) 271-276.

Yaffe, M. J., et al. (2009). 1. Introduction. *Journal of the ICRU* 9(2) 7.

Authors Full Publication List

2005

Quantification of disease activity in patients undergoing leucocyte scintigraphy for suspected inflammatory bowel disease, 2005, H.K. *Cheow, D.D. Voutnis, J.W. Evans,* ***K.R. Szczepura****, E.A. Swift, N.J. Bird, P. Ruparelia, C.K. Solanki, J.R. Ballinger, E.R. Chilvers, S.J. Middleton, A.M. Peters* European Journal of Nuclear Medicine and Molecular Imaging (32, 3) DOI: 10.1007/s00259-004-1617-7

2007

Healthy passive cigarette smokers have increased pulmonary alveolar permeability, 2007, *C. Beadsmoore, H.K. Cheow,* ***K. Szczepura****, P. Ruparelia, A. Michael Peters,* Nuclear Medicine Communications (28, 2) DOI: 10.1097/MNM.0b013e328013eb1e

2008

Pulmonary elimination rate of inhaled 99mTc-sestamibi radioaerosol is delayed in healthy cigarette smokers, 2008, *P. Ruparelia, H.K. Cheow, J.W. Evans, L. Banney, S. Shankar,* ***K.R. Szczepura****, A.E. Swift, J.R. Ballinger, N.G. Hartman, E.R. Chilvers, A.M. Peters*. British Journal of Clinical Pharmacology, (65, 4) DOI: 10.1111/j.1365-2125.2008.03099.x

2009

Does P-glycoprotein have a role in the lung clearances of inhaled 99mTc-sestamibi and 99mTc-tetrofosmin?, 2009, *H.K. Cheow, P. Ruparelia, S. Shankar,* ***K.R. Szczepura****, J.R. Ballinger, N.G. Hartman, E.R. Chilvers, A.M. Peters,* Nuclear Medicine Communications (30, 8) DOI: 10.1097/MNM.0b013e32832b9a2d

99mTechnetium-labelled neutrophil scanning in pneumonia, 2009, *P. Ruparelia,* ***K. Szczepura,*** *C. Summers, A.M. Peters, E.R. Chilvers*, Thorax (64, 1) DOI: 10.1136/thx.2007.089573

2011

Measuring whole-body neutrophil redistribution using a dedicated whole-body counter and ultra-low doses of ^{111}Indium, 2011, ***K.R. Szczepura****, P. Ruparelia, C.K. Solanki, K. Balan, P. Newbold, C. Summers, E.R. Chilvers, A.M. Peters*, European Journal of Clinical Investigation (41, 1) DOI: 10.1111/j.1365-2362.2010.02382.x

Quantification of neutrophil migration into the lungs of patients with chronic obstructive pulmonary disease, 2011, *P. Ruparelia,* ***K.R. Szczepura****, C. Summers, C.K. Solanki, K. Balan, P. Newbold, D. Bilton,*

A.M. Peters, E.R. Chilvers, European Journal of Nuclear Medicine and Molecular Imaging (38, 5) DOI: 10.1007/s00259-010-1715-7

2012

The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units, 2012, *I.H.R. Hauge, P. Hogg, **K. Szczepura**, P. Connolly, G. McGill, C. Mercer*, Medical Physics (39, 1) DOI: 10.1118/1.3663579

ROCView: Prototype software for data collection in jackknife alternative free-response receiver operating characteristic analysis, 2012*, J. Thompson, P. Hogg, S. Thompson, D. Manning, **K. Szczepura***, British Journal of Radiology (85, 1017) DOI: 10.1259/bjr/99497945

Analysis of CT acquisition parameters suitable for use in SPECT/CT: A free-response receiver operating characteristic study, 2012, *J. Thompson, P. Hogg, **K. Szczepura**, D. Manning*, Radiography (18, 4) DOI: 10.1016/j.radi.2012.05.004

Analysis of CT acquisition parameters suitable for use in SPECT/CT: A free-response receiver operating characteristic study, 2012, *J. Thompson, P. Hogg, **K. Szczepura**, D. Manning*, Radiography (18, 4) DOI: 10.1016/j.radi.2012.05.004

Blurred digital mammography images, 2012, *P. Hogg, **K. Szczepura**, J. Kelly, M. Taylor,* Radiography (18, 1) DOI: 10.1016/j.radi.2011.11.008

2013

Pressure and breast thickness in mammography - An exploratory calibration study, 2013, *P. Hogg, M. Taylor, **K. Szczepura**, C. Mercer, E. Denton*, British Journal of Radiology (86, 1021) DOI: 10.1259/bjr.20120222

Pressure and breast thickness in mammography - What about physics? Author reply, 2013, *P. Hogg, M. Taylor, **K. Szczepura**, C. Mercer, E. Denton*, British Journal of Radiology (86, 1027) DOI: 10.1259/bjr.20130267

A method to measure paddle and detector pressures and footprints in mammography, 2013, *P. Hogg, **K. Szczepura**, A. Darlington, A. Maxwell*, Medical Physics (40, 4) DOI: 10.1118/1.4792720

Practitioner compression force variation in mammography: A 6-year study, 2013, *C.E. Mercer, P. Hogg, **K. Szczepura**, E.R.E. Denton,* Radiography (19, 3) DOI: 10.1016/j.radi.2013.06.001

Optimising the number of thermoluminescent dosimeters required for the measurement of effective dose for computed tomography attenuation correction data in SPECT/CT myocardial perfusion imaging, 2013, *A.K. Tootell, **K.R. Szczepura**, P. Hogg,* Radiography (19, 1) DOI: 10.1016/j.radi.2012.11.001

Fact or fiction: An analysis of the 10kVp 'rule' in computed radiography, 2013, *E. Allen, P. Hogg, W.K. Ma, **K. Szczepura***, Radiography (19, 3) DOI: 10.1016/j.radi.2013.05.003

Tissue bulge during stereotactic core biopsy, 2013*, L. Hackney, S. Williams, P. Hogg, **K. Szczepura,*** Radiography (19, 4) DOI: 10.1016/j.radi.2013.06.007

2014

A free-response evaluation determining value in the computed tomography attenuation correction image for revealing pulmonary incidental findings. A phantom study, 2014, J.D. Thompson, P. Hogg, D.J. Manning, K. Szczepura, D.P. Chakraborty, Academic Radiology (21, 4) DOI: 10.1016/j.acra.2014.01.003

An overview of measuring and modelling dose and risk from ionising radiation for medical exposures, 2014, *A. Tootell, **K. Szczepura**, P. Hogg*, Radiography (20, 4) DOI: 10.1016/j.radi.2014.05.002

Comparison of effective dose and lifetime risk of cancer incidence of CT attenuation correction acquisitions and radiopharmaceutical administration for myocardial perfusion imaging, 2014, *A.K. Tootell, **K. Szczepura**, P. Hogg,* British Journal of Radiology (87, 1041) DOI: 10.1259/bjr.20140110

Breast tissue bulge and lesion visibility during stereotactic biopsy - A phantom study, 2014, S. *Williams, L. Hackney, P. Hogg, **K. Szczepura***, Radiography (20, 3) DOI: 10.1016/j.radi.2014.04.006

2015

A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, 2015, *C.E. Mercer, **K. Szczepura,** J. Kelly, S.R. Millington, E.R.E. Denton, R. Borgen, B. Hilton, P. Hogg*, Radiography (21, 1) DOI: 10.1016/j.radi.2014.07.004

A phantom-based JAFROC observer study of two CT reconstruction methods: The search for optimisation of lesion detection and effective dose, 2015, *J.D. Thompson, D.P. Chakraborty, **K. Szczepura**, I. Vamvakas, A. Tootell, D.J. Manning, P. Hogg,* Progress in Biomedical Optics and Imaging - Proceedings of SPIE (9416), DOI: 10.1117/12.2081632

Does elevating image receptor increase breast receptor footprint and improve pressure balance?, 2015, *H. Smith,* ***K. Szczepura****, C. Mercer, A. Maxwell, P. Hogg,* Radiography (21, 4) DOI: 10.1016/j.radi.2015.02.001

An audit to investigate the impact of false positive breast screening results and diagnostic work-up on re-engagement with subsequent routine screening, 2015, *J.M. Nightingale, R. Borgen, L. Porter-Bennett,* ***K. Szczepura,*** Radiography (21, 1) DOI: 10.1016/j.radi.2014.05.005

2016

Effect of reconstruction methods and x-ray tube current-time product on nodule detection in an anthropomorphic thorax phantom: A crossed-modality JAFROC observer study, 2016, *J.D. Thompson, D.P. Chakraborty,* ***K. Szczepura,*** *A.K. Tootell, I. Vamvakas, D.J. Manning, P. Hogg,* Medical Physics (43, 3) DOI: 10.1118/1.4941017

A JAFROC study of nodule detection performance in CT images of a thorax acquired during PET/CT, 2016, *J.D. Thompson, A. Wareing,* ***K.R. Szczepura****, S. Vinjamuri, P. Hogg,* Radiography DOI: 10.1016/j.radi.2017.03.001

Effective Dose and Effective Risk from Post-Single Photon Emission Computed Tomography Imaging of the Lumbar Spine, 2016, *A. Tootell, M. McEntee,* ***K. Szczepura****, P. Hogg*, Journal of Medical Imaging and Radiation Sciences, DOI: 10.1016/j.jmir.2016.04.012

Validated novel software to measure the conspicuity index of lesions in DICOM images, 2016, ***K.R. Szczepura****, D.J. Manning*, Progress in Biomedical Optics and Imaging - Proceedings of SPIE (9787) DOI: 10.1117/12.2216211

2017

Analysis of effective and organ dose estimation in CT when using mA modulation: A single scanner pilot study, 2017, *A. Tootell,* ***K. Szczepura****, P. Hogg*, Radiography, (23, 2) DOI: 10.1016/j.radi.2017.02.006

Effects of increased compression with an ultrasound transducer on the conspicuity of breast lesions in a phantom, 2017, ***K. Szczepura****, T. Faqir, D. Manning*, Progress in Biomedical Optics and Imaging - Proceedings of SPIE (10136) DOI: 10.1117/12.2254263

Hounsfield unit inaccuracy in computed tomography lesion size and density, diagnostic quality vs attenuation correction, 2017, ***K. Szczepura****, J. Thompson, D. Manning*, Progress in Biomedical Optics and Imaging - Proceedings of SPIE, (10136) DOI: 10.1117/12.2254275

Impact of tube current modulation on lesion conspicuity index in hi-resolution chest computed tomography, 2017, ***K. Szczepura****, D. Tomkinson, D. Manning*, Progress in Biomedical Optics and Imaging - Proceedings of SPIE (10136) DOI: 10.1117/12.2254256

A JAFROC study of nodule detection performance in CT images of a thorax acquired during PET/CT *John David Thompson, Amy Wareing,* ***Katy Szczepura****, Peter Hogg* March 2017 Radiography 23 (3), DOI: 10.1016/j.radi.2017.03.001

An Investigation of Pressure Ulcer Risk, Comfort, and Pain in Medical Imaging Seth K. Angmorterh, *Andrew England, Jo Webb,* ***Katy Szczepura****, Melanie Stephens, Judith Anaman-Torgbor, Eric K. Ofori, Peter Hogg,* August 2018, Journal of Medical Imaging and Radiation Sciences, DOI: 10.1016/j.jmir.2018.07.003

Appendix 1
Personal contribution summary

Paper 1 (Hauge et al., 2012)

This was my first research paper within mammography and, as such, I had a smaller contribution in comparison to the other included papers. Within this research I was involved in the design of the measuring device and developing the analysis and data display in collaboration with the first author. Within this project I was involved in all research discussions, which formed the research project focus, the methodological design, data analysis and article writing. This research was key to the progression of the next pieces of research and involved me collaborating with a new team and building key relationships.

Paper 2 (Hackney et al., 2013)

This paper arose due to an issue that the clinical co-authors (Sue Williams and Lisa Hackney) observed in practice. They came to the research team at the University of Salford for advice on how to approach the research. From this point I led the research project.

Within the work I designed the data collection method, and once the data was acquired by qualified mammographers I performed all analysis. Once analysis was performed I met with the co-authors to discuss and agree the structure of the paper, I led these discussions as the main focus of the work was to inform the phantom design for the subsequent paper (Williams et al., 2014). My contribution to the presented work was the results section and the discussion. I also contributed to all other sections of the paper.

Paper 3 (Williams et al., 2014)

This was a follow-on phantom project based on the previous work to demonstrate the impact on lesion visibility from the issue of the extra breast tissue.

Again, I led this work. My previous experience in Radiation Protection meant I had experience in phantom design (Smith et al., 2011) and a knowledge of the requirements for phantom construction (Szczepura et al., 2006).

I used this experience to design and construct the phantom to ensure a reproducible and reliable research tool. I designed the method, and data collection was performed in the clinical setting by Sue Williams and Lisa Hackney (co-authors).

Once the data was collected I performed all the analysis. This was a more complex task as there was the combination of the standard number of discs that could be seen, as well as 2AFC visual grading. I combined these data sets by creating the “visibility score” within the work, this was helped by my previous work in 2AFC and visual grading (Thompson et al., 2009, Thompson et al., 2010, Thompson et al., 2012a, Thompson et al., 2012c, Thompson et al., 2012b, Thompson et al., 2014, Thompson et al., 2015, Thompson et al., 2016). Once this was performed it was essential to indicate the increased risk due to the inaccurate function of the AECs as well as the reduction in visibility score. My previous experience in radiation protection and dosimetry (Tootell et al., 2013, Tootell et al., 2014b, Tootell et al., 2014a) led me to create the “optimisation score” where optimisation is a legal term used in IR(ME)R regulations to define the balance between image quality and dose. Following on from the creation of these two novel metrics I again led the discussions of the structure of the paper, and wrote the results, analysis, discussion and conclusion sections of the paper, and contributed to all other sections.

Paper 4 (Mercer et al., 2013)

My main involvement in this work came after the initial data had been collected, and the team were struggling to analyse it in the way they needed to.

I quickly realised that the way the data was being organised meant that the task was going to be laborious and complex, and that by designing a more appropriate data analysis technique this task could be simplified.

Instead of using Excel, I used my experience in Access and Visual Basic to design a database that meant the data could be interrogated in multiple ways by the use of Queries and appropriate coding.
I designed the Access database so that the data could be exported into Excel with the click of a button on a form, depending on the required analysis (i.e. by patient or by operator) this enabled a quick and easy way for Claire Mercer (first author) to access the data in a structured, safe and controlled way.
Once the data was exported I worked with Claire Mercer on the analysis and discussed the appropriate interpretation of the data with the team. Once analysis was completed I contributed to all sections of the paper.

Paper 5 (Mercer et al., 2015)

This work was similar to the work in paper 4, but it was a multi-centre study to see if the same practitioner behaviour occurred at different sites. Within this paper I was key to the discussions around data collection and methodology from the beginning and was key to inform the research at all stages.
As this study was multi-centre, this led to a more complex data collection task, as the data was being collected by various people at various clinical sites. Therefore I adapted the Access database designed in (Mercer et al., 2013) to ensure accuracy and control of the data input as well as the data output.
I created an input form to ensure consistency of all variables, especially practitioner name as this was an important part of the research, and we needed to ensure that the input was consistent as well as controlled from a data protection point of view.

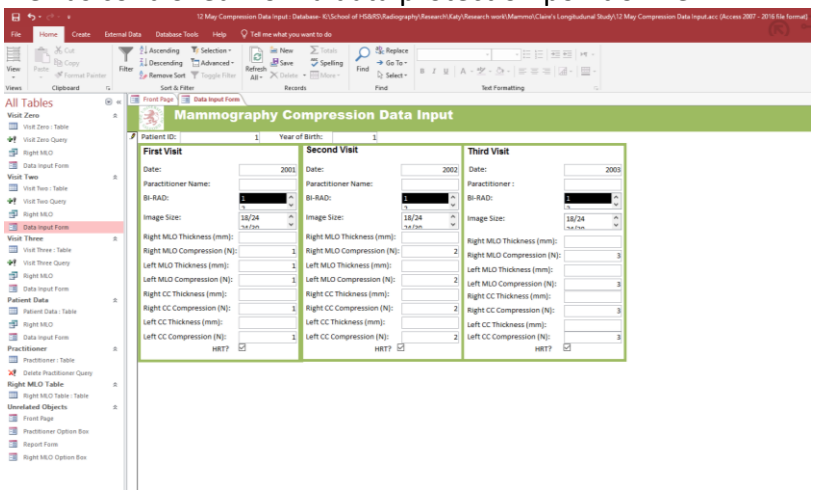


Figure 1: Data input page example, demonstrating the number of Forms, Tables and Queries required to interrogate the data appropriately.
Once the data was exported I collaborated very closely with Claire Mercer to guide the analysis and directed the appropriate interpretation of the data with the team and contributed to all sections of the journal article.

Papers 6 and 7 (Hogg et al., 2013a) (Smith et al., 2015)

These papers were a phantom study and follow on human participant study that demonstrated the issue of positioning in mammography; how this impacts on the effectiveness of the applied compression force. My role and input in both studies was very similar, and so have been combined for the purposes of this appendix.
Within this project I acquired the pilot data with Peter Hogg and Alison Darlington, and from the pilot data I designed the main method and wrote an instruction manual for the data acquisition using the pressure mat to enable reliable and consistent data acquisition. It is noted that this

instruction manual has now been used to support other studies with pressure mat use (Angmorterh et al., 2019, Mercer et al., 2016)
Once the data had been acquired I performed all analysis. To answer the research question it was required to indicate the balance between the pressure on top and below the breast. To do this I used my previous nuclear medicine experience (Cheow et al., 2005, Ruparelia et al., 2006, Szczepura et al., 2006, Beadsmoore et al., 2006, Beadsmoore et al., 2007, Ruparelia et al., 2008a, Cheow et al., 2009, Ruparelia et al., 2009b, Szczepura et al., 2011a, Ruparelia et al., 2011) to develop and define the "Uniformity Index" which was based on the uniformity assessment performed on gamma cameras in nuclear medicine. Although the physical premise is very different, the data in terms of RGB were similar and so the application of this technique worked to demonstrate the balance of force above and below the breast.
Once all analysis was completed I wrote the method, analysis and results section and contributed to all other sections of these papers

Paper 8 (Hogg et al., 2013b)

Within this work I developed the method along with the co-authors. The data was acquired by qualified mammographers, and once the data was acquired I performed all the analysis and defined the 3 "colour coded" sections on the graphs after discussion with Peter Hogg.
Once all data was acquired and analysed, I wrote the results section of the paper and contributed to all other aspect of the paper. Once the paper was published there were letters to the Editor, and Peter Hogg and I responded to those letters on behalf of the co-authors at the questions were focused on the physics terminology within the paper.

Papers 9 and 10 (Szczepura and Manning, 2016, Szczepura et al., 2017a)

Within this work my contribution was similar, so they have been combined for the purposes of this appendix.
This work was following on from work that Manning et al had done on analog images in 2002. (Manning et al., 2004) on discussion with David Manning I proposed recreating this work due to the introduction of digital imaging, and the possibility of performing the analysis with software rather than manually as with the previous work.
My previous experience of image analysis and software design (Szczepura et al., 2006, Szczepura et al., 2011b, Ruparelia et al., 2006, Ruparelia et al., 2008b, Ruparelia et al., 2009a, Thompson et al., 2010) enabled me to develop this work.
The initial stages of the work were to design and develop the software. At each iteration I performed quality control and validation studies. Version 8 of the software satisfied all the requirements for image analysis and so a full validation study was performed (Szczepura and Manning, 2016). For this work I performed all data collection and analysis. For the paper I wrote the first draft of the paper, David Manning then provided input and edited the paper, I then finalised and submitted the paper. I also presented the work at the SPIE conference in San Diego in 2016.
Paper 10 then used the software in a different imaging modality, ultrasound breast imaging. This was a phantom study that was to demonstrate the issue with over compression of tissue breast ultrasound. The idea arose I am module lead for the Physics of Ultrasound module where I teach breast sonographers, and we were discussing imaging techniques and how if you "press too hard" the lesion becomes less visible, so I wanted to quantify this clinical observation.
Within this work I designed the method, the phantom, the data acquisition, and performed the data analysis. For the paper I wrote the first draft of the paper, then received feedback from the co-authors, finalised and submitted the paper. I also presented the work at the SPIE conference in Orlando in 2017, where I also presented additional work using this software to assess lesion visibility in CT with mA modulation (Szczepura et al., 2017b).

References

ANGMORTERH, S. K., ENGLAND, A., WEBB, J., SZCZEPURA, K., STEPHENS, M., ANAMAN-TORGBOR, J., OFORI, E. K. & HOGG, P. 2019. An Investigation of Pressure Ulcer Risk, Comfort, and Pain in Medical Imaging. *Journal of Medical Imaging and Radiation Sciences,* 50**,** 43-52.

BEADSMOORE, C., CHEOW, H., SZCZEPURA, K., RUPARELIA, P. & PETERS, A. 2006. P16 Healthy passive cigarette smokers have increased pulmonary alveolar permeability. *Nuclear Medicine Communications,* 27**,** 1027-1028.

BEADSMOORE, C., CHEOW, H. K., SZCZEPURA, K., RUPARELIA, P. & PETERS, A. M. 2007. Healthy passive cigarette smokers have increased pulmonary alveolar permeability. *Nucl Med Commun.* England.

CHEOW, H., VOUTNIS, D., EVANS, J., SZCZEPURA, K., SWIFT, E., BIRD, N., RUPARELIA, P., SOLANKI, C., BALLINGER, J., CHILVERS, E., MIDDLETON, S. & PETERS, A. 2005. Quantification of disease activity in patients undergoing leucocyte scintigraphy for suspected inflammatory bowel disease. *European Journal of Nuclear Medicine and Molecular Imaging,* 32**,** 329-337.

CHEOW, H. K., RUPARELIA, P., SHANKAR, S., SZCZEPURA, K., BALLINGER, J., HARTMAN, N., CHILVERS, E. & PETERS, A. M. 2009. Does P-glycoprotein have a role in the lung clearances of inhaled Tc-99m-sestamibi and Tc-99m-tetrofosmin? *Nuclear Medicine Communications,* 30**,** 617-621.

HACKNEY, L., WILLIAMS, S., HOGG, P. & SZCZEPURA, K. 2013. Tissue bulge during stereotactic core biopsy. *Radiography,* 19**,** 366-368.

HAUGE, I. H., HOGG, P., SZCZEPURA, K., CONNOLLY, P., MCGILL, G. & MERCER, C. 2012. The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Med Phys,* 39**,** 263-71.

HOGG, P., SZCZEPURA, K., DARLINGTON, A. & MAXWELL, A. 2013a. A method to measure paddle and detector pressures and footprints in mammography. *Med Phys,* 40**,** 041907.

HOGG, P., TAYLOR, M., SZCZEPURA, K., MERCER, C. & DENTON, E. 2013b. Pressure and breast thickness in mammography--an exploratory calibration study. *Br J Radiol.* England.

MANNING, D. J., ETHELL, S. C. & DONOVAN, T. 2004. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol,* 77**,** 231-5.

MERCER, C. E., HOGG, P., SZCZEPURA, K. & DENTON, E. 2013. Practitioner compression force variation in mammography: A 6-year study. *Radiography,* 19**,** 200-206.

MERCER, C. E., REIS, C., SZCZEPURA, K. & HOGG, P. 2016. Pressure based mammography compression: a feasibility study to determine operational level(s). *Symposium Mammographicum.* Liverpool.

MERCER, C. E., SZCZEPURA, K., KELLY, J., MILLINGTON, S., DENTON, E., BORGEN, R., HILTON, B. & HOGG, P. 2015. A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography,* 21**,** 68-73.

RUPARELIA, P., CHEOW, H. K., EVANS, J., BANNEY, L., SHANKAR, S., SZCZEPURA, K., SWIFT, A., BALLINGER, J., HARTMAN, N., CHILVERS, E. & PETERS, A. M. 2008a. Pulmonary elimination rate of inhaled Tc-99m-sestamibi radioaerosol is delayed in healthy cigarette smokers. *British Journal of Clinical Pharmacology,* 65**,** 611-614.

RUPARELIA, P., CHEOW, H. K., EVANS, J. W., BANNEY, L., SHANKAR, S., SZCZEPURA, K. R., SWIFT, A. E., BALLINGER, J. R., HARTMAN, N. G., CHILVERS, E. R. & PETERS, A. M. 2008b. Pulmonary elimination rate of inhaled 99mTc-sestamibi radioaerosol is delayed in healthy cigarette smokers. *Br J Clin Pharmacol.* England.

RUPARELIA, P., SZCZEPURA, K., BILTON, D., PETERS, A. M. & CHILVERS, E. 2006. Granulocyte trafficking in chronic obstructive pulmonary disease. *Thorax,* 61**,** II108-II109.

RUPARELIA, P., SZCZEPURA, K., SUMMERS, C., PETERS, A. & CHILVERS, E. 2009a. 99mTechnetium-labelled neutrophil scanning in pneumonia. *Thorax,* 64**,** 92-92.

RUPARELIA, P., SZCZEPURA, K., SUMMERS, C., PETERS, A. M. & CHILVERS, E. R. 2009b. 99mTechnetium-labelled neutrophil scanning in pneumonia. *Thorax.* England.

RUPARELIA, P., SZCZEPURA, K., SUMMERS, C., SOLANKI, C., BALAN, K., NEWBOLD, P., BILTON, D., PETERS, A. M. & CHILVERS, E. 2011. Quantification of neutrophil migration into the lungs of patients with chronic obstructive pulmonary disease. *Eur J Nucl Med Mol Imaging,* 38**,** 911-9.

SMITH, H., SMITH, J., HOGG, P., MERCER, C. & SZCZEPURA, K. 2011. Elastically deformable anthropomorphic breast phantom foruse in mammographic imaging research. *UK Radiological Congress.* Manchester, UK.

SMITH, H., SZCZEPURA, K., MERCER, C., MAXWELL, A. & HOGG, P. 2015. Does elevating image receptor increase breast receptor footprint and improve pressure balance? *Radiography,* 21**,** 359–363.

SZCZEPURA, K., FAQIR, T. & MANNING, D. Effects of increased compression with an ultrasound transducer on the conspicuity of breast lesions in a phantom. SPIE Medical Imaging, 2017a. SPIE, 11.

SZCZEPURA, K., RUPARELIA, P., BILTON, D., CHILVERS, E. & PETERS, A. M. 2006. A16 Using whole-body counting to determine neutrophil trafficking and loss in COPD. *Nuclear Medicine Communications,* 27**,** 285.

SZCZEPURA, K., RUPARELIA, P., SOLANKI, C., BALAN, K., NEWBOLD, P., SUMMERS, C., CHILVERS, E. & PETERS, A. M. 2011a. Measuring whole-body neutrophil redistribution using a dedicated whole-body counter and ultra-low doses of 111Indium. *European journal of clinical investigation,* 41**,** 77-83.

SZCZEPURA, K., TOMKINSON, D. & MANNING, D. 2017b. *Impact of tube current modulation on lesion conspicuity index in hi-resolution chest computed tomography*, SPIE.

SZCZEPURA, K. R. & MANNING, D. J. 2016. Validated novel software to measure the conspicuity index of lesions in DICOM images. 9787**,** 978703-978703-15.

SZCZEPURA, K. R., RUPARELIA, P., SOLANKI, C. K., BALAN, K., NEWBOLD, P., SUMMERS, C., CHILVERS, E. R. & PETERS, A. M. 2011b. Measuring whole-body neutrophil redistribution using a dedicated whole-body counter and ultra-low doses of 111Indium. *Eur J Clin Invest,* 41**,** 77-83.

THOMPSON, J., CHAKRABORTY, D., SZCZEPURA, K., VAMVAKAS, I., TOOTELL, A., MANNING, D. & HOGG, P. A phantom-based JAFROC observer study of two CT reconstruction methods: the search for optimisation of lesion detection and effective dose. SPIE Medical Imaging, 2015. International Society for Optics and Photonics, 94160B-94160B-6.

THOMPSON, J., HOGG, P., SZCZEPURA, K. & MANNING, D. 2012a. Analysis of CT acquisition parameters suitable for use in SPECT/CT: A free-response receiver operating characteristic study. *Radiography,* 18**,** 238-243.

THOMPSON, J., HOGG, P., THOMPSON, S., MANNING, D. & SZCZEPURA, K. 2012b. ROCView: prototype software for data collection in jackknife alternative free-response receiver operating characteristic analysis. *Br J Radiol.* England.

THOMPSON, J., SZCZEPURA, K., MANNING, D. & HOGG, P. 2012c. Lesion detection in the CT attenuation correction image of 5 different low resolution SPECT/CT systems: a multi-centre study. *40th Annual Meeting of the British Nuclear Medicine Society*.

THOMPSON, J., SZCZEPURA, K., TOOTELL, A., SIL, J., MANNING, D. & HOGG, P. 2010. Determination of optimal CT exposure factors for lung lesions using an anthropomorphic chest phantom for SPECT-CT. *European Journal of Nuclear Medicine and Molecular Imaging,* 37**,** S494.

THOMPSON, J., TOOTELL, A., DRIVER, J., GRIFFITHS, M., KANE, T., SZCZEPURA, K., HOLMES, K., MOUNTAIN, V. & HOGG, P. 2009. CPD in Focus-Focussing in on SPECT-CT. *Synergy Representing Radiographers Promoting Radiography***,** 26.

THOMPSON, J. D., CHAKRABORTY, D. P., SZCZEPURA, K., TOOTELL, A. K., VAMVAKAS, I., MANNING, D. J. & HOGG, P. 2016. Effect of reconstruction methods and x-ray tube current-time product on nodule detection in an anthropomorphic thorax phantom: A crossed-modality JAFROC observer study. *Med Phys,* 43**,** 1265-74.

THOMPSON, J. D., HOGG, P., MANNING, D. J., SZCZEPURA, K. & CHAKRABORTY, D. P. 2014. A Free-response Evaluation Determining Value in the Computed Tomography Attenuation Correction Image for Revealing Pulmonary Incidental Findings: A Phantom Study. *Academic Radiology,* 21**,** 538-545.

TOOTELL, A. K., SZCZEPURA, K. & HOGG, P. 2013. Optimising the number of thermoluminescent dosimeters required for the measurement of effective dose for computed tomography attenuation correction data in SPECT/CT myocardial perfusion imaging. *Radiography,* 19**,** 42-47.

TOOTELL, A. K., SZCZEPURA, K. & HOGG, P. 2014a. An overview of measuring and modelling dose and risk from ionising radiation for medical exposures. *Radiography,* 20**,** 323-332.

TOOTELL, A. K., SZCZEPURA, K. & HOGG, P. 2014b. Comparison of effective dose and lifetime risk of cancer incidence of CT attenuation correction acquisitions and radiopharmaceutical administration for myocardial perfusion imaging. *Br J Radiol,* 87**,** 20140110.

WILLIAMS, S., HACKNEY, L., HOGG, P. & SZCZEPURA, K. 2014. Breast tissue bulge and lesion visibility during stereotactic biopsy–A phantom study. *Radiography,* 20**,** 271-276.