**PhD by Published Works**

# Practitioner Variation of Applied Breast Compression Force in Mammography


School of Health Sciences

College of Health and Social Care

University of Salford, Salford, UK

**Claire Elizabeth Mercer**


This thesis is submitted in part fulfilment for the requirement of the Degree of Doctor of Philosophy by Published Works, September 2014

**Roll number/Student number: @00113846**

| Word Count | |
|---|---|
| **Thesis** | 23,043 |
| **Publications** | 26,688 |
| **Total** | 49,731 |

# CONTENTS

# LIST OF TABLES

**Chapter Three**:



**Chapter Six**:

# LIST OF FIGURES

**Chapter Two:**



**Chapter Three**:



**Chapter Four**:



**Chapter Six**:

# TERMINOLOGY

| Term | Definition |
|---|---|
| Analogue | A mammogram image recorded on film |
| Digital | A mammogram image recorded on digital medium |
| Observers | Qualified mammography practitioners or Radiologists who graded images for their BI-RADS density |
| University | University of Salford, Salford, UK |
| Practitioner | Either an assistant practitioner, a qualified mammographer or an advanced practitioner specialising in mammography. |

# LIST OF ABBREVIATIONS

| Acronym | Definition |
|---|---|
| ACR | American College of Radiology |
| AGR | Average glandular dose |
| ANOVA | Analysis of variance |
| BI-RADS | Breast Imaging-Reporting and Data System |
| CC | Cranio-caudal view |
| daN | decaNewtons |
| DoH | Department of Health |
| EU | European Union |
| IQ | Image quality |
| kg | Kilogram |
| kPa | Kilopascal |
| kVp | Kilovolt peak |
| MGD | Mean glandular dose |
| mGy | milligray |
| MLO | Medio-lateral oblique view |
| N | Newtons |
| NHS | National Health Service |
| NHSBSP | National Health Service Breast Screening Programme |
| Pa | Pascal |
| RG | Research Gate |

| | |
|---|---|
| TMD | Thickness measuring device |
| UK | United Kingdom |
| QA | Quality Assurance |

# ACKNOWLEDGMENTS

During my career, I have been encouraged and inspired by some incredibly talented individuals who have wholeheartedly supported me to develop my career; I am indebted to such individuals.

To my husband Miles - I would like to sincerely thank you for your unfailing support, guidance and never ending encouragement throughout my MSc and PhD research. Your continued patience and tolerance have kept me going and I could not have done any of this without you.

To my daughter Johanna - I hope that I am an inspiration to you and I am sorry that my research work has meant that I have not always been there. To my sister, who directed me to the poem 'Don't you quit' when I was a teenager; this has kept me going more than you will ever know. To my parents, for their support and encouragement throughout this whole process.

To my supervisors Professor Peter Hogg, Katy Szczepura and Julie Nightingale, you have been my inspiration and motivation throughout. Peter, for all your support, advice, critique, wisdom and opportunities that you have directed my way, you have unrelentingly supported me throughout my research journey - you truly are an inspiration. Katy, for that incredible 'Access Database' without which none of this research would have been possible, and for your continued friendship and support throughout. Julie, thank you for the final push, support and guidance I required.

To all my friends, peers, colleagues and researchers who have encouraged me throughout this whole process and managed to sustain my motivation during the challenging times of balancing work and research.

# AUTHOR SYNOPSIS

The author works at the University Hospital of South Manchester NHS Foundation Trust as Lead Radiographer and Programme Manager for Breast Imaging and concurrently and is an honorary research fellow at the University of Salford where this PhD research was based.

The research within this thesis commenced in June 2011 immediately following the author's award of MSc with Distinction in Advanced Medical Imaging (Radiology). Research work for this thesis was undertaken part time whilst the author was in full time employment; the research for this thesis completed in September 2014.

Supervision for the research work was directed by Professor Peter Hogg and Katy Szczepura at the University of Salford. Research teams were developed in collaboration with NHS Trust colleagues, University staff and other researchers.

The author of this thesis is passionate about the application of a stringent quality breast service; motivating her team, striving hard, to ensure that staff within her service are driven to promote high quality care.

## THESIS STRUCTURE

To guide the reader and ease navigation through the manuscript the chapters are colour coded on the bottom right corner of the script.

Chapter One

The first chapter of this thesis provides the research abstract. This is a concise overview of the research work, illustrating the research rationale with aims and objectives, together with a concise summary of the key research findings.

Chapter Two

The intellectual ownership and declaration for the published work contained within this thesis is included in this chapter. This chapter also demonstrates the overall contributions the author made to each publication.

Chapter Three

The third chapter provides a contextual background of the thesis. The aims and objectives are discussed in full, along with the research rationale. The rationale is supported by a succinct review of the literature pertaining to the research context.

## Chapter Four

The systematic acquisition and distribution of the acquired and novel knowledge are discussed in the fourth chapter. This chapter illustrates the concepts and design for each research project in turn and discusses the various research teams constructed. This chapter contains a critical review of the published research, together with key research findings and future research work in development.

## Chapter Five

The fifth chapter contains the published research work. This includes eight journal articles, seven of which are peer reviewed, one of which was commissioned for an annual radiography journal of the Society and College of Radiographers.

## Chapter Six

This chapter is based on research development, performance metrics and publication impact. Analysis of these metrics demonstrates the research works originality and impact. This chapter illustrates the development of the researcher and the new research teams and pathways that have arisen as a direct outcome of this thesis. This chapter also provides a succinct summary of the thesis findings and the development of future projects arising from this research.

# CHAPTER ONE: ABSTRACT


**Rationale**

Mammography practitioners control the amount of compression force applied to the breast. There are no quantifiable recommendations for optimal compression force levels for practitioners to follow. Clients report variations in pain and discomfort when compression force is applied. Until now practitioner compression force variability has not been investigated; even though this might lead to variations in client pain and discomfort. The primary purpose of this thesis was to investigate whether practitioner compression force variability exists.

**Method**

Three research papers investigated practitioner compression force variability: one used a cross sectional design; two used longitudinal designs, one was single centre and the other was multicentre. Three further research papers investigated important issues which might confound practitioner variability results: the first investigated compression paddle bend and distortion; the second investigated how breast thickness and compression force vary; the third evaluated practitioner ability to grade breast density, visually. The final research paper was a 'within client' investigation to determine how image quality varied with breast thickness and compression force.

**Key findings**

The research firmly demonstrates that practitioner compression force variability exists. Multicentre analysis (4500 client visits) confirmed two out of three screening

sites with significant practitioner variability, with the third screening site having a minimum dictate of compression force at 100N. As displayed by MLO/CC projections clients underwent a 55%/57% (site one), 66%/60% (site two) and 27%/26% (site three) change in compression force through their three screening visits.

The research confirmed that the compression force received by a client was highly dependent upon the practitioner, and not the client. Within an individual clients screening pathway the research has demonstrated that clients could receive significantly different compression force levels over time.

**Conclusion and further research**

For the first time practitioner compression force variability has been identified. Novel methods for reducing breast thickness need investigating; an example of a novel method is the use of pressure rather than force.

# CHAPTER TWO: DECLARATION AND CONTRIBUTIONS

## 2.1 Declaration and intellectual property rights

I, Claire Elizabeth Mercer, hereby declare that this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes eight original papers, seven published in peer reviewed journals, one commissioned and published in an annual radiography journal. The core theme of the thesis is centred on the original research of practitioner variation in the application of breast compression force in mammography. The ideas, development and script of all the papers in the thesis were the principal responsibility of the author of this thesis, the candidate, working within the University of Salford under the direct supervision of Professor Peter Hogg and Katy Szczepura. The research work was developed from active partnership between the candidate, researchers and academics together with the acknowledgement of input into collaborative based research.

The research within this thesis is based upon the following published papers which will be referred to throughout this thesis by roman numerals. These papers are included in Chapter 5 of this thesis:

**Paper I** Hauge, I.H.R., Hogg, P., Szczepura, K., Connolly, P., & Mercer C.E. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Medical physics*, 39 (1), 263–271. doi:10.1118/1.3663579

**Paper II** Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E.R.E. (2013). Pressure and breast thickness in mammography- an exploratory calibration study. *The British journal of radiology*, 86 (1021), 20120222. doi:10.1259/bjr.20120222

**Paper III** Mercer, C.E., Hogg, P., Lawson, R., Diffey, J., & Denton, E.R.E. (2013). Practitioner compression force variability in mammography: a preliminary study. *The British journal of radiology*, 86 (1022), 20110596. doi:10.1259/bjr.20110596

**Paper IV** Mercer, C.E., Hogg, P., Kelly, J., Borgen, R., Millington, S., Hilton, B. … Whelehan, P. (2014). A mammography image set for research purposes using BI-RADS density classification. *Radiologic technology*, 85 (6), 609–613.

**Paper V** Mercer, C.E., Hogg, P., Szczepura, K., & Denton, E.R.E. (2013). Practitioner compression force variation in mammography: A 6-year study. *Radiography*, 19 (3), 200–206. doi:10.1016/j.radi.2013.06.001

**Paper VI** Mercer, C.E., Hogg, P., Cassidy, S., & Denton, E.R.E. (2013). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography*, 19 (4), 363–365. doi:10.1016/j.radi.2013.07.002

**Paper VII** Hogg, P., Mercer, C., Maxwell, A., Robinson, L., Kelly, J., & Murphy F. (2013). Controversies in compression, *Imaging and oncology*, 28-36, ISBN 9871 871101581

**Paper VIII** Mercer, C.E., Szczepura, K., Kelly, J., Millington, S.R., Hilton, B., & Hogg, P. (2014). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography*, Published online: July 29, 2014. doi:10.1016/j.radi.2014.07.004

| Candidate Name | Signature | Date |
|---|---|---|
| Claire Elizabeth Mercer | | |

## 2.2 Overall contribution and declaration for PhD thesis based on peer reviewed published work

Overall contributions to each published article with individual declarations for each paper are detailed in Figure 2.1.

The extent of the researcher's contribution to each paper is defined in accordance with these descriptors:

- Research concepts and experimental / method design
- Organisational roles
- Data contribution
- Ethics
- Input into research paper

**Paper I**
(page 79)
15%

- Hauge, I.H.R., Hogg, P., Szczepura, K., Connolly, P., & Mercer C.E. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Medical physics*, 39 (1), 263–271. doi:10.1118/1.3663579
- ***Assisted with the experimental design of the breast phantom by contextualising the phantom design by incoperating the use of breast prosthesis***
- ***Active participation in all group project meetings***
- ***Contributed to result evaluation, paper write up and review***

**Paper II**
(page 89)
20%

- Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E.R.E. (2013). Pressure and breast thickness in mammography- an exploratory calibration study. *The British journal of radiology*, 86 (1021), 20120222. doi:10.1259/bjr.20120222
- ***Designed the data gathering tool and performed a trial of its use in department.***
- ***MSc supervisor for the main researcher for this project***
- ***Active participation in all group project meetings***
- ***Contributed to result evaluation, paper write up and review***
- ***Participated in ethics approval***

**Paper III**
(page 97)
65%

- Mercer, C.E., Hogg, P., Lawson, R., Diffey, J., & Denton, E.R.E. (2013). Practitioner compression force variability in mammography: a preliminary study. *The British journal of radiology*, 86 (1022), 20110596. doi:10.1259/bjr.20110596
- ***Responsible for the research group and co-ordination of group project meetings***
- ***Main responsibility for:***
  - ***ethics approvals***
  - ***planning, design and undertaking of the research work***
  - ***data gatheirng, analysis and evaluation of the results***
  - ***writing the paper and co-ordination of co-authors and review stages***

**Paper IV**
(page 107)
65%

- Mercer, C.E., Hogg, P., Kelly, J., Borgen, R., Millington, S., Hilton, B. … Whelehan, P. (2014). A mammography image set for research purposes using BI-RADS density classification. Radiologic technology, 85 (6), 609–613.
- ***Main responsibility for:***
- ***ethics approvals and co-ordination at each location***
- ***planning, design and undertaking of the research work***
- ***planning and implementing of the research group, resources, meetings, presentations with the team***
- ***data gatherng, analysis and evaluation of the results***
- ***writing the paper and co-ordination of co-authors and review stages***

**Paper V**
(page 113)
65%

- Mercer, C.E., Hogg, P., Szczepura, K., & Denton, E.R.E. (2013). Practitioner compression force variation in mammography: A 6-year study. *Radiography*, 19 (3), 200–206. doi:10.1016/j.radi.2013.06.001
- ***Responsible for the co-ordination of group project meetings***
- ***Main responsibility for:***
  - ***ethics approvals***
  - ***the planning, design and undertaking of the research work***
  - ***data gatheirng, analysis and evaluation of the results***
  - ***writing the paper and co-ordination of co-authors and review stage***

**Paper VI**
(page 121)
70%

- Mercer, C.E., Hogg, P., Cassidy, S., & Denton, E.R.E. (2013). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography,* 19 (4), 363–365. doi:10.1016/j.radi.2013.07.002
- ***Main responsibility for:***
  - ***ethics approvals and co-ordination at each location***
  - ***the planning, design and undertaking of the research work***
  - ***the planning and implementing of the research group, resources and feedback meetings/ presentations with the team***
  - ***data gathering, analysis and evaluation of the results***
  - ***writing the paper and co-ordination of co-authors and review stages***

**Paper VII**
(page 125)
25%

- Hogg P, Mercer C, Maxwell A, Robinson L, Kelly J and Murphy F, Controversies in Compression, Imaging and Oncology, 2013, Imaging and Oncology, 28-36, ISBN 9871 871101581
- ***Participated in the planning phase - research to date was a key contributor to this paper***
- ***Assisted in paper write up and review***

**Paper VIII**
(page 133)
65%

- Mercer, C.E., Szczepura, K., Kelly, J., Millington, S.R., Hilton, B., & Hogg, P. (2014). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography*, Published online: July 29, 2014. doi:10.1016/j.radi.2014.07.004
- ***Main responsibility for:***
  - ***ethics approvals and co-ordination at each location***
  - ***the planning, design and undertaking of the research work***
  - ***the planning and implementing of the research group, resources and feedback meetings/ presentations with the team***
  - ***data gatherng, analysis and evaluation of the results***
  - ***writing the paper and co-ordination of co-authors and review stages***

Figure 2.1: Overall contribution to research publications

The undersigned hereby certifies that the above declaration correctly reflects the nature and extent of the candidate's contributions to the published research work.

| Designation / Name | Signature | Date |
|---|---|---|
| Candidate: Claire E Mercer | | |
| Main supervisor: Peter Hogg | | |

# CHAPTER THREE: CONTEXTUAL BACKGROUND

## 3.1 Aims and nature of research

Rigorous maintenance of a quality assurance (QA) programme is crucial in upholding the effectiveness of the National Health Service Breast Screening Programme (NHSBSP). The QA guidelines for mammography established to facilitate the NHSBSP objectives of a long term contribution in the reduction in breast cancer mortality. Within the QA guidelines for mammography there are no guidelines for optimal compression force levels for practitioners to follow when performing mammography.

The overall concept of this research was to identify the range and extent of compression forces used in order to investigate practitioner variation of breast compression force within the NHSBSP.

The objective of this research was to establish:

- if practitioner variation in the application of compression force existed
- if so, to establish the range of that variation
- if establishing a range, to realise potential consequences to image quality and identify possible client effects over sequential screening within the NHSBSP

## 3.2 Research context

To establish the research framework, contextual information on the NHSBSP, mammography and quality assurance (QA) guidelines follow.

### 3.2.1 National Health Service Breast Screening Programme (NHSBSP)

The incidence of breast cancer is increasing on a global scale and the disease remains significantly high in public health issues. Mammography is, at present, the best method for the detection of clinically non-palpable breast cancer. The aim of the NHSBSP to detect cancer at an early stage [NHS Cancer Screening Programmes, 2013] when there is a greater chance of it being treated successfully [Health and Social Care Information Centre, 2013]; prognosis of the disease being directly related to the disease stage at diagnosis, with early detection leading to better prognosis [Ng & Muttarak, 2003].

The Forrest Report in 1986 [Department of Health and Social Security, 1986] recommended screening asymptomatic women aged 50-64 years in a three yearly cycle. Since its introduction in 1988, the NHSBSP has seen many procedural and structural changes; all aimed at increasing cancer detection rates. Two projection mammography (one in the cranio-caudal projection known as the CC and one in the medio-lateral oblique projection known as the MLO) was introduced at prevalent (first) round screening in 1995, followed by incident (subsequent) round introduction in 2002. In 2004 the upper age range for screening increased to 70years and in 2010 the NHSBSP commenced a randomised control trial of age expansion to those aged 47 to 73, with expected completion in 2016 [NHS Cancer Screening Programmes, 2013].

Key facts about the NHS breast screening programme are highlighted in Table 3.1, with 'uptake' being the percentage of women who are invited for screening in the year and screened within six months of their invitation.

| Year | Number of women screened aged 45+over (Millions) | Uptake (50-70years) (Percent) | Referred for assessment (45+over) (Thousands) | Cancers detected (45+over) (Thousands) | Cancers detected per 1,000 women screened (45+over) (Rate) |
|---|---|---|---|---|---|
| **2010** | 1.79 | 73.2 | 74.3 | 14.2 | 7.9 |
| **2011** | 1.88 | 73.4 | 75.0 | 14.7 | 7.8 |
| **2012** | 1.94 | 73.1 | 80.5 | 15.7 | 8.1 |
| **2013** | 1.97 | 72.2 | 81.9 | 16.4 | 8.3 |

[Health and Social Care Information Centre, 2014]

Table 3.1: NHSBSP Key facts

The risks and harms of breast screening in the NHSBSP is a continued controversial debate, in terms of lives saved it has been demonstrated that the benefit of mammographic screening is deemed greater than the harm in terms of over diagnosis, with between two to two and a half lives saved for every one over diagnosed case [Duffy et al 2010]. A research group, the Strategic Review of Health Inequalities in England [2012], was assigned to fully assess the risks and harms of screening for breast cancer; their key aim to establish how effective the screening programme was at saving lives. This review, led by Professor Marmott, summarised that for each breast cancer death prevented, about three over diagnosed cases will be identified and treated. The conclusion of the review concluded that the UK breast screening programmes had a significant benefit and should continue.

### 3.2.2 Mammography

Mammography has been long established (over 40years) as the leading modality for the detection of breast cancer [Strategic Review of Health Inequalities in England, 2012]. In

2005, it was recognised that digital mammography was significantly better than analogue mammography at detecting breast cancer in women aged 50years or younger, or in women at any age with very or extremely dense breast tissue [Pisano, Gatsonis, Hendrick, & Yaffe, 2005]. Whitman and Haygood [2013], demonstrated that digital screening is similar in efficacy or slightly better than film screen. The Department of Health (DoH) [2007] decreed that full field digital mammography was to be made available for women in the screening assessment process within the NHSBSP in the 47-50 age range, together with a roll out of digital equipment in all screening services, with digital mammography for all clients screened within the NHSBSP by 2012.

A mammogram, be it analogue or digital, consists of two projections of each breast; one in the cranio-caudal projection known as the CC and one in the medio-lateral oblique projection known as the MLO. 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 3.1). For the MLO the arm of the mammography gantry is tilted from the vertical and angled to be parallel to the pectoral muscle angle of the client; the angle determined by the practitioner in accordance with the client body habitus (Figure 3.2).

Figure 3.1: The cranio-caudal (CC) mammogram

Figure 3.2: The medio-lateral oblique (MLO) mammogram

Imaging the breast is challenging due to the large variations in breast volume and morphology. Client anatomical variations, particularly within the sternum and spine pose challenges for practitioners and require adapted techniques. For a successful mammography image the practitioners require client co-operation and must ensure accurate breast positioning; it is recognised that an optimum image can be achieved by employing the '3 Cs; carefully, correctly and consistently' [Simmons, Chavez, & Barke, 2012].

It is identified that the application of breast compression force that is required prior to image acquisition is one of the most essential parts of the mammography process. When compression force is applied it reduces breast thickness; though the exact relationship between compression force and reduction in breast thickness is neither linear nor clear-cut [Hogg, Taylor, Szczepura, Mercer, & Denton, 2013; Poulos, McLean, Rickard,

& Heard, 2003]. It is clearly recognised that breast thickness reduction minimises radiation burden, lessens superimposition of breast structures and decreases geometric and motion unsharpness [Bentley, Poulos, & Rickard, 2008; Long, 2000; Poulos, & McLean, 2004].

The importance of sustained and consistent high standards of practitioners who perform mammography and apply breast compression force are essential in maintaining the efficacy of the NHS breast screening service.

### 3.2.3 Principles of quality assurance (QA)

The main aim of the NHSBSP is to reduce mortality from breast cancer. Quality assurance within the NHSBSP facilitates that objective by providing robust standards to ensure focus and adherence with this key goal [NHS Cancer Screening Programmes, 2006].

When the NHSBSP service was introduced, the Forrest Report, stressed that all service aspects would have to be of highest quality [Breast Cancer Screening, 1987]. From this point onwards QA became a central, fundamental and integral part of the service; the first QA guidelines for mammography being published in 1989 [Department of Health, 1989]. This, the Pritchard Report, set out key standards, objectives and intrinsic elements of staff training, responsibilities and key lines of reporting frameworks [NHS Cancer Screening Programmes, 2006].

The DoH in 1991 provided an advisory committee report for breast cancer screening which highlighted the evidence and experiences since the introduction of the Forrest Report [Breast Cancer Screening, 1991]. In 1997 a further review of QA services was requested from the Secretary of State for Health [Calman & Hine, 1997]. The executive letter, EL(97)67, fully clarified relationships between breast screening units, host Trusts and regional QA teams; stating that adherence to national standards and a rigorous QA programme were key prerequisite elements in high quality breast screening services [NHS Executive, 1997].

Specific to the context of mammography and practitioners employed within screening services, guidelines exist which managers of breast services have responsibility to ensure compliance. NHS Cancer Screening Programmes [2006] in their publication 63, establish the QA guidelines for mammography staff including quality control. These objectives concern the whole aspect of the service and equipment. Two specific objectives are concerned with the achievement and sustainment of optimum image quality with as low a radiation dose as practicable. Specific guidance is directed at 'minimum standards' for specific high contrast spatial resolution and minimal detectable contrast levels on images. The guidance for the minimum standard for mean glandular dose per film for a standard breast at clinical settings is ≤ 2.5 mGy [NHS Cancer Screening Programmes, 2006]. Other than this there are no further dictates or guidance in this area.

Within these standards a section on mammographic techniques deals with 'appropriate compression' [NHS Cancer Screening Programmes 2006]. The standards state that:

*"Compression is important in reducing radiation dose, movement blur, geometric unsharpness and overlapping tissue shadows. The compression should be applied slowly and gently to ensure the breast is held firmly in position. The breast should be lifted and the tissue separated while compression is applied to enable better visualisation on the mammogram. The force of the compression on the x-ray*

*machine should not exceed 200 Newtons or 20 kilograms".* [NHS Cancer Screening Programmes, 2006, p.42]

Practitioners employed within the NHSBSP have no further specific guidance on compression force application. It is apparent that there is scope and potential for significant variations in practice with the application of breast compression force.

## 3.3 Literature analysis

In the development of this research the identification of a key theme was established; practice within and between practitioners in the application of breast compression force. Literature analysis surrounding this key theme is presented within this section of the thesis.

### 3.3.1 Image quality

The aim in mammography is to clearly visualise breast tissue structures in order to aid cancer detection. A criterion of NHSBSP guidelines focuses on image evaluation systems in order to guide staff to ensure optimum image quality. Taplin et al [2002], highlighted a positive correlation between poor image quality and the occurrence of breast cancer within two years of a negative screening mammogram and highlighted the importance of image quality within the NHSBSP. The experience of the practitioners and the standardised training of such staff is therefore of upmost importance in order that image quality is maintained to the highest clinical standards.

Challenges, as described by Bentley et al [2008], are in the quantification of image quality of the mammography image and the skill of the practitioners in breast immobilisation and positioning prior to compression.

In relation to image quality of a mammogram, breast compression to reduce breast thickness is deemed to be one of the most important factors. Any reduction in breast thickness with adequate breast compression reduces the radiation dose required for exposure and improves image contrast by reducing radiation scatter [Chida et al., 2009; Pisano et al., 2005].

Individual practitioners involved in mammography service provision are evaluated through self-assessment. Rigorous three yearly quality assurance visits to a service encompass assessment of image quality through evaluation [NHS Cancer Screening Programmes, 2000]. Image quality is measured by a tool produced by the NHSBSP [NHS Cancer Screening Programmes, 2006] which directly relates to the amount of breast tissue included on a mammogram in both the MLO and CC views. It is important to recognise that this tool is not derived from evidenced research base nor does it monitor compression forces. Practitioners are continuously monitored by this tool with self-assessment and peer review to monitor their standards. This tool is imperative to maintain standards; it was recognised back in 2003 that 10% to 30% of cancers can be missed through poor mammography screening [Majid, Shaw de Paredes, Doherty, Sharma, & Salvador, 2003], highlighting the importance of strict image quality standards.

It is clear that mammography image quality is dependent on numerous interlinked components including equipment, client positioning, compression force, viewing conditions, patient tolerance and practitioner skill. Comparisons of image quality have to take into account these factors as these will lead to the ultimate indicator in the performance of the NHSBSP; the success or failure in the detection of non-palpable breast cancer. Taplin et al [2002] suggested that little was known about exactly which of these image quality parameters affect cancer detection and no research has been further established in this field. A systematic review in 2010 [Li, Poulos, McLean & Rickard, 2010], noted that when image quality was rated higher, the lesion detection rate did not alter and further studies were suggested to be carried out to explore the relationship between image quality components. At this time no further details emerge.

This review of clinical image quality evaluation methods in 2010 [Li et al., 2010] assessed the EU [European Commission., 1996] and the ACR [Committee ACR., 1999] guidelines on image quality, highlighting an expectation that similar research studies with similar aims would use similar image quality evaluation methods. The review demonstrated this was not the case and overall, although the rating methods for image quality in all these studies varied considerably, it is acknowledged that all but one study utilised the BI-RADS classification scale.

Li et al [2010] strongly suggested that it was essential that research focussing on mammography image quality evaluates the inter - reader reliability together with an evaluation of breast density and an overall impression of image quality. The article noted that more importantly, the method should permit simple, reproducible evaluation of clinical components.

In summary, it is apparent that the term 'image quality' can be addressed from varying perspectives and that analysis of visual image quality is complex and multifactorial. For the direct monitoring of image quality standards within the NHSBSP, together with research activities, it is essential that criteria to assess image quality have a sound evidence base and remain consistent.

### 3.3.2 Compression and pressure force

Within mammography generalisation of the terms force, pressure and weight are often used by practitioners interchangeably and there is a recognised lack of understanding in the terminology. For clarity:

- Force is an interaction which causes the change of motion of an object, measured in Newtons (N). With ten decaNewtons of force (daN) being equivalent to one Newton (N); 10daN =1N.

- The term pressure is referred to as the force per unit area which is applied in a direction perpendicular to an objects surface, measured in Newtons per square metre measured Pascal (Pa). With one Pascal relating to one unit per square metre.

- The weight of an object (kg) is the force generated by the gravitational attraction of the earth on an object; 1kg is equal to 9.80655N. The weight of an object in kg is generally taken to be the force (N) of an object due to gravity.

Compression force is applied to the breast tissue during mammography and the readout in N or daN of force often visualised by the practitioners on the mammography machine; practitioners commonly refer to this interchangeably as a ‘pressure’ and a ‘force’.

In standard mammography practice, breasts are compressed until adequate thickness reduction is induced. Deciding when enough compression force has been applied is the remit of the practitioners and various descriptors have been proposed to indicate when enough compression force has been applied [Eklund, Cardenosa, & Parsons, 1994; Kopans, 2007; Long, 2000; Poulos & McLean, 2004; Wentz, 1992]. There are no evidence-based agreed guidelines for practitioners to identify optimal compression force levels.

The NHSBSP published a set of imaging criteria providing clear guidelines for the 'ideal' mammogram. These guidelines refer to the compression force being applied to the breast "slowly and gently to ensure the breast is held firmly in position" [NHS Cancer Screening Programmes, 2006] and also allude to the fact that the compression should not be in excess of 20kilograms. Mammography machine readouts are in Newtons or decaNewtons of force and the guidelines should reflect these measures.

The key competency framework [Skills for Health, 2013] directs training centres to the criteria which practitioners should meet upon qualification; occupational standards to position individuals and produce radiographic images of the breast state that the breast should be 'compressed to ensure the whole breast is included' [Skills for Health, 2013] and do not offer further guidance on the compression force values.

Through recent research it is clear [Murphy F, et al., 2014] that many practitioners do not refer to the numerical value of compression being applied, but make a decision to cease compression related to the look and feel of the breast. Within this research Murphy and colleagues [2014] noted that some practitioners used compression force as a final check prior to exposure and some practitioners involve the client. They also found that the speed of the application of compression force varied and practitioners demonstrated self-doubt about their practice. In another study it was noted that clients could often compare their experience with a previous examination [Robinson, Hogg, & Newton-Hughes, 2013].

It is clear that positioning the client for a mammography image requires a great deal of skill. The application of breast compression force that is required prior to image acquisition is one of the most essential parts of the process. It is well established that the development of a relationship of trust with the client will assist with their relaxation and that effective communication is essential in order for the client to understand the positioning required and the use of compression force [Lee, Strickland, Wilson, & Rickard, 2002; Simmons et al., 2012]. Doyle and Stanton [2002] referred to breast compression application as an 'art' and discussed the challenges that practitioners faced in communicating effectively with clients whilst applying compression.

New technologies are coming into play [De Groot, Broeders, Branderhorst, den Heeten, & Grimbergen, 2013] which change the focus from compression 'force' to compression 'pressure'. In a newly designed compression paddle, the paddle indicates the pressure applied to the breast rather than assessing the overall force. Such new technology could lead to consistency of pressure for each screening attendance; that is, if standards were in place.

### 3.3.3 Mammography pain

In practice breasts are compressed until "the breast is taut at the sides", "the skin blanches" [Bragg, 1986; Long, 2000; Wentz, 1992]. Poulos et al [2004] highlighted that the application of compression influenced pain with potentially no associated benefits with breast thickness reduction.

Poulos' studies [Poulos et al., 2003; Poulos, & McLean, 2004] utilised experimental mammographic compression, with no exposure, noting down the point at which blanching of the tissue and/or tautness at the sides together with minimal thickness and patient discomfort. The study results reported that “blanching/tautness” occurred at a wide range of compression forces and breast thicknesses and have the potential to create variation in the application of compression force in breasts. It was shown that in practice it is possible to assess whether the breast is firm or soft in the first 30N of application; the specific requirements for each breast can then be applied by the practitioner.

In 2004 Poulos and McLean called for a “..,new perspective on breast compression..,” discussing the fact that essential focus is required on training in mammography with regards to the effects of breast compression focusing on the minimisation of breast thickness rather than the amount of compression applied. To date, no further work is apparent in this field.

It is imperative to maximise the number of women who attend for routine breast screening in order to reduce breast cancer mortality. A systematic review by Whelehan, Evans, Wells & McGillivray [2013] confirmed the effect of pain on repeated attendance for screening. Though it is stated that there is a complexity between the phenomena of pain and screening behaviour, the research was able to firmly conclude that there was sufficient evidence that painful mammography contributed to non-re-attendance. 25-46% of women cited pain as a reason for non-re-attendance; in real terms between 47,000 and 87,000 women each year in England. Their research concluded with an appeal for pain reducing interventions in mammography.

### 3.3.4 Radiation risks

Mammography is required to be performed to a high standard with a low breast radiation dose; mean glandular dose per image for a standard breast at clinical settings being ≤ 2.5 mGy [NHS Cancer Screening Programmes, 2006]. The breast tissue should be adequately penetrated with radiation in order that the fibrous strands can be visualised through the breast parenchyma tissue; it was acknowledged by Cheung [2006] that underexposure resulted in a marked risk for missing breast lesions. As the various breast tissue types (fat, glandular and fibrous) have similar atomic numbers they have little inherent density differences; high contrast images are required for subsequent high quality mammography using the lower kVp ranges [Eklund et al., 1994].

The risks and health effects after radiation exposure with such low doses is a topic that remains under debate today; "For every 14 000 women in the age range of 50-70 screened by the NHSBSP three times over a 10 year period, the associated exposure to x-rays will induce one fatal breast cancer" [NHS Cancer Screening Programmes., 2006]. In order for the breast screening service to be justified, in radiation protection terms, then the benefit of screening must far outweigh the risk of inducing breast cancer. The benefit of screening maximised by the number of cancers detected, which is increased by improvements to image quality. It is to note that diagnostic performance is not solely dependent on image quality, but on other powerful parameters such as observer decision making, expertise, workload and experience.

In summary it is noted that the breast tissue is a radiosensitive structure and the radiation risk is considered to be acceptable compared to the benefits for a screening programme such as the NHSBSP.

### 3.3.5 Breast density

The breast tissue itself is made of soft tissue structures with two different densities; adipose (fatty) and fibroglandular tissue. Breast density refers to the relative composition of this fibroglandular and fatty tissue; glandular tissue having a high density with fatty tissue a lower density.

In early reproductive life the breast consists of around 20% fatty tissue, with 20% being epithelium and 60% connective tissue. Breast density represents this epithelium and connective tissue (fibroglandular). The proportions of these alter with age, the amount of fatty tissue increasing with decreasing proportions of epithelium and connective tissues [Howell et al., 2005].

The association between increased breast density and an elevation in risk of breast cancer is well established. In basic terms, dense breast tissue contains less fat with more breast cells and connective tissue, therefore a greater proportion of breast cells associated with an increased risk of breast cancer [Assi, Warwick, Cuzick, & Duffy 2011; Boyd et al., 1995; McCormack, & dos Santon Silva, 2006; Wolfe, 1976]. Howell et al [2005] described multifaceted and interrelated associations with breast cancer; finding that some are unavoidable, such as inherited genes, and some are modifiable such as diet, alcohol and exercise. Other associations, such as late menopause and early menarche (longer lifetime exposure to the hormones oestrogen and progesterone), having increased associations with breast cancer development [Howell et al., 2005].

Though most risk factors associated with breast cancer are unable to be altered such as age, family history and parity; breast density can be altered by diet and exercise. Body weight is linked with breast cancer. After the menopause increased oestrogen levels are linked with the amount of body fat; an increased oestrogen level is associated with an increased risk of breast cancer. Higher amounts of fat in the diet also increase the risk of developing breast cancer. Research studies are attempting to ascertain individual's breast cancer risk within screening programmes and therefore the methods for the measurement of breast density need to be accurate [Sergeant et al., 2012].

Breast classification models have been utilised in order to ascertain magnitudes of risk of breast cancer in accordance with breast density. The classifications of breast tissue were first defined by John Wolfe MD in 1976 and, as such, are referred to Wolfe patterns. The density of the breast tissue progressively increases throughout the patterns [Heine & Malhotra, 2002] and Wolfe's classification took into account both quantitative and qualitative considerations of breast tissue [Byng et al., 1998].

The Breast Imaging Reporting and Data System (BI-RADS®) [Sickles, D'Orsi, & Bassett 2013; D'Orsi, Bassett, & Berg, 2003] reported by the American College of Radiology, is a tool for the standardisation of mammography reporting. It consists of a lexicon of standardised terminology, a reporting organisation with an assessment structure, together with a coding and data collection structure. The BI-RADS® breast density classification provides a means of breast pattern density classification and again highlights four progressively dense mammographic patterns; almost entirely fat (A), scattered fibroglandular densities which could obscure lesions (B), heterogeneously dense which may lower the sensitivity of mammography (C) and extremely dense which lowers the sensitivity of mammography (D). Most descriptions of breast density used for clinical purposes today use this classification system. It is noted that there is variability in using the BI-RADS® system and few studies have evaluated such variability between film readers in screening mammography [Heine & Malhotra, 2002].

In 2004, Hershe discussed a further way of classification of breast density by computer software programmes; restrictions and inconsistencies in reader classification of breast density in subjective ways uphold the use of computer software programmes on a continuous scale. In many research studies now undertaken to ascertain breast density, volumetric density estimation is provided by raw full field digital images from screening being processed through Quantra™ or Volpara™ software [Sergeant et al., 2012].

# CHAPTER FOUR: CRITICAL APPRAISAL

## 4.1 Research pathway

Following a literature review several themes were identified which directly contributed to the research rationale; forming the aims, objectives, methods and moulding the research.

The following statements were established in the development stages of the research:

- Compression force: there were no set directed quality control standards for mammography practitioners in relation to compression force application, other than a maximum force set at 200 Newtons [NHS Cancer Screening Programmes, 2006]
- Resultant pain: re-attendance into the NHSBSP was being affected by breast pain following mammography [Whelehan et al., 2013]
- Radiation risks: radiation doses should be kept as low as reasonably possible in the radiosensitive breast tissue
- Image quality: comparison of images over time through sequential screening is imperative to detect small, subtle breast changes and improve breast cancer detection

The main concern for the researcher was that even though very strict quality assurance and control guidelines were apparent through the NHSBSP, one area was deficient - the guidance and resultant standards regarding the application of compression force.

As the application of compression force directly affects the breast thickness, radiation dose levels and image quality to the breast, there required an edict to guide practitioners in this field. Research within this area was therefore essential.

The research objectives were clarified (Figure 4.1):

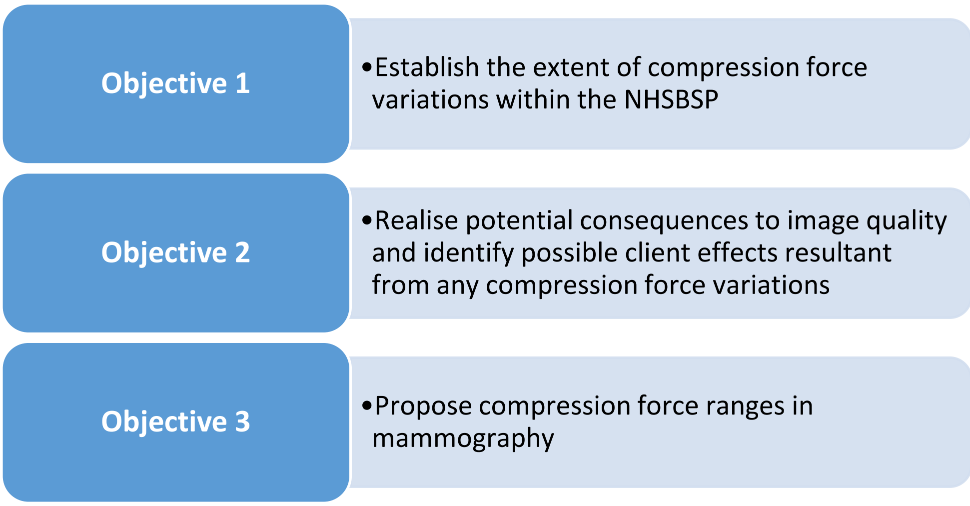



Figure 4.1: Research objectives

The researcher focused on designing and establishing research pathways which had direct significance to the research objectives above. In turn these could likely establish significance in clinical practice. Several research pathways were designed. Development of these pathways and resultant research papers are summarised, for clarity, in Figure 4.2

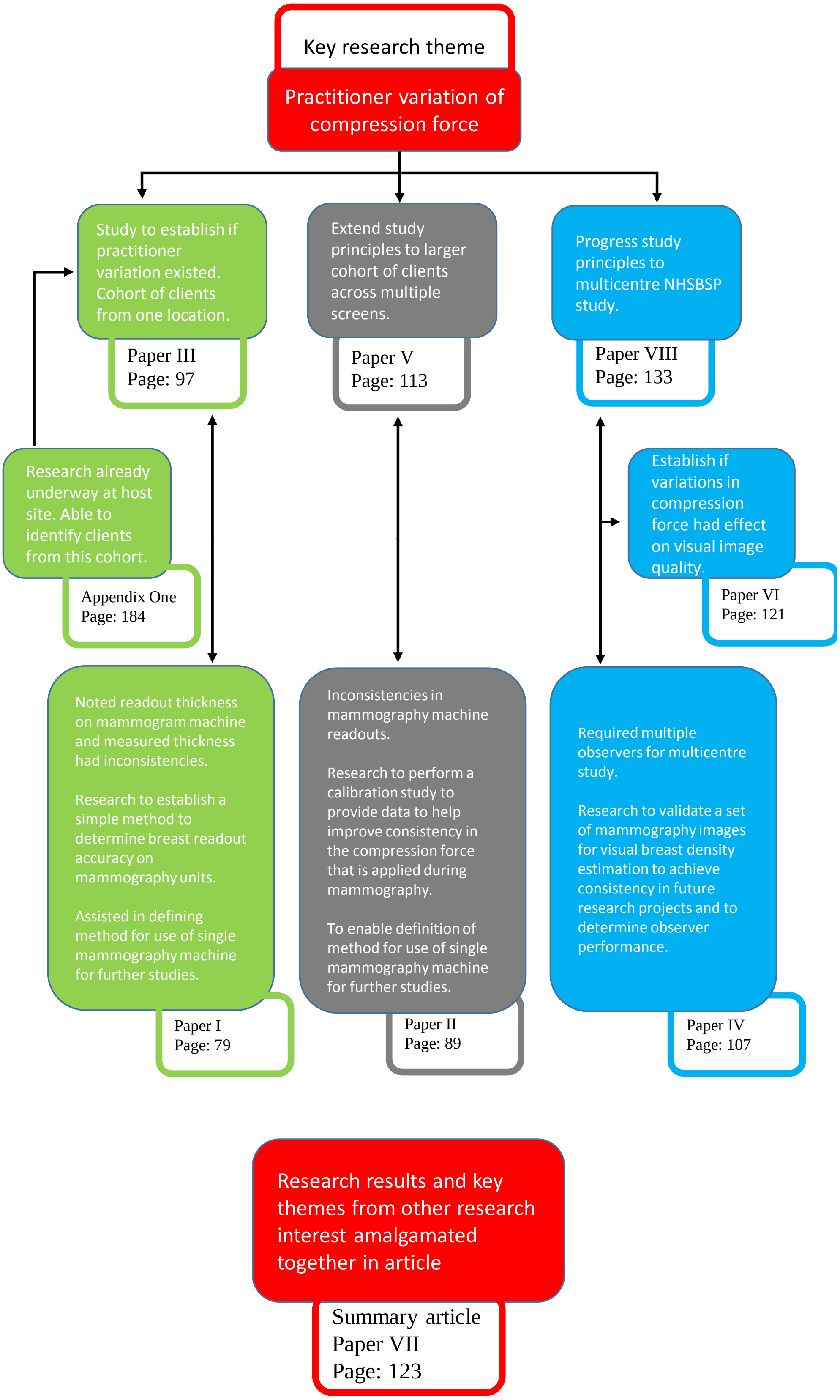



Figure 4.2: Development of research pathways and research papers

Within this chapter the published works inter-relationships are highlighted and the established research objectives referred to as:

- ***Objective 1***: *Establish the extent of compression force variations within NHSBSP*

- ***Objective 2***: *Realise potential consequences to image quality and identify possible client effects resultant from any compression force variations*

- ***Objective 3***: *Propose compression force ranges in mammography*

The development of research teams for all the research projects discussed within this chapter are considered in Chapter 6 of this thesis.

## 4.2 Establishing if compression force variation existed (objective 1)

The first key research objective was to:

- *Establish the extent of compression force variations within the NHSBSP*

In order to investigate this, a preliminary study was designed to determine whether the absolute amount of compression force in mammography varied between and within practitioners (Paper III).

The researcher was previously involved in a feasibility study (Appendix One) to assess the practicality of using a step-wedge based technique for measuring breast density from mammograms and to determine if additional information (relevant to breast cancer risk) could be collected by questionnaire. As part of this feasibility study, a cohort of clients from the NHSBSP was utilised from one screening service, taken from one static mammography centre which utilised one mammography system. As the cohort was a non-randomised consecutive group of NHSBSP clients, imaged on one mammography machine, they were considered ideal for sampling purposes for the initial study (Paper III).

Additional and coincidental, volumetric data was available from the feasibility study (Appendix One) which was utilised. Together with this, during the planning stage, it was acknowledged that breast density should be assessed. Following the literature search in this area the four-point Breast Imaging Reporting and Data System (BI-RADS) scale was utilised for this purpose [D'Orsi et al., 2003]. It is acknowledged that this reporting system was updated in 2013 [Sickles et al., 2013] and the reporting scale is now classified as A-D instead of 1-4. The research contained within this thesis was conducted prior to the publication of the new guidelines and as such refers to BI-RADS 1-4.

Exclusion criteria were set to ensure the data sample was not compromised by clients who had undergone surgery or had breast implants; anything which would purposefully alter the practitioners application of compression force. The data from retrospective mammography images from 500 clients was collated in an Excel spreadsheet. Following advice from two separate statisticians the client number to be sampled was not derived by a power calculation, due to the many factors involved for analysis. Instead a number of 500 was clarified with the statisticians to enable enough data within each BIRADS breast density for representative sampling.

### 4.2.1 Data interpolation

Data interpolation for this paper (Paper III) followed through the Excel spreadsheet. It was clear at the onset of data interpolation that a large number of confounding variables existed which could affect compression force application, for example; client tolerance, client habitus, practitioner experience in positioning, practitioner skill in positioning, breast volume and breast density. It was realised, very rapidly, that confounding factors had to be excluded and a clear focus was required on the research aim – practitioner variation.

The results from this study (Paper III) demonstrated a highly significant difference in mean compression force used by different practitioners ($p<0.0001$ for each BI-RADS density). It also demonstrated that practitioners applied compression in one of three ways using either low, intermediate or high compression force, with no significant difference in mean compression within each group (p=0.99, p=0.70, p=0.54, respectively) but a significant difference between each group (p<0.0001). When compression was analysed by breast volume there was a wide variation in compression for a given volume. The general trend was the application of higher compression to larger breast volumes by all three practitioner groups.

The conclusions from this study highlighted that practitioners did vary in the amount of compression force they applied during mammography, and the same variation existed in each BI-RADS grade. It was essential that this work now progressed.

### 4.2.3 Recognised shortcomings

Prior to the development of the next research study it was essential that a critique of the previous work was undertaken and any shortcomings recognised and resolved prior to the development of the next research project.

Within the initial stages of this project [(Paper III)], the first shortcoming was highlighted. It was acknowledged that a large proportion of time was squandered collecting data which was not required. This was due to the lack of understanding for the removal of the confounding variables at the developmental stage. It took the researcher some time to gain a full understanding of the requirement to understand which variables were not required in the dataset. Taking this forwards into the next research project, this shortcoming was resolved.

The second shortcoming of this research [(Paper III)] was in the design. The design utilised analogue mammogram images as they were readily available and a retrospective study data set was easily collated. In 2010 very few breast screening services had digital mammography; though in hindsight, it may have been possible to gain retrospective information from a screening centre who had been digital for nearly a year. If the researcher had done this in the early stages it could have stabilised the research in the digital screening arena. Instead, the data gathered within this research is from analogue images.

### 4.2.4 Focus and definition of method for future projects (objective 1)

During the development of the first study on practitioner variation [Paper III] it was acknowledged by practitioners within the department that when performing mammography the compression force 'dropped off'. This occurred during the time of the application of compression force on the client, to the practitioner returning to the control panel. It was also realised that resultant thickness changes may also occur during this time. Practitioners also suggested that some mammography machines were more affected than others.

This was considered to be an important area to develop in order to ensure stabilisation of the design and research outcomes from any future studies. As such, a study was designed to assess the measured thickness and the readout thickness measurements on mammogram machines [Paper I]. It was important when planning this study that a breast phantom was designed to mimic the compression characteristics of the breast. It was recognised in the early stages of design that actual, real effects would not be gained from a typical perspex phantom used for equipment testing and a realistic breast phantom was essential.

The method for this study comprised of three stages. Firstly the design of a clinically realistic breast phantom and rigid torso [Smith, Smith, Hogg, Mercer, & Szczepura, 2011]. Secondly, a device to measure breast thickness (TMD) and finally, the breast phantom and breast thickness measuring device utilised to assess several mammography units/paddle combinations.

Several different mammography machine manufacturers and varying compression paddles were used for the study. The results from this study demonstrated a difference between the readout thickness and the measured thickness, which varied between units of the same model and between manufacturers.

The results of this study assisted in defining the methods for the next part of the research, ensuring that machine variables were taken into consideration and confirming that the same mammography machine was used in the research studies. The next part of the research was to establish if practitioner variation in the application of breast compression was actual and not just defined to the client sample in the first study.

## 4.3 Confirming existence of compression force variation and identifying possible client effects (objective 1 & 2)

The research so far had identified that variations may exist in the application of compression force between and within practitioners. It was important to replicate the method and develop a further single centre study (Paper V), whilst addressing the shortcomings from the previous study (Paper III), to enable the researcher to substantiate the research outcomes.

### 4.3.1 Design phase

In the early stages of the development it was important to recognise that the results of the breast thickness readout study (Paper I) defined the method for this research, by ensuring that the same mammography machine that was used in the previous study (Paper III) was utilised to negate any variables in equipment.

During the design phase for this research it was deemed essential that the focus remain on the clients; variation of compression force over a period of time within the NHSBSP could potentially have profound effects on a client's experience and may affect the uptake rates to the service. In order for this to be factored in, instead of increasing the number of clients for analysis, the study was designed to progress longitudinally over sequential screening mammography in order to specifically take compression effects over time into account.

It was decided that a different cohort of clients would be utilised and three consecutive screening images analysed. This was deemed essential in the method definition at this stage. The NHSBSP requires serial imaging to occur at regular intervals, with images reviewed to assess for subtle changes, if compression force variability between practitioners existed then a comparison between images over time may become more challenging and cancer detection may be compromised.

The same exclusion criteria were applied as the single centre study (Paper III) and included 500 clients over 3 screening rounds (Paper V). The sample was gathered retrospectively from the same mammography system as the first study to enable direct comparison and minimise design error with machine given thickness and readouts (Paper I).

### 4.3.2 Results

The results from this longitudinal study (Paper V) highlighted that practitioners had similar compression force means as the single centre study (Paper III) (rank sum correlation coefficient = 0.9). The practitioners performed similarly in their compression force behaviours for both client datasets. This highlighted that practitioners were not altering their compression behaviours for clients with different breast sizes, but applying compression force within their own set 'tolerances'.

Importantly, the study results demonstrated that compression force varied over time and this was dependent upon the practitioner who imaged the client. For a client who was imaged with a practitioner from a different compression group on each attendance, breast compression force values were significantly different ($p<0.0001$). The breast thickness reduction was also significantly different between groups, suggesting that there was significance to the application of compression force in the reduction of breast thickness.

Significantly, as this was a retrospective sample, mean glandular doses were retrospectively analysed for clients who had been imaged from practitioners in different compression groups on each attendance. It was highlighted that in certain cases, the larger thickness reductions resulted in lower mean glandular doses (MGDs). Though ‘t tests’ indicated that these were not statistically significant in some cases, there has to be consideration of clinical importance – doses should be kept as low as practical.

### 4.3.4 Recognised shortcomings

As with the previous research, the design of this research was limited due to the utilisation of analogue mammogram images as they were readily available and a retrospective study data set was easily collated. In 2011 it would have been impossible to undertake a retrospective study of 6 years with digital imaging. It may though have been possible to design a prospective study commencing in late 2010 at some centres. Data collection for this though for such a study would have continued into 2016.

The key shortcoming of this project was in the data collection tool utilised. It was only apparent, upon data extraction, that the Excel spreadsheet was inappropriate to enable the generated reports required. Accordingly an Access Database was developed in support of this, and future work, within this area. It was apparent to the researcher that this was the second time that data gathering and analysis had been a key issue with the research and it was essential that this was addressed prior to any further research work was undertaken.

### 4.3.5 Further developments

Practitioner compression force variation in application was apparent within a different dataset and over a period of time, highlighting potential client and image quality effects (objective two) within one screening service. Further research was required to now clarify objective one within other screening centres and to establish:

- ***Objective 2***: *Realise potential consequences to image quality and identify possible client effects resultant from any compression force variations.*

and

- ***Objective 3***: *Propose compression force ranges in mammography*

## 4.4 Compression force ranges through optimisation (objectives 2&3)

On reviewing the literature it was clear there was little published evidence on the optimisation of compression force in mammography; almost no empirical data was available to describe how the breast behaved under compression force. This may begin to explain why the NHSBSP guidance is deficient and why this aspect of practice is, in the opinion of the researcher, inadequately quality assured.

### 4.4.1 Design phase

A research study was therefore designed in an attempt to determine a method for compression force cessation (when to stop applying breast compression force). This would hope to establish local compression force standards (Paper II) which may then be taken forwards to develop future mammography practice.

The research team was developed and included an MSc student who would lead this study in respect to data collection. The study was carried out within a symptomatic breast unit and consisted of 250 clients who had compression force and breast thickness levels recorded during compression application prior to imaging. It was decided that this would commence at 50N (5daN) and increased through 10N (1daN) increments until the practitioner had reached the termination compression force and thickness for the client's mammogram. The termination force was chosen subjectively by the practitioner taking into account client tolerance.

It was established that assessments would again be made on BI-RADS breast density as in the previous studies that encompassed this thesis. It was deemed imperative to do this within this study as it was envisaged that the compression forces required for breasts of different densities would be different.

### 4.4.2 Results

The results established that there were almost no differences between compression forces in all BI-RADS densities up to 110N (11daN). It was recognised that this may be due to the machine's limited precision for thickness measurements (minor compressibility differences may exist but the machine cannot differentiate them).

Differences were highlighted between the small and the medium/large compression paddles. The small paddle was used exclusively on small breasts and was non-tilting; for these breasts there tends to be less mobility with a much smaller compression capability range. The medium and large paddles utilised did tilt and the previous study (Paper I) noted that larger thickness readout errors were associated with tilting paddles. The differences therefore could be partly owing to precision.

The key findings from this study were that three different gradient zones were identified (the gradient being the amount of reduction in tissue thickness per unit of compression force). The three zones established concurred with high, medium or low rates of changes/gradients. In the high gradient zone a high level of thickness reduction is achieved with relatively small amounts of compression force. In the low gradient zone the amount of thickness reduction was relatively small compared with the compression force required to effect that change. In this zone the resistance increases rapidly, and the potential for discomfort thought also likely to increase per applied

Newton of compression force. The benefit of applying additional compression force from the point of entering this zone by practitioners ought to be questioned (Paper II).

The important factor from these results were that there was a lack of difference in gradient zones between BI-RADS scores. This meant that any application of compression force cessation models could be applied in the clinical setting without any adaptation for breast density. It was also important to note that the previous compression force research (Paper III) had established that compression force levels and thickness levels were not statistically different between BI-RADS breast density grades.

The results of this study demonstrated that practitioners, given latitude for clients who experience pain/discomfort, should enter the middle gradient zone and attempt to reach, but not necessarily enter, the low gradient zone before ceasing the application of compression. For this one machine termination of compression force application was to begin approaching 130N.

The results of this work (Paper II) provide a strong indication that there is the ability to provide practitioners with the required guidance and standards in the application and cessation of compression force. It is clearly acknowledged that for this one mammography machine in this study, terminations of compression force at 130N would be accepted as the most beneficial to thickness reduction, termination in the high gradient zone would not be acceptable (when breast thickness levels were highly affected by compression force). Termination in the middle zone could be acceptable.

As such a range of termination from 90 to 130N could be provided as guidance for practitioners on this machine.

### 4.4.3 Summary

In summary a method has been identified to minimise practitioner variability. It is important to recognise that population specific resistance scales would have to be completed at NHSBSP screening service and for different manufacturers. These resistance scales would help to standardise local practice and serve as an audit tool for QA standards.

### 4.4.4 Recognised shortcomings

The first shortfall in this study was that it was conducted in a symptomatic service. This was due to the fact that the researcher gathering the information was based within a symptomatic service. The research within this thesis mainly focused on screening units. Though the two services go hand in hand, it would have been beneficial to either centre the work on screening clients or use two cohorts of clients; one symptomatic and one screening.

It is clear, following the research outcomes, that this research would have been best conducted on a variety of mammography units in different locations simultaneously. Though this research has ascertained a very strong outcome, it would have been beneficial to have compared this to results from many other mammography systems. Though this is a recognised shortfall in this study it can be taken forward for future research in this field.

The main shortcoming of this paper developed into the main advantage of this paper. The paper had inadvertently misused the word ‘pressure’, a colloquialism common in clinical practice, for ‘compression force’. A letter to the editor of the journal, to which this paper was published, was generated and a response given from the authors. This ‘colloquial error’ was indeed the making of a new research relationship and work is now ongoing between the two research groups developing research proposals centered on a pressure based compression application system.

## 4.5 The effect of varying compression force upon image quality (objective 2)

Following the research outcomes from the single centre (Paper I) and longitudinal study (Paper V) together with the outcomes from the compression force cessation paper (Paper II), it was deemed necessary to begin evaluation on the effects of compression force upon image quality.

### 4.5.1 Study design

The longitudinal study (Paper V) had defined thirty nine clients within its data outcomes, who had received markedly different compression forces on each successive screen (low 6 to 7.4daN, intermediate 7.95 to 9.6daN and high 11.45 to 14daN). A study was therefore designed (Paper VI) to evaluate the image quality of the mammogram images of these clients for their three screening episodes. Due to the variation in scoring image quality (IQ) scales the study method utilised three different IQ scales, two of which were not evidence based; the validity of these scales was assessed in the method design. One of these three IQ scales was a new scale developed through psychometrics at the University.

### 4.5.2 Results

The results of this study (Paper VI) highlighted that the three image quality scales were positive and highly correlated (0.82, 0.9 and 0.85) indicating that they evaluated similar image parameters. Even though the mammograms, from an individual client, had statistically significantly different compression forces, the image quality scores did not vary significantly.

Correlating the results of this study [Paper VI] with the cessation of compression force study [Paper II] support the requirement of standards to guide practitioners in compression force application. It has been demonstrated, although only from a small sample, that visual image quality was not affected by changes in compression force from 6 to 14daN. This is an important finding which could have far reaching implications; though it is very clear that research into lesion visibility at different compression forces is required.

### 4.5.3 Recognised shortcomings

The dataset for paper VI was directly sourced from the outcomes of the longitudinal study dataset [Paper V] and, as such, did not have a formalised study design. As image quality descriptors are subjective, it would have been more advised to formulate research based on a clinically realistic breast phantom assessing visual image perception and lesion visibility to validate this outcome. This, however, can be taken forwards for future research in this field.

## 4.6 Confirming the existence of practitioner compression force variation in multiple screening centres (objective1, 2 and 3)

Prior to confirming if there was any practitioner variation in compression force within other centres, it was recognised in the design phase for this research, that multiple users would be scoring mammogram images for BI-RADS breast density. Though this tool is well recognised and established within mammography, it was important to ensure that inter and intra operator validity was acceptable prior to the research being established. As such, a study was designed in order to determine observer performance for breast density estimation and to achieve consistency in the following research projects.

### 4.6.1 BI-RADS consistency across multi-centres

In accordance with Li et al [2010], the method for this research was designed in order to be able to provide simple, reproducible evaluation for observer performance and to achieve consistency in additional research projects. Fifty mammogram images were scored for density grade by eight observers (Paper IV) at the three sites which were to be used for the multi-centre study (Paper VIII), together with one observer from the original study site who scored the previous research images (Paper III and V).

*Design phase*

During the design phase of this study advice was sought from a breast researcher at another University who had recently carried out a similar project [Eadie A et al., 2011]; she was bought into the research team and had an effect on how the project was steered. Fifty film-screen mammogram images were drawn from an anonymised University film library. Images 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 scored on two iterations with an interval of at least two weeks, to minimise recall bias.

Data analysis comprised of *within* observer variability (intra-observer variability), using Cohen's Kappa and delta variance, and *between* observer variability (inter-observer variability) by using Cohen's and Fleiss's Kappa. Cohen's Kappa measures agreements between two observers; Fleiss's Kappa measures the overall agreements between all the observers.

*Results*

Identifying the level of agreement which is acceptable for research purpose was difficult with no defining system in place. The baseline for acceptance of this research was set at strong agreement or above (0.61). It was also established that the delta variance between readers should be 1 or lower. The results demonstrated six of eight observers achieved strong intra-observer agreement (Kappa >0.81) with no observers demonstrating a delta variance above 1. Inter-observer variability was analysed twice and Fleiss' Kappa was used to evaluate concordance between all observers on first and second iteration; first scoring Fleiss kappa =0.64, second iteration =0.56. It was highlighted that each time an observer was paired with observer 7, who had low agreement, correlations reduced, observer 7 was extracted for the purpose of this analysis in order to set an acceptable baseline level at strong or above. All other observers were thus accepted for participation into further research studies together utilising BI-RADS breast density grading [Paper IV].

### 4.6.2 Multi-Centre study progression

Established very early in the research design were the NHSBSP centres to be used and the 'observers' who would grade the mammographic density and take readings from the mammography images. The observers defined by the previous study (Paper IV) as having strong inter and intra reliability; this deemed essential by Li et al [2010].

*Design phase*

The multicentre study assessed 3 consecutive analogue screens of 500 clients from each location. The same tested method and exclusion criteria applied as the previous single centre longitudinal study (Paper V). As it was well established that clients often compared experiences from previous examinations [Robinson et al., 2013], consecutive screening images were again deemed essential in method design.

975 clients met the inclusion criteria across three sites; 2925 mammography images. Data analysis focused on compression force (N) and breast thickness (mm) variation over 3 sequential screens to determine whether compression force and breast thickness were affected by practitioner variations (Paper VIII).

*Results*

The results from this study demonstrated that compression force over 3 consecutive screens varied significantly at each site. It was demonstrated that site three had a dictate of a minimum value of compression force application to its practitioners (100N) whereas site one and site two did not.

Site one and two demonstrated no significant difference in both the mean values for the CC (p>0.5) and MLO projections (p>0.1), though site one and three, together with site two and three did (p<0.0001). Variation was highly dependent upon the practitioner who performed the mammogram. At site one practitioners fell into one of three practitioner compression groups by their compression force mean values; high (mean 126N), intermediate (mean 89N) and low (mean 67N) (Paper VIII).

Minimum and maximum compression force values in the CC projection ranged from: Site one 47N to 122N (75N), site two 42N to 114N (72N), site three 103N to 158N (55N). For the MLO projection: site one 65N to 136N (71N), site two 48N to 139N (91N), site three 103N to 163N (60N). ANOVA of percentage changes were calculated for MLO and CC views. In the MLO view sites one and three, together with 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 held true for each BI-RADS grade (Paper VIII).

Breast thickness levels demonstrated the same themes; in both the CC and MLO views across each BI-RADS grade site one and two demonstrated no significant difference (>0.5) whilst site one and three together with site two and three did (p<0.0001). This held true for mean values and first and third quartile values (Paper VIII).

*Recognised shortcomings*

This research project was large with a significant number of data sets (a potential of 6000 data sets from each location) for analysis. It was recognised very early on in data collection that a more robust method was required for analysis. As such a member of the research team from the University designed a more robust method of data collection utilising an access database system. This system was then tested for use and then rolled out for use in the other two centres. This system allowed for ease of data manipulation and data findings, which would not have been possible previously. It was a shortcoming that this was not identified in the design phase of the previous longitudinal study (Paper V) and this significantly held up data collection and the start of the project.

*Summary*

In summary, this research (Paper VIII) firmly established that the amount of breast compression force applied by practitioners was not consistent across three NHSBSP screening sites, nor was the resultant breast thickness. This research clearly demonstrates that the practitioners from the breast screening units behave differently in the application of compression force when undertaking mammography. Greater consistency between practitioners in the application of compression force for clients is exhibited when guidance dictates a minimum compression force. This may have a positive impact on image quality comparisons over time, radiation dose and potentially cancer detection. The large variation could negatively impact on client experience; resulting in varying pain on each attendance, potentially reducing rates of re-attendance and cancer detection.

## 4.7 Research integration

The research was integrated and summarised into a final review paper requested by the editor of an annual radiographic journal (Paper VII). Unfortunately, this paper was produced prior to the multicentre study results being available (Paper VIII). This paper collated the key elements of research work contained within this thesis and within the University developed mammography research teams. It was intended to have an insightful impact on the mammography field. It was published immediately following the Francis report [Francis, 2013] and as such the readers were reminded of the importance of quality and standards in healthcare.

Firstly, the paper articulated that mammography is well-established, though there is little published empirical research into practitioner compression force application. It then recognised that literature within this field provides viewpoints, though few are based upon quality evidence based results. The paper summarised that compression behaviours amongst practitioners have been explored which may influence compression force practice [Robinson et al., 2013] and suggested cessation guidelines based upon the work carried out within this thesis. It summarised the variability in compression forces by practitioners, highlighting the work completed within this thesis and confirmed the research carried out by Poulos and McLean [2004].

### 4.7.1 Recognised shortcomings

It was unfortunate that this paper was published prior to the results of the multicentre study (Paper VIII) being available; this was mainly due to the fact that the design of the database for the multicentre study was not effectively planned in the initial stages and this slowed down data collection considerably.

## 4.8 Key research findings

The key research findings from this thesis work are discussed in line with the research objectives:

- ***Objective 1***: *Establish the extent of compression force variations within NHSBSP*

To date the research contained within this thesis is the only focused work within this field of breast compression. The research, performed by the author of this thesis, firmly concluded that there is compression force variation amongst practitioners within the NHSBSP. Multicentre analysis (4500 client visits) confirmed variation of compression force values 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 held true within each BI-RADS density classification.

- ***Objective 2***: *Realise potential consequences to image quality and identify possible client effects resultant from any compression force variations.*

It is clear from this research that compression force variations were not reflected in any measured change in visual image outcome on the grading scales used. In a cohort of clients (1500) widely variant compression force levels over longitudinal screens were demonstrated; as displayed by MLO/CC projections clients underwent a 55%/57% (site one), 66%/60% (site two) and 27%/26% (site three) change in compression forces through their three screening visits. This research demonstrated that measured differences in image quality scores were not reflected with large variations in compression forces; the IQ scores not varying significantly even though different compression levels were applied (Kappa: 0.92, 0.89, 0.89) [Paper VI]. It is recognised though that image quality is a complex area; having to assess and score with multiple confounding factors. Lesion visibility research linked to image quality has yet to be established within this field.

It is apparent from this research that variation in compression force over sequential screening attendances has been recognised and this could have an impact on client experience.

- ***Objective 3***: *Propose compression force ranges in mammography*

Importantly, it has been established from this research, that there is a compression cessation scale that can be developed on an individual mammography unit level suggested between 90-130N of force [Paper II]. Practitioners would have a guided scale for compression force and cessation of force; this being the same for both the CC and the MLO projection in all BI-RADS scales. Such a scale could standardise local practice and serve as an audit tool for QA standards. It is recognised that the scale may have to be developed on a site by site basis and for individual manufacturers and could then be utlised to form cessation guideline standards for mamography compression force.

### 4.8.1 Key research outcomes

For the first time in the NHS breast screening service, this evidenced based research has defined that there are practitioner variations in breast compression application. Across three screening sites the compression force variations were defined in the CC projection as: site one 47N to 122N (75N), site two 42N to 114N (72N), site three 103N to158N (55N). In the MLO projection as: site one 65N to 136N (71N), site two 48N to 139N (91N), site three 103N to 163N (60N).

Implications for successive client screens have been noted with clients seeing different percentage changes in compression forces across three successive screens dependent on the screening site they attend, the MLO projection: site one 55%, site two 66%, site three 27% and the CC projection: site one 57%, site two 60% and site three 26%.

Cessation guidelines have been proposed (between 90 and 130N of force) for the very first time since the breast screening services introduction in 1988. These guidelines are now being introduced within the national mammography training centre that the researcher manages and within a new mammography academic book that the researcher has co-edited and co-authored.

Though clients experience compression force variations, both over time and in different screening locations, with significant differences demonstrated ($p<0.0001$), there are no subsequent significant difference in visual image quality.

## 4.9 Future work

Future work in this field is now being developed in three ways.

Firstly the researcher is one of three editors in a mammography evidenced based academic book due to be published in early 2015. This book has an international authorship and is aimed at an international audience. It is hoped to have high impact on practitioner trainees and current practitioners in the future and guide practitioners in new evidenced based principles. As a co-author, key themes from this thesis on compression behaviours have been introduced, together with the introduction of cessation guidelines of 90-130N of force and the importance of standardisation over sequential screening.

The research that has arisen as a direct outcome of the research contained within this thesis is also contained within this mammography book (Chapter 6, Figure 6.1). It is considered that this academic mammography book would not have been achievable without the research contained within this thesis.

Secondly, discussions are underway with a company [VolparaAnalytics™ and VolparaDensity™, Volpara Solutions, Wellington, New Zealand] whose software has been developed, not only to estimate breast density, but to collate a number of factors ascertaining to the digital mammogram image which can be analysed to provide reports on practitioners. This research has an aim to run for a number of years to establish practitioner behaviours in more detail. This will be the first large scale research in this field.

Finally, and considered most important research development in this field, is the collaboration with researchers from the Netherlands who have designed a compression paddle based upon pressure force [De Groot et al., 2013; De Groot, Broeders, Branderhorst, den Heeton, & Grimbergen, 2014]. The design is complete, though no clinical trials have been undertaken with this pressure paddle within the UK. Discussions are underway to plan several research projects in this area to run from 2015 and 2016 with the researcher being the principle investigator.

## 4.10 Summary and recommendations

It is important to identify the effect that this research will have for clients within screening and symptomatic services. Identification that practitioners vary in the compression force they apply over sequential screening attendances could have an impact on client experience and potentially reduce re-attendance rates and cancer detection. Establishing guidance at 90-130N of force to allow a set range of compression forces may have a positive impact, over time, on image assessment together with potential cancer detection.

This research demonstrates that practitioners in some breast screening units behave differently in the application of compression force when undertaking mammography with significant differences in mean compression values between practitioners ($p<0.0001$ for each BI-RADS density). Where guidance dictates a minimum force to be applied this results in greater consistency between practitioners in the application of compression force for clients. This may have a positive impact on image quality, radiation dose reduction and potentially cancer detection; though may also have a negative impact on client experience.

Though it is recognised that effects on client experience are multifactorial, there is potential for this large variation in compression force in certain breast screening units to negatively impact on client experience by resulting in varying discomfort / pain on each attendance. This could therefore potentially reduce rates of re-attendance and therefore reduce cancer detection. As variation between some screening sites is apparent, a client moving location could have strikingly different experiences.

In summary, this research has firmly established that practitioners vary in the amount of compression force applied during mammography over sequential screens and in different mammography units. The compression force that it applied is not consistent through screening cycles. As such, correlation between previous images could be impaired. It has also been established that there are three compression force gradients, enabling the development of compression force cessation guidelines.

These key research findings can define that change is required within the NHSBSP within the compression force field. No standards are available to guide practitioners on the amount of compression force to apply; this research has established a need for such guidance to prevail. If standards are established then the effects on repeated client experience over time may become apparent; expectantly an increase in re-attendance at screening could be established as the client will have similar compression force experiences throughout the screening programme.

Dissemination of these cessation guidelines and the importance of standardisation through successive screens is ongoing by the researcher though the academic text book to be published in 2015, conference proceedings, and directly to new mammographers practitioners through the national training centre that the researcher manages.

# CHAPTER FIVE: PUBLICATIONS

The research within this thesis is based upon the following published papers which are contained, in full, within this chapter.

**Paper I** Hauge, I.H.R., Hogg, P., Szczepura, K., Connolly, P., & Mercer C.E. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Medical physics*, 39 (1), 263–271. doi:10.1118/1.3663579

**Paper II** Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E.R.E. (2013). Pressure and breast thickness in mammography- an exploratory calibration study. *The British journal of radiology*, 86 (1021), 20120222. doi:10.1259/bjr.20120222

**Paper III** Mercer, C.E., Hogg, P., Lawson, R., Diffey, J., & Denton, E.R.E. (2013). Practitioner compression force variability in mammography: a preliminary study. *The British journal of radiology*, 86 (1022), 20110596. doi:10.1259/bjr.20110596

**Paper IV** Mercer, C.E., Hogg, P., Kelly, J., Borgen, R., Millington, S., Hilton, B. … Whelehan, P. (2014). A mammography image set for research purposes using BI-RADS density classification. *Radiologic technology*, 85 (6), 609–613.

**Paper V** Mercer, C.E., Hogg, P., Szczepura, K., & Denton, E.R.E. (2013). Practitioner compression force variation in mammography: A 6-year study. *Radiography*, 19 (3), 200–206. doi:10.1016/j.radi.2013.06.001

**Paper VI** Mercer, C.E., Hogg, P., Cassidy, S., & Denton, E.R.E. (2013). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography*, 19 (4), 363–365. doi:10.1016/j.radi.2013.07.002

**Paper VII** Hogg, P., Mercer, C., Maxwell, A., Robinson, L., Kelly, J., & Murphy F. (2013). Controversies in compression, *Imaging and oncology*, 28-36, ISBN 9871 871101581

**Paper VIII** Mercer, C.E., Szczepura, K., Kelly, J., Millington, S.R., Hilton, B., & Hogg, P. (2014). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography*, Published online: July 29, 2014. doi:10.1016/j.radi.2014.07.004

## Paper I

Hauge, I.H.R., Hogg, P., Szczepura, K., Connolly, P., & Mercer C.E. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Medical physics*, 39 (1), 263–271. doi:10.1118/1.3663579

# 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 Østerås, 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)

**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. 
[DOI: 10.1118/1.3663579]

Key words: mammography, breast thickness, breast compression

## I. 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.

## II. 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.

### II.A. 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

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

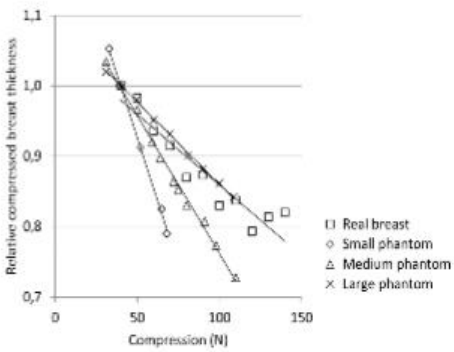



FIG. 2. Compressed breast thickness (mm) as a function of compression force (N) for real breasts and the three breast phantoms.

compression 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.

### II.B. 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.

### II.C. How the study was conducted

The measurements were performed on different mammography units from three different manufacturers [General

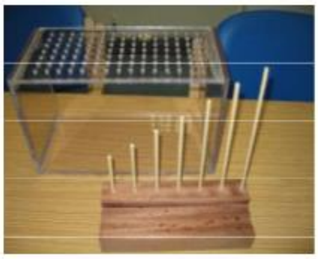


FIG. 3. Thickness measuring device (TMD) and rods.

TABLE I. Mammographic units included in this study.

| Location | Manufacturer/Model | SFM/FFDM | Compressed breast thickness accuracy (specified by manufacturer) | QC: maximum difference in measured and readout thickness[b] | Paddle size | Flexible/Nonflexible paddle | Tilting/Nontilting |
|---|---|---|---|---|---|---|---|
| A | GE Senographe 800T | SFM | ±10 mm | ±0.4 cm | 18 cm × 24 cm | Nonflexible | Nontilting |
| | | | ±10 mm | | 24 cm × 30 cm | Nonflexible | Nontilting |
| A | GE Senographe DMR+ | SFM | ±10 mm | +0.5 cm | 18 cm × 24 cm | Nonflexible | Nontilting |
| | | | ±10 mm | | 24 cm × 30 cm | Nonflexible | Nontilting |
| B | GE Senographe DMR+ | SFM | ±10 mm | +0.5 cm | 18 cm × 24 cm | Nonflexible | Nontilting |
| | | | ±10 mm | | 24 cm × 30 cm | Nonflexible | Nontilting |
| C | Siemens Mammomat Inspiration | FFDM | 39–45 mm[a] | −0.1 cm | 18 cm × 24 cm | Nonflexible | Nontilting |
| | | | | | 24 cm × 30 cm | Nonflexible | Nontilting |
| B | GE Senographe Essential | FFDM | ±10 mm | −0.3 cm | 19 cm × 23 cm[d] | Nonflexible | Nontilting |
| | | | ±10 mm | | 19 cm × 23 cm[d] | Flexible | Tilting |
| | | | ±10 mm | | 24 cm × 31 cm | Flexible | Tilting |
| D | Hologic Lorad Selenia | FFDM | ±0.5 cm | −0.1 cm | 18 cm × 24 cm | Flexible | Tilting |
| | | | ±0.5 cm | | 24 cm × 30 cm | Flexible | Tilting |
| D | Hologic Selenia Dimensions | FFDM | ±0.5 cm | −0.1 cm | 18 cm × 24 cm[d] | Flexible | Tilting |
| | | | ±0.5 cm | | 24 cm × 29 cm[d] | Flexible | Tilting |
| E | Hologic Lorad Selenia | FFDM | ±0.5 cm | −0.4 cm[c] | 18 cm × 24 cm | Flexible | Tilting |
| | | | ±0.5 cm | | 24 cm × 30 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 cm × 24 cm (18 × 24) and 24 cm × 30 cm (24 × 30) in the figures.

Electric (GE Medical Systems, Buc, France), Hologic Inc. (Bedford, MA) and Siemens (Siemens Healthcare, Erlangen, Germany)]. Both screen film mammography (SFM) and full-field 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 cm × 24 cm (18 × 24)] and large [approximately 24 cm × 30 cm (24 × 30)] were used (Table I).

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 prosthesis 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).

## II.D. 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:

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

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:

$$M = H - t - p - l \quad (2)$$

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 caliper, 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-

$$\text{Percentage} = \frac{(\text{Average/min/max measured breast area}) - \text{Readout thickness}}{\text{Readout thickness}} \quad (3)$$

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.

## II.E. 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, centered 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 $-0.04 \pm 0.12$ 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 $-0.17 \pm 0.07$ mm (95% confidence interval). Concluding from this their precision and repeatability was more than adequate for this study.

## II.F. 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 ±5 mm.[11] Each machine was tested every six months (Table I); all units were operating within manufacturer specification.

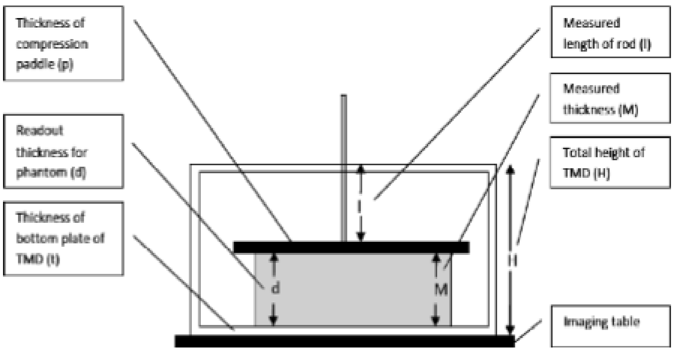



FIG. 5. Diagram to illustrate the measurements performed to calculate readout and measured thickness of the object.

### II.G. 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 $\pm 10$ N (in accordance with IPEM 89 Ref. 11) for all the units.

## III. 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.

### III.A. 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 $18 \times 24$ flexible paddle (smallest difference: 12 mm and largest difference: 19 mm) and Fig. 7 for the $18 \times 24$ 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 $18 \times 24$: 5.0/16.0 mm, nonflexible/flexible $24 \times 30$: 5.3/10.0 mm). Figure 7 illustrates that the compression paddle may be uneven in the left to right direction.

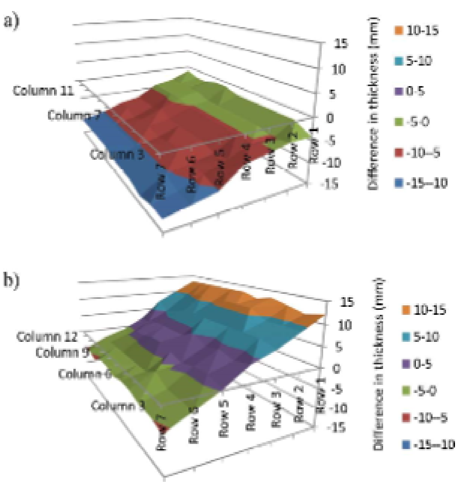



FIG. 6. Map of differences in thickness for the whole area for 18 cm × 24 cm flexible compression paddle for (a) Hologic Selenia Dimensions, which had the smallest (12 mm) difference in thickness across the whole area and (b) Hologic Lorad Selenia, which had the largest (19 mm) difference in thickness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

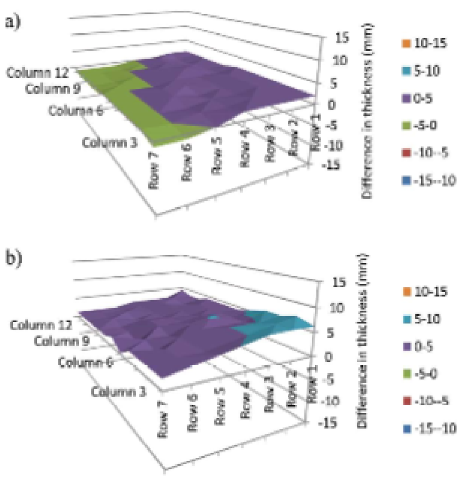



FIG. 7. Map of differences in thickness for the whole area for 18 cm × 24 cm nonflexible compression paddle for (a) Siemens Mammomat Inspiration, which had the smallest (3 mm) difference between measured and readout thickness across the whole area and (b) GE Senographe 800 T, which had the largest (7 mm) difference in measured and readout thickness across the whole area, when applying 100 N compression force.

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 $18 \times 24$ 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.

### III.B. 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.

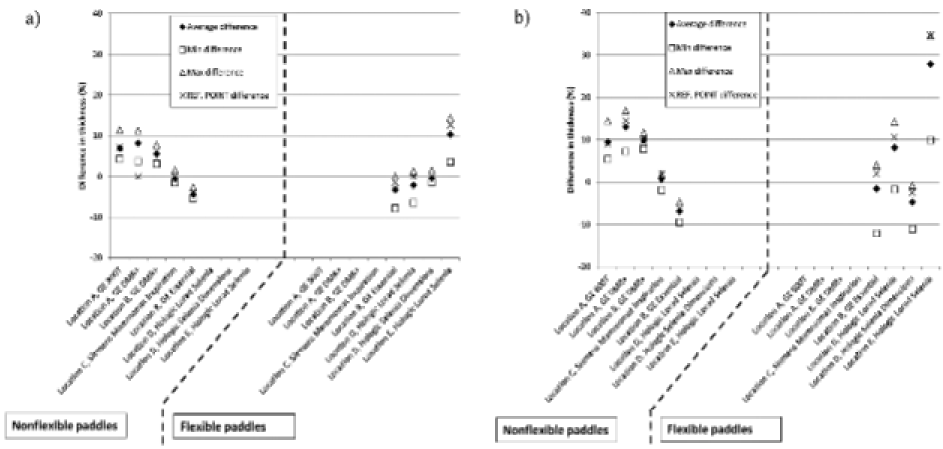



FIG. 8. The percentage difference between measured thickness and readout thickness for the breast area for 18 cm × 24 cm nonflexible and flexible compression paddle for (a) 60 N and (b) 100 N applied compression force.

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 (18 × 24 paddle: +2.6% ($p < 0.01$), 24 × 30 paddle: +0.7% ($p = 0.05$)) and largest for GE Senographe DMR+ (18 × 24 paddle (location A): +11.2% ($p < 0.01$), 24 × 30 paddle (location B): +14.3% ($p < 0.01$)). For the 18 × 24 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 24 × 30 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% (18 × 24), 0.7 vs 14.3% (24 × 30), Hologic Lorad Selenia: 6.8 vs 26.0% (18 × 24), 3.0 vs 8.3% (24 × 30)].

### III.C. 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 (18 × 24: 17.8 ± 1.4%, 24 × 30: 17.7±5.4%) when using nonflexible paddles. When using flexible paddles a larger decrease in measured compressed breast thickness can be observed for the 18 × 24 paddles (18.6 ± 2.6%) versus the 24 × 30 paddles (17.1 ± 1.9%).

### III.D. 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 −0.7 ± 0.2 mm (in percentage: −1.4 ± 0.5%).

## IV. 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 18 × 24 paddle, only 2 (Location B GE Essential/18 × 24 flexible; Location C, Siemens Mammomat Inspiration/18 × 24/24 × 30 nonflexible) out of 9 machines (22%) gave reference point and average values for the breast area that were within ±5% of the readout thickness. Flexible paddles had greater departure from measured thickness when compared with nonflexible paddles.

### IV.A. 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

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, 18 × 24 | | | | | | | | |
| 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, 24 × 30 | | | | | | | | |
| 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, 18 × 24 | | | | | | | | |
| 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, 24 × 30 | | | | | | | | |
| 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).

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 18 × 24 and/or 24 × 30 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 ±10 mm stated in the operator manuals. Hologic Lorad Selenia user manual states that compression thickness accuracy should be ±0.5 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 (18 × 24)], when the 18 × 24 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 (±10 mm). Had the tolerance limit been ±5 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 18 × 24 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.

### IV.B. Reference point

The difference between readout and measured thickness for the reference point and the average breast area values are similar [$-0.7 \pm 0.2$ mm (in percentage: $-1.4 \pm 0.5\%$)], 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.

### IV.C. 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: $18 \times 24$ and $24 \times 30$) and different breast compression forces (in this study: 60 and 100 N).

### IV.D. 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.

### IV.E. 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.

## V. 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.

## ACKNOWLEDGMENTS

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, "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 II

Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E.R.E. (2013). Pressure and breast thickness in mammography- an exploratory calibration study. *The British journal of radiology*, 86 (1021), 20120222. doi:10.1259/bjr.20120222

Cite this article as: Hogg P, Taylor M, Szczepura K, Mercer C, Denton E. Pressure and breast thickness in mammography—an exploratory calibration study. Br J Radiol 2013;86:20120222.

# 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*

**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 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 [$\leq -2$ (light grey); $-1.99$ to $-1$ (mid-grey); and $\geq -0.99$ (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.

Received 24 April 2012
Revised 3 September 2012
Accepted 1 October 2012

DOI: 10.1259/bjr.20120222



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

Address correspondence to: Professor Peter Hogg, School of Health Sciences, University of Salford, The Crescent, Salford M5 4WT, UK. E-mail: P.Hogg@Salford.ac.uk

(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™ 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 (15×29 cm), 2—medium (18×24 cm) and 3—large (24×30 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

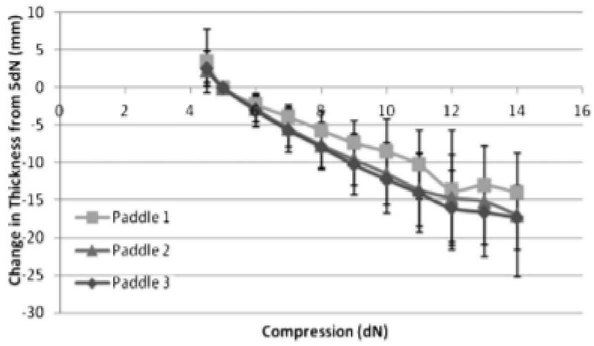



**Figure 1** Paddle comparison—craniocaudal view. Paddle 1, small (15×29 cm); Paddle 2, medium (18×24 cm); Paddle 3, large (24×30 cm).

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 $\sqrt{N}$, 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 $N$ ($\sqrt{N}$), 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.

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 BI-RADS 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). Maximum compression forces derived from the composite graphs

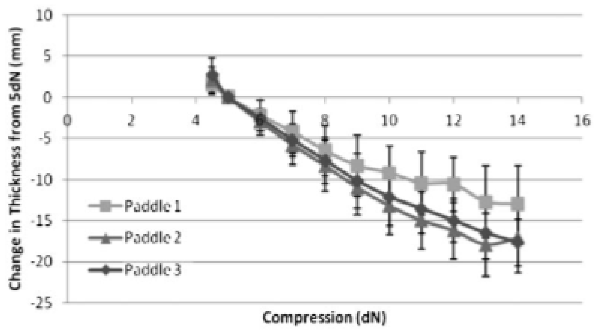



**Figure 2** Paddle comparison—mediolateral oblique view. Paddle 1, small (15×29 cm); Paddle 2, medium (18×24 cm); Paddle 3, large (24×30 cm).

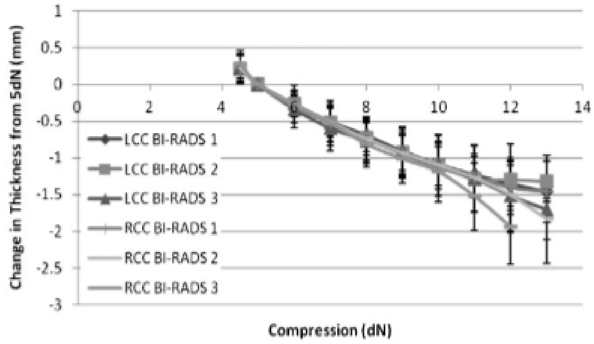



**Figure 3** Breast Imaging Reporting and Data System (BI-RADS) comparison—craniocaudal (CC) view. LCC, left CC; RCC, right CC.

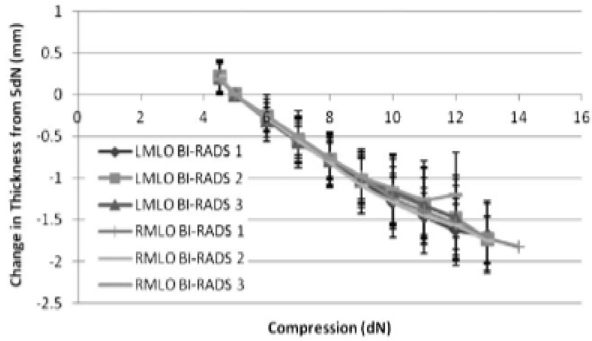



**Figure 4** Breast Imaging Reporting and Data System (BI-RADS) comparison—mediolateral oblique (MLO) view. LMLO, left MLO; RMLO, right MLO.

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

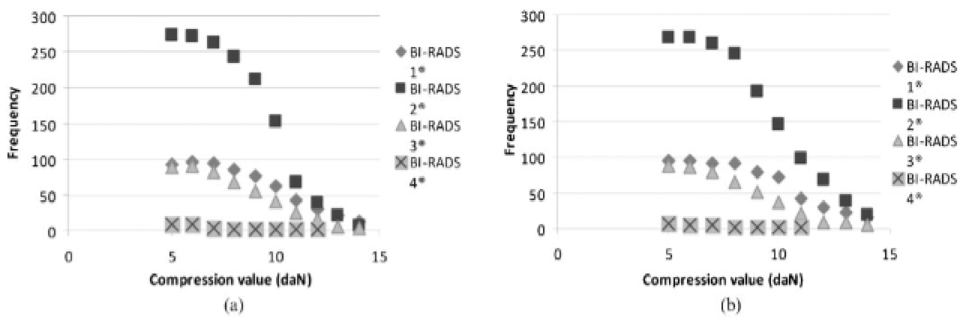



**Figure 5** (a) Craniocaudal compressions; (b) mediolateral oblique compressions. BI-RADS, Breast Imaging Reporting and Data System.

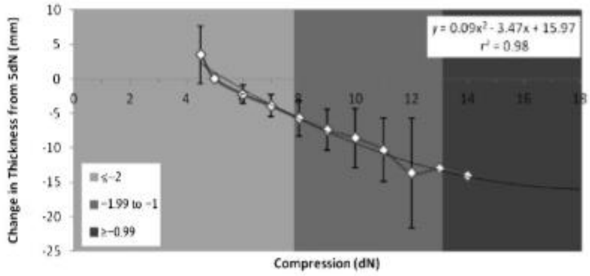



**Figure 6** Small paddle—average craniocaudal.

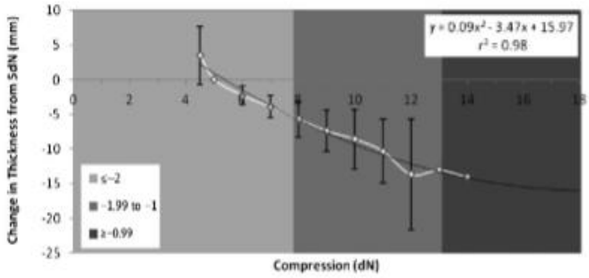



**Figure 7** Small paddle—average mediolateral oblique.

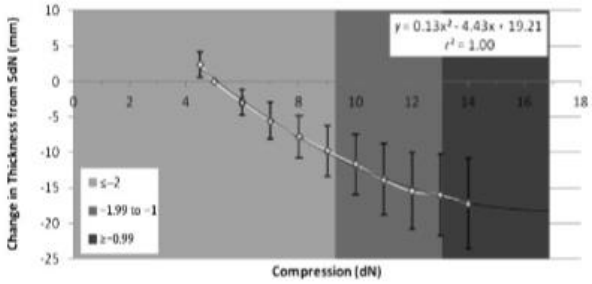



**Figure 8** Medium and large paddles—average craniocaudal.

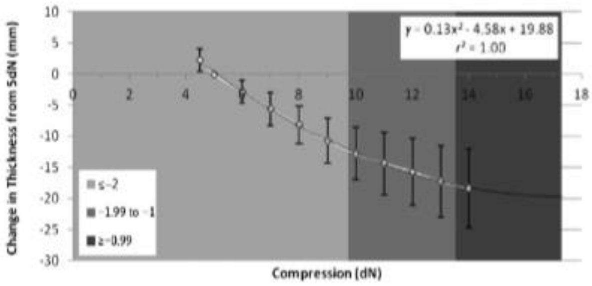



**Figure 9** Medium and large paddles—average mediolateral oblique.

≥−0.99 (dark grey). The use of this gradient calculation and the colour coding is described in the discussion section below.

## 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 mid-grey 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 13 daN.

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 NHS breast screening programme. Publication no. 60. Sheffield, UK: NHS Cancer Screening Programmes; 2005.
3. NHSBSP. Quality assurance guidelines for mammography including 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, scatter control, 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 and the 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 compression in mammography: a new perspective. Radiography 2004;10: 131–7.
13. Mercer CE, Hogg P, Lawson R, Diffey J. Radiographer variability 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'': an objective technique for measuring capsular fibrosis. Plast Reconstr Surg 1993;92:1078–84.
17. acr.org [homepage on the internet]. Reston, VA: American College of Radiography; 2003 [accessed 19 November 2012]. Available from: http://www.acr.org/Quality-Safety/Resources/BIRADS/Mammography

## Paper III

Mercer, C.E., Hogg, P., Lawson, R., Diffey, J., & Denton, E.R.E. (2013). Practitioner compression force variability in mammography: a preliminary study. *The British journal of radiology*, 86 (1022), 20110596. doi:10.1259/bjr.20110596

ARTICLE INFORMATION


Received: 18 July 2011 | Revised: 12 August 2012 | Accepted: 15 October 2012

doi: 10.1259/bjr.20110596

Cite this article as:
Mercer CE, Hogg P, Lawson R, Diffey J, Denton ERE. Practitioner compression force variability in mammography: a preliminary study. Br J Radiol 2013;86:20110596.

# Practitioner compression force variability in mammography: a preliminary study

**C E MERCER, MSc, BSc (Hons), P HOGG, MPhil, PGC, R LAWSON, PhD, MInstP, J DIFFEY, MSc, MIPEM and E R E DENTON, FRCP, FRCR**

The Nightingale Centre and Genesis Prevention Centre, University Hospital of South Manchester, Manchester, UK

Address correspondence to: Mrs Claire E. Mercer
E-mail: *mercer.claire@gmail.com*

**Objective:** This preliminary study determines whether the absolute amount of breast compression in mammography varies between and within practitioners.

**Methods:** Ethics approval was granted. 488 clients met the inclusion criteria. Clients were imaged by 14 practitioners. Collated data included Breast Imaging Reporting and Data System (BI-RADS) density, breast volume, compression and practitioner code.

**Results:** A highly significant difference in mean compression used by different practitioners ($p<0.0001$ for each BI-RADS density) was demonstrated. Practitioners applied compression in one of three ways using either low, intermediate or high compression force, with no significant difference in mean compression within each group ($p=0.99$, $p=0.70$, $p=0.54$, respectively). Six practitioners showed a significant correlation ($p<0.05$) between compression and BI-RADS grade, with a tendency to apply less compression with increasing BI-RADS density. When compression was analysed by breast volume there was a wide variation in compression for a given volume. The general trend was the application of higher compression to larger breast volumes by all three practitioner groups.

**Conclusion:** This study presents an insight into practitioner variation of compression application in mammography. Three groups of practitioners were identified: those who used low, intermediate and high compression across the BI-RADS density grades. There was wide variation in compression for any given breast volume, with trends of higher compression demonstrated for increasing breast volumes. Collation of further studies will facilitate a new perspective on the analysis of practitioner, client and equipment variables in mammography imaging.

**Advances in knowledge:** For the first time, it has been practically demonstrated that practitioners vary in the amount of compression applied to breast tissue during routine mammography.

Breast cancer is the second most common cause of cancer death in England for females, and mammography plays a critical role in its detection [1]. The clinical efficacy of mammography is dependent on the production of high-quality images and many factors contribute to this; one example being the application of adequate breast compression [2,3]. Compression is applied to reduce breast thickness; however, it should be noted that the exact relationship between compression and reduction in breast thickness is neither linear nor clear cut [4]. Thickness reduction minimises radiation burden, lessens superimposition of breast structures and decreases geometric and motion unsharpness [5–7]. Overall, thickness reduction is said to improve image quality, thereby heightening the chance of detecting cancer [8–11].

Various compression guidelines exist. National Health Service Breast Screening Programme (NHSBSP) guidelines [12] indicate that compression should be applied slowly and gently to ensure that the breast is held firmly in position [13–15] and that 20 kg (20 daN) of force should not be exceeded [14–15]. The NHSBSP has no exact guidelines for the application of breast compression; therefore, potential exists for practitioner variation. Anecdotally, variability is said to exist between practitioners and some publications have alluded to this [4,7]. If variability between and within practitioners does exist, in order to ensure that each client has a similar experience over time and that image quality differences are minimised, more detailed guidelines regarding compression may well be advantageous.

To date, research on breast compression has focused on the effects on the client on application. No robust research has been published to determine whether the amount of compression applied is dependent on the practitioner. In an attempt to start to address this literary deficiency, this preliminary study used a cross-sectional design to establish whether compression variability exists "within" and "between" practitioners.

## METHODS AND MATERIALS

The study followed the principles and ethics of the UK Department of Health Research Governance Framework [16]. Ethics approval was granted from the University of Salford, UK, together with the hospital research ethics committee (National Research Ethics Service, Bolton Research Ethics Committee, Manchester, UK). The study was performed in a regional breast screening service in the north of England. This service comprised two static and two mobile sites. One static site was selected and from that site a sample of 500 clients' mammogram images was drawn. The sample was opportunistic and derived from a previous research study [17]. It was retrospective and consecutive; factors such as socioeconomic, educational and menopausal status, breast tenderness and tolerance of compression could not therefore be assessed. We acknowledge that some of these factors could have influenced the amount of compression applied by the practitioners. In future prospective studies this information would be taken into account.

Mammograms were carried out by 14 trained practitioners who rotated through the department at the time of the study (the staff comprised advanced practitioners, mammographers and assistant practitioners). Craniocaudal and mediolateral oblique projections were acquired using an analogue DMR+ mammography machine (GE Healthcare, Chalfont St. Giles, UK).

Compression and practitioner details (name, number of years' experience and grade) of those who performed the imaging were noted for all images. Each practitioner was assigned a unique code to conceal their identity. Volumetric data (available from a previous research study [17]) were noted and breast density was assessed and recorded for each image using the four-point Breast Imaging Reporting and Data System (BI-RADS) scale [18–20].

Breast volume and density were evaluated in relation to compression applied by practitioners in order to determine whether relationships existed. One of the authors reviewed and scored all mammogram images for density assessment. For 20 mammogram images, this author was assessed against 4 experienced readers for interobserver BI-RADS scoring variability. When compared with each of the other four readers, Cohen's kappa test gave kappa values of 0.83, 0.92 and 0.83, demonstrating good agreement. Intraobserver characteristics determined by Cohen's kappa test gave a value of 0.92.

Statistical analysis comprised several steps. First, the sample was characterised to ascertain any distribution variations in BI-RADS grades between the mammograms

(Pearson's $\chi^2$ test). Second, the relationship between the amount of breast compression applied by different practitioners was analysed using analysis of variance (ANOVA). Third, data for each BI-RADS grade were analysed separately (ANOVA) to test whether practitioners applied the same compression to breasts with the same BI-RADS grade. Fourth, quantification of the correlations between compression and BI-RADS grade for each individual practitioner was calculated using Spearman's rank correlation. Next, using ANOVA, the sample was analysed to ascertain any variation in breast volumes between practitioners and what effect this may have upon the amount of compression that was applied. Finally, the employment grade and time since the mammography qualification of the practitioners were also assessed.

Of the sample, 12 clients did not have compression and/or practitioner information available and were therefore excluded, leaving 488 clients (1952 images) for analysis.

## RESULTS AND ANALYSIS

Images were seperated into BI-RADS grades regardless of practitioner. The following distribution of grades was ascertained: BI-RADS 1 (11%), BI-RADS 2 (64%), BI-RADS 3 (21%) and BI-RADS 4 (4%).

It was necessary to establish whether there were any BI-RADS differences between the clients that the practitioners imaged from the whole client sample using Pearson's $\chi^2$ test; it would have been unwise to compare practitioners if some had inadvertently imaged all clients who had breasts from just one BI-RADS category.

For the purposes of the Pearson's $\chi^2$ test, combination of BI-RADS 1 and 2 (referred to as Group A) and also of BI-RADS 3 and 4 (referred to as Group B) was required, owing to a low number of images in BI-RADS Catagories 1 and 4. Pearson's $\chi^2$ test compared the number of images in BI-RADS Groups A and B for all practitioners. Pearson's $\chi^2$ of 99.79 ($p<0.0001$) indicated a significant difference in the distribution of images within BI-RADS Groups A and B between these groups of practitioners. Table 1 demonstrates that there were similar groupings (percentages) of clients in BI-RADS Groups A and B for each practitioner. We could therefore be sure that each practitioner imaged clients of similar groupings of BI-RADS densities.

Table 1. Pearson's $\chi^2$ test with Breast Imaging Reporting and Data System (BI-RADS) groups and all practitioners

| Practitioners | Number of images in Group A (BI-RADS 1 and 2) | Group A % of total images | Number of images in Group B (BI-RADS 3 and 4) | Group B % of total images | Total number of images |
|---|---|---|---|---|---|
| A | 162 | 84 | 31 | 16 | 193 |
| B | 53 | 65 | 28 | 35 | 81 |
| C | 173 | 80 | 44 | 20 | 217 |
| D | 83 | 81 | 20 | 19 | 103 |
| E | 38 | 61 | 24 | 39 | 62 |
| F | 103 | 94 | 7 | 6 | 110 |
| G | 61 | 91 | 6 | 9 | 67 |
| I | 180 | 65 | 97 | 35 | 277 |
| J | 150 | 63 | 89 | 37 | 239 |
| L | 33 | 77 | 10 | 23 | 43 |
| M | 91 | 87 | 14 | 13 | 105 |
| N | 173 | 69 | 78 | 31 | 251 |
| P | 87 | 77 | 26 | 23 | 113 |
| Q | 73 | 80 | 18 | 20 | 91 |
| Total | 1460 | | 492 | | 1952 |

The relationship between the amount of compression applied by different practitioners was analysed using ANOVA (Table 2). The low $p$-value ($<0.0001$) demonstrates that practitioners did not use the same mean compression force. This could be because the practitioners were imaging breasts with different BI-RADS grades and potentially different breast volumes. Therefore, further analysis was performed to identify whether associations existed between compression and practitioners if BI-RADS grades and breast volume were taken into account.

Data for each BI-RADS grade were analysed separately to test whether practitioners applied the same compression to breasts with the same BI-RADS grade. ANOVA showed a significant difference within the mean compression values used by different practitioners within each BI-RADS grade (BI-RADS 1, $p<0.0001$; BI-RADS 2, $p<0.0001$; BI-RADS 3, $p<0.0001$; and BI-RADS 4, $p<0.002$). Taking the practitioner group as a whole, there were significant differences between practitioners in the application of breast compression within each BI-RADS category.

Table 2. Analysis of variance of breast compression (in decanewtons) for all practitioners

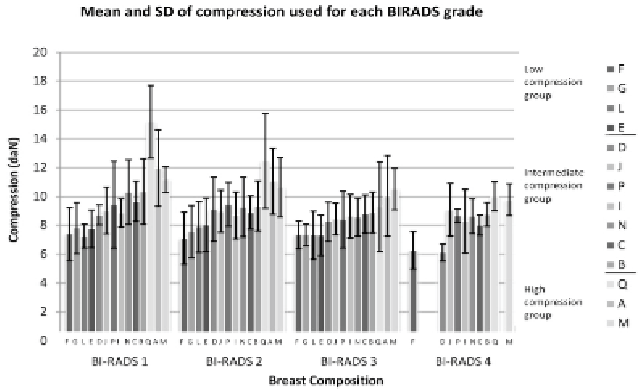
| Practitioner | Number | Mean | SD |
|---|---|---|---|
| A | 193 | 11 | 2.5 |
| B | 81 | 9.3 | 1.8 |
| C | 217 | 8.9 | 1.2 |
| D | 103 | 8.7 | 2.1 |
| E | 62 | 7.7 | 1.7 |
| F | 110 | 7.1 | 1.8 |
| G | 67 | 7.6 | 1.8 |
| I | 277 | 8.6 | 1.6 |
| J | 239 | 8.8 | 1.4 |
| L | 43 | 7.7 | 1.7 |
| M | 105 | 10.6 | 1.9 |
| N | 251 | 9.2 | 2.0 |
| P | 113 | 9.2 | 1.8 |
| Q | 91 | 12.2 | 3.5 |
| **Source of variation** | **Sum squares** | **DF** | ***p*-value** |
| Practitioner | 2722.4 | 13 | |
| Residual | 7173.2 | 1938 | <0.0001 |
| Total | 9895.5 | 1951 | |

DF, degrees of freedom; SD, standard deviation.

The mean and standard deviation of compression used by each practitioner for each BI-RADS grade were assessed. This clearly demonstrated that there was a large variation in compression used by each practitioner, with a tendency to apply less compression for higher BI-RADS grades. Compression data for BI-RADS Grade 3 were analysed by mean compression and practitioners could be separated into three distinct compression groups: those with low practitioner mean compression (7.33, 7.33, 7.30 and 7.29 daN), those with intermediate practitioner mean compression (8.25, 8.42, 8.39, 8.63, 8.56, 8.78 and 8.88 daN) and those with high practitioner mean compression (9.29, 10.03 and 10.5 daN). Clarification of these "groups" was acquired by analysing data by mean compression for the four BI-RADS groups; similar groupings by mean were highlighted (Figure 1). Further analysis described by ANOVA demonstrated no significant difference in compression within each of the practitioner groups for BI-RADS Grade 3 ($p=0.99$, $p=0.70$, $p=0.54$). Thus, three groups of practitioners can be defined according to whether they used low compression, intermediate compression or high compression.

ANOVA was also used to evaluate BI-RADS Grades 1, 2 and 4, following separation of the practitioners by the practitioner groups identified above, to determine whether practitioners remained consistent with their group (Table 3). For BI-RADS Grade 1, there was no significant difference in mean compression for the low ($p=0.91$) and intermediate ($p=0.08$) compression practitioners; for BI-RADS Grade 2, there were significant differences for the low and intermediate compression groups only ($p<0.05$ and $p<0.01$, respectively); and for BI-RADS Grade 4, there were significant differences in the intermediate compression group only ($p<0.02$). Only the high compression group of practitioners failed to maintain their consistency in BI-RADS Grades 1 and 2 ($p<0.0005$ and $p<0.0001$, respectively). This suggests that all groups of practitioners performed similarly within their group apart from the group which used higher compression forces.

Figure 1. The means and standard deviations of compression used for each Breast Imaging Reporting and Data System (BI-RADS) grade by each practitioner.




Quantification of the correlations between compression force and BI-RADS grade in each individual practitioner was then performed using Spearman's rank order correlation (Table 4). This demonstrates that only 6 out of 14 practitioners (A, C, D, P, Q and N, as demonstrated in bold in Table 4) showed significant correlation between the amount of compression applied and the BI-RADS grade of breast tissue. For these practitioners, there was a negative correlation between applied breast compression force and the BI-RADS grade of breast tissue—*i.e.* compression force decreased with increasing breast density. However, the remaining eight practitioners (F, G, L, M, B, E, I and J) showed no correlation between breast compression force and BI-RADS grade.

Overall, it was concluded that there was no consistency between practitioners in the amount of compression applied for breasts with the same composition (BI-RADS grade); there were, however, three groups of practitioners who maintained a degree of consistency between themselves.

Assistant practitioners were found in the low- and the high-compression groups, advanced practitioners within the low and intermediate groups and more experienced practitioners (>10 years) were found in all three groups. The less experienced practitioners (<3 years) were found in the low- and intermediate-compression groups. Dispersal of practitioner grade and length of experience across the three compression groups appeared to

Table 3. Results using analysis of variance to test whether practitioners in each group use the same mean compression

| | Low-compression group | Intermediate-compression group | High-compression group | |
|---|---|---|---|---|
| Practitioners included in group | F, G, L and E | D, J, P, I, N, C and B | Q, A and M | All practitioners |
| BI-RADS 1 | NS | NS | $p<0.0005$ | $p<0.0001$ |
| BI-RADS 2 | $p<0.05$ | $p<0.01$ | $p<0.0001$ | $p<0.0001$ |
| BI-RADS 3 | NS | NS | NS | $p<0.0001$ |
| BI-RADS 4 | No data | $p<0.02$ | Insufficient data | $p<0.0001$ |
| All grades | NS | $p<0.0001$ | $p<0.0001$ | $p<0.0001$ |

BI-RADS, Breast Imaging Reporting and Data System; NS, not significant.

Table 4. Spearman's rank correlation between compression force and Breast Imaging Reporting and Data System grades for each practitioner

| | | | 95% CI | | | | |
|---|---|---|---|---|---|---|---|
| Spearman's rank order correlation | *n* | R/S statistic | From | To | *t* statistic | DF | Two-tailed *p*-value |
| **Low-compression group** | | | | | | | |
| F | 110 | −0.06 | −0.25 | 0.12 | −0.67 | 108 | NS |
| G | 97 | 0.12 | −0.12 | 0.35 | 1.01 | 65 | NS |
| L | 43 | −0.07 | −0.36 | 0.23 | −0.46 | 41 | NS |
| E | 62 | −0.19 | −0.42 | 0.07 | −1.47 | 60 | NS |
| **Intermediate-compression group** | | | | | | | |
| **D** | 103 | −0.31 | −0.47 | −0.12 | −3.27 | 101 | <0.002 |
| J | 239 | −0.11 | −0.23 | 0.02 | −1.7 | 237 | NS |
| **P** | 113 | −0.23 | −0.4 | −0.05 | −2.5 | 111 | <0.02 |
| I | 277 | −0.07 | −0.19 | 0.05 | −1.21 | 275 | NS |
| **N** | 251 | −0.25 | −0.37 | −0.13 | −4.13 | 249 | <0.0001 |
| **C** | 217 | −0.29 | −0.41 | −0.16 | −4.47 | 215 | <0.0001 |
| B | 81 | −0.19 | −0.39 | 0.03 | −1.73 | 79 | NS |
| **High-compression group** | | | | | | | |
| **Q** | 91 | −0.49 | −0.63 | −0.32 | −5.31 | 89 | <0.0001 |
| **A** | 193 | −0.2 | −0.33 | −0.06 | −2.82 | 191 | <0.01 |
| M | 105 | −0.13 | −0.32 | 0.06 | −1.37 | 103 | NS |

CI, confidence interval; DF, degrees of freedom; NS, not significant.
Letters in bold indicate practitioners who had a significant correlation.

demonstrate no particular trend for the purposes of this study.

Characterisation of the client sample was important to ascertain variation of breast volume between practitioners. ANOVA was used to compare the volumes of the breasts imaged by each practitioner. A significant difference between the mean breast volume imaged by different practitioners was noted ($p<0.0001$). Further analysis of breast compression in relation to breast volume was undertaken. The low-, intermediate- and high-compression groups of practitioners were compared. Figures 2–4 illustrate the relationship between compression and breast volumes within the three practitioner subgoups. These graphs illustrate that there was wide variation in the compression used for any given breast volume, even for practitioners who used similar compression values. They do, however, all follow the same trend, which indicates that higher compression is applied with increasing breast volume. The slopes of regression lines in all three practitioner groups were similar: low compression, 1.45±0.18; intermediate compression, 1.44±0.08; and high compression, 2.22±0.31. However, each compression group had significantly

Figure 2. Correlation of compression and breast volume in practitioner group "low compression" (Practioners F, G, L and E).

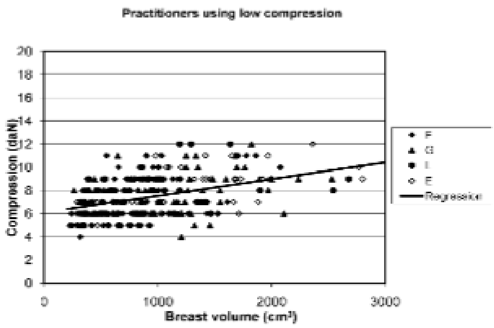
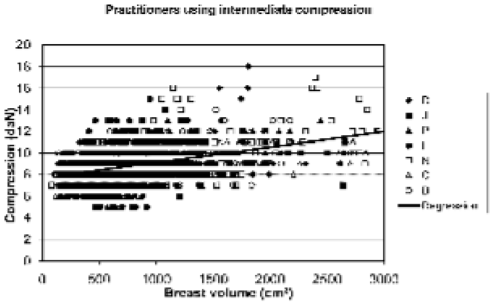


Figure 3. Correlation of compression and breast volume in practitioner group "intermediate compression" (Practitioners D, J, P, I, N, C and B).

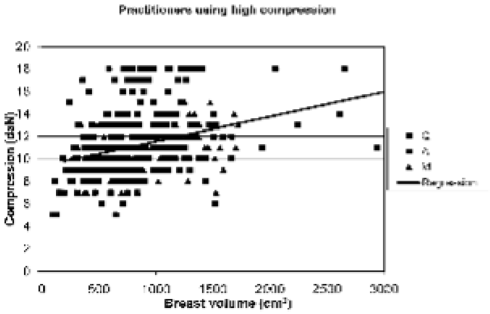



different intercepts (low, 6.1±0.2; intermediate, 7.7±0.1; and high, 9.3±0.3). The intermediate-compression group used an average of 1.6 daN more than the low-compression group, and the high-compression group used 1.6 daN more than the intermediate-compression group over all breast volumes.

## DISCUSSION

Factors which influence compression can be threefold. They can be attributed to client effects, practitioner effects and/or equipment effects. In 2004, a new perspective on breast compression was called for [7]; in turn, our group undertook preliminary work in order to establish whether practitioner variability did have cause to affect the amount of compression that is applied to breast tissue during mammography. Once any relationship has been recognised, linking practitioner variation with client and equipment variables will be essential in order to establish consistency within the NHSBSP.

Figure 4. Correlation of compression and breast volume in practitioner group "high compression" (Practitioners Q, A and M).




The main limitations of this study (retrospective and consecutive sampling) have been highlighted. Factors such as socioeconomic/educational status, breast tenderness and tolerance of compression could not be assessed owing to the nature of the sample. We acknowledge that some of these factors could affect the amount of compression applied by practitioners and in future prospective studies this information would be taken into account.

Consistency in the application of breast compression for females attending NHSBSP mammography is important to maintain high standards of image quality throughout the programme [6,7]. Within our study, compression used by practitioners was analysed for different BI-RADS densities and breast volumes to ascertain whether any relationships existed. Neither showed consistancy for all practitioners, although it clearly identified three distinct groups by compression means: those using low, intermediate and high compression. A relationship has been demonstrated between compression and BI-RADS density evaluation, with lower compression being applied to higher BI-RADS grades. Further research into this area is required.

This study shows that some practitioners perform similarly within themselves and against others; this does not, however, necessarily equate to good or bad practitioner practice. As the NHSBSP has rigorous processes for quality assurance and consistency for clients, this area may merit further research together with a focus on training process.

Practitioners in this study can be grouped into the low-, intermediate- and high-compression users. This may be of concern, given that this lack of consistency in the application of compression could have an impact on the consistency of image quality together with client experience over sequential attendences. The grade or experience of practitioners within the three groups did not have any statistical relationship to these findings and there appears to be no correlation between the experience (in number of years) of the practitioners or their grade.

For different BI-RADS catagories, some practitioners are consistent in their application of compression force while others are not. In clinical practice, such variation of compression application may be evidence of the practitioner adapting her technique to individual client characteristics and may not be a sign of inconsistent practice. This study did not assess client characteristics such as tolerance of compression. To address this deficiency, further prospective work would be required using a combination of quantitative and qualitative approaches.

## CONCLUSION

Several preliminary conclusions may be drawn from this research. Practitioners do not use the same mean compression when undertaking mammography and they can be grouped into low, intermediate and high compressors. There was a general tendency to apply less compression for higher BI-RADS grades, although this was only statistically significant in 6 out of 14 practitioners. Higher compression values were applied to breasts of larger volume. In addition to this, neither the experience nor the grade of the practitioners had any effect on their use of compression.

This study presents some insight into practioner variability for mammography and it is acknowledged that a combination of both client and practitioner effects on compression go hand in hand. Being preliminary in nature, this study had low client numbers. A larger sample from more imaging centres would be required to determine whether the findings demonstrated in this study could be replicated elsewhere. As a follow-up to this study, we have completed a single-centre longitudinal study of practitioner variability, in order to determine whether practitioners vary in their application of compression over time. This study will also demonstrate whether client compression values vary over sequential attendances. In conjunction with this, development of a breast phantom [21] and analysis of breast thickness readouts on a range of mammography machines [22] will lead to a new perspective on the analysis of practitioner, client and equipment variables in mammography imaging.

## REFERENCES


1. nhs.uk [homepage on the internet]. Leeds, UK: The NHS Information Centre for Health and Social Care; 2012 [accessed 22 March 2011]. Available from: www.ic.nhs.uk/searchcatalogue?productid=1472&dwd=Breast+cancer&sort=Relevance&size=10&page=1#top
2. Tabar L, Duffy SW, Vitak B, Chen H, Prevost TC. The natural history of breast carcinoma: what have we learned from screening? Cancer 1999;86:449–62.
3. Eklund GW. Mammographic compression: science or art? Radiology 1991;181:339–41.
4. Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? Australas Radiol 2003;47:121–6.
5. nhs.uk [homepage on the internet]. Sheffield, UK: NHS Cancer Screening Programmes; 2006 [accessed 22 March 2011]. Available from: www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp61.pdf
6. Spuur K, Hung WT, Poulos A, Rickard M. Mammography image quality: model for predicting compliance with posterior nipple line criterion. Eur J Radiol 2011;80:713–8. doi: 10.1016/j.ejrad.2010.06.026
7. Poulos A, McLean D. The application of breast compression in mammography: a new perspective. Radiography 2004;10:131–7.
8. Long SM. The handbook of mammography. 4th edn. Edmonton, AB: Mammography Consulting Services; 1990.
9. 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. Eur J Radiography 2009;1:66–72.
10. Bentley K, Poulos A, Rickard M. Mammography image quality: analysis of evaluation criteria using pectoral muscle presentation. Radiography 2007;14:189–94.
11. Chida K, Komatsu Y, Sai M, Nakagami A, Yamada T, Yamashita T, et al. Reduced compression mammography to reduce breast pain. Clin Imaging 2009;33:7–10. doi: 10.1016/j.clinimag.2008.06.025
12. nhs.uk [homepage on the internet]. Sheffield, UK: NHS Cancer Screening Programmes; 2006 [accessed 22 March 2011]. Available

from: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp63.pdf
13. Bragg J. The radiographers guide to film-screen mammography. Newcastle, UK: Magazine Associates; 1986.
14. Wentz V. Mammography for radiologic technologists. New York, NY: McGraw Hill; 1992.
15. skillsforhealth.org [homepage on the Internet]. Bristol, UK: Skills for Health; 2011 [accessed 30 December 2008]. Available from: http://www.skillsforhealth.org.uk
16. Department of Health. Research governance framework for health and social care. 2nd edn. London, UK: Department of Health; 2005.
17. Diffey J, Hufton A, Astley S, Mercer C, Maxwell A. Estimating individual cancer risks in the UK national breast screening programme: a feasibility study 2006. In: Krupinski EA, ed. IWDM 2008. LNCS, vol. 5116. Heidelberg, Germany: Springer; 2008. pp. 651–8.
18. D'Orsi CJ, Bassett LW, Berg WA. Breast Imaging Reporting and Data System: ACR BI-RADS-mammography. 4th edn. Reston, VA: American College of Radiology; 2003.
19. Swann CA, Kopans DB, McCarthy KA, White G, Hall DA. Mammographic density and physical assessment of the breast. AJR Online [serial on the internet]. March 1987 [accessed 30 March 2010];148:525–6. Available from: www.ajronline.org/content/148/3/525.full.pdf
20. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA. Accuracy of assigned BI-RADS breast density category definitions. Acad Radiol 2006;13:1143–9.
21. Smith HR, Smith J, Hogg P, Mercer C, Szczepura K. Elastically deformable anthropomorphic breast phantom for use in mammographic imaging research: 2011. In: Proceedings of the UK Radiological Congress 2011; 6–8 June 2011; Manchester, UK. London, UK: The British Institute of Radiology; 2011.
22. Hogg P, Hauge I H, Szczepura K, Connolly P, McGill G, Mercer C. An analysis of readout compressed breast thickness versus measured thickness on a range of mammography machines. Med Phys 2012;39:263. doi: 10.1118/1.3663579

## Paper IV

Mercer, C.E., Hogg, P., Kelly, J., Borgen, R., Millington, S., Hilton, B. … Whelehan, P. (2014). A mammography image set for research purposes using BI-RADS density classification. *Radiologic technology*, 85 (6), 609–613.

# A Mammography Image Set for Research Purposes Using BI-RADS Density Classification

Claire Elizabeth Mercer, MSc, BSc(Hons)
Peter Hogg, MPhil, PGC, BSc(Hons), FCR
Judith Kelly, MSc, DCR
Rita Borgen, MSc, FAETC
Sara Rachael Millington, PGC, BSc(Hons)
Beverley Hilton, BSc(Hons)
David Enion, MD
Patsy Whelehan, DCR, MSc, SRR(R)

**Purpose** Breast density categorization consistency is important when performing research, and minimization of interoperator and intraoperator variability is essential. This research aimed to validate a set of mammography images for visual breast density estimation to achieve consistency in future research projects and to determine observer performance.

**Methods** Using the Breast Imaging Reporting and Data System (BI-RADS) as the visual grading scale, 50 mammography images were scored for density grade by 8 observers.

**Results** Six of 8 observers achieved near-complete intraobserver agreement (kappa > 0.81). Strong agreement among observers (kappa = 0.61-0.8) was found in 10 of 28 paired observation episodes on the first iteration and 12 of 28 on the second. No observers demonstrated a delta variance above 1. Fleiss' kappa was used to evaluate concordance among all observers on the first and second iterations (first iteration, 0.64; second iteration, 0.56).

**Discussion** This research illustrates the difficulties of comparing observer visual performance scores because differences can exist when studies are repeated by and among individuals.

**Conclusion** We confirmed that the 50 images were suitable for research purposes. Some variability existed among observers; however, overall density classification agreement was strong. Future research should include repeating this study with digitally acquired images.

Breast density, which refers to the relative composition of glandular and fatty tissue, can be estimated from mammography images. Glandular tissue has a high density, whereas fatty tissue has a lower density. Breast density can be estimated using computerized methods or through visual analysis; both methods use the 2-D data.[1] Mammographic density estimation by visual analysis is subjective; therefore, it is important to identify and limit intraobserver and interobserver variability before using density data in research or clinical work.[2-4] Some imaging centers routinely assess breast density visually because density can be used as a predictor of risk for developing cancer.[1,5,6]

## Literature Review

If the density of the breast could be precisely measured, observer performance could accurately be evaluated against this true density using a visual method. Currently, it is impossible to precisely measure breast density, so observer accuracy cannot be judged with high confidence.[4,7] Computer-based approaches also can be prone to error, such as errors in machine-provided thickness readouts.[8] Therefore, it appears that at present, researchers can only judge whether observers using visual grading methods are obtaining acceptable performance on the grounds of observer consistency.

Various visual grading scales for density estimation for observers analyzing mammography images exist within the literature. These include Wolfe's 4-category system (now rarely used because of inconsistency issues)[9]; the Breast Imaging Reporting and Data System (BI-RADS) 4-point scale (A through D) developed by the American College of Radiology[10]; the percentage density score (visual analog scale[11]); and Boyd's 6-category classification.[12]

The visual analog scale is considered a reproducible and pragmatic way to estimate breast density and a recognized way to identify density on a continual scale. However, visual analog scales also have been reported to have problems, such as poor observer consistency.[7,11] Because breast density estimation lexicons have varied widely, breast density classification errors occur.[13] The BI-RADS lexicon was developed to standardize mammographic reporting and has recently been updated to catagories A, B, C, and D.[10]

The use of a standard set of mammography images as a tool for assessing observer performance of visual breast density estimation has previously been suggested.[7] A 2008 study conducted by Gao et al assessed interobserver and intraobserver reliability of visual mammographic density estimations.[3] The authors concluded that visual estimations of breast density were highly reproducible in research studies, assuming that appropriate training had been given to the observers. In 2011, Heine et al noted significant variations between observers in breast density estimation,[14] although a previous study by Ooms et al in 2007 noted substantial interobserver agreement.[15] Variation among observers is thus known and reported in the literature.

This article reports on the validation of a set of mammography images for visual breast density estimation for research purposes. Concurrently, we report on intraobserver and interobserver variability of experienced mammography staff. We routinely conduct research on various aspects of mammography, and the research often requires mammography images to be classified into BI-RADS categories. Therefore, it is necessary to determine observer performance in density classification prior to permitting an observer to participate in a study.

## Methods

The BI-RADS density grading scheme was selected as the visual grading scale for this research. The observers had experience using this scale, and the coauthors believed it to be suitable for film-screen mammography images. The accuracy of BI-RADS is limited because of its relatively broad categories and inherent subjectivity (eg, a small change in density can be difficult to detect).

Fifty film-screen mammography images, comprising left and right craniocaudal and mediolateral oblique views, were drawn from a university film library and identifying information was removed. Film-screen mammograms were used because they were readily available from the film library. We acknowledge that not using digital images was a study limitation because many breast imaging units have changed to digital imaging.

The image sets were scored by 8 observers who worked in the mammography field in 3 separate hospitals. Each set was allocated an identification number. Images were scored by each observer independently, under the same viewing conditions, and blinded to the findings of the other observers. To provide data to assess intraobserver variability, mammography image sets were scored (first iteration) and then rescored (second iteration) within an interval of at least 2 weeks to minimize recall bias.

Data analysis included within-observer variability (intraobserver variability) using Cohen's kappa and delta variance and between-observer variability (interobserver variability) using Cohen's kappa and Fleiss' kappa.[4,7] Cohen's kappa measures agreements between 2 observers; Fleiss' kappa measures the overall agreements among all the observers.

Identifying the level of agreement that is acceptable for research purposes is difficult, not only within the mammography setting but also in other research settings.[16] The baseline for acceptance was set at strong agreement or above (ie, 0.61). It also was established that the delta variance among observers should be 1 or lower.

## Results

Intraobserver variability was analyzed using weighted Cohen's kappa (see **Figure 1**). The following kappa scores showed agreement as follows:

- Less than 0.2 was considered poor agreement.
- 0.21 to 0.4 was considered fair.
- 0.41 to 0.6 was considered moderate.
- 0.61 to 0.8 was considered strong.
- Greater than 0.81 was considered near complete.

The values from this study for weighted kappa indicate near-complete intraobserver agreement for observers 1, 2, 3, 4, 5, and 6, with observer 8 showing strong agreement and observer 7 showing moderate agreement.

The delta variances within observers for 2 BI-RADS scores for the same images also were categorized. A delta of 0 indicated the same BI-RADS score on each iteration of the pair. A delta of 1 indicated a difference of 1 BI-RADS category between the first and second iteration. No observers demonstrated a delta variance above 1.

Repeated pairwise interobserver variability was analyzed by Cohen's weighted kappa (see **Figure 2**). The kappa scores showed agreement as follows:

- Less than 0.2 was considered poor agreement.
- 0.21 to 0.4 was considered fair.
- 0.41 to 0.6 was considered moderate.
- 0.61 to 0.8 was considered strong.
- Greater than 0.81 was considered near complete.

Interobserver variability was analyzed twice—once across the first set of scores given by the observers and once for the second set of scores—providing 56 (2 × 28) kappa values in total.

On the first iteration, of 28 observer correlations, 10 near-complete absolute agreements were found, 12 strong agreements, 5 moderate, and 1 fair. The second iteration returned 7 near-complete absolute agreements, 11 strong agreements, 8 moderate, and 2 fair. Each time an observer was paired with observer 7, the level of absolute agreement decreased to moderate or less.

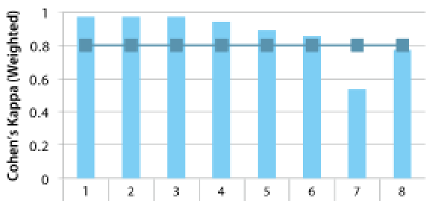
Fleiss' kappa was then used to evaluate concordance among all observers on the first and second iterations. Fleiss' kappa was 0.64 for the first iteration and 0.56 for the second. As previously noted, each time an observer was paired with observer 7, correlations decreased; when observer 7 was removed, Fleiss' kappa statistic rose to 0.77 for the first iteration and 0.65 for the second. This suggested that observer 7 should be extracted for the purpose of this analysis to set an acceptable baseline level of strong agreement or greater.

## Discussion

The primary purpose of this research was to validate a set of mammography images for use as a breast density assessment tool for observers prior to participation in our research. The baseline for acceptance for this purpose was set at strong agreement. Our data demonstrated that agreement was strong among all observers on the first iteration and moderate on the second iteration. After removing observer 7, the scores rose to strong on both iterations. We believe the baseline for acceptance of delta variance should be 1 or lower. This means there would be a difference of one BI-RADS category or less between the first and second observer scores. Within our study, none of the observers demonstrated a delta variance above 1. On this basis, we propose that our image set is suitable for determining whether an observer can participate in a research study that involves scoring BI-RADS density.

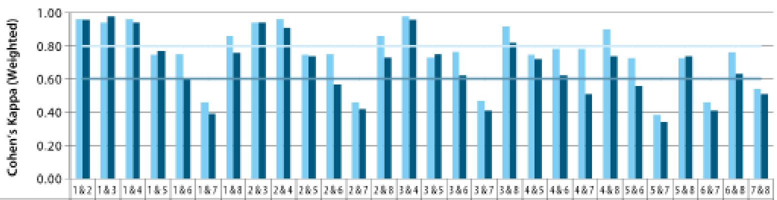
The purpose of this research also was to report intraobserver and interobserver variability. Six of 8 observers achieved near-complete intraobserver agreement. The 2 observers who did not have near-complete agreement had strong or moderate agreement. When the observer with moderate agreement was excluded, the agreement among the remaining 7 observers was near complete. This aspect of the discussion illustrates the difficulties of comparing observer visual performance scores because differences can exist when studies are repeated by and among individuals. However, our findings are similar to other perceptual studies in which visual appreciation of image quality was determined using a BI-RADS scale.[15]

As previously mentioned, a limitation of this study was that the images used were analog and not digital; therefore, the findings might not be applicable to future clinical and research practice. However, a large study of mammographic density comparing BI-RADS density between film-screen and digitally acquired images showed that




| Observer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Cohen's kappa weighted | 0.971 | 0.971 | 0.97 | 0.94 | 0.893 | 0.855 | 0.536 | 0.771 |
| Near-complete agreement line | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 |

**Figure 1.** *Intraobserver variance: Cohen's kappa (weighted).*




| | 1&2 | 1&3 | 1&4 | 1&5 | 1&6 | 1&7 | 1&8 | 2&3 | 2&4 | 2&5 | 2&6 | 2&7 | 2&8 | 3&4 | 3&5 | 3&6 | 3&7 | 3&8 | 4&5 | 4&6 | 4&7 | 4&8 | 5&6 | 5&7 | 5&8 | 6&7 | 6&8 | 7&8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cohen's kappa weighted (first score) | 0.96 | 0.94 | 0.96 | 0.75 | 0.75 | 0.46 | 0.86 | 0.94 | 0.96 | 0.75 | 0.75 | 0.46 | 0.86 | 0.98 | 0.73 | 0.76 | 0.47 | 0.92 | 0.75 | 0.78 | 0.78 | 0.90 | 0.73 | 0.38 | 0.73 | 0.46 | 0.76 | 0.54 |
| Cohen's kappa weighted (second score) | 0.96 | 0.98 | 0.94 | 0.77 | 0.6 | 0.39 | 0.76 | 0.94 | 0.91 | 0.74 | 0.57 | 0.42 | 0.73 | 0.96 | 0.75 | 0.62 | 0.41 | 0.82 | 0.72 | 0.62 | 0.51 | 0.74 | 0.56 | 0.34 | 0.74 | 0.41 | 0.63 | 0.51 |
| Strong agreement | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 | 0.61 |
| Near-complete agreement | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 |

Observer

**Figure 2.** *Interobserver variance: Cohen's kappa (weighted).*

regardless of image acquisition type, reported BI-RADS categories were similar.[17] Therefore, it could be argued that if observers demonstrate strong agreement when analyzing analog images, the same could be demonstrated when viewing digital images. However, we recommend repeating this study using the same methodology with digital images to establish whether observers have similar interobserver and intraobserver agreement levels compared with film-screen mammographic density assessment.

## Conclusion

We identified a set of images suitable for research purposes and assessed intraobserver and interobserver agreement characteristics for classifying mammographic breast density using the BI-RADS classification system. Our results demonstrated strong observer consistency. Future research should repeat this study using digitally acquired images.

*Claire E Mercer, MSc, BSc(Hons), is lead radiographer for breast imaging at the Nightingale Centre & Genesis Prevention Centre at the University Hospital of South Manchester in the United Kingdom. Concurrently, she is working on a doctoral degree at the University of Salford in the United Kingdom and is an honorary research fellow there. Her research work is centered on practitioner variation in breast compression, and she has published several papers in this area.*

*Peter Hogg, MPhil, PGC, BSc(Hons), FCR, is a professor of radiography and the associate head of research for the School of Health Sciences at the University of Salford.*

*Judith Kelly, MSc, DCR, is a consultant radiographer in breast imaging at the Countess of Chester Hospital NHS Foundation Trust in the United Kingdom.*

*Rita Borgen, MSc, FAETC, is a consultant radiographer in breast imaging at the East Lancashire Breast Screening Unit at Burnley General Hospital in the United Kingdom.*

*Sara R Millington, PGC, BSc(Hons), is an advanced practitioner in mammography at the Countess of Chester Hospital NHS Foundation Trust.*

*Beverley Hilton, BSc(Hons), is a mammographer at the East Lancashire Breast Screening Unit at Burnley General Hospital.*

*David Enion, MD, is a consultant radiologist at the East Lancashire Breast Screening Unit at Burnley General Hospital.*

*Patsy Whelehan, DCR, MSc, SRR(R), is a senior research radiographer in breast imaging at the Medical Research Institute at the University of Dundee in the United Kingdom.*

*The authors would like to thank Anne Eadie, MB, ChB, for sharing her work on this topic.*

*Reprint requests may be mailed to the American Society of Radiologic Technologists, Communications Department, at 15000 Central Ave SE, Albuquerque, NM 87123-3909, or e-mailed to communications@asrt.org.*

## References

1. Sergeant JC, Warwick J, Evans G, et al. Volumetric and area-based breast density measurement in the predicting risk of cancer at screening (PROCAS) study. In: *Lecture Notes in Computer Science.* Berlin, Germany: Springer-Verlag; 2012;7361:228-235.
2. Assi V, Warwick J, Cuzick J, Duffy SW. Clinical and epidemiological issues in mammographic density. *Nat Rev Clin Oncol.* 2012;9(1):33-40.
3. Gao J, Warren R, Warren-Forward H, Forbes JF. Reproducibility of visual assessment on mammographic density. *Breast Cancer Res Treat.* 2008;108(1):121-127.
4. Ciatto S, Houssami N, Apruzzese A, et al. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. *Breast.* 2005;14(4):269-275.
5. Astley S, Swayamprakasam C, Berks M, et al. Assessment of change in breast density: reader performance using synthetic mammographic images. In: *SPIE Medical Imaging.* Bellingham, WA: SPIE; 2012:8318.
6. Ting C, Astley SM, Morris J, et al. Longitudinal change in mammographic density and association with breast cancer risk: a case controlled study. In: Maidment ADA, Bakic PR, Gavenonis S, eds. *Lecture Notes in Computer Science.* Berlin, Germany: Springer-Verlag; 2012;7631:205-212.
7. Eadie A, Whelehan P, Baker L, Berg J, Evans A. Improving the validity of breast density assessment. *Breast Cancer Res.* 2011;13(suppl 1):24.
8. 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(1):263-271.
9. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer.* 1976;37(5):2486-2492.
10. Sickles, EA, D'Orsi CJ, Bassett LW, et al. ACR BI-RADS® Mammography. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA, American College of Radiology; 2013. http://www.acr.org/~/media/ACR/Documents/PDF/QualitySafety/Resources/BIRADS/01%20Mammography/02%20%2BIRADS%20Mammography%20Reporting.pdf. Accessed April 26, 2014.
11. Sergeant JC, Walshaw L, Wilson M, et al. Same task, same observers, different values: the problem with visual assessment of breast density. In: *SPIE Medical Imaging.* Bellingham, WA: SPIE; 2013:SPIE 8673:86730T.
12. Boyd NF, Byng JW, Jong RA, et al. Quantification classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *J Natl Cancer Inst.* 1995;87(9)670-675.
13. Boyer B, Canale S, Arfi-Rouche J, Monzani Q, Khaled W, Balleyguier C. Variability and errors when applying the BIRADS mammography classification. *Eur J Radiol.* 2013;82(3):388-397.
14. Heine JJ, Fowler EE, Flowers CI. Full field digital mammography and breast density: comparison of calibrated and noncalibrated measurements. *Acad Radiol.* 2011;18(11):1430-1436.
15. Ooms EA, Zonderland HM, Eijkemans MJ, et al. Mammography: interobserver variability in breast density assessment. *Breast.* 2007;16(6):568-576.
16. Ma WK, Hogg P, Norton S. Effects of kilovoltage, milliampere seconds, and focal spot size on image quality. *Radiol Technol.* 2014:85(5):479-485.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

## Paper V

Mercer, C.E., Hogg, P., Szczepura, K., & Denton, E.R.E. (2013). Practitioner compression force variation in mammography: A 6-year study. *Radiography,* 19 (3), 200–206. doi:10.1016/j.radi.2013.06.001

Radiography 19 (2013) 200–206

Contents lists available at SciVerse ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

# 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]

[a] University 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
[c] University of East Anglia and Norfolk & Norwich University Hospital, Norfolk & Norwich University Hospital Trust, Colney Lane, Norwich NR4 7UY, UK

ARTICLE INFO

*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.[1–5] 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

* Corresponding author. Tel.: +44 0161 291 4466.
*E-mail addresses:* claire.mercer@uhsm.nhs.uk (C.E. Mercer), p.hogg@salford.ac.uk (P. Hogg), k.szczepura@salford.ac.uk (K. Szczepura), erika.denton@nnuh.nhs.uk (E.R.E. Denton).

1078-8174/$ – see front matter © 2013 The College of Radiographers. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.radi.2013.06.001

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 ($p = 0.99$, $p = 0.70$, $p = 0.54$, 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 – 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 DMR+ 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 (cranial–caudal) and left/right MLO (medio-lateral oblique)). For each client the following information was recorded – 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]) – BI-RADS® 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 – 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.

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 – the practitioners had similar compression force means as the previous study[9] (rank sum correlation coefficient = 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

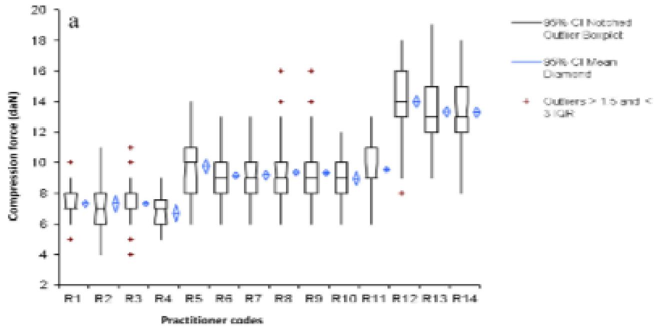



**Figure 1a.** Compression force variation (daN) on MLO mammography images per practitioner.

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.

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 BI-RADS 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

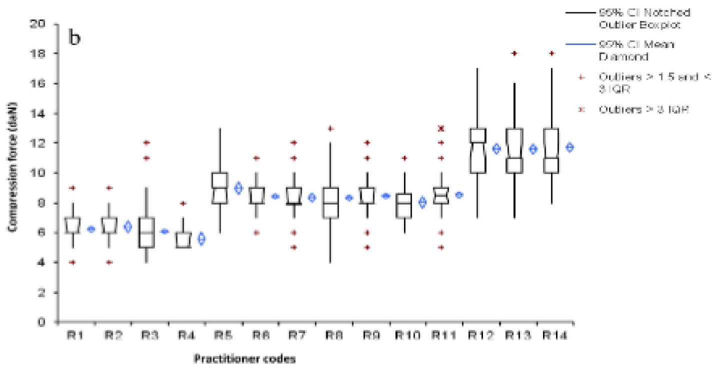



**Figure 1b.** Compression force variation (daN) on CC mammography images per practitioner.

CHAPTER FIVE: PAPER V

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

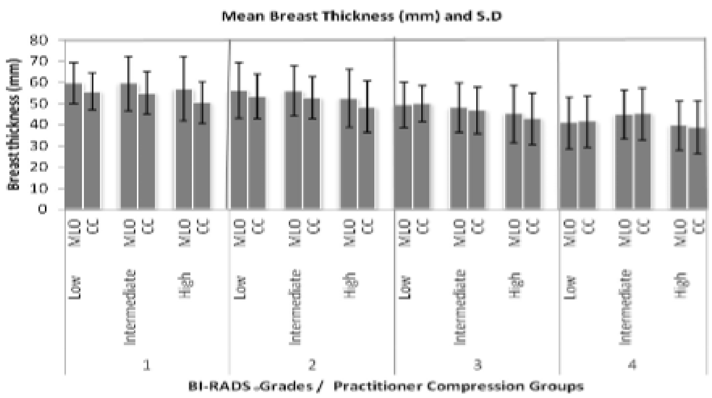



**Figure 2.** Mean breast thickness (mm) and SD within BI-RADS® grades and practitioner group.

CHAPTER FIVE: PAPER V

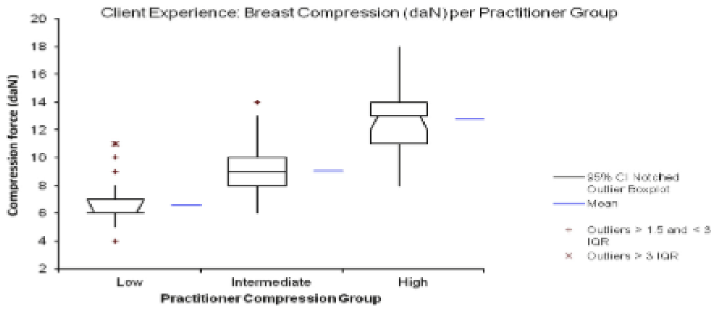



**Figure 3.** Breast compression force changes (daN) for clients imaged by different practitioner compression group – time independent.

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.

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.

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.

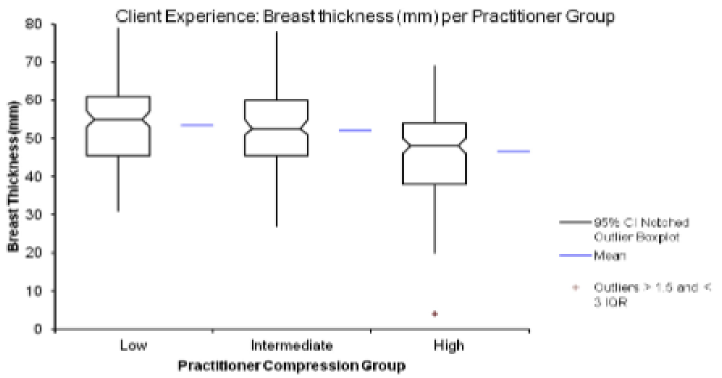



**Figure 4.** Breast thickness changes (mm) for clients imaged by different practitioner compression group – time independent.

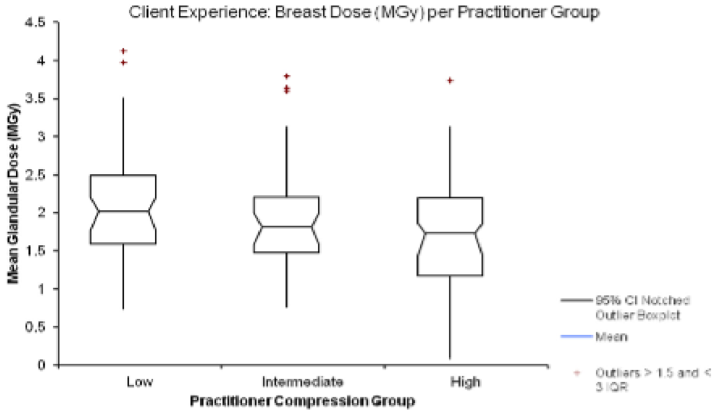



**Figure 5.** Breast dose changes (mGy) for clients imaged by different practitioner compression group – time independent.

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 – 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[14–18] 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**:21–53.
5. Poulos A, McLean D. The application of breast compression in mammography: a new perspective. *Radiography* 2004;**10**:131–7.
6. Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? *Australasian Radiology* 2003;**47**:121–6.
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**:10–5. 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 Imaging* 2009;**33**:7–10.
15. Sullivan DC, Beam CA, Goodman SM, Watt DL. Measurement of force applied during mammography. *Radiology* 1991;**181**:355–7.
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**:66–72.
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**:1773–83.
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**:574–7. 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 VI

Mercer, C.E., Hogg, P., Cassidy, S., & Denton, E.R.E. (2013). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography*, 19 (4), 363–365. doi:10.1016/j.radi.2013.07.002

Radiography 19 (2013) 363–365

Contents lists available at SciVerse ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Technical note

# Does an increase in compression force really improve visual image quality in mammography? — An initial investigation

C.E. Mercer [a,*], P. Hogg [b], S. Cassidy [c], E.R.E. Denton [d]

[a] The Nightingale Centre & Genesis Prevention Centre, University Hospital South Manchester, Manchester M23 9LT, UK
[b] Directorate of Radiography, University of Salford, UK
[c] School of Health Science, University of Salford, UK
[d] Department of Health, University of East Anglia and Norfolk & Norwich University Hospital, UK

ARTICLE INFO

Article history:
Received 23 December 2012
Received in revised form
7 February 2013
Accepted 2 July 2013
Available online 21 July 2013

Keywords:
Mammography
Breast compression
Image quality

ABSTRACT

*Objective:* Literature speculates that visual image quality (IQ) and compression force levels may be directly related. This small study investigates whether a relationship exists between compression force levels and visual IQ.

*Method:* To investigate how visual IQ varies with different levels of compression force, 39 clients were selected over a 6 year screening period that had received markedly different amounts of compression force on each of their three sequential screens. Images for the 3 screening episodes for all women were scored visually using 3 different IQ scales.

*Results:* Correlation coefficients between the 3 IQ scales were positive and high (0.82, 0.9 and 0.85). For the scales, the IQ scores their correlation does not vary significantly, even though different compression levels had been applied. Kappa IQ scale 1: 0.92, 0.89, 0.89. ANOVA IQ scale 2: $p = 0.98$, $p = 0.55$, $p = 0.56$. ICC IQ scale 3: 0.97, 0.93, 0.91.

*Conclusion:* For the 39 clients there is no difference in visual IQ when different amounts of compression are applied. We believe that further work should be conducted into compression force and image quality as 'higher levels' of compression force may not be justified in the attainment of suitable visual image quality.



## Introduction

31% of breast cancers are detected through the National Health Service Breast Screening Programme (NHSBSP).[1] Early detection of breast cancer through mammography reduces mortality and its success is said to be dependent upon the production of consistently high quality images.[2,3] It has been estimated that 10%–30% of cancers are missed through poor quality mammograms,[4] with one explanation being inadequate compression force. Literature therefore speculates that visual image quality (IQ) and compression force levels may be directly related.[5–7] Examples of aspects of image quality which are said to improve with increasing compression force include: separation of internal breast structures; better visualisation of skin surface and better visualisation of cancer lesions. However, no robust empirical study has been published to affirm or refute this relationship. Consequently, our small study investigates whether a relationship exists between compression force levels and visual IQ.

## Materials/method

This study investigates how visual IQ varies with different levels of compression force within the same woman (client) over a 6 year period. Due to radiation risks clients could not be imaged 3 times under different compression force levels. This study therefore reviews images from clients with 3 consecutive screening mammograms within the NHSBSP. The clients were drawn from a retrospective single centre longitudinal study[8] that had been imaged sequentially over a 6 year period; for each client this represented 3 mammography screens. Audit approval to conduct this study was granted from the Hospital Trust.

* Corresponding author.
E-mail addresses: claire.mercer@uhsm.nhs.uk, mercer.claire@gmail.com (C.E. Mercer), p.hogg@salford.ac.uk (P. Hogg), S.Cassidy@salford.ac.uk (S. Cassidy), erika.denton@nnuh.nhs.uk (E.R.E. Denton).

1078-8174/$ — see front matter © 2013 The College of Radiographers. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.radi.2013.07.002

*Client selection*

Progressing from a single centre study focussing on compression,[9] the retrospective single centre longitudinal study[8] analysed 500 clients' mammography images for compression and breast thickness values. All images were from one analogue mammography machine using the same film–screen combination.

Exclusion criteria for these clients included: the inability to tolerate compression force, breast pain, previous breast surgery, breast implants or cysts/abscesses, disabled clients, clients with arm/shoulder movement limitations or change in breast density over the six year period. Inclusion criteria included: the same BI-RADS density over the three screens together with no change in breast volume, as denoted by a change in film size or a visually observable change in breast size on the film. Within this study,[8] 39 clients received markedly different amounts of compression force on each of their three sequential screens; on each screen these clients were placed in a different 'compression force group' which were statistically significantly different to each other. Consequently these 39 clients were identified for inclusion within the current study to determine whether a relationship exists between compression force levels and visual measures of IQ.

The 39 clients' breast density was assessed by BI-RADS density grade: BI-RADS 1: 23.1%. BI-RADS 2: 48.7%. BI-RADS 3: 20.5%. BIRADS 4: 7.7%. We did not attempt to sample clients that represented a larger population; the sample was therefore stratified and convenience.

For the 39 clients, on each successive screen they had four images - left and right medio lateral oblique (LMLO, RMLO) and left and right cranio-caudal (LCC, RCC). Across the three screens an individual client experienced a maximum compression force variation of 12 daN (daN) in the MLO views and 11 daN in the CC views. Table 1 illustrates the compression force data for the 39 clients across three consecutive screens.

*Assessment of visual image quality*

Images for the 3 screening episodes for all women were scored visually using 3 different IQ scales. In all cases the scorer was blinded to IQ scores from the other scales, the level of applied compression force and the breast thickness.

The first scale (PGMI (IQ1)) was developed from the descriptors set out in the Pritchard Report in 1989[10] and categorises mammography images as (P) Perfect, (G) Good, (M) Moderate and (I) Inadequate. The second scale (IQ2)[11] comprised of 24 items and it had acceptable internal reliability (cronbachs alpha 0.975, 0.806 and 0.853, for mammography image sets of known quality). It contained three content domains - positioning quality (cronbach's alpha: 0.852), image quality (cronbach's alpha: 0.862) and breast composition (cronbach's alpha: 0.876). The third scale (IQ3) was a 5 point scale[12] which evaluates images on the basis of exposure, contrast and sharpness.

IQ2 and IQ3 scales gave numerical results; the PGMI scale (IQ1) was not numerical and consequently a numerical sequence was assigned to allow for statistical analysis. IQ scores for all the scales were calculated by summing the values from each of scale items.

An experienced mammographer scored all the images. It is recognised that the assessment of visual IQ can be subjective[13] and because of this the mammographer evaluated a subset of 30 images along with 4 experienced mammography image readers to assess variability. For the mammographer this demonstrated minimal intra-observer variability (within 95% limits of agreement Bland-Altman[14]) and for all 5 it revealed minimal inter observer variability (Pearsons' correlation <0.95).[15] The mammographer assessed the images at a similar time of the day under the same viewing conditions for a maximum individual evaluation period of 2 h (to minimise fatigue). The viewing conditions were those used clinically when reporting mammogram images (mammography light film boxes with dimmed ambient lighting).

## Results

Spearmans' rank correlation was performed between total scale scores for each of the 3 scales. Correlation coefficients between the scales were positive and high (0.82, 0.9 and 0.85). It should be noted that each of the 3 scales measure different attributes of visual image quality, as indicated in the method.

Visual IQ scores for the 39 clients are displayed in Fig 1. The *X* axis illustrates the 3 compression force levels (low, intermediate and high) and the 3 IQ scales. The *Y* axis demonstrates the IQ scale results. All 3 assessments gave similar results for low, intermediate and high levels of compression force.

Statistical tests were then applied as shown in Table 2. For all 3 visual IQ scales, the IQ scores do not vary significantly, even though different compression force levels have been applied.

## Discussion

We have found that IQ1 (PGMI) and IQ3 (5 point scale) perform in a similar way to the scale with known psychometric properties (IQ2).

For our 39 cases, the data indicates that compression force values have little or no impact on visual IQ. This suggests that mammography imaging could be performed with less compression force, whilst still achieving comparable IQ to that achieved at higher compressions force levels. However, it should be remembered that lesion visibility may also be improved by effective compression force. The impact of compression force levels on lesion visibility is not considered in this paper and further research will be needed to assess this.

**Table 1**
Compression force data for 39 clients across three consecutive screening mammograms.

| Compression force (daN) | CC view | | | MLO view | | |
|---|---|---|---|---|---|---|
| | Low | Intermediate | High | Low | Intermediate | High |
| Minimum level applied | 4 | 7 | 9 | 4 | 7 | 10 |
| Maximum level applied | 9 | 12 | 18 | 9 | 12 | 18 |
| Mean | 6 | 8.4 | 11.9 | 7.2 | 9.6 | 13.6 |
| Median | 6 | 8 | 12 | 7 | 10 | 14 |
| 95% CI | 5.8 to 6.2 | 8.1 to 8.7 | 11.5 to 12.3 | 6.9 to 7.5 | 9.3 to 10 | 13.1 to 14.1 |
| S.D. | 1 | 1.4 | 1.9 | 1.3 | 1.5 | 2.1 |
| | Low/Intermediate | Intermediate/High | Low/High | Low/Intermediate | Intermediate/High | Low/High |
| T-Tests | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 |
| ANOVA | <0.01 | <0.01 | <0.01 | <0.01 | 0.011 | 0.011 |

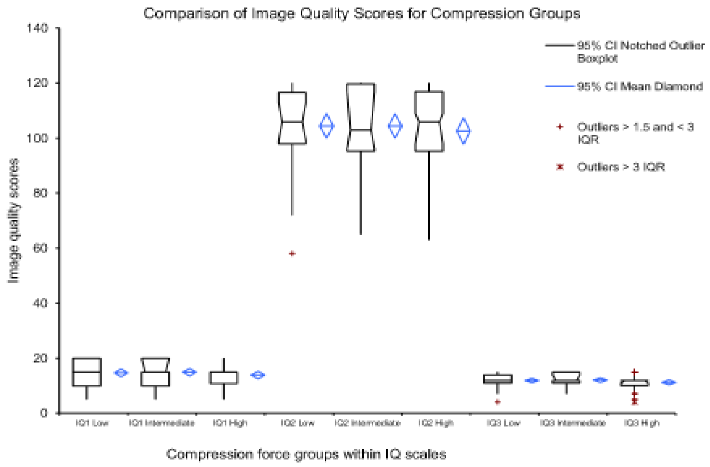



**Figure 1.** Image quality scores for compression force groups.

**Table 2**
Image quality statistical test results for compression force groups.

| | Low/Intermediate | Intermediate/High | Low/High |
|---|---|---|---|
| IQ 1 (KAPPA) | 0.92 | 0.89 | 0.87 |
| IQ 2 (ANOVA) | $P = 0.98$ | $P = 0.55$ | $P = 0.56$ |
| IQ 3 (Inter Class Correlation) | 0.97 | 0.93 | 0.91 |

We acknowledge that our study had some limitations - it is a retrospective study from a small client sample (39) and only assessed analogue images. Together with this the analysis of IQ, in relation to breast compression force, within the digital image setting is considered essential.

## Conclusion

For 39 cases, we have determined that there is no significant difference in visual IQ when different amounts of compression force are applied. On this basis we speculate that it may not be necessary to use high levels of compression force when lower amounts may suffice. Further research is needed to confirm this finding. Further research is also required to determine the relationship between digital image quality and compression force, together with determining any associations between breast compression force and lesion visibility.

## Conflict of interest

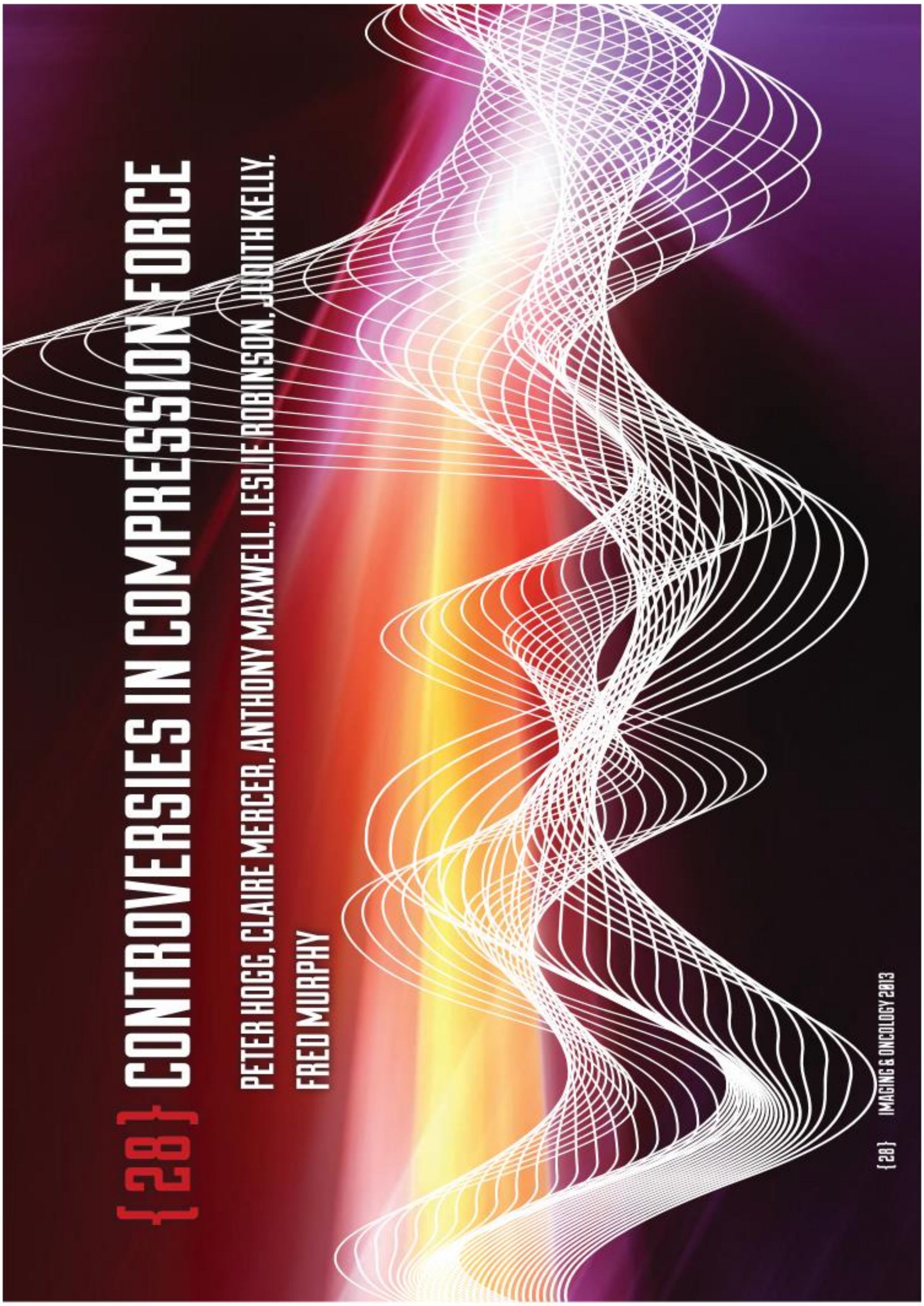
We can confirm that there are no financial and personal relationships with other people or organisations that could inappropriately influence (bias) our work.

## References


1. nhs.uk; 2012 [homepage on the internet]*The NHS information centre for health and social care*. Available from: http://www.ic.nhs.uk/statistics-and-data-collections/screening/breast-screening; 2012.
2. 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**:449–62.
3. Eklund GW, Cardenosa G. The art of mammographic positioning. *Radiologic Clinics of North America* 1992;Vol 30(1):21–53.
4. Majid AS, Shaw de Paredes E, Doherty D, Sharma NR, Salvador X. Missed breast carcinoma: pitfalls and pearls. *Radiographics* 2003;**23**(4):881–95.
5. Sullivan DC, Beam CA, Goodman SM, Watt DL. Measurement of force applied during mammography. *Radiology* 1991;**181**:355–7.
6. Helvie MA, Chan HP, Adler DD, Boyd PG. Breast thickness in routine mammograms: effect on image quality and radiation dose. *AJR* 1994;**163**:1371–4.
7. Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? *Australasian Radiology* 2003;**47**(2):121–6. PMID: 12780439.
8. Mercer CE, Hogg P, Szczepura K, Denton ERE. Practitioner compression force variation in mammography: a 6 year study. *Radiography* 2013;(19):200–6. http://dx.doi.org/10.1016/j.radi.2013.06.001.
9. Mercer CE, Hogg P, Lawson R, Diffey J, Denton ERE. Practitioner compression force variability in mammography: a preliminary study. *Breast Journal of Radiology* 2013;**86**:20110596.
10. NHSBSP (1989). *Quality assurance guidelines for mammography. Pritchard Report*. Oxford: NHS Breast Screening Programme; 1989.
11. Hogg P, Moran B, Cassidy C. *The development of a psychometric scale for assessing visual image quality in mammography*. UK Radiological Congress; 2012. p. 138.
12. NHSBSP. *Guidance notes for the evaluation of mammographic equipment: Publication number 51. Equipment evaluation sub-group of the National Co-ordinating Group for Equipment*. UK: NHS; 2002.
13. Bentley K, Poulos A, Rickard M. Mammography image quality: model for predicting compliance with posterior nipple line criterion. *Radiography* 2007;**14**:189–94.
14. Dewitte K, Fierens C, Stockl D, Thienpont LM. Application of the Bland-Altman plot for interpretation of method-comparison studies: a Critical investigation of its practice. *Clinical Chemistry* 2002;**48**:799–801.
15. Mercer CE, Hogg P, Diffey J. *The impact of breast compression on mammographic image quality: initial findings*. UK Radiological Congress; 2010. BJR supplement, p9, ISBN.10: 0-905749-69-3.

**Paper VII**

Hogg, P., Mercer, C., Maxwell, A., Robinson, L., Kelly, J., & Murphy F. (2013). Controversies in compression, *Imaging and oncology*, 28-36, ISBN 9871 871101581


{28} CONTROVERSIES IN COMPRESSION FORCE
PETER HOGG, CLAIRE MERCER, ANTHONY MAXWELL, LESLIE ROBINSON, JUDITH KELLY,
FRED MURPHY
{28} IMAGING & ONCOLOGY 2013

## For many years mammography has played a key cancer detection role in the assessment of screening and symptomatic populations. Controversy has surrounded screening[1] but new European work has re-emphasised its value[2].

A recent independent review within the UK indicated that screening does reduce breast cancer mortality, but over-diagnosis also occurs; it concluded by indicating that information for women should be more transparent and objective so that informed decisions can be taken[3]. For the moment we can assume that mammography will continue in widespread use, indeed its use is likely to increase because of the screening age extension[4] and the expectation that breast cancer incidence will increase considerably by 2040[5].

Given that mammography is well established, there is surprisingly little published empirical research into the technique of performing it. This is particularly true for the application of compression force. During the examination the breast is compressed to reduce breast thickness as this is said to have several values, including radiation dose reduction[6], enhancement of image quality[7] and lesion visibility, and minimisation of patient motion[8]. Literature exists to describe how compression force might be applied[9,10] and by what amount. Based on such literature, guidelines have been drawn up to inform practice[8]. However, most literature provides anecdotal or theoretical perspectives and few are based upon quality empirical observations. Perhaps a consequence of this is that compression force guidelines tend to be highly general and detailed information is lacking.

For the breast screening service as a whole, quality assurance is rigorous and it has been an integral part of the service since its introduction in 1988. All screening programmes have to adhere to stringent quality standards, with strict quality control procedures which are monitored by regional quality assurance reference centres[9]. Whilst such standards have rightly focused on patient outcomes, such as cancer detection rates and screening to result targets, the independent monitoring of practitioner standards is achieved through tight monitoring of programmes' technical repeat and recall rates. No direct quality standards are associated with practitioner compression force performance.

This article draws on new research, conducted in collaboration with the University of Salford, which has a particular focus on breast compression force.

### Variability in compression force

In 2004 Poulos and McLean[11] published a small research study which required two practitioners to compress a woman's breast independently. Compression force differences were noted between the two practitioners for the same woman. From this, Poulos and McLean predicted that variations between practitioners could exist, however further work needed to be conducted to verify this.

Their prediction was not confirmed until 2013, when Mercer et al[12] concluded that variability between practitioners does exist. Mercer's cross sectional study, involving 500 clients and 14 practitioners, identified three types of 'compressor' – low, intermediate and high, with significant compression force differences existing between these groups. Following on from this, Mercer et al[13] reported on a six year longitudinal study, over three screening rounds. A different patient cohort was selected for this study but the practitioners remained the same. The 'compressor' trend

> **"There is a subset of women who are permanently put off reattending after their first screen"**

was sustained and implications for clients were determined. For the same client across the three screening rounds, considerable compression force variations existed, breast thickness differences were noted and variations in mean glandular dose also occurred. In both studies it was suggested that the amount of applied compression force was more dependent upon the practitioner than the client and this might be due to the poor quality of published evidence and the lack of detailed clinical guidelines. Not surprisingly, Mercer et al concluded by suggesting that better evidence needed generating so that guidelines can be more explicit and image quality and patient experience would be more consistent.

## Client experience

Mercer's work leads us to question what the attitude of the service user (eg client) might be to practitioner compression force variability. Women find compression force uncomfortable and many studies report that a number of patients experience moderate to severe pain[14,15,16,17]. Although Myklebust et al[18] showed women tolerate pain as a necessary aspect of the examination for maximising diagnostic yield, Dibble et al[19] have shown up to 8% of women delayed or missed appointments because of the pain experienced at previous examinations. Clearly, breast compression force variability has the potential to influence women's behaviour and attitude towards mammography. Consequently, Robinson and Newton-Hughes[20] explored what practitioner variability might mean to National Health Service Breast Screening Programme (NHSBSP) service users. This study investigated how service users would interpret Mercer's practitioner variability work and whether this might influence their behaviour.

Employing a feminist research methodology, Robinson and Newton-Hughes' study comprised three focus group interviews involving 14 women. All participants indicated that Mercer's findings were interesting, although some of them were not surprised as they had experienced variation in compression force between visits themselves. One participant said she had even experienced variation in compression force between breasts during one visit. Another participant found the variability research quite alarming, suggesting it was unethical to compress the breast more than was necessary.

Despite a logical interpretation of Mercer's findings (eg that variation in practice might suggest high levels of compression force may not be necessary), only two participants said they felt empowered to change their behaviour at future mammography. The majority did not believe this new knowledge would influence their behaviour because they viewed the practitioner as the expert and to question them was inappropriate. It appears that the provision of evidence may not be sufficient on its own to change behaviour. Robinson and Newton-Hughes also showed that involving service users in evaluating research is essential, because it offers an alternate lens through which to consider the findings. Furthermore, because Robinson and Newton-Hughes employed a feminist approach, the participants were empowered to talk about what was important to them; it provided a more holistic understanding of how compression force was viewed and its relative importance in the context of mammography more generally. This way, like Poulos and Llewellyn[21], Robinson and Newton-Hughes showed that compression force related pain is not necessarily the chief concern for service users and other aspects of the screening experience could have a profound influence on discomfort.

A recent ongoing and unpublished study of breast screening reattendance rates has demonstrated that women who have been screened only once previously (and had a normal screening result) are six percentage points less likely to reattend than women who have been previously screened more than once. This suggests that there is a subset of women who are permanently put off reattending after their first screen, presumably because they view it as a painful or otherwise negative experience.

## Why does practitioner variability occur?

Aside from breast composition and volume changing over time, explanations have been offered that might explain why compression force variations within women might exist, a noteworthy one being breast tenderness in relation to the menstrual cycle[22]. However, Mercer's findings cannot be explained by breast tenderness, composition or volume change as her two studies clearly point to variability being practitioner-focused. Consequently, a national qualitative project to investigate the compression force behaviours of practitioners, and how they were taught to apply compression force, was conducted[23]. This project investigated compression force behaviours and explored individual and collective beliefs and values that influence compression force practice. Ultimately, it sought to identify the 'why' and 'how' of breast compression force. Concurring with Mercer, early findings showed variability on both how and why compression force is applied.

Many practitioners never referred to the numerical value of compression force being applied, but made a decision on the look and feel of the breast. In contrast, others used the compression force level as a final check rather than a primary assessment before making an exposure. The speed with which the compression force was applied also varied. Some mammographers tended to use the (fast) foot control until the final part of the process and then preferred the (slow) manual hand control. It was felt they could apply additional force whilst being sensitive to the patients' discomfort if this approach was adopted. Another method involved client empowerment; here control was given to the client for the final level of compression force. However, this was not always the case and some admitted to using white lies in order to attain that final level. It was recognised that clients would often compare their experience with a previous examination (if they had one) and staff were very mindful of this fact. Practitioners were concerned that a poor experience may result in a subsequent non-attendance.

Practitioners felt that for those clients attending for the first time, expectations were quite diverse but generally never as bad as they had anticipated. When speaking about compression force, practitioners demonstrated a good deal of self-doubt about their practice, and with no evidence of peer observations. They indicated it was impossible for them to know if they were performing mammography in the same manner as their colleagues. Values and beliefs specific to each centre were also evident, adding to the variability of the mammographic experience.

## Towards minimising variability

### Existing literature

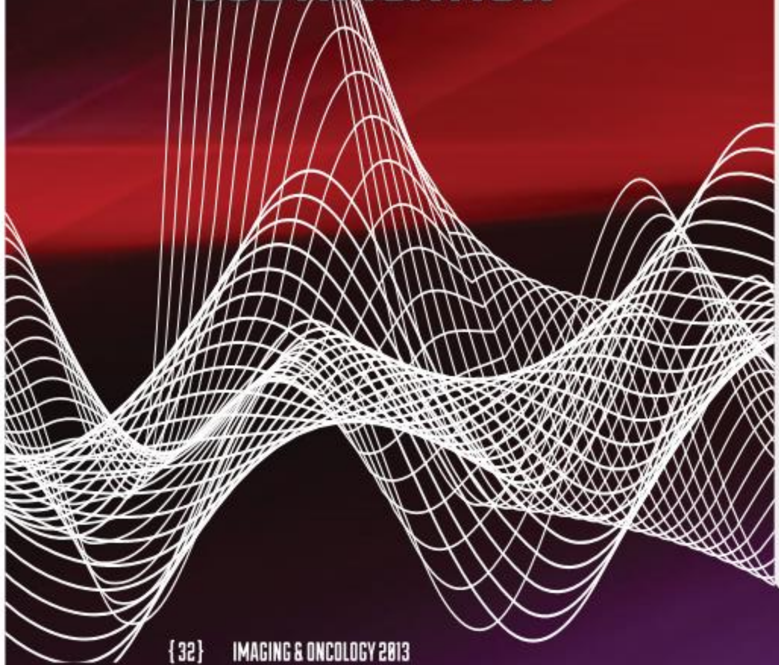
Various compression force guidelines and publications exist to help guide practitioners. For instance the NHSBSP indicates that the compression force should be applied slowly and gently to ensure the breast is held firmly in position and that 20kg (20daN) of force should not be exceeded. Generally speaking, these publications offer descriptive accounts with limited or no evidence-base and, not surprisingly, translating them into practice is likely to give rise to variations. However, they do provide a base upon which to build.

### Clinical image quality

Routine quality control tests are conducted on mammography equipment to ensure dose is consistent and adequately controlled and that mechanical safety is assured[9]. Ultimately these tests ensure that imaging and display equipment is fit for purpose. Beyond this the assessment of (visual) clinical image quality is the key indicator to determine the success of the mammogram. Clinical image quality should be considered in two ways – visibility of lesions and overall quality of the image. FROC (free-response receiver operating characteristic) analysis[24] would assist for the former and some form of visual grading would be ideal for the latter. Two visual grading scales have gained widespread application.

**"A ROBUST CONSISTENT ROUTINE METHOD FOR EVALUATING OVERALL CLINICAL IMAGE QUALITY DOES NOT EXIST"**

## "MANY TECHNIQUES USED IN RADIOGRAPHY DO NOT HAVE ADEQUATE JUSTIFICATION"



PGMI[25] categorises mammography images as (P) perfect, (G) good, (M) moderate and (I) inadequate; the second scale[26] evaluates images on the basis of exposure, contrast and sharpness. This second scale is used when evaluating images for the introduction of new equipment into the NHSBSP. Surprisingly neither of these scales has been validated – which brings into question the confidence with which they can be used. One scale has reached a level of development such that some validation data have been published[27], but this scale is not used in practice. The lack of published evidence about image quality scale performance severely confounds visual grading for research purposes. For clinical purposes mammograms tend to be reviewed for quality, based on simple checklists or by using one of the two scales that have gained popular use. A robust, consistent, routine method for evaluating overall clinical image quality does not exist.

### The relationship between compression force, breast thickness and clinical image quality

If the relationship between compression force and clinical image quality/ lesion visibility is true then it is important to understand how in-vivo breast thickness varies with increasing amounts of compression force so that a more informed decision can be taken about its application. Until recently no robust information has been published about this.

In 2013[28] Hogg et al published a study to describe the relationship between breast thickness and compression force. Knowing that data from different mammography units cannot be combined because of variability errors,[29] this research drew data from one machine (235 women; 940 compression force sets) for craniocaudal (CC) and mediolateral oblique (MLO) projections. Breast thickness/compression force change graphs were derived and the data demonstrated a good fit to polynomial equations. Differentiating the curves for gradients, critical junctures were determined at which the rate of change of pressure/thickness occurred. Through extrapolation the maximum amount of compression force which (theoretically) could be applied was identified. Interestingly, although this fell within the NHSBSP maximum compression force parameter, it was below the maximum. The point at which little thickness reduction was achieved for further increasing pressure was also identified, as illustrated in figure 1.

The graph is divided into zones: rapid rate of change (light grey) – here small amounts of compression force give rise to large thickness reductions; medium rate of change (mid grey) – here the rate of change is slower, indicating that more compression force is required for thickness reduction; and finally, slow rate of change (dark grey). The practitioner should aim to enter the mid grey zone, but not progress into the dark grey zone because little thickness reduction is achieved within that zone. Hogg et al[28] propose that the graphs could be used to help guide compression force practice on first presentation. For subsequent presentations the practitioner would consider the graphs, along with the recorded compression force values from the previous visit.

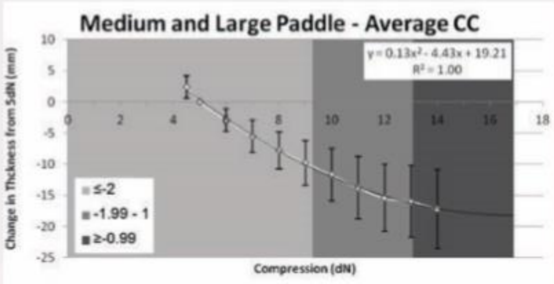



*Figure 1: Craniocaudal projection.*

## Relative position of inframammary fold and image receptor

Many techniques used in radiography do not have adequate justification and determining best practice can be difficult because of lack of quality evidence, notably empirical research. A clear example of this concerns how the image receptor (IR) might be positioned relative to the inframammary fold (IMF). Some literature supports breast elevation in relation to the IMF in order to perform the CC projection[30]. Theory indicates that IMF elevation would bring the object (lesion and breast) closer to the IR, and this should enhance image quality. Coincidentally, elevating the breast relative to the IMF might also balance the pressure exerted from the paddle and receptor onto the breast and improve the patient experience. Until recently, no empirical data within the literature affirms or refutes elevation of the breast relative to the IMF. Nevertheless, despite the lack of evidence, a variety of different breast positioning techniques have been taught for many years.

Using a breast phantom and pressure mapping system, work recently reported by Hogg et al[31] provided the first evidence to show that the amount of breast in contact with the IR is likely to increase as the receptor is elevated. Similarly the pressure from paddle and receptor onto the breast becomes better balanced when the receptor is elevated. Figure 2 demonstrates how 'pressure balancing and area on IR' (Uniformity Index) vary with relative IR and IMF positions. Hogg et al propose an elevation of 1-2cm in order to give the best pressure balance along with the largest amount of breast in contact with the receptor. However, caution should be exercised with this work because it is phantom-based, although work on women participants using the same method has recently been completed[32]. Here 16 female volunteers each received two compressions to each breast. The CC was positioned with IR at IMF, then 2cm above the IMF. Initial analysis indicates that in all cases raising the IR 1-2cm vertically relative to the IMF increases the area of breast in contact with the IR.

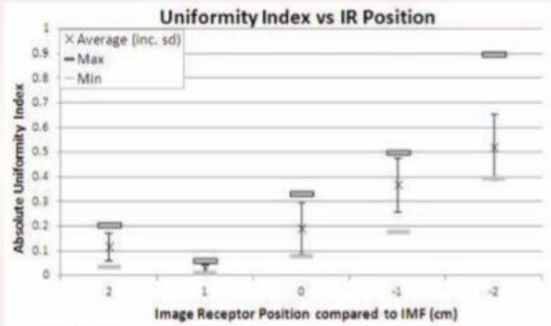



*Figure 2: Uniformity Index: pressure balancing versus area on IR.*

## Conclusion

Research has shown that compression force variability can exist between practitioners. On reflection it is apparent that this problem is likely to occur because there is almost no empirical evidence upon which to base compression force practice. This is quite surprising because mammography is a highly common procedure conducted on a large number of women per year through screening programmes or the investigation of symptoms.

Reflecting on the current state of evidence provided to assist practitioners in their endeavour to perform mammography to an acceptable and consistent level of practice, we suggest that considerably more research is required into the fundamentals of the technique.

## References


1. Mayor S, Mammography screening has little or no effect on breast cancer deaths, Swedish data indicate, British Medical Journal 2012; 345:e4847.
2. EUROSCREEN Working Group, Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet, Journal of Medical Screening, 2012; 19, Supplement 1, 5–13.
3. Independent UK Panel on Breast Cancer Screening, The benefits and harms of breast cancer screening: an independent review, The Lancet, 2012; 380, 9855, 1778-1786.
4. Moser K, Sellars S, Wheaton M, Cooke J, Duncan A, et al, Extending the age range for breast screening in England: pilot study to assess the feasibility and acceptability of randomization, Journal of Medical Screening, 2011; 18, 2, 96-102.
5. Breast cancer timebomb fear. http://www.dailymail.co.uk/health/article-2218077/Breast-cancer-NHS-faces-crisis-number-women-living-disease-set-treble-1-7m.html; Last accessed 8-1-2013.
6. Barnes G. Mammography equipment: compression, scatter control, and automatic exposure control. In: Haus A, Yaffe M, editors. Syllabus: a categorical course in physics. 1993. Oak Brook: RSNA Publications.
7. Tucker A K, Ng Y Y, Textbook of Mammography (2nd Edition). 2001, Churchill Livingstone, London, ISBN-13: 978-0443063404.
8. Kopans D, Breast Imaging, 3rd edition, 2007, Lippioncott Williams and Wilkins, Philadelphia, ISBN 13 978 0 7817 4768 4.
9. National Quality Assurance Coordinating Group for Radiography, National Health Service Breast Screening Programme (NHSBSP), Quality assurance guidelines for mammography including radiographic control, Publication no 63, 2006, NHS Cancer Screening Programmes, Sheffield, ISBN 1 84463 028 5.
10. Poulos A, Rickard M, Compression in mammography and the perception of discomfort, Australasian Radiology, 1997; 41, 3, 247-252.
11. Poulos A, McLean D, The application of breast compression in mammography: a new perspective, Radiography, 2004; 10, 2, 131-137.
12. Mercer CE, Hogg P, Lawson R, McGill G and Denton ERE, Practitioner Compression Force Variability in Mammography: a preliminary study, Br J Radiol, 2013; 000: 20110596 (in press).
13. Mercer CE, Hogg P, Szczepura K, Denton ERE and McGill G., A six year longitudinal study of compression force in mammography, UK Radiological Congress 2012, pg 15.
14. Rutter D R, Calnan M, Vaile M S B, Field S, Wade K A, Discomfort and pain during mammography: Description, prediction, and prevention. BMJ, 1992; 305(6851): 443-445.
15. Kashikar-Zuck S, Keefe F J, Kornguth P et al, Pain coping and the pain experience during mammography: A preliminary study. Pain, 1997, 73(2), 165-172.
16. Feig S A, Adverse effects of screening mammography. Radiologic Clinics of North America, 2004; 42(5), 807-819.
17. Miller D, Livingstone V, Herbison P, Interventions for relieving the pain and discomfort of screening mammography. Cochrane Database of Systematic Reviews, 2008, (1).
18. Myklebust A M, 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(2), 66-72.
19. Dibble S L, Israel J, Nussey B et al, Mammography with breast cushions. Women's Health Issues, 2005; 15(2), 55-63.
20. Robinson L, Newton-Hughes A, Feeling the Pressure, Research Seminar, 2012, University of Salford.
21. Poulos A, Llewellyn G, Mammography discomfort: a holistic perspective derived from women's experiences. Radiography, 2005; 11 (1), 17-25.
22. RadiologyInfo.org The radiology information resource for patients. Developed jointly by Radiological Society of North America and American College of Radiology. http://www.radiologyinfo.org/en/info.cfm?pg=mammo Last accessed 8-1-2013.
23. Murphy F, Nightingale J, Mackay S et al, Under Pressure, Research Seminar, 2012, University of Salford.
24. Free-response Receiver Operating Characteristic web site. http://www.devchakraborty.com/; Last accessed 8-1-2013.
25. Quality Assurance Guidelines for Mammography (Pritchard Report). Sheffield: NHS Breast Screening Programme, 1989 | ISBN 1 87199 771 2 | Archived http://www.cancerscreening.nhs.uk/breastscreen/publications/qa-01.html Last accessed 3.2.2013.
26. NHSBSP, 2002, Guidance notes for the evaluation of mammographic equipment: Publication number 51. Equipment evaluation sub-group of the National Co-ordinating Group for Equipment: NHS Cancer Screening Programmes Sheffield, UK, ISBN 1 871997 64 X, available from: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp51.pdf Last accessed 3.2.2013.
27. Hogg P, Moran B and Cassidy C, The development of a psychometric scale for assessing visual image quality in mammography, 2012, UK Radiological Congress, Manchester p138.
28. Hogg P, Taylor M, Szczepura K, Mercer CE, Denton ERE, Pressure and breast thickness in mammography – an exploratory study, Br J Radiol 2013;86:2012022.
29. Hauge I, Hogg P, Szczepura K et al, The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units, Medical Physics, 2012, 39, 1, 263-271.
30. Lee L, Strickland V, Wilson R, Roebuck E, Fundamentals of mammography, 1995, WB Saunders Comp. Ltd, London.
31. Hogg P; Szczepura K, Darlington A, Maxwell A, A method to measure paddle and detector pressures and footprints in mammography, Medical Physics, in press, 2013.
32. Smith H, Hogg P, Maxwell A, Mercer C and Szczepura K, An analysis of the compressed breast area and image receptor/compression paddle pressure balance in different mammographic projections. Poster, UKRC 2013.



Peter Hogg is Professor of Radiography, University of Salford
Claire Mercer is Lead Radiographer in breast imaging; Honorary Research Fellow, University of Salford.
Anthony Maxwell is a Consultant Radiologist; Regional QA Radiologist; Honorary Senior Research Fellow, University of Salford.
Leslie Robinson is a Senior Lecturer, University of Salford.
Judith Kelly is a Consultant Radiographer (breast); Honorary Senior Research Fellow, University of Salford.
Fred Murphy is a Senior Lecturer, University of Salford.

## Paper VIII

Mercer, C.E., Szczepura, K., Kelly, J., Millington, S.R., Hilton, B., & Hogg, P. (2014). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography*, Published online: July 29, 2014. doi:10.1016/j.radi.2014.07.004

ARTICLE IN PRESS

Radiography xxx (2014) 1–6

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

# 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 INFO

*Article history:*
Received 26 April 2014
Received in revised form
21 June 2014
Accepted 7 July 2014
Available online xxx

*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), 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.



## 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.

* Corresponding author.
*E-mail address:* mercer.claire@gmail.com (C.E. Mercer).

http://dx.doi.org/10.1016/j.radi.2014.07.004


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

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

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.

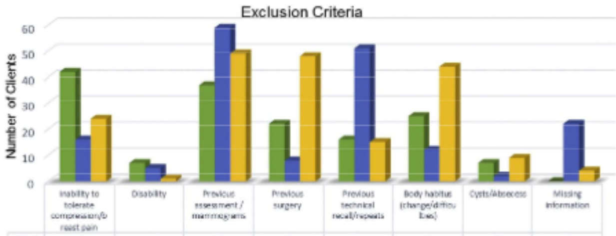
Breast density was established by 5 observers in the three screening units using the 4 point BI-RADS® scale (Breast Imaging Reporting and Data System)[17] – BI-RADS® 1 < 25% dense, BI-RADS® 2: 25%–50% dense, BI-RADS® 3: 51%–75% dense, and BI-RADS® 4 > 75% dense. In order to establish inter and intra 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 intra-observer agreement (Kappa >0.81) and strong or above inter-observer 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/Abscess | 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.

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

*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 (10–146); site two (10–155); site three (12–139). 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 BI-RADS® 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 BI-RADS® 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 1–4 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.

**Table 1**
Practitioner grade per site.

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

| 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 |

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 42N–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.

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

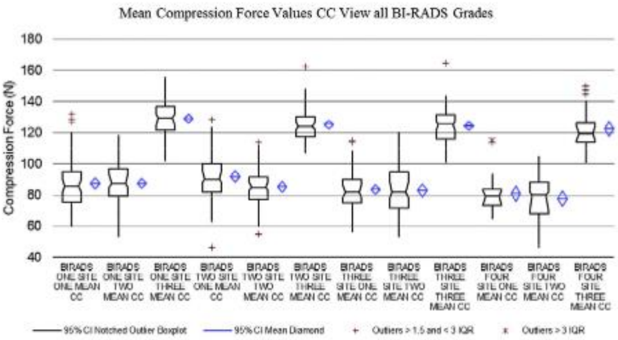



**Figure 2.** Mean compression force values CC view.

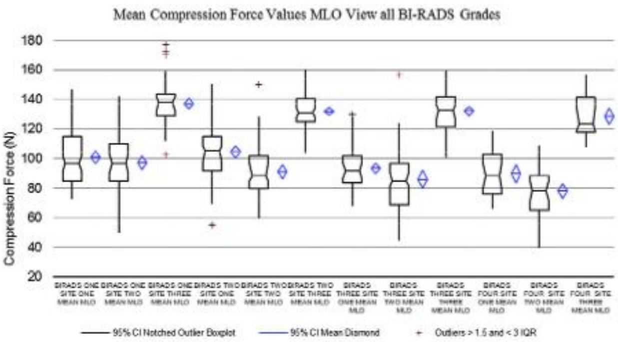



**Figure 3.** Mean compression force values MLO view.

## 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

**Table 3**
First and third quartile compression forces all sites.

| | | | 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 |

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

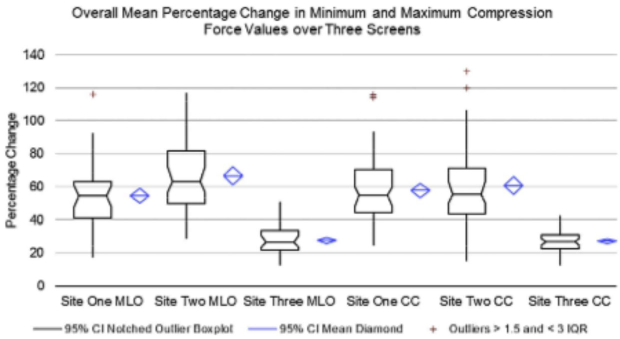



**Figure 4.** Overall mean percentage change in minimum and maximum compression force values over three screens.

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 study[19] 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, identified that practitioners vary in their approach to the application of compression force. This current study is therefore likely to reflect behaviour nationally.

*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 – especially when sequential imaging comparison is considered. On this basis site three might be considered superior for consistency and dose minimisation.

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

| 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 |

*National standards*

From this and prior research[13,14] there is a need for the NHSBSP to consider the introduction of national guidance on compression force levels. Hogg and colleagues[21] 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 colleagues[22] 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

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

| Site | First quartile MLO thickness (mm) | S.D | CC thickness (mm) | S.D | Third quartile 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 |

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

screening cycle ($p < 0.001$). Reducing breast thickness has potentially quantifiable reductions in radiation risks to clients within the screening programme.

*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

1. 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):1778–86.
2. Weller DP, Campbell C. Uptake in cancer screening programmes: a priority in cancer control. *Br J Cancer* 2009 Dec 3;**101**(Suppl. 2):S55–9. http://dx.doi.org/10.1038/sj.bjc.6605391.
3. 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):389–94.
4. 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].
5. NHS Breast Screening Programme. *Annual review published 2012*. ISBN 978-1-84463-093-6; 2012.
6. NHSBSP 33. *Quality assurance guidelines for medical physics services*. 2nd ed. June 2005. ISBN 1 84463 016 1.
7. Long S. *The handbook of mammography*. 4th ed. Edmonton: Mammography Consulting Services Ltd; 2000.
8. Wentz G. *Mammography for radiologic technologist*. New York: Mc Graw Hill; 1992.
9. Kopans D. *Breast imaging*. 3rd ed. Lippincott: Williams and Wilkins; 2007.
10. Eklund GW, Cardenosa G. The art of mammographic positioning. Breast imaging: current status and future directions. *Radiol Clin North Am* 1992;**30**:21–53.
11. Poulos A, McLean D. The application of breast compression in mammography: a new perspective. *Radiography* 2004;**10**:131–7.
12. 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.
13. Mercer CE, Hogg P, Szczepura K, Denton ERE. Practitioner compression force variation in mammography: a 6-year study. *Radiography* 2013;**19**:200–6.
14. 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.
15. *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.
16. NHSBSP 63. *Quality assurance guidelines for mammography*, **42**; April 2006. ISBN 1 84463 028 5.
17. 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.
18. 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**:609–13.
19. 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**:363–5.
20. Murphy FJ, Nightingale JM, Mackay SM, Robinson L, Seddon D, Hogg P. *Compression behaviours – an exploration of the beliefs and values influencing the application of breast compression during screening mammography*. UKRC; 2013.
21. Hogg P, Taylor M, Szczupera K, Mercer C, Denton E. Pressure and breast thickness in mammography—an exploratory calibration study. *Br J Radiol* 2013;**86**. 20120222.
22. 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).
23. 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.

Please cite this article in press as: Mercer CE, et al., A 6-year study of mammographic compression force: Practitioner variability within and between screening sites, Radiography (2014), http://dx.doi.org/10.1016/j.radi.2014.07.004

# CHAPTER SIX: RESEARCH DEVELOPMENT, PERFORMANCE AND IMPACT

An important factor in establishing collaboration, whilst being facilitated and empowered to progress research forward, was in the development of key relationships with University and multi-professional colleagues, which enabled progressive impact of this research into the mammography arena.

## 6.1 Development

In late 2010 the research team at the University had diversified their research programme into three key areas, one of which had a mammography focus. It was quickly ascertained that this work on compression force sat comfortably within this theme. The first project, centred on compression force practitioner variability, was established and this soon developed into a whole new research programme in the mammography arena for the University with both qualitative and quantitative approaches. This saw the creation of new research teams being formed to support this research in areas such as; compression paddle bend, paddle motion, image blurring, a focus on practitioner behaviours, emotional intelligence and the development of breast image phantoms. The formation of these new research directions and teams within the University was as a direct result of the introduction of the research contained within this thesis.

## 6.2 Research originality

The central ambition of this research was to provide new, evidence based knowledge with a possible resultant change in compression force application in mammography. There was an absence of evidenced based research within this arena and available guidelines would seem to allow for considerable variations to occur between practicing practitioners.

The only evidenced based research available, related to this research area, was conducted by Poulos and McLean [2004]. They called for a "..,new perspective on breast compression..," following conduction of a small scale study which concluded that large variations between practitioners could occur with the same client. In 2010, when commencing this research, no further work had been published in this field and practitioners had no evidence-based agreed guidelines to identify optimal compression force levels. It was clear, therefore, that this research field was novel.

## 6.3 Research collaboration

The researcher identified that collaboration and development was essential; strategic links were required with several multi-professional teams to establish key, specific outcomes (Table 6.1), these included Consultant Radiographers and Research Radiographers from other Hospital Trusts and links with other professionals within the University community in Psychology and Statistics. Several research teams were developed which worked concurrently, yet somewhat independently within their own research foci.

Given that breast screening directly, or indirectly, affects a large proportion of the population (1.97 million women screened in 2013 [Health and Social Care Information Centre, 2014], it was acknowledge by the research teams that this research could have widespread value, not just within the UK but with possible international reaches.

| University Professionals | Collaborations |
|---|---|
| Professor at University | Developed links and worked as part of a collaborative team on his project on Emotional Intelligence within Radiography |
| University Radiographic Lecturers (various) | Developed links to form the basis on new research themes within the University (multiple, with both qualitative and quantitative elements) |
| Programme Leader Applied Psychology | Forged links for the collaboration in Paper VI |
| Research Radiographer at University of Dundee | Forged links for the participation in Paper VIII and Paper IV |
| PhD student | Bought into the team at the University to work on the compression bend and distortion research (Paper I). Worked closely with this student in the initial stages of her work on the design and testing phase |
| PhD student | Bought into the team at the University to work on breast phantom design. Worked closely with this student at the beginning of her research in the design and testing phase |
| PhD student | Bought into the team to work primarily with breast research, from 2015 to mentor this PhD student |
| **Scientists/Physicists** | **Collaborations** |
| Consultant Clinical Scientist at Central Manchester NHS Trust | Developed new working relationship to assist in statistical knowledge. Collaborated on the first research on practitioner variability (Paper III) |
| Medical Physicist/Clinical Educator at Queensland Health | Continued and developed existing working relationship in the medical physics arena. Support during the first practitioner variation research (Paper III) |
| Senior Medical Physics Specialist at John Hunter Hospital, Australia | Continued and developed existing working relationship (Appendix 1, Paper A) in the medical physics arena. Provided support during the whole research work |
| **Consultant Radiographers** | **Collaborations** |
| Consultant Radiographer (First NHS Trust) | Developed discussions for Initial Research (Paper V) and then developed key relationship for support throughout this research and collaboration with Papers: IV, VI, VIII and the Mammography Academic Book |
| Consultant Radiographer (Second NHS Trust) | Forged links for participation in the Multicentre Research (Paper VIII) and also the BI-RADS research (Paper IV) |
| Consultant Radiographer (Third NHS Trust) | Working links established and the development of Pressure Map Research was as a result of the work within this thesis |
| **Radiographers** | **Collaborations** |
| Radiographer (First NHS Trust) | Assisted in the data gathering stage for the Multicentre Research (Paper VIII). Progressed into co-author in one book chapter of the Mammography Book |
| Radiographer (Second NHS Trust) | Assisted in the data gathering stage for the Multicentre research (Paper VIII). Progressed into co-author in one book chapter of the Mammography Book |
| Radiographer (Third NHS Trust) | Mentored throughout her MSc project and worked collaboratively with her on Paper II |

Table 6.1 Research teams

## 6.4 Overall research impact

Figure 6.1 represents the development of research in this field arising from the research theme within this thesis. It directly illustrates the impact of the thesis publications, not only in developing new research themes within the University, but in developing research interests outside the University and also outside of the UK. It is unequivocal that further development of research in this arena was a direct outcome from the mammography breast compression force research developed and contained within this thesis.

It is important for the practitioners when performing mammography that they gain compliance of the client with effective interactions throughout the process. The practitioner is also required to respond to the emotional and physical needs of the client to be able to produce a high quality image. Qualitative research has been established seeking to understand why practitioners behave the way they do when they apply breast compression force [Murphy F et al., 2014; Robinson et al., 2013] and to understand if there is a process that practitioners follow when applying compression. Both these projects developed as a direct result from the work contained within this thesis on practitioner variation.

Development in the Netherlands into a focus on pressure based compression instead of force based compression is rapidly progressing [De Groot et al., 2014]. The development of this pressure based research was directly influenced by the research outcomes of this thesis as they saw the requirement for standardisation. A new project is currently being established with the thesis author and this group of researchers to progress pressure based compression within the UK.

The culmination of this research is to be published in a new mammographic academic book (Springer expected published date 2015). The aims of which are to provide a single holistic and evidence-based publication to cover mammography and mammography based techniques; currently not present in the mammography field. The author of this thesis is one of three editors of this book, and also the author of one of the book chapters. The key themes from this thesis research are contained within the book; specifically the requirement for standardisation at a suggested 90-130N of force.

It is clearly demonstrated (Figure 6.1) that further projects are developing outside the University teams. This highlights the impact and significance of this research, illustrating the contextual impact for future clinical practice.

Code: Figure 6.1 -

> *Within Figure 6.1 the full lines indicate research developed within the University as a direct result from the research themes.*
>
> *The dashed lines indicates established research developed from the research themes outside the University.*
>
> *The numbers adjacent to the text refer to the publication details; these are detailed in Table 6.2.*

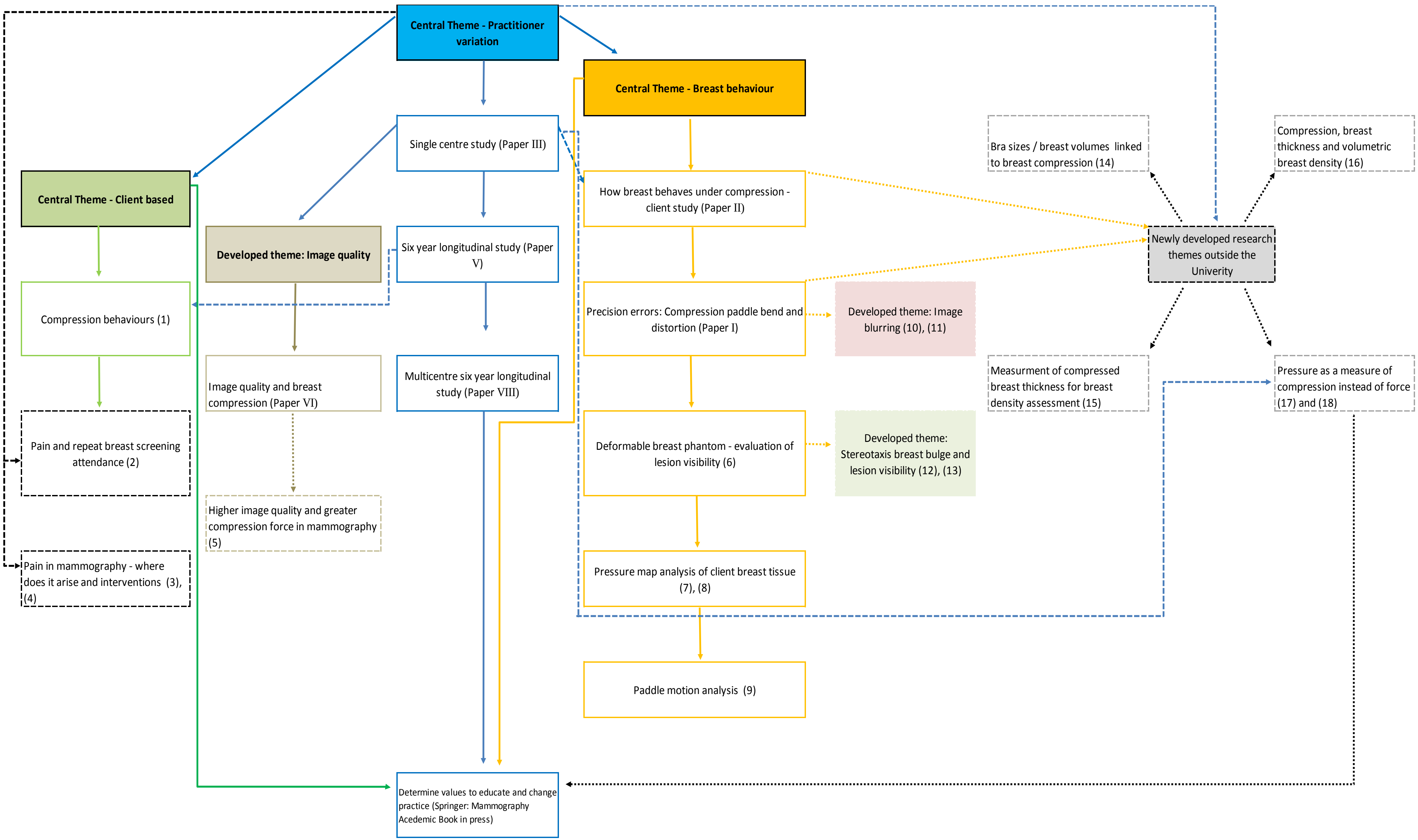



Figure 6.1: Research developments

| Publication details for Figure 6.1 |
|---|
| 1. Murphy, F., Nightingale, J., Hogg, P., Robinson, L., Seddon, D., & Mackay, S. (2014). Compression force behaviours: An exploration of the beliefs and values influencing the application of breast compression during screening mammography. *Radiography*, Published Online: June 13, 2014. doi:10.1016/j.radi.2014.05.009 |
| 2. Whelehan, P., Evans, A., Wells, M., & McGillivray, S. (2013). The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. *Breast (Edinburgh, Scotland)*, 22 (4), 389–394. doi:10.1016/j.breast.2013.03.003 |
| 3. O'Leary, D., & Al Maskari, Z. (2014). Pain in mammography: where and why does it arise? Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014 |
| 4. O'Leary, D., & Al Maskari Z. (2014). Silicon cushions: a pain reduction intervention for mammography. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014 |
| 5. O'Leary, D., & Rainford, L. (2014). Higher image quality and greater compression force in mammography are linked. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014 |
| 6. Smith, H., Smith, J. J., Hogg, P., Mercer, C., & Szczepura, K. (2011). *Elastically deformable anthropomorphic breast phantom for use in mammographic imaging research.* Paper presented at the UK Radiological Congress 2011. ISBN 10: 0-905749-72-3; ISBN 13: 978-0-905749-72-3 |
| 7 Smith, H., Hogg, P., Szczepura, K., Mercer, C., & Maxwell, A. (2013). *An analysis of compressed breast area and image receptor/compression paddle pressure balance in different mammographic projections*. Paper presented at the UK Radiological Congress 2013. |
| 8. Darlington, A., Hogg, P., Szczepura, K., & Maxwell, A. (2013). Optimising paddle and detector pressures and footprints in mammography. *Medical physics*, 40 (4), 041907. |
| 9. Kelly J., Hogg P., Millington S., Sanderud A., Wilcock C., McGeever G., … Kelly S. (2012). *Paddle motion analysis preliminary research*. Paper presented at the Symposium Mammographicum 2012. ISBN 10:0-905749-77-4. ISBN 13:978-0-905749-77-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2012 |
| 10. Hogg, P., Szczepura, K., Kelly, J., & Taylor, M. (2012). Blurred digital mammography images. *Radiography*, 02/2012, 18(1), 55–56. doi: 10.1016/j.radi.2011.11.008 |
| 11. Ma, W.K., Hogg, P., Kelly, J., & Millington, S. (2014). A method to investigate image blurring due to mammography machine compression paddle movement. *Radiography*, Published Online: June 21, 2014. doi:10.1016/j.radi.2014.06.004 |
| 12. Williams, S., Hackney, L., Hogg P., Szczepura K. Breast tissue bulge and lesion visibility during stereotactic biopsy – A phantom study. Radiography. Volume 20, Issue 3, Pages 271–276, August 2014 |
| 13. Williams, S., Hackney, L., Hogg P., & Szczepura K. (2013). Tissue bulge during stereotactic core biopsy. *Radiography*, 19, (4) 366–368. |
| 14. O'Leary, D., & Rainford L. (2014). Patient bra size as a gauge of compression required in mammography. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014 |
| 15. Hewes, H., Williamson, A., Noonan, P., Sergeant, J.C., Dunn, T., Haste, S. … Astley, S. (2013). *Measurement of compressed breast thickness for breast density assessment using a games console input device*. Paper presented at the 6th International Breast Densitometry and Breast Cancer Risk Assessment Workshop 6-7 June 2013 San Francisco, USA. Retrieved from: http://dx.doi.org/10.1594/ecr2013/C-2199 |
| 16. Khan-Perez J., Mercer C., Bydder M., Sergeant J., Morris J., Maxwell A., … Astley S. (2013). Breast compression, compressed breast thickness and volumetric breast density. *Breast cancer research*, 15 (1). |
| 17. De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. *Medical physics*, 40 (8), 081901. doi:10.1118/1.4812418 |
| 18. De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2014). Mammographic compression after breast conserving therapy: controlling pressure instead of force. *Medical physics*, 41 (2), 023501. doi:10.1118/1.4862512 |

Table 6.2: Publication details for research developments

## 6.5 Research impact overview

The research for this thesis was completed between 2010 and 2014, with peer reviewed publications between 2012 and 2014. Posters, proffered papers and invited speaker papers at conferences commenced in 2010 in order to promote the research topic and stimulate peer and expert debate within this research arena.

### 6.5.1 Conference proceedings

Enthusiasm for the research topic was rapidly fortified; the first initial presentation within a United Kingdom (UK) conference on the subject in 2010 progressed into developed proffered posters papers in 2011-13, cumulating to an invited speaker in an international conference on mammography in 2014. Table 6.3 summarises the contributions to conferences and peer reviewed education for the research area; recognising the presentation awards, seminars and invited speaker invitations. Appendix Two contains the poster/speaker abstracts and awards.

| Conference Information | UKRC 2010 | UKRC 2011 | UKRC 2012 | UKRC 2013 | UKRC 2014 |
|---|---|---|---|---|---|
| | Symposium Mammographicum 2010 | | Symposium Mammographicum 2012 | | Symposium Mammographicum 2014 |
| Paper I | | Poster | | | |
| Paper II | | | Poster<br>Proffered paper | | |
| Paper III | Proffered paper*<br>Proffered paper | | | | |
| Paper IV | | | | | Poster<br>Poster |
| Paper V | | | Poster<br>Proffered paper** | | |
| Paper VI | | | | Poster | |
| Paper VII | | | Distributed article to all delegates at conference | | |
| Paper VIII | | | | | Poster<br>Invited speaker |

Table 6.3: Conference proceedings

Proffered paper* Awarded Alan Nichols Award

Proffered paper** Awarded Best Oral Presentation in the session entitled: Challenges of Screening

UKRC was selected as one of the main conferences for presentation of this work as it is the leading and largest diagnostic annual imaging event in the UK. It covers all aspects of diagnostic imaging and oncology and consists of a three day multidisciplinary conference with technical exhibition. It attracts between 3,000 and 4,000 people each year and as such it was hopeful that presentations of research work at this conference would gain an impact not just within mammography, but in the radiography field as a whole.

Symposium Mammographicum was also selected for presentation of the research works. It is a registered charity and aims to stimulate and support research and disseminate knowledge in the area of breast cancer diagnosis and treatment. This is a biennial Symposium and attracts both UK and international delegates. It was seen essential that the work was disseminated within this conference to gain both UK and international impact.

### 6.5.2 Dissemination of research at a local level

Together with UKRC and the Symposium Mammographicum it was essential that this research was disseminated locally within the mammography arena. The first research was disseminated at the inaugural University of Salford Breast Research evening seminar in 2012 and had very encouraging feedback, this followed further dissemination in 2014 at a Mammography Update day within the Hospital Trust. Finally, at a local setting, the thesis work was presented at the NHS Research &

Development North West Conference in September, 2014. Together with the presentation of the research findings the author also presents how collaborative work is essential when undertaking a PhD by published works. The author was invited to be part of the conference organising committee and was involved in peer reviewing the publication abstracts for this conference.

## 6.6 Summary

The research within this thesis has made a significant, initial contribution within the UK and international mammography field. It has highlighted a considerable issue of lack of standards within current mammography compression force practice and, with an advanced line of enquiry, has provided evidenced based research to effect a change in the professional mammography landscape by suggesting such standards (90-130N) and highlighting the requirement for practitioner standardisation. It has done this by educating peers and experts within the field to the changes that are required to create an evidenced based quality standard for future mammography practice.

## 6.7 Publication metrics

It was important to clarify the contributions of the peer reviewed publications to the research field and this was done by assessing the publications with citation metrics. Google Scholar [http://scholar.google.co.uk/citations] was utilised to illustrate citations for all the published publications from 2011 to 2013 (Figure 6.2), with Research Gate [https://www.researchgate.net] utilised to demonstrate publication views and full text downloads since 2011 (Figure 6.3). This research established immediate interest following publications in 2011 and has continued within the following years to date. Citations arise from Papers I to III and VI only, with current citations limited due to the publication dates of most of this research work being in 2013 and 2014.

It is recognised that the citation metrics shown from Research Gate underestimate the total downloads, such as within journal websites and the Society of Radiographers website. Research Gate was used as a tool to compare journal articles within the same research forum. Though Research Gate is an essential distributer of research throughout its networking site and claims to have 3 million users, it is not clear how many of these have active accounts and it is recognised that the forum is open to manipulation. The use of Research Gate has enabled distribution of these thesis publications to places such as Malaysia, Denmark and New Zealand and has directly resulted in the formation of new research collaborations in the Netherlands.

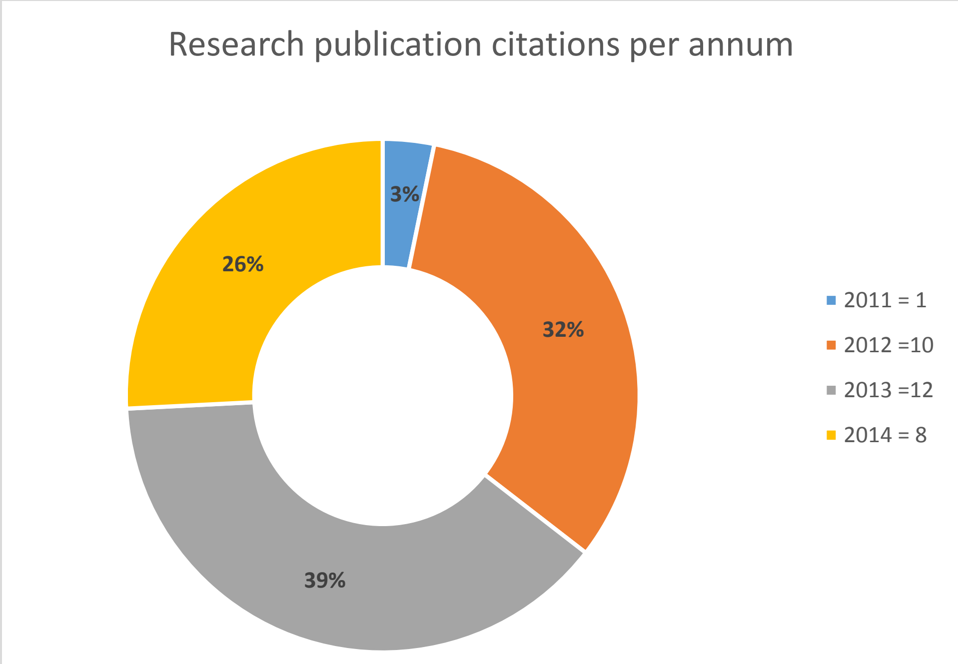



Figure 6.2: Research publication citations per annum [Google scholar citations, 2014]

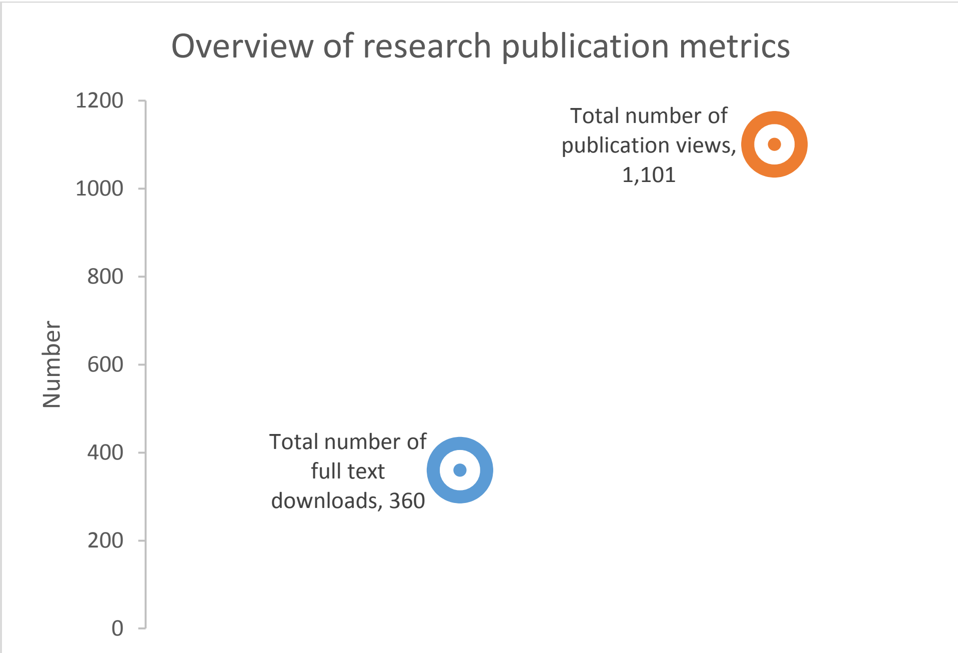



Figure 6.3: Overview of research publication metrics [Research Gate, 2014]

Illustrated in Figure 6.4 Research Gate provides an overall impact score for an author (RG) and summarises this over time. The key aim of this score is to assist researchers in measuring standing within the scientific community; the RG algorithm works by not only assessing how the researchers and peers receive and evaluate the research but by assessing who those peers are. The higher the scores of those researchers that the researcher interacts with, the higher their RG score, the published research is then factored in and the RG score calculated. It is acknowledged that this score can be subject to misuse and manipulation and there is also known differences in impact scores for different research genres; cancer research being quite high. Aside from this, Research Gate has been an essential forum for the distribution of research within this thesis work and has enabled collaboration with other researchers, both within and outside, the field of mammography.

Research Gate indicates an RG score of 14.03 for the author, being in the 55% of research gate members, demonstrating that this research is having a substantial contribution to the research field within the RG arena.

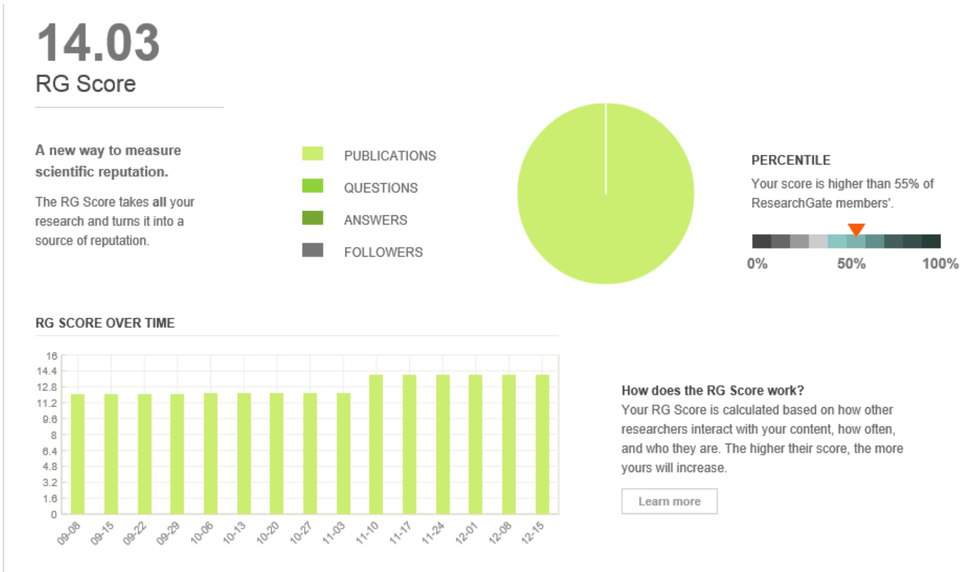



Figure 6.4: Author’s impact score [Research Gate, 2014]

## 6.8 Individual publication metrics

In order to demonstrate the standing of the research, the papers within this thesis are considered for journal metrics, critical performance metrics and citation mapping representing:

- publication journal metrics (journal performance)
- critical performance metrics (article performance)
- citation mapping (qualitative review of citations received)

The referred 'h index' indicates the productivity and impact of the published work of the journal based upon citations. The h5 index this demonstrates the index for the articles published within the last 5 complete years for the journal authors.

### 6.8.1 Journal metrics

The papers within this thesis are published in five journals, the metrics for each journal are demonstrated (Table 6.4) followed by an overview on the journal selection for each individual paper.

| Journal | h5-index [Google Scholar, 2014] | h5-median [Google Scholar, 2014] | Impact factor | 5 year Impact factor |
|---|---|---|---|---|
| Medical Physics | 60 | 78 | 2.911 | 3.138 |
| British Journal of Radiology | 32 | 39 | 2.11 | 1.938 |
| Radiologic Technology | 11 | 19 | 1.08 | - |
| Radiography | 14 | 16 | No impact factor rating. Official quarterly press journal for the Society and College of Radiographers and distributed to each member of the Society | |
| Imaging and Oncology | 14 | 16 | No impact factor rating. Included in every delegate pack in 2013 at both UKRC and UKRO, as well as at College of Radiographers' seminars, study days and events | |

Table 6.4: Journal publication metrics

**Medical Physics:** Paper I is published within Medical Physics, a scientific journal which publishes research concerned with the application of physics and mathematics in the solution of problems in medicine and human biology. Manuscripts concerning theoretical or experimental approaches are published within this journal [Medical physics journal online, 2014]. This high impact journal was selected as the research was physics based and considered an appropriate fit within this journal.

**The British Journal of Radiology (BJR):** Paper II and III are published within BJR**,** an international research journal of the British Institute of Radiology. It is essential reading for radiologists, medical physicists, radiotherapists, radiographers and radiobiologists. The journal publishes original research papers from centres internationally together with editorials, review articles, communications and letters to the editor. Articles cover a wide range of subjects, including diagnostic radiology, radiotherapy, oncology, nuclear medicine, ultrasound, radiation physics, radiation protection and radiobiology [British Journal of Radiology Publications, 2014]. This journal was selected for these papers as it was a high impact journal, featuring novel research with a wide reading audience. The subject matter of these papers was highly original and was therefore considered an ideal base for these publications.

**Radiologic Technology:** Paper IV was published within Radiologic Technology, the official scholarly journal of the American Society of Radiologic Technologies. It is award winning and publishes bi-monthly; it has published continuously since 1929 and circulates to more than 145,000 readers worldwide. It covers all disciplines within medical imaging and in addition to peer reviewed articles features educational articles and columns of interest to the profession Research Gate 2014].

**Radiography:** Papers V, V1 and VIII are published within the international peer reviewed journal of Radiographic Imaging and Radiation Therapy, Radiography. It is the official quarterly press journal for the Society and College of Radiographers and is published by Elsevier Ltd. Its aims are to publish high quality clinical, scientific and educational material on all aspects of radiographic imaging and all aspects of radiation therapy. The journal includes original research, review articles, technical notes, evaluations and case studies.

Radiographic society members can directly access journals from the Radiography web site, together with this each society member has the journal directly distributed to them. The Canadian society members also have full access to the journal articles which increases potential circulation. In order to ensure that the mammographers themselves had sight of this research, this journal was considered to be the most desirable way of disseminating this work having direct readership with the mammographers who practiced in the field.

**Imaging & Oncology:** Paper VII was commissioned on request for 2013 Imaging and Oncology. This annual title publication coincides with the United Kingdom Radiology Congress (UKRC) [UKRC, 2014]. It is widely circulated and sent to all radiologists, oncologists and heads of education centres. It is also circulated to clinical radiology, radiotherapy and medical physics departments in the UK. In 2013 when this article was published Imaging & Oncology was included free in every delegate pack at both UKRC and UKRO [UKRO, 2013], as well as at College of Radiographers' seminars, study days and events. This article summarises all the research carried out in the thesis together with further work by other research groups driven and developed as a direct result of the authors' initial work. The journal editors requested this work illustrating the importance of the work in this field at this time.

### 6.8.2 Critical performance metrics

The specific performance metrics for each paper are discussed and illustrate the impact from a perspective of Research Gate (publication views and downloads) and Google Scholar (citations). Publication views and downloads for each paper (Fig 6.5) illustrate the immediate interest of this research within the field.

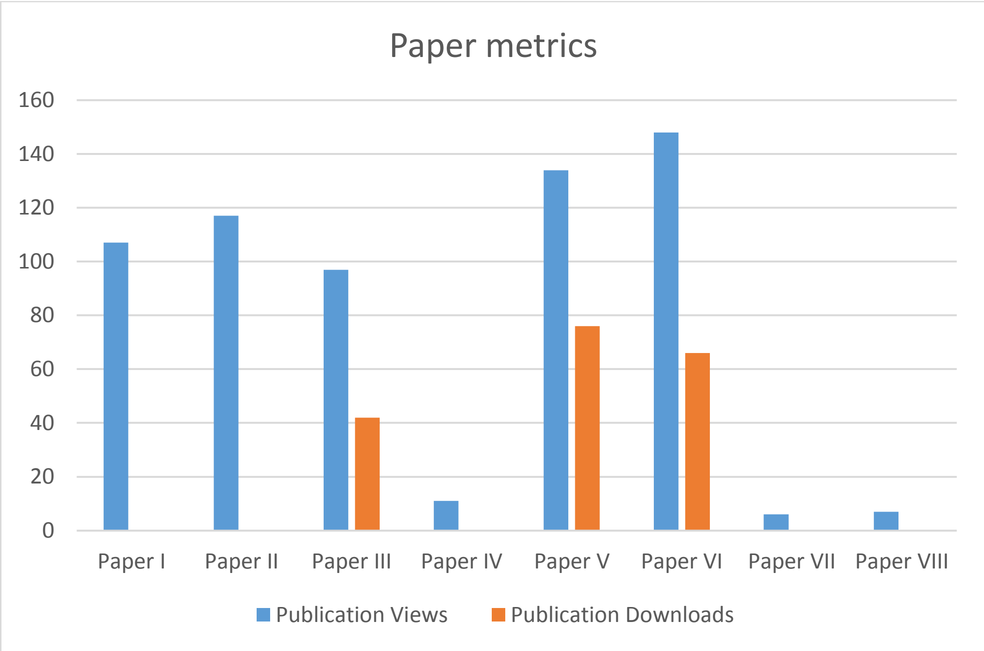



Figure 6.5: Paper metrics

Publication citations are low; 23 in total from all papers, namely due to the recent publication year of the research (2012-2014) and the highly novel research theme with few other researchers working in this field. Newly developed research within the mammography arena is being established through the University of Salford following the introduction of the research within this thesis and, as such, citations established from work at the University form a 26% proportion of the total citations (6/23). Self-citations form a 39% proportion (9/23), and citations from other established researchers in the field 35% (8/23). Within this 35%, table 6.5 illustrates indicative impact factor (IF) ratings assigned to the journals which the papers are cited.

| Paper | Journal title | Journal details | Journal impact factor |
|---|---|---|---|
| I | Geeraert, N., (2014).Comparison of volumetric breast density estimations from mammography and thorax CT. | *Phys med biol*. 7; 59 (15):4391-409. doi: 10.1088/0031-9155/59/15/4391. Epub 2014 Jul 22. | 2.992 |
| | De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. | *Medical physics*, 40 (8), 081901. doi:10.1118/1.4812418 | 2.91 |
| | Groobe, A, et al., (2012). Spectral Volumetric Glandularity Assessment | Breast Imaging Lecture Notes in Computer Science. Volume 7361, 2012, pp 529-536 | No IF |
| | Alonzo-Proulx, R.A., Jong & Yaffe, M.J. (2013). Volumetric breast density characteristics as determined from digital mammograms | Physics and Medicine in Biology. Vol 57 No:22 | 2.992 |
| II | De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. | *Medical physics*, 40 (8), 081901. doi:10.1118/1.4812418 | 2.91 |
| | De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2014). Mammographic compression after breast conserving therapy: controlling pressure instead of force. | *Medical physics*, 41 (2), 023501. doi:10.1118/1.4862512 | 2.91 |
| III | De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. | *Medical physics, 40 (8), 081901. doi:10.1118/1.4812418* | 2.91 |
| | De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2014). Mammographic compression after breast conserving therapy: controlling pressure instead of force. | *Medical physics*, 41 (2), 023501. doi:10.1118/1.4862512 | 2.91 |

Table 6.5: Citation overview

### 6.8.3 Originality of publications and new lines of enquiry

Paper I was the first publication to detail the non-concordance of the readout thickness display on the mammography machine and the actual breast thickness, together with specifying the compression paddle bend and distortion on numerous mammography machines. This paper challenged current beliefs in regards to the accuracy of readouts on mammography systems. Within the context of the author's work this research established a solid framework for the continuation of the research within this thesis. It ensured that ongoing research utilised one mammography machine, to limit the variability of inaccuracies in data gathering and analysis using multiples mammography machines. In this way it added stability to the rest of the research framework.

Together with this, this publication assisted with research in the Netherlands, De Groot et al [2013] took into account this empirical research (Paper 1) and ensured that they extensively calibrated their mammography unit to accommodate for compression plate bend and distortion. Instigation of new research themes focused on image blurring as a direct result of compression paddle movement have also been resultant from this paper.

Paper II was the first publication to detail the correlation between breast thickness and compression force on a sample of patients from a mammography service. This paper was highly novel; it had not been researched since Poulos and colleagues in their small study in 2004 [Poulos, A. and McLean, D 2004] who then called for more research within this field. This paper had a direct impact on the work of researchers in the Netherlands; they were working on observations of pressure instead of compression force. The paper used the word ‘pressure’, a colloquialism common in clinical practice for ‘compression force’ (Section 3.2.2). As this was not the correct physical terminology a letter to the editor of this journal was generated and a response given. This ‘colloquial misuse’ was indeed the making of a new research relationship and work is now being developed between the two research groups to develop research themes for the future.

Paper III was the very first publication within the research arena to demonstrate practitioner variation within and between mammography practitioners within a patient centred study. As such, this paper was highly novel and followed on from the work in the small cohort of clients by Poulos and Mclean in 2003 and 2004, who had demonstrated in their research outcomes that there was an element of practitioner variation. This research saw the introduction of new lines of enquiry and the development of further work in this field. It had coherently demonstrated that practitioner variation did exist in a cohort of practitioners; development of this work was essential to further cement this theory.

Paper IV was not ground-breaking research, however it was imperative to underpin the continuing lines of research and allowed for the continuation of research within the multicentre sites. Without the knowledge of the good intra and inter reliability in the scoring of images for BIRADS breast density, the rest of the research would have been flawed.

Paper V was a continuation of the research findings of Paper III; and was the first published research to demonstrate a significant demonstration of practitioner variation of compression force over a six year period. This research illustrated a developed research theme, generated interest following on from conference presentations, and highlighted the requirement for a multicentre study in this area. This research also gave rise to the development of new qualitative breast research focusing on practitioners behaviours within mammography [Murphy et al., 2014; Robinson et al., 2013]. This was the first time that practitioner behaviour had been directly linked to variation in breast compression force.

Paper VI was a continuation of the research findings of Paper V, it was empirical research and to date the paper has over a hundred and fifty views and sixty six downloads. This research established that visual image quality was not effected by a change in compression force. It was recognised that this is a small cohort and further research is required into lesion visibility and breast thickness; ideally within a breast phantom. A PhD research student at the University is now developing this theme.

Paper VII was novel within this journal. Previous issues of this journal from 2005 to 2014 inclusive had only 3 articles based in breast cancer care which focused on treatment, sentinel node imaging and brachytherapy. This paper was the first within this journal focusing on empirical mammography research and was seen to be highly novel for this journal. Following the Francis report [Francis, 2013] it was recognised within the forward of this journal that patients must come first and that professionals must take a responsibility to ensure this; it acknowledges this paper as contemporary within its field.

Paper VIII was the accumulation of research so far and was the first paper to be published which demonstrated practitioner variation in breast compression force in a multicentre study across 6 years. As such this research was highly novel and has been acknowledged by a conference forum that this research could have widespread value with the potential to change compression force protocols in the future. As a direct outcome of this research, further research teams are being established to work on practitioner variation in the digital mammography field using software called Volpara™ which instantly enables direct reports of practitioner compression force values linked to breast thickness readouts.

The author was asked, as an invited speaker, to present the findings of this research at the Symposium Mammographicum conference in 2014. This was considered to be of substantial importance for this research and considered a development of the researcher (Figure 6.6). The impact of presenting the research findings at this conference were high, with well esteemed colleagues in attendance. Together with this the research findings are being presented at the European Society of Radiology conference ECR in 2015.


Figure 6.6: Development of researcher at conferences

## 6.9 Overview of journal metric impact

For all articles published in peer reviewed journals for this thesis the impact per publications (IPP), the measure of the scientific influence of each journal (SJR) and the source normalised impact per paper (SNIP) is compared. This illustrates the published journals quality and reputation within the subject field and allows for a direct comparison of the journals which these thesis publications are published [Journal Metrics, 2014].

Figures 6.7-6.9 compare the SNIPP, IPP and SJR creating an accurate overview and comparison of the citation impact of the journals to which these thesis articles are published. SNIP, IPP and SJR are known to form good correlation with current impact factors of journals. This is useful for the Radiography journal, which has no impact factor, to which several of the main articles for this thesis are published within.

It is demonstrated that articles within Radiography are being increasingly cited each year; with a more notable increase from 1999 to 2013 than the other journals within the same time period. The IPP (Figure 6.8) illustrate a steady increase over the last 15 year period for all journals in which these research papers have been published; with the exception of Radiologic Technology whose journal metrics were not available until 2013. It is of interest to note that, though not impact factor rated, the SNIP, IPP and SJR of Radiography is higher than Radiologic Technology which has an impact factor rating of 1.08.

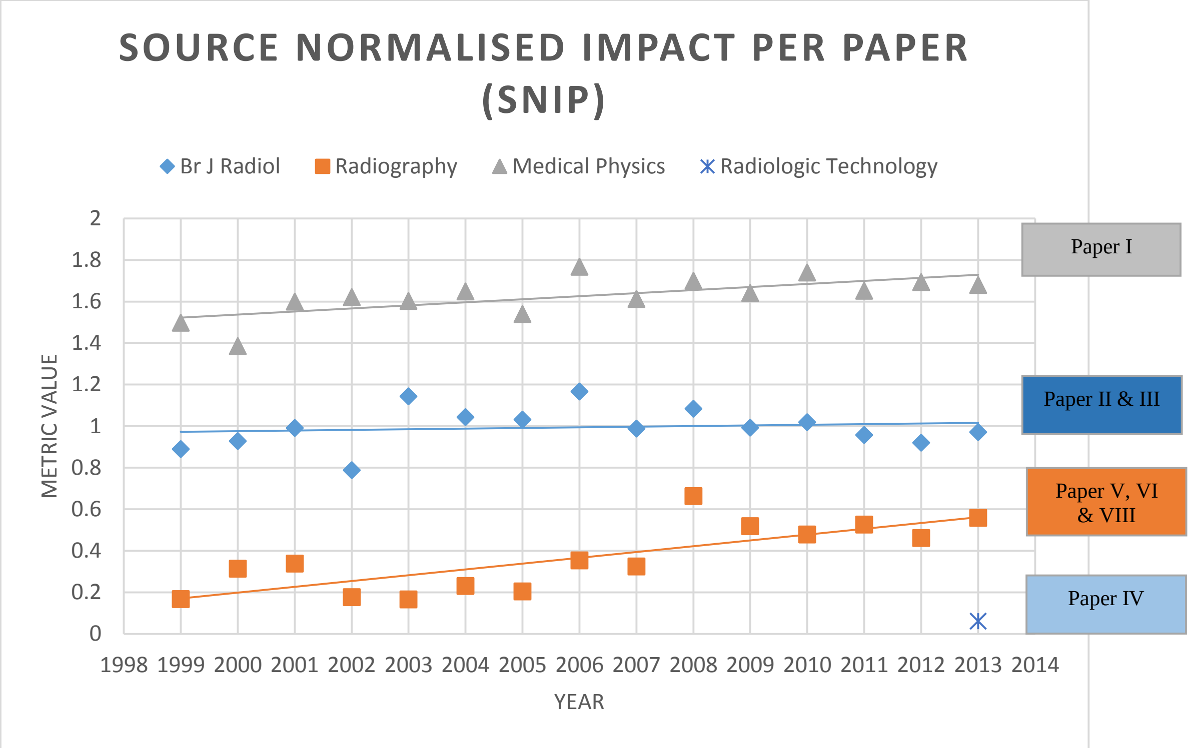



Figure 6.7: Source normalised impact per paper [Elsevier Connect, 2014]

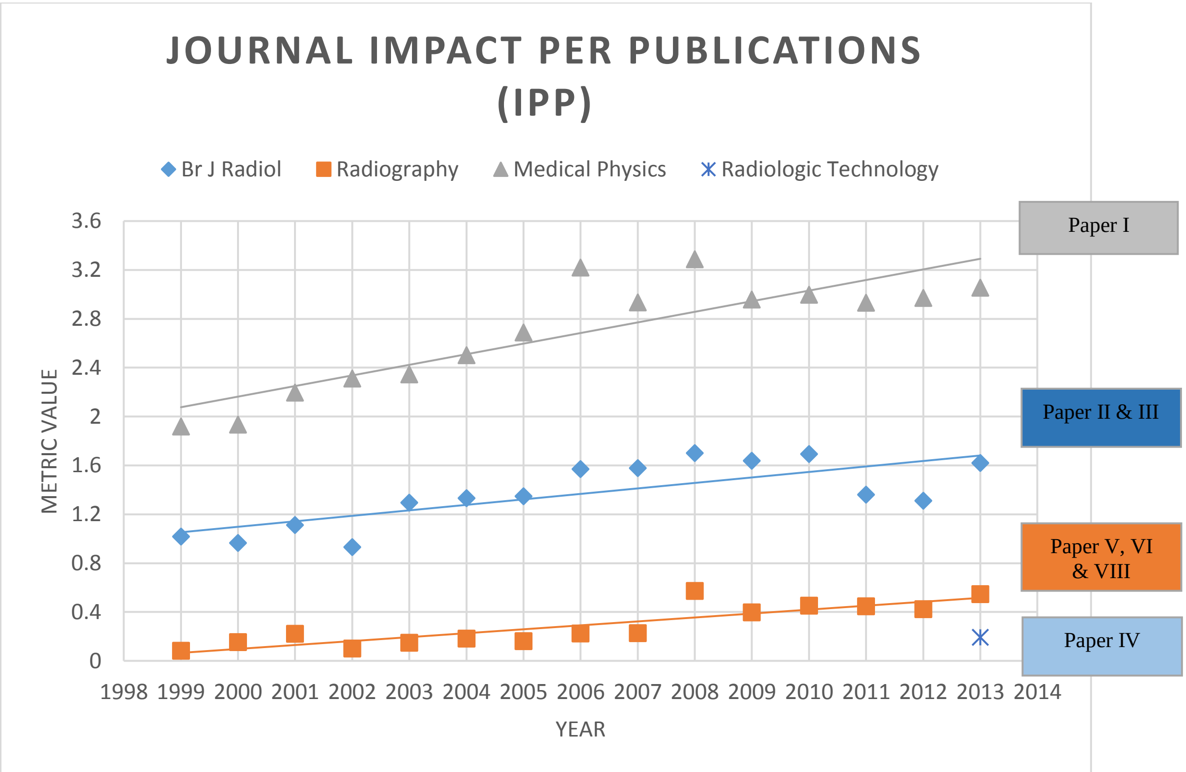



Figure 6.8: Journal impact per publications [Elsevier Connect, 2014]

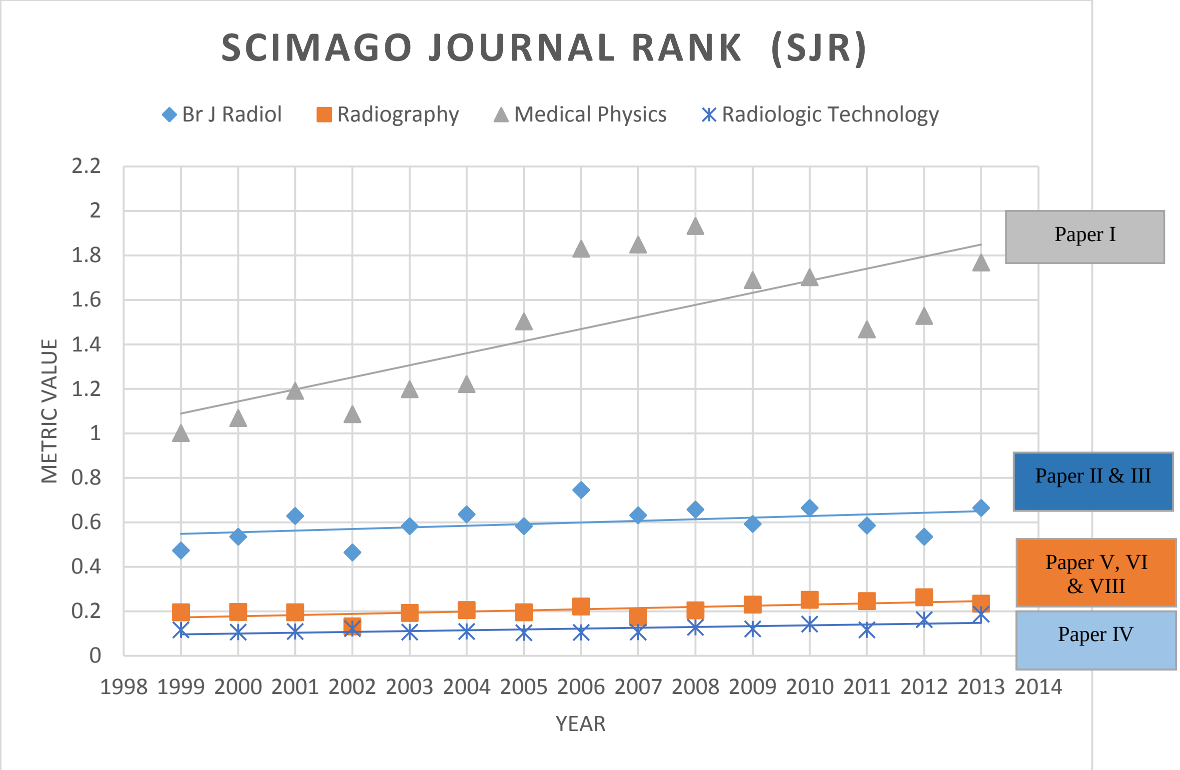



Figure 6.9: SCImago Journal Rank [Elsevier Connect, 2014]

## 6.10 Conclusion of research development, performance and impact

The research contained within this thesis has demonstrated the creation and interpretation of new academic knowledge through original research, which has merited both peer reviewed publications and invitations to present at well-established conference proceedings.

Through systematic acquisition of new knowledge and the formation of new research teams with developed research links, the researcher has demonstrated the ability to conceptualise study design and structure and process ethics approvals; both within the University and within Hospital Trusts. The researcher has demonstrated the ability to establish new research groups and lead research teams, formulate and action issues, analyse and interpret data, and edit and structure research papers.

The researcher has demonstrated that the work contained within this thesis has had a direct and novel impact in the mammography arena. It has effected new research pathways within the UK and internationally. The research has been published in peer reviewed journals with established metrics and wide reaching audiences. The author's performance is recognised and the peer reviewed articles and conference articles are being viewed in established research forums. The research is also disseminated into a new international academic mammography book.

In summary the researcher has shown progressive and influential research impact into the mammography arena with the published work contained within this thesis.

# REFERENCES


Assi, V., Warwick, J., Cuzick, J., & Duffy, S.W. (2011). Clinical and epidemiological issues in mammographic density. *Nature Reviews Clinical Oncology*, 9, 33–40. doi:10.1038/nrclinonc.2011.173

Barentsz, J., Takahashi, S., Oyen, W., Mus, R., De Mulder, P., Reznek, R., … Mali, M. (2006). Commonly used imaging techniques for diagnosis and staging. *Journal of clinical oncology : official journal of the American society of clinical* o*ncology*, 24, 3234–3244. doi:10.1200/JCO.2006.06.5946

Bentley, K., Poulos, A., & Rickard, M. (2008). Mammography image quality: Analysis of evaluation criteria using pectoral muscle presentation. *Radiography*, 14,189–194. doi:10.1016/j.radi.2007.02.002

Berg, W., Birdwell, R., & Gombos, E. (2006). *Diagnostic imaging breast.* Amirsys Inc, Salt Lake City, Utah, USA.

Boyd, N.F., Byng, J.W., Jong, R.A., Fishell, E.K., Little, L.E., Lockwood, G.A., … Yaffe, M.J. (1995). Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *Journal of the national cancer institute*, 87(9), 670–675. doi:10.1093/jnci/87.9.670

Bragg, J. (1986). *The radiographers guide to film-screen mammography*. Magazine Associates, Newcastle.

Breast Cancer Screening. (1987). *Report to the Health Ministers of England, Wales, Scotland and Northern Ireland.* Working Group Chaired by Professor Sir Patrick Forrest. HMSO.

Breast Cancer Screening. (1991). *Evidence and Experience since the Forrest Report*. *A Report of the Department of Health Advisory Committee*. Sheffield, NHSBSP.

British Journal of Radiology Publications. (2014). Retrieved from: http://www.birpublications.org/

Byng, J.W., Yaffe, M.J., Jong, R., Shumak, R.S., Lockwood, G.A., Math, M., … Boyd, N.F. (1998). Analysis of mammographic density and breast cancer risk from digitized mammograms. *RadioGraphics,* 18, 1587–1598. Retrieved from: http://radiographics.rsna.org/content/18/6/1587.short

Calman, K.C., & Hine, D. (1997). Breast Cancer Services in Exeter and Quality Assurance for Breast Services. *Report to the Secretary of State*. Department of Health.

Cheung, H.S. (2006). Mammographic quality assurance – the radiologist's point of view. *Biomedical imaging and intervention journal, 2*(4), 45-48. Retrieved from: http://www.biij.org/2006/4/e45/abstract.asp?ID=119

Chida, K., Komatsu, Y., Sai, M., Nakagami, A., Yamada, T., Yamashita, T., … Zuguchi, M. (2009). Reduced compression mammography to reduce breast pain. *Clinical imaging*, 33, 7–10. doi:10.1016/j.clinimag.2008.06.025

Committee ACR. (1999). Mammography Quality Control Manual. *American college of radiology*. Retrieved from:

*http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0CDkQFjAD&url=http%3A%2F%2Fwww.acr.org%2F~%2Fmedia%2FACR%2FDocuments%2FAccreditation%2FMammography%2FRequirements.pdf&ei=HecKVO3yA8PLaP7JgMAC&usg=AFQjCNEEu1RRLKRwLol5OkFwXL7AIy6kIA*

D'Orsi, C.J., Bassett, L.W., & Berg, W.A. (2003). *Mammography*, 4th Edition. Breast Imaging Reporting and Data System: ACR BI-RADS®. Reston VA: American College of Radiology.

Darlington, A., Hogg, P., Szczepura, K., & Maxwell, A. (2013). Optimising paddle and detector pressures and footprints in mammography. *Medical physics*, 40 (4), 041907.

De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2013). A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. *Medical physics*, 40 (8), 081901. doi:10.1118/1.4812418

De Groot, J.E., Broeders, M.J.M., Branderhorst, W., den Heeten, G.J., & Grimbergen CA. (2014). Mammographic compression after breast conserving therapy: controlling pressure instead of force. *Medical physics*, 41 (2), 023501. doi:10.1118/1.4862512

Department of Health and Social Security. (1986). The Forrest Report Breast Cancer Screening. *Report to the Health Ministers of England, Wales, Scotland and Northern Ireland* by a working group chaired by Professor Sir Patrick Forrest. Retrieved from: http://www.cancerscreening.nhs.uk/breastscreen/publications/forrest-report.html

Department of Health. (1989). Guidelines on the Establishment of a Quality Assurance Scheme for the Radiological Aspects of Mammography used for Breast Cancer

Screening. *Report of a Sub-committee of the Radiological Advisory Committee of the Chief Medical Officer*. Department of Health.

Department of Health. (2007). Cancer reform strategy. *Clinical oncology (Royal College of Radiologists (Great Britain))*, 20,144. doi:10.1016/j.clon.2008.01.014

Diffey, J., Hufton, A., Astley, S., Mercer, C., & Maxwell A. (2008). Estimating individual cancer risks in the UK national breast screening programme: a feasibility study. EA Krupinski (ED.), *IWDM 2008. LNCS*, 5116, 469-479, 2008 Springer-Verlag Berlin Heidelberg.

Ding, J., & Warren, R. (2004). Update on the importance of parenchymal patterns and breast cancer risk. *Breast cancer online*, 7 (12), e36. Retrieved from: http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=288203

Doyle, C.A., & Stanton, M.T. (2002). Significant factors in patient satisfaction ratings of screening mammography. *Radiography*, 8, 159–172. doi:10.1053/radi.2002.0379

Duffy, S.W., Tabar, L., Olsen, A.H., Vitak, B., Allgood, P.C., Chen, T.H.H., … Smith, R.A. (2010). Absolute numbers of lives saved and over diagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *Journal of medical screening*, 17, 25–30. doi:10.1258/jms.2009.009094

Eadie, A., Whelehan, P., Baker, L., Berg, J., & Evans, A. (2011). Improving the validity of breast density assessment. *Breast cancer research*, 13 (1), 24.

Eklund, G.W., Cardenosa, G., & Parsons, W. (1994). Assessing adequacy of mammographic image quality. *Radiology*, 190, 297–307.

Elsevier Connect, 2014. Retrieved from: http://www.elsevier.com/connect/with-2013-journal-rankings-no-one-metric-rules-them-all

European Commission. (1996). *European Guidelines on Quality Criteria for Diagnostic Radiographic Images.* EUR 16260.

Feig, S.A. (2002). Image quality of screening mammography: effect of clinical outcome. *American journal of radiology*, 178(April), 150-153.

Francis R. (2013). *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationary Office; Feb 2013. doi:10.1002/yd.20044

Google scholar citations, 2014. Retrieved from: http://scholar.google.co.uk/citations

Hauge, I.H.R., Hogg, P., Szczepura, K., Connolly, P., & Mercer C.E. (2012). The readout thickness versus the measured thickness for a range of screen film mammography and full-field digital mammography units. *Medical physics*, 39 (1), 263–271. doi:10.1118/1.3663579

Health and Social Care Information Centre. (2013). Retrieved from: http://www.hscic.gov.uk/catalogue/PUB10339

Health and Social Care Information Centre. (2014). Retrieved from: http://www.hscic.gov.uk/article/3943/Breast-cancer-fall-in-proportion-of-women-screened-in-England

Heine, J.J., & Malhotra, P. (2002). Mammographic tissue, breast cancer risk, serial image analysis, and digital mammography. *Academic radiology*, 9. Pp 298–316-335. doi: 10.1016/S1076, 6332(03)80373-2, 6332(03)80374-4

Hershe, M.R. (2004). Imaging the dense breast, *Applied radiology*, *33*(1). Retrieved from: http://medgenmed.medscape.com/viewarticle/468145

Hewes, H., Williamson, A., Noonan, P., Sergeant, J.C., Dunn, T., Haste, S. … Astley, S. (2013). *Measurement of compressed breast thickness for breast density assessment using a games console input device*. Paper presented at the 6th International Breast Densitometry and Breast Cancer Risk Assessment Workshop 6-7 June 2013 San Francisco, USA. Retrieved from: http://dx.doi.org/10.1594/ecr2013/C-2199

Hogg, P., Szczepura, K., Kelly, J., & Taylor, M. (2012). Blurred digital mammography images. *Radiography,* 02/2012, 18(1), 55–56. doi: 10.1016/j.radi.2011.11.008

Hogg, P., Mercer, C., Maxwell, A., Robinson, L., Kelly, J., & Murphy F. (2013). Controversies in compression, *Imaging and oncology*, 28-36, ISBN 9871 871101581

Hogg, P., Taylor, M., Szczepura, K., Mercer, C., & Denton, E.R.E. (2013). Pressure and breast thickness in mammography- an exploratory calibration study. *The British journal of radiology*, 86 (1021), 20120222. doi:10.1259/bjr.20120222

Howell, A., Sims, A.H., Ong, K.R., Harvie, M.N., Evans, G.R., & Clarke, R.B. (2005). Mechanisms of disease: prediction and prevention of breast cancer--cellular and molecular interactions. *Nature clinical practice. Oncology*, 2, 635–646. doi: 10.1038/ncponc0361

Journal Metrics. (2014). Retrieved from: http://www.journalmetrics.com/values.php

Kelly J., Hogg P., Millington S., Sanderud A., Wilcock C., McGeever G., … Kelly S. (2012). *Paddle motion analysis preliminary research*. Paper presented at the Symposium Mammographicum 2012. ISBN 10:0-905749-77-4. ISBN 13:978-0-905749-77-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2012

Khan-Perez J., Mercer C., Bydder M., Sergeant J., Morris J., Maxwell A., … Astley S. (2013). Breast compression, compressed breast thickness and volumetric breast density. *Breast cancer research*, 15 (1).

Kopans D. (2007). *Breast imaging*, 3rd ed. Lippincott: Williams and Wilkins.

Lee, L., Strickland, V., Wilson, R., & Evans, E. (2002). *Fundamentals of mammography*. 2nd ed. Churchill Livingstone.

Li, Y., Poulos, A., McLean, D., & Rickard, M. (2010). A review of methods of clinical image quality evaluation in mammography. *European journal of radiology*, Jun; *74*(3):122-31. doi: 10.1016/j.ejrad.2009.04.069.

Long, S. (2000). *The handbook of mammography*. 4th ed, Mammography Consulting Services Ltd, Edmonton.

Ma, W.K., Hogg, P., Kelly, J., & Millington, S. (2014). A method to investigate image blurring due to mammography machine compression paddle movement. *Radiography*, Published Online: June 21, 2014. doi:10.1016/j.radi.2014.06.004

Majid, A.S., Shaw de Paredes, E., Doherty, D., Sharma, N.R., & Salvador, X. (2003). Missed Breast Carcinoma: pitfalls and pearls. *Radiographics*, 23(4), 881-895.

McCormack, V.A., & dos Santos Silva, I. (2006). Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 15, 1159 - 1169. doi:10.1158/1055-9965.EPI-06-0034

MDLinx. (2014). Retrieved from: http://www.mdlinx.com/

Medical Physics Journal Online, 2014. Retrieved from: http://scitation.aip.org/content/aapm/journal/medphyss

Mercer, C.E., Hogg, P., Cassidy, S., & Denton, E.R.E. (2013). Does an increase in compression force really improve visual image quality in mammography? – An initial investigation. *Radiography*, 19 (4), 363–365. doi:10.1016/j.radi.2013.07.002

Mercer, C.E., Hogg, P., Lawson, R., Diffey, J., & Denton, E.R.E. (2013). Practitioner compression force variability in mammography: a preliminary study. *The British journal of radiology*, 86 (1022), 20110596. doi:10.1259/bjr.20110596

Mercer, C.E., Hogg, P., Szczepura, K., & Denton, E.R.E. (2013). Practitioner compression force variation in mammography: A 6-year study. *Radiography*, 19 (3), 200–206. doi:10.1016/j.radi.2013.06.001

Mercer, C.E., Hogg, P., Kelly, J., Borgen, R., Millington, S., Hilton, B. … Whelehan, P. (2014). A mammography image set for research purposes using BI-RADS density classification. *Radiologic technology*, 85 (6), 609–613. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/25002640

Mercer, C.E., Szczepura, K., Kelly, J., Millington, S.R., Hilton, B., & Hogg, P. (2014). A 6-year study of mammographic compression force: Practitioner variability within and between screening sites. *Radiography*, Published online: July 29, 2014. doi:10.1016/j.radi.2014.07.004

Murphy, F., Nightingale, J., Hogg, P., Robinson, L., Seddon, D., & Mackay, S. (2014). Compression force behaviours: An exploration of the beliefs and values influencing the application of breast compression during screening mammography. *Radiography*, Published Online: June 13, 2014. doi:10.1016/j.radi.2014.05.009

Ng, K.H., & Muttarak, M. (2003). Advances in mammography have improved early detection of breast cancer. *J HK Coll Radiol, 6,* 126-131.

NHS Cancer Screening Programmes. (2000). *Guidelines on quality assurance visits.* NHSBSP 40. Retrieved from: http://www.cancerscreening.nhs.uk/breastscreen/publications/qa-12.html

NHS Cancer Screening Programmes. (2006). NHSBSP 63. *Quality Assurance Guidelines for Mammography*. April (42). ISBN 1 84463 028 5.

NHS Cancer Screening Programmes. (2013). Public Health England. Retrieved from: http://www.cancerscreening.nhs.uk/breastscreen/screening-programme.html

NHS Executive (1997). EL(97)67.*Cancer Screening: Quality Assurance and Management*. NHS Executive.

O'Leary, D., & Al Maskari, Z. (2014). Pain in mammography: where and why does it arise? Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8.
Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014

O'Leary, D., & Al Maskari Z. (2014). Silicon cushions: a pain reduction intervention for mammography. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8.
Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014

O'Leary, D., & Rainford, L. (2014). Higher image quality and greater compression force in mammography are linked. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014

O'Leary, D., & Rainford L. (2014). Patient bra size as a gauge of compression required in mammography. Symposium Mammographicum 2014 meeting abstracts. ISBN 10:0-905749-80-4 ISBN 13: 978-0-905749-80-8. Retrieved from: http://www.birpublications.org/doi/book/10.1259/conf-symp.2014

Obenauer, S., Luftner-Nagel, S., von Heyden, D., Munzel, U., Baum, F., & Grabbe, E. (2002). Screen film vs full-field digital mammography: Image quality, detectability and characterization of lesions. *European radiology*, 12(7), 1697–1702. doi: 10.1007/s00330-001-1269-y

Oliveira, M., Nogueira, M.S., Guedes, E., Andrade, M.C., Peixoto, J.E., … Castro, J.G. (2007). Average glandular dose and phantom image quality in mammography. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 580, 574–577. doi:10.1016/j.nima.2007.05.228

Pisano, E.D., Gatsonis, C., Hendrick, E., & Yaffe, M. (2005) Diagnostic performance of digital versus film mammography for breast-cancer screening. *The New England journal of medicine*, 353, 1773–1783. doi:10.1097/01.ogx.0000201894.94616.71

Poulos, A., McLean, D., Rickard, M., & Heard, R. (2003). Breast compression in mammography: How much is enough? *Australasian radiology*, 47, 121–126. doi:10.1046/j.0004-8461.2003.01139.x

Poulos, A., & McLean, D. (2004). The application of breast compression in mammography: A new perspective. *Radiography*, 10, 131–137. doi:10.1016/j.radi.2004.02.012

Research Gate. (2014). Retrieved from: https://www.researchgate.net/profile/Claire_Mercer/stats?ev=prf_stats

Robinson, L., Hogg, P., & Newton-Hughes, A. (2013). The power and the pain: Mammographic compression research from the service-users' perspective. *Radiography*, 19 (3), 190–195. doi:10.1016/j.radi.2013.04.005

Schueller, G., Riedl, C.C., Mallek, R., & Eibenberger, K. (2008). Image quality, lesion detection, and diagnostic efficacy in digital mammography: Full-field digital mammography versus computed radiography-based mammography using digital storage phosphor plates. *European journal of radiology*, 67 (3), 487–496. doi:10.1016/j.ejrad.2007.08.016

Sergeant J.C., Warwick J., Evans D., Howell A., Berks M., Stavrinos P., … Astley S.M. (2012). Volumetric and area-based breast density measurement in the predicting risk of cancer at screening (PROCAS) study. *IWDM* 2012, *LNCS* 7361, 228-235. Springer-Verlag Berlin Heidelberg.

Sickles, E.A., D'Orsi C.J., & Bassett L.W. (2013). ACR BI-RADS® Mammography. In: *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. Reston, VA, American College of Radiology; 2013. Retrieved from: http://www.acr.org/~/media/ACR/Documents/PDF/QualitySafety/Resources/BIRADS/01%20Mammography/02%20%20BIRADS%20Mammography%20Reporting.pdf

Simmons, M.A., Chavez, J., & Barke, L.D. (2012). Patient Positioning and Mammography. *ERad imaging*.
Retrieved from: http://www.eradimaging.com/site/article.cfm?ID=787

Skills for Health. (2013). Retrieved from: https://tools.skillsforhealth.org.uk/competence_search/?search=mammography&pagesize=25&search__operator=B_OR&suite%5B%5D=0&ksfdimension%5B%5D=0&level%5B1%5D=1&level%5B2%5D=2&level%5B3%5D=3&level%5B4%5D=4&search_btn.x=37&search_btn.y=13

Smith, H., Smith, J. J., Hogg, P., Mercer, C., & Szczepura, K. (2011). *Elastically deformable anthropomorphic breast phantom for use in mammographic imaging research.* Paper presented at the UK Radiological Congress 2011. ISBN 10: 0-905749-72-3; ISBN 13: 978-0-905749-72-3

Smith, H., Hogg, P., Szczepura, K., Mercer, C., & Maxwell, A. (2013). *An analysis of compressed breast area and image receptor/compression paddle pressure balance in different mammographic projections*. Paper presented at the UK Radiological Congress 2013.

Strategic Review of Health Inequalities in England post-2010. (2012). Fair society, healthy lives: The Marmot review. *Public health*, 126 Suppl, p.S4-10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22784581.

Taplin, S.H., Rutter, C.M., Finder, C., Mandelson, M.T., THaun, F., & White, E. (2002). Screening mammography: clinical image quality and the risk of interval breast cancer. *AJR. American journal of roentgenology*, 178, 797–803. doi:10.2214/ajr.178.4.1780797

Thomas Reuters. (2014). Journal Citation Reports 2013. Available from: http://www.cityu.edu.hk/ap/nru/ifactors.htm

UKRC (2014). Exhibition and sponsorship guide. Available from: http://www.ukrc.org.uk/about-exhibition

UKRO. (2013). Exhibition and sponsorship guide. Available from: http://www.ukro.ac.uk/aboutukro/conference/Pages/index.aspx

Williams, S., Hackney, L., Hogg P., & Szczepura K. (2013). Tissue bulge during stereotactic core biopsy. *Radiography*, 19, (4) 366–368.

Williams, S., Hackney, L., Hogg P., & Szczepura K. (2014). Breast tissue bulge and lesion visibility during stereotactic biopsy – A phantom study. *Radiography*, 20, (3,) 271–276.

Wentz, G. (1992). *Mammography for Radiologic Technologist.* Mc Graw Hill, New York.

Whelehan, P., Evans, A., Wells, M., & McGillivray, S. (2013). The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. *Breast (Edinburgh, Scotland),* 22 (4), 389–394. doi:10.1016/j.breast.2013.03.003

Whitman, G.J., & Haygood, T.M. (Ed.). (2013). *Digital mammography a practical approach*, 65-67. Cambridge University Press 2013. ISBN 978-0-521-76372.

Wolfe, J.N. (1976). Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, 37 (5) 2486–2492. Retrieved from: http://onlinelibrary.wiley.com/doi/10.1002/1097-0142(197605)37:5%3C2486::AID-CNCR2820370542%3E3.0.CO;2-8/abstract

# APPENDICES

## Appendix One

Diffey, J., Hufton, A., Astley, S., Mercer, C., Maxwell A. (2008). Estimating individual cancer risks in the UK national breast screening programme: a feasibility study. EA Krupinski (ED.): IWDM 2008. LNCS 5116, pp 469-479, 2008 Springer-Verlag Berlin Heidelberg. 2008 http://link.springer.com/chapter/10.1007%2F978-3-540-70538-3_65?LI=true

# Estimating individual cancer risks in the UK national breast screening programme: a feasibility study

Jennifer Diffey[1], Alan Hufton[1], Susan Astley[2], Claire Mercer[3] and Anthony Maxwell[3]

[1] North Western Medical Physics, Christie Hospital, Withington, Manchester M20 4BX
jenny.diffey@physics.cr.man.ac.uk
[2] Division of Imaging Science and Biomedical Engineering, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT
sue.astley@manchester.ac.uk
[3] Bolton, Bury and Rochdale Breast Screening Service
Royal Bolton Hospital, Bolton, Lancashire, BL4 OJR
anthony.maxwell@rbh.nhs.uk
claire.mercer@rbh.nhs.uk

**Abstract.** Conventional risk models for the development of breast cancer use inputs such as age, weight, hormonal factors and family history to compute individual breast cancer risk. These are employed in the management of women at high risk. The addition of breast density as an input has been shown to improve the accuracy of such models. An improved risk model could facilitate risk-based population screening. However, in order to use breast density in risk models there is a need to employ objective methods for measuring the density. A feasibility study has been carried out to assess the practicality of using a stepwedge-based technique for measuring breast density from mammograms in the UK National Health Service Breast Screening Programme and to determine whether additional information, relevant to risk, can be collected by questionnaire. Preliminary results suggest that it is practical to use such a technique in the screening environment. In a sample of 100 women, the mean density was 27% (range 2 - 81%). A negative trend in breast density was observed with Body Mass Index.

**Keywords:** Breast density, risk factors, prediction models

## 1 Introduction

A number of established mathematical models exist for the estimation of individual risk, including the Gail [1], Claus [2] and Ford [3, 4] models. The earlier models are limited in that they do not integrate information on family history, hormones and benign breast disease in a comprehensive fashion [5]. The model developed by Tyrer et al [6] incorporates such risk factors, for example, age at menarche, parity, age at first childbirth, age at menopause, atypical hyperplasia, lobular carcinoma in-situ, height and body mass index. However, hormone therapy and breast density were not considered.

Recently, Barlow et al [7] developed a model including breast density and the use of hormone therapy as additional inputs. The Breast Imaging Reporting and Data System (BI-RADS) [8] was used to classify density into one of four categories, based on subjective assessment of appearance. Breast density was found to be a statistically significant risk factor for breast cancer diagnosis in pre- and post-menopausal women and it is thought that its inclusion in risk prediction models may offer improved accuracy in the identification of women at high risk of developing breast cancer. It is possible that an improved risk model could facilitate risk-based screening. This means that the frequency of screening is determined by the level of risk, for example, women at low risk would be screened less frequently than women at high risk.

We have carried out a feasibility study with the primary aim of assessing the practicality of using a stepwedge-based method to measure breast density from mammograms in the UK National Health Service Breast Screening Programme (NHSBSP). An additional aim was to ascertain whether information relevant to individual breast cancer risk could be obtained from women attending for routine screening, using a questionnaire.

## 2 Method

6,000 women invited for routine breast screening were invited to participate in the study. These women had their mammograms taken at either a screening van or a static site. An information sheet, covering letter and brief questionnaire were sent to each woman prior to her appointment. Invitations for screening appointment in this breast screening programme are sent out three weeks before the appointment; the study information was sent separately two days after the invitation as we were limited to two sheets of paper per envelope using the automatic folding machine.

The questionnaire gathered information on date of birth, height, weight, date of first pregnancy, ages of menarche and menopause, ethnicity and family history of breast cancer (mother or sister only) including the age at which breast cancer was diagnosed in this relative. With the exception of age, this information is not available in the patient's notes. Further relevant information, including use of hormone replacement therapy and details of present or previous symptoms, was recorded in the patient's notes. For each question, the woman was asked to tick a box to state how certain she was about her answer ('don't know', 'not sure', 'quite sure', 'certain'). It was hoped that this would give us an approximate indication of the reliability of data. In order to quantify the error associated with the weight data provided, it was anticipated that a sample of the women attending at the static site would be weighed using scales calibrated on a monthly basis and the actual weight recorded alongside the reported value.

Women willing to participate in the study took their completed questionnaire with them when they attended for their routine mammogram. Informed consent was taken during the appointment by a radiographer or receptionist trained in doing so.

Mammograms were taken as usual. An aluminium stepwedge remained clipped to the breast support platform alongside the view markers, and an acetate with magnification markers used to measure breast thickness remained on the compression paddle at all times. This is shown in Figure 1a; a mammogram from the study is shown in Figure 1b. It was anticipated that the only time the stepwedge would be moved out of the way was if it was too tall to fit under the compression plate for a small breast, or there was too little space for it to fit alongside a large breast. In either case, the radiographer would note this on the consent form. The test objects remained in place for all women imaged during the course of the study to enable retrospective analysis of the anonymised images of those subsequently diagnosed with cancer, with consent.

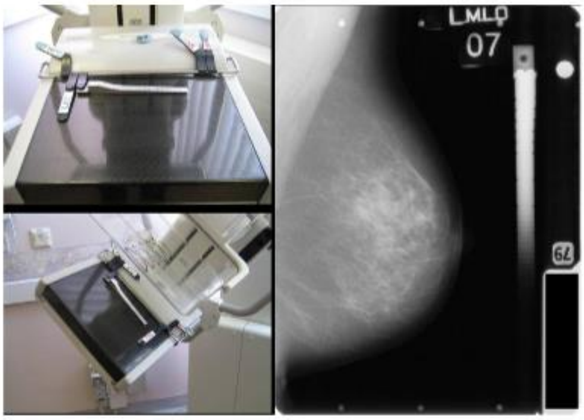



**Figure 1a.** Stepwedge clipped to the breast support platform. **Figure 1b.** Left medio-lateral oblique mammogram taken with the stepwedge and markers in position

Following the reading of films, they were collected from the breast screening unit in order to be digitized and analysed using a semi-automatic procedure that computes the volumes of glandular and fatty tissue in each breast. Data from the questionnaire, the patient's notes and breast density information are recorded in a database.

## 3 Results

The study commenced on 9 July 2007 and data collection was completed in December 2007. Analysis commenced in August 2007. However, it became apparent that the quality of digitization was unacceptable, resulting in a delay prior to the acquisition of a new digitizer in April 2008. Results are presented for a sample of 100 women, with density having been scored by a radiologist. For every view, the percentage of dense glandular tissue was estimated by the radiologist and marked on a 10cm visual analogue scale.

### 3.1 Consent Rate

Out of the 6,000 women invited to participate, 1,414 provided informed consent, giving a consent rate of only 24%. Of these, 1,055 were screened at the static site and 359 were screened on the van. Given that an equal number were invited to each site, the consent rate at the static site was 35% compared with just 12% at the mobile site.

### 3.2 Weight Estimation

The patient information leaflet did not inform women that we wanted to weigh a sample of them even though we had obtained consent for this. As a result, almost all of them refused to be weighed and some said they would not participate in the study if weighing was required. We therefore used their reported weight and made use of results from the literature to estimate the error associated with this value.

### 3.3 Questionnaire Data

Results are presented for the first 600 cases. The ethnicity of this sample is predominantly 'white British' (89%) with the remainder being mainly 'white Irish' or 'Asian/British'. 518 women have attended for previous mammograms of which 12 have had previous breast cancer. 79 women have a first degree relative who has developed breast cancer. Table 1 below shows the number of completed entries and the certainty levels associated with each statement.

**Table 1.** Summary of questionnaire data

| Statement | Number of completed entries | Number of entries for each level of certainty | | | | |
|---|---|---|---|---|---|---|
| | | Certain | Quite Sure | Not Sure | Don't Know | Left blank |
| Height | 588 | 273 | 234 | 28 | 0 | 56 |
| Weight | 571 | 228 | 244 | 51 | 7 | 62 |
| Age at first period | 582 | 234 | 225 | 97 | 9 | 30 |
| Age at menopause | 535 | 180 | 217 | 99 | 38 | 61 |
| Age at first child | 576 | 493 | 33 | 0 | 0 | 64 |

A blank entry for 'age at menopause' or 'age at first child' could mean that the woman has not started the menopause or has not had children, rather than that she has ignored the question.

### 3.4 Correlation of breast density with risk factors

The mean percentage breast density in the sample is 27% (range 2 - 81%). Taking the mean of the cranio-caudal and medio-lateral oblique density for each side, it was found that on average, right breasts were denser than left breasts by 0.3% with a maximum of 8%. This excludes one woman who had previous cancer in her right breast resulting in the right breast being 25.5% more dense than the left breast. The cranio-caudal view was found to be denser than the medio-lateral oblique view by an average of 0.5% for left breasts and 0.8% for right breasts.

The correlation of breast density with previous cancer, family history, age, weight and body mass index (BMI) was examined. In the sample of 100 women, there were nine cases of previous cancer with no family history. Each of these women was matched to a single control by age (within 3 years), weight (within 6kg), ethnicity and HRT status. The average density in the cancer free breast was 22.8% (±10.2%) in the cases compared with 27.6% (±20.5%) in the controls.

Within the same sample of 100 women, there were 20 cases of family history with no previous symptoms. Each of these women was matched to a single control by age (within 1 year), weight (within 10kg), ethnicity and HRT status. The average density in the cases was 24.0% (±18.2%) compared with 25.3% (±24.3%) in the controls.

It is impossible to draw any conclusions from a sample of this size with such large standard deviation on the mean. Similarly, there appeared to be little correlation between density and age (matched by weight within 5kg) and density and weight (matched by age within 5 years). However, there did appear to be a negative trend between breast density and BMI as shown in Figure 2.

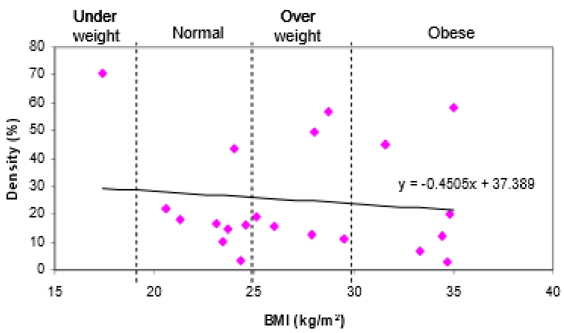



**Figure 2.** Relationship between breast density and BMI

## 4 Discussion

There are several aspects involved in assessing whether a stepwedge-based method of breast density measurement is practical for use in the screening programme. One factor is purely to determine if the method is convenient to use in an environment where appointments last a maximum of 6 minutes. This includes time taken to confirm the identity of the woman, time taken for the woman to get changed and time taken to make two exposures of each breast in both the cranio-caudal and medio-lateral oblique views, bearing in mind that the women attending may be anxious and may require verbal reassurance from the radiographer carrying out the examination. Feedback from the radiographers involved in the study suggests that the stepwedge-based method was suitable for use in these circumstances.

A separate aspect is to determine the adequacy of the method in measuring the breast density of women in the breast screening programme where breasts will vary in thickness from approximately 2cm to 10cm and from predominantly fatty to predominantly glandular in composition. Analysis is underway to determine the proportion of cases where the method fails, either because the stepwedge does not contain a sufficient range of pixel values or because the stepwedge and markers are incorrectly imaged or missing from the film. Preliminary results indicate that it is impossible to use in < 0.8% of cases.

The consent rate of 24% is disappointing as the anticipated rate was 50% and radiographers questioned at various intervals throughout the study felt that 60 – 70% of women were consenting. The effect of the low consent rate is that it may not be possible to correlate density with risk factors with a high enough degree of statistical confidence. However, a sample size of 1,414 should still be adequate to gain information about the breast density distribution in the screening population.

The reason for the relatively poor uptake rate is thought to be due to forgotten questionnaires rather than an objection to the use of the stepwedge and markers which are non-invasive. However, it was stated in the patient information leaflet that the markers may overly tissue on women with larger breasts and a small number of women provided this as the reason for not consenting. Previous work has shown that the markers only overly the breast tissue in 2.5% of cases and do not overly the same area of tissue on both views [10].

It is interesting to note that the consent rate at the static site was almost three times that on the screening van. This can most likely be attributed to the fact that there is a large reception area at the static site where several women can wait for their appointment and complete an additional questionnaire if they have forgotten to bring theirs. In addition, there is a receptionist available at all times to answer any questions or concerns they have about the study. The van is staffed by only two radiographers who move between the reception area, the changing rooms and the examination room.

The issue of women refusing to be weighed came as a surprise. If a study of this nature was to be repeated, it would have to be stated clearly in the patient information leaflet that weighing was required as this is likely to be the parameter with the greatest degree of error associated with it. However, there is an issue of practicality associated with weighing every woman, especially on a van. We would expect a systematic error in the estimate of weight and would ideally like to weigh a sufficient number of women to produce a correction curve for weight.

Based on the 600 cases considered, the mean weight is 69.3kg (±13.5kg) with a range 41.3 - 139.9kg. Using BMI as an indicator, it was found that 1.3%, 40.9%, 36.7% and 21.1% of women were found to be underweight, normal, overweight and obese respectively which appears to be consistent with population data [9]. This suggests that the low consent rate is not linked to weight and that there is no bias introduced in this sample as a result.

The number of adequately completed questionnaires is encouraging and suggests that this is a suitable method for collecting information of this nature on a large scale. The level of certainty associated with each statement is high although in future it may be worth employing more rigorous methods of assessing the reliability of data, at least for a sample. However, the overall number of completed questionnaires is disappointing and as stated above, is likely to be the reason for the poor consent rate. Women attending for breast screening are anxious and for this reason, it would be easy to forget the questionnaire. They may be too worried about the examination to want to complete the questionnaire on-site and on the screening van they would not have enough time to do this. An alternative might be to send the questionnaire by post, with a return envelope, after the screening appointment and possibly after the results. There are limitations with this method. A woman who has been diagnosed with cancer or recalled for further investigation may be too upset to complete a questionnaire; additionally, women may not be interested in participating in a study after their mammogram has been taken and they have received their results.

In conclusion, we have shown that it is feasible to use a stepwedge-based method to measure breast density in the UK NHSBSP. The method had previously only been assessed on a small sample of women taking part in a lifestyle study. Although it was shown to be feasible [9, 10], the stepwedge used was made of Teflon and was far larger and more difficult to attach to the breast support platform making it unsuitable for use in the screening programme where the appointment time is much shorter. The advantages of a stepwedge method are that it is an objective technique and provides a measure of the volumes of glandular and fatty tissue, which should provide more relevant correlations with factors such as weight and BMI than a subjective assessment of percentage density based on area.

## Acknowledgements

We are very grateful to John Lewis and Genesis for providing funding for this study. We would like to thank all staff involved at Bolton, Bury and Rochdale Breast Screening Service, in particular Susan Butler and Jackie Kirby. Thanks also go to Tom Hamnett at the Christie Hospital for constructing the stepwedges and marker sheets required for the study.

## References

1. Gail, M.H., Brinton, L.A., Byar, D.P., Corle, D.K., Green, S.B., Schairer, C., Mulvhill, J.J.: Projecting individualised probabilities of developing breast cancer for white females who are being examined annually. J. Nat. Can. Inst. 81(24), 1879--1886 (1989)
2. Claus, E.B., Risch, N., Thompson, W.D.: Autosomal dominant inheritance of early-onset breast cancer: implications for risk prediction. Cancer. 73(3), 643--651 (1994)
3. Ford, D., Easton., D.F.: The genetics of breast and ovaraian cancer. Br. J. Cancer. 72(4), 805--812 (1995)
4. Ford, D., Easton., D.F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop D.T., Weber, B., Lenoir, G., Chang-Claude,, S. J., H., Teare, M.D., Struewing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T.R., Tonin, P., Neuhausen, S.,, and R. Barkardottir, Eyfjord, J., Lynch, H., Ponder, B.A., Gayther, S.A., Zelada-Hedman, M and the Breast Cancer Linkage Consortium.: Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. Am. J. Hum. Genet. 62, 676--689 (1998)
5. Amir, E., Evans, D.G., Shenton, A., Lalloo, F., Boggis, C., Wilson, M., Howell, A.: Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. J. Med. Genet. 40, 807--814 (2003)
6. Tyrer, J., Duffy, S.W., Cuzick, J. , A breast cancer prediction model incorporating familial and personal risk factors.: Statisitics in Medicine. 23, 1111--1130 (2004)
7. Barlow, W.E., White, E., Ballard-Barbash, R., Vacek, P.M., Titus-Ernstoff, L., Carney, P.A., Tice, J.A., Buist, D.S.M., Geller, B.M., Rosenberg, R., Yankaskas, B.C., Kerlikowske, K.: Prospective breast cancer risk prediction model for women undergoing screening mammography. J. Nat. Can. Inst. 98(17), 1204--1212 (2006)
8. American College of Radiology. Illustrated Breast Imaging Reporting and Data System (BI-RADS $^{TM}$), 4$^{th}$ ed. Reston (VA): ACR (2003)
9. UK Office of National Statistics: Obesity among adults: by sex and NS-SeC, 2001 http://www.statistics.gov.uk/StatBase/ssdataset.asp?vlnk=7447&More=Y
10. Berks, M., Diffey, J., Hufton, A., Astley, S.: Feasibility and Acceptability of Stepwedge-Based Density Measurement. In: Digital Mammography, Proceedings of 8$^{th}$ International Workshop, IWDM 2006, pp. 355--361. Ed. Astley, S.M., Brady, M., Rose, C. and Zwiggelaar, R. Lecture Notes in Computer Science 4046 (Springer-Verlag, Berlin & Heidelberg)
11. Patel, H.G., Astley, S.M., Hufton, A.P., Harvie, M., Hagan, K., Marchant, T.E., Hillier, V., Howell, A.: Automated breast tissue measurement of women at increased risk of breast cancer. In: Digital Mammography, Proceedings of 8$^{th}$ International Workshop, IWDM 2006, pp. 131--136. Ed. Astley, S.M., Brady, M., Rose, C., and Zwiggelaar, R. Lecture Notes in Computer Science 4046 (Springer-Verlag, Berlin & Heidelberg)

# Appendix Two: Support of conference proceedings

## A.2.1 Symposium Mammographicum Conference 2010



detection performance for the three imaging systems operational in our screening program.

### 4.1b 14.02-14.12 The Impact of Breast Compression on Mammographic Image Quality: initial findings

**Principal Author: C E Mercer**
*Royal Bolton Hospital NHS Foundation Trust, UK*
**Contributing Authors:** P Hogg, *University of Salford, UK* & J Diffey, *Christie Hospital, UK*

**Purpose/Background/Objectives:** In mammography it is considered that compression reduces radiation dose and improves image quality. No guidelines exist on how much pressure should be applied for different breast types and volumes, consequently for 'similar patients' there can be variation in imaging practice. This research seeks to establish whether any relationship exists between compression and image quality.
**Methods**: Ethical approval obtained for 2,000 images. Following parameters recorded: applied pressure; breast type (BIRADS); radiation dose, breast thickness, breast volume. Pilot study assessed operator variability. Inferential statistical tests (ANOVA, Kruskall-Wallis, Spearman's Rank) were applied to data.
**Results:** Spearman's rank revealed no relationship between breast compression and image quality score in all density categories under these study determinants. ANOVA / Kruskall-Wallis showed that all radiographers do not use the same mean compression value. Strong positive correlations were found between breast volume/breast thickness and breast dose/ breast thickness.
**Conclusion:** Findings suggest no correlation between breast compression and overall image quality grade; further research is strongly suggested to determine a more robust technique for the assessment of image quality for breast compression analysis. We are currently developing a new psychometric scale for assessing image quality and intend to apply this scale to the same 2,000 images.

### 4.2b 14.02-14.12 MRI for lobular breast carcinoma; is it likely to be useful?

**Principal Author: T Hanna**
*Derriford Hospital, Plymouth, UK*
**Contributing Authors:** R Watkins & S Andrews, Derriford Hospital, Plymouth, UK.

**Introduction:** Invasive lobular cancer (ILC) is often multifocal with implications for surgical treatment. MRI more accurately detects multifocality than standard imaging and NICE guidelines recommend MRI for patients considering breast conserving therapy (BCT). Our aim was to determine the potential benefit of MRI.
**Methods:** Women diagnosed with ILC between 1996 and 2009 who did not have MRI were identified. The preoperative diagnosis and surgical treatment were recorded.
**Results:** 366 women underwent surgery. 159 (43%) initially received BCT and 207 (57%) mastectomy. Of 159 having BCT, only 94 had a preoperative diagnosis of ILC and would now warrant MRI. Of these 64 had no further surgery. 18 required completion mastectomies, 9 had repeat BCT and 3 needed repeat BCT and completion mastectomy. The maximum theoretical advantage from MRI would be avoidance of 33 repeat operations but at a cost of approx £500 for each patient with ILC eligible for BCT.
**Conclusions:** Preoperative MRI staging of ILC could potentially reduce repeat operations by 26%. Unless MRI can reduce reoperation rates to below 5% its costs may be difficult to justify.

### 4.3b 14.02-14.12 Comparing the accuracy of digital breast tomosynthesis with full field digital mammography

**Principal Author: R K Wasan**
*King's College Hospital, UK*
**Contributing Authors:** A. Iqbal, D.R. Evans, C. Peacock, J.C. Morel, A. Douiri, C.P. Lawinski & M.J. Michell, *King's College Hospital, UK*

**Purpose:** To compare the accuracy of Digital Breast Tomosynthesis (DBT) with 2D Full-Field Digital Mammography (FFDM) in women recalled for mammographic abnormalities found on routine screening.
**Methods:** Ethics approval for the study was obtained in December 2008. Entry into the study was offered to all women recalled for further assessment of a mammographic abnormality found on routine film-screen mammography. Study participants underwent bilateral 2D FFDM and DBT in the cranio-caudal and medio-lateral oblique projections using the Hologic Dimensions unit. Mammographic features, mammography score using the RCR Breast Group classification 1 to 5, breast parenchymal density and outcome for assessment were recorded. Receiver Operating Characteristic (ROC) analysis was applied.
**Results:** Results of the first 450 study participants are presented. The participation rate was 91.3% of eligible women. 107 (23.8 %) were diagnosed as malignant (in situ or invasive cancer), 156 (34.6%) as benign and 187 (41.6 %) as normal. ROC analysis demonstrates a significant improvement in diagnostic accuracy of DBT compared to FFDM (0.9649 and 0.9125, respectively; $p=0.0001$) and the effect is significantly greater for soft tissue lesions compared to microcalcification.
**Conclusions:** DBT is more accurate compared to 2D FFDM in the assessment of mammographic abnormalities detected on routine film-screen mammography.

### 4.1c 14.14-14.24 False positive mammographic screening; factors influencing re-attendance

**Principal Author: P Fitzpatrick**
*National Cancer Screening Service, Ireland, University College Dublin, Ireland*
**Contributing Authors:** P Fleming, S O'Neill, D Kiernan & T Mooney, *National Cancer Screening Service, Ireland.*

PROFFERED PAPERS

# A.2.2 Symposium Mammographicum Conference 2012

*Symposium Mammographicum Conference 2012*

## Oral presentations

### *Session 4.1 – Challenges of screening*

#### 4.1 (1) Six year longitudinal study of pressure force in screening mammography

**Mercer, C.E.[1]; Hogg, P.[2]; Szczepura, K.[2]; Denton, E.[3]; McGill, G.[4]**

*[1]Royal Bolton Hospital NHS Foundation Trust, United Kingdom; [2]University of Salford, United Kingdom; [3]Norfolk and Norwich University Hospital Trust, United Kingdom; [4]The Christies NHS Foundation Trust, United Kingdom*

**Background** Previous research[1] identified compression is more heavily influenced by practitioners than clients. This retrospective longitudinal study (6 years) assessed three consecutive screening attendances to determine how pressure varied within and between practitioners and clients.

**Method** Retrospective selection of 500 clients; commencing at 50 years. One centre, GE DMR+ analogue. Recorded data: practitioners, compression, breast thickness, BI-RADS density, dose estimations. Exclusion criteria included: breast surgery, previous/ongoing interventions, and implants.

**Results** Significant pressure variations over 3 screens noted for the same client. Amount of pressure applied highly dependent upon the practitioner. 3 practitioner compressor groups demonstrated: high (mean 126N), intermediate (mean 89N) and low (mean 67N). When the same practitioner performed the 3 screens, pressure variation was low (–40N to +25N). When practitioners from different compressor groups performed 3 screens variations were higher (–20N to +100N). Retrospective dose analysis demonstrate mean reductions of 0.07 mGy (MLO), 0.05 mGy (CC) from an image taken by low compressors compared to an image taken by high compressors.

**Conclusions** Amount of pressure used highly dependent upon practitioner factors. Implications for radiation dose, image quality consistency and client experience over sequential attendances.

**Reference**

1. Mercer. C. E, Hogg P, Lawson R, Diffey J. Practitioner variability of breast compression in Mammography, UK Radiological UK Radiological Congress 2011 ISBN 10: 0–905749–72–3; ISBN 13: 978–0–905749–72–3.

#### 4.1 (2) Identification of potential under-performance in breast screening interpretation

**Chen, Y.; Gale, A.G.; Dong, L.**

*Loughborough University, United Kingdom*

**Purpose** To determine whether under-performing breast screeners can be identified quickly.

**Methods** UK breast screeners voluntarily undertake the PERFORMS scheme where they read the same set of challenging cases. From the data, any outlier (an individual who is performing significantly lower than their peers) can be identified. However, this can take several months. To see if potential under-performers can be quickly identified the anonymous data of 283 participants were re-analysed by bootstrapping the information from 1,000 groups (of sizes 4–50 individuals). From this, a distribution of 1,000 estimated outlier threshold values was constructed. Then the accuracy of these estimations was determined by calculating the median value and standard error of this distribution and comparing it with the known actual outlier threshold value.

**Results** Using data from as few as 50 individuals allowed a good approximation of the known outlier cut off value.

**Conclusions** Individuals who are performing markedly differently to their peers can be identified by examining the data on the PERFORMS scheme of groups of 50 individuals. This approach is being implemented in the PERFORMS scheme which enables individuals who have difficulties to be identified very early after taking part and then helped to improve their performance.

#### 4.1 (3) The effect of pain in mammography on participation in breast cancer screening: a systematic review

**Whelehan, P.; MacGillivray, S.**

*University of Dundee, United Kingdom*

**Purpose/Background** Despite the common perception that mammography is painful, known relevant studies vary in their assessments of any deterrent effect on breast screening participation. We therefore undertook a systematic literature review to assess the current evidence and quantify the proportions of women dissuaded from breast screening by mammography pain.

**Methods** Searches were run in 10 online databases. Articles were included if they contained: data from a





ORAL PRESENTATIONS

Symposium Mammographicum

This is to certify that

*Mrs Claire Mercer*

Royal Bolton Hospital NHS Foundation

has been awarded the prize for the

**Best Oral Presentation**

in the session entitled Challenges of Screening

at Symposium Mammographicum 2012

for the presentation entitled

*Six year longitudinal study of pressure force in screening mammography*

Dr Michael J Michell
President
Symposium Mammographicum 2012

## A.2.3 Symposium Mammographicum Conference 2014

### Bournemouth International Centre

### 29 June - 1 July 2014

### 2014 Programme

We are delighted that Professor Richard Sullivan, Director of the new Kings Health Partners Institute of Cancer Policy, has agreed to deliver The Sir John Stebbings Lecture on Age and Affordability.

Professor Elizabeth Morris will be joining us from the United States where she is Chief of the Breast Imaging Service at the Memorial Sloan-Kettering Cancer Centre. We are really pleased as she will be presenting two talks; the first of which will focus on imaging and new technology and the second will look at MRI screening in the younger high risk woman and avoiding over diagnosis.

Other confirmed speakers include Mrs Claire Mercer (University Hospital of South Manchester), Mrs Claire Borelli (St George's Hospital) and Dr Sian Taylor-Phillips (University of Warwick), who will be focussing on topics such as Compression, Implants and Fatigue and Changing Case Order in Breast Screening Radiology: The CO-OPS Trial.

Other highlights will include Dr Elizabeth O'Flynn discussing MRI Parameters, Professor Andy Evans presenting Shear Wave Elastography Prediction and Professor Fiona Gilbert looking at Tomosynthesis.

There will be a session dedicated to Tailored Treatments and Professor Carlos Caldas (Cancer Research UK), Miss Adele Francis (University Hospitals Birmingham NHS Trust) and Professor Lesley Fallowfield (Brighton and Sussex Medical School) will focus on Biology of Breast Cancer, The DCIS Trial – initial experience and Our Esoteric Breast Cancer World.

Other sessions planned will include Mammographic Fundamentals, Imaging; Optimising Current Tools and Age and Breast Cancer.

Dr. Ros Given-Wilson
Chair, Organising Committee
Symposium Mammographicum 2014 http://conferencesympmamm.org.uk/

**SYMPOSIUM MAMMOGRAPHICUM 2014**
**PROGRAMME**

| **Sunday 29 June 2014** | |
|---|---|
| 15.00-20.00 | Registration open |
| 17.00-19.00 | Welcome reception and buffet supper in the Exhibition area |

| **Monday 30 June 2014** | |
|---|---|
| 08.45 – 08.50 | Introduction – Dr Michael J Michell, President of Symposium Mammographicum |
| 08.50 – 09.20 | Session 1 – The Sir John Stebbings Lecture<br>Chaired by Dr Michael J Michell |
| | 1.1 Affordability<br>Professor Richard Sullivan, King's College London, UK |
| 09.20 – 10.50 | Session 2 - Imaging: new technology<br>Chaired by Dr Michael J Michell and Professor Ken Young |
| 09.20-09.50 | 2.1 New technology<br>Dr Elizabeth Morris, Memorial Sloan-Kettering Cancer Centre, New York, USA |
| 09.50 – 10.10 | 2.2 Dual energy mammography<br>Dr Matthew Wallis, Addenbrooke's Hospital, Cambridge, UK |
| 10.10 – 10.30 | 2.3 Tomosynthesis<br>Professor Fiona Gilbert, University of Cambridge, UK |
| 10.30 – 10.50 | Discussion |
| 10.50-11.35 | Morning coffee and exhibition |
| 11.35-12.55 | Session 3 – Mammographic Fundamentals<br>Chaired by Ms Zebby Rees and Dr Barbara Dall |
| 11.35-11.50 | 3.1 Practitioner variation in breast compression<br>Mrs Claire Mercer, University Hospital of South Manchester, UK |
| 11.50-12.05 | 3.2 Women's experiences of mammography<br>Ms Patsy Whelehan, University of Dundee, UK |
| 12.05-12.20 | 3.3 Imaging the augmented breast<br>Mrs Claire Borrelli, St George's Healthcare Trust, London, UK |
| 12.20-12.35 | 3.4 Fatigue and changing case order in Breast Screening Radiology: The CO-OPS Trial<br>Dr Sian Taylor Phillips, Warwick Medical School, University of Warwick, UK |
| 12.35-12.45 | Discussion |
| 12.45-14.15 | Lunch and exhibition viewing |
| 14.15-15.30 | Session 4 – Predicting response to treatment<br>Chaired by Dr Elizabeth Morris and Ms Jenny Rusby |
| 14.15-14.35 | 4.1 MRI parameters predicting response<br>Dr Elizabeth O'Flynn, Institute of Cancer Research & The Royal Marsden Hospital, UK |
| 14.35-14.55 | 4.2 Shear wave elastography prediction<br>Professor Andy Evans, University of Dundee, UK |
| 14.55-15.05 | 4.3 Proffered papers from submitted abstracts<br>Preoperative MRI for invasive lobular cancer: not a panacea<br>J Parikh[1], J Scudder[1], A Spence[1], F Worth[1], M Selmi[2], M Charles-Edwards[2]<br>[1]Guys and St Thomas NHS Foundation Trust; UK [2]Division of Imaging Sciences and Biomedical Engineering Kings College, UK |
| 15.05-15.15 | 4.4 Proffered papers from submitted abstracts<br>Preoperative Role of Breast MRI in High Grade Ductal Cancer In Situ<br>Hajaj M<br>Bedford Hospital NHS Trust, UK and Kettering General Hospital Foundation Trust NHS Trust, UK |
| 15.15-15.30 | Discussion |
| 15.30-16.00 | Afternoon tea and exhibition |
| 16.00-17.30 | Session 5 – Imaging: optimising current tools<br>Chaired by Ms Patsy Whelehan and Dr Anna Murphy |

POSTER PRESENTATIONS

explore and resolve their barriers to communication. The 5 courses so far have attracted very high ratings from participants with further developments planned to enhance user experience.

**References:**
1. Fontes, LA (2008) Interviewing Clients across Cultures: a Practitioners Guide. Guilford Press.
2. Jain, AK (2010) NS Evidence – Ethnicity and Health.
3. Kagan, N (1980) influencing human interaction-Eighteen years with IPR. In A.K. Hess (Ed Psychotherapy supervision: Theory, research and practice. New York: Wiley.
4. De Maesschalck S, Willems S, De Maeseneer J, Deveugele M. (2010) Development and validation of EMP-3: an instrument to measure physician's attitudes toward ethnic minority patients. Fam Med. 2010 Apr;42(4):262–7.
5. Tervalon M and Murray-Garcia, J (1998) Cultural humility versus Cultural Competence: a critical distinctio Journal of Health Care for the Poor and Underserved. Vol 9, No.2 117–125.

## P.44
### A National Digital Mammography Image Database and Associated Observer Study Software

M Patel, P Looney, K Young, M Halling-Brown
*Royal Surrey County Hospital, UK*

Research into digital medical imaging requires the large-scale collection and anonymisation of images/data. To conduct such research we have developed a flexible image database, which prospectively collects images and data from UK sites for research.

**Methods:** The database contains unprocessed and processed images, associated data and expert-determined ground truths. Currently, the associated data is made up of radiological, clinical and pathological information extracted from the National Breast Screening System (NBSS). The process of collection, annotation and storage is fully automated and adaptable. All images and data are anonymised. Furthermore, a software application MedXviewer has been developed which allows radiologists to annotate clinical features and participate in observer studies.

**Results:** At present we have collected 2,623 patient cases, consisting of 34,014 2D images of which 680 are normal cases, 1,836 malignant and 107 benign. These images are being utilised in multi-site observer studies.

**Conclusions:** A valuable database have been developed which holds both processed and unprocessed mammographic images. The provision of unprocessed images enables a multitude of potential research possibilities that utilise the images. Furthermore, the availability of associated data and expertly determined ground truth can facilitate other research applications, such as big data analysis.

## P.45
### A mammography image set for observer training and assessment in BI-RADS density classification

C Mercer[1], P Hogg[2], J Kelly[3], R Borgen[4], S Millington[3], S Hilton[4], S Enion[4], P Whelehan[5]
*[1]University Hospital of South Manchester, UK; [2]University of Salford, UK; [3]Countess of Chester Hospital NHS Foundation Trust, UK; [4]East Lancashire Breast Screening Unit, Burnley General Hospital, UK; [5]Division of Cancer Research, Medical Research Institute University of Dundee, UK*

**Purpose:** Breast density categorisation consistency is important when performing research where density is a relevant variable. Minimisation of inter and intra-operator variability is essential if findings are to be meaningful. Aim: validate a set of mammography images for visual breast density estimation to achieve consistency in research and to determine observer agreement.

**Methods:** 50 mammograms scored twice by eight observers (BI-RADS – four category density scale). Scoring agreement within and between observers was assessed. Film screen utilised as research was been carried out on film images. Further work includes repeating this study for digitally acquired images.

**Results:** Six of eight observers achieved strong intra-observer agreement (Cohens' Kappa >0.81). Strong agreement between paired observers was demonstrated in 10 of 28 pairs (first scoring round) and 12 of 28 on second. No observers demonstrated delta variance above 1. Fleiss' Kappa used to evaluate concordance between all observers on first and second scoring rounds (0.64 and 0.56 respectively).

**Conclusion:** We have set a gold-standard score for 50 images and enabled evaluation of observers' scoring. This will facilitate rigour in future research where BIRADS mammographic density scores are relevant.





## P.46
### Comparison of calcification cluster detection by CAD and human observers at different image quality levels

P Looney, L Warren, S Astley, K Young
*NCCPM, UK*

**Purpose:** To compare calcification detection by human observers and CAD at different image quality (IQ) levels.

**Method:** 162 normal screening cases were collected from a Hologic Selenia and subtle calcification clusters inserted. By transforming these images produce four other sets of images with different IQ were created: simulating CR images at the same dose and at half dose and DR images at half and quarter dose. These images were read by seven expert observers in a previous study and by a commercial CAD using the "certainty of finding" (1-100) in the DICOM structured report with JAFROC analysis.

**Results:** At normal dose DR the figure of merit for the CAD was 0.82 and 0.84 for the humans. At the lowest IQ level the figure of merit for the CAD and humans were 0.62 and 0.55 respectively. At each IQ level there was no significant difference ($p>0.05$). The IQ defined by threshold gold thickness had a significant correlation with both human and CAD figures of merit.

**Conclusion:** The performance of the CAD and humans were both significantly degraded by changes in IQ. There was no significant difference in calcification detection between the CAD algorithm and the human observers at each IQ level.

## P.47
### Word of Mouth Mammography e Network (WOMMeN): a feasibility study

L Robinson[1], K Burke[2], M Griffiths[1], C Ure[1], J Stein-Hodges[2], J Shires[3], J Hill[3], J Wray[1], J Murphy[1]
*[1]University of Salford, UK; [2]Royal Bolton Hospitals, UK; [3]Nightingale Centre, UK*

**Purpose:** We plan to develop a Digital Social Network (DSN) app for women attending breast screening. This approach to communication with service users reflects current NHS strategy (DOH 2010). It also responds to women's preferences for eliciting information about mammography using word-of-mouth (Poulos and Llewellyn 2005) and their growing interest in social networking (Brenner 2012).

**Aim**: This paper presents the findings of a feasibility survey.

**Method:** A survey was sent to a convenience sample comprising university employees ('service-users') and employees at two NHS Trusts ('health professionals'). It explored: current social media usage; preferred health information format; preferred modes of interaction with other users and health professionals; and whether WOMMeN was a good idea.

**Results:** 88 surveys were completed:
- ~75% use DSNs, mainly for 'social chat'.
- DSNs are not frequently used for information/networking related to health, but Twitter is the preferred choice.
- Women want to view uploaded videos but not to upload their own.
- Health professional involvement on the DSN is desirable.
- There may be some reticence by health professionals to engage in health-related DSNs.

**Conclusions:** Generally the app was thought to be a good idea. We intend to extend the survey to include a wider cross-section of UK communities.

**References:**
1. Brenner Joanna (2012) Pew-Internet: Social Networking [available at] http://pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx [accessed 5/2/13].
2. DOH (2010) Equity and excellence: Liberating the NHS. TSO.
3. Poulos, A. Llewellyn, G. (2005) Mammography discomfort: a holistic perspective derived from women's experiences. Radiography11 (1): p.17–25.

## P.48
### How do the results of virtual clinical trials using simulated cancers relate to cancer detection in screening

L Warren[1], L Dummott[1], M Wallis[2], R Given-Wilson[3], J Cooke[4], J Dance[1], J Young[1]
*[1]Royal Surrey County Hospital, UK; [2]Royal Surrey County Hospital, UK ;[2]Cambridge Breast Unit, UK; [3]St Georges Hospital, UK; [4]Jarvis Breast Screening and Diagnostic Centre, UK*

POSTER PRESENTATIONS

### A.2.4 UKRC Conference 2010

## Proffered presentation awards

The winners of the proffered presentation competition at this year’s United Kingdom Radiology Conference (UKRC) were awarded as the celebrations continued.

The awards were given as follows:

**The Alan Nichols Award** went to Claire Mercer from the Royal Bolton Hospital for her paper, ‘The impact of breast compression on mammography image quality:  ‘Initial Findings’.

In a gesture to mark what would have been Alan’s 100th year had he still been alive, this award was presented by his three children.

Claire’s win marked the second consecutive year that someone from the Royal Bolton had picked up this particular prize.

**The Forder Memorial Award** was given to Tienne Lockwood from the University of Bradford for her paper, ‘Diffusion tensor imaging and schizophrenia’ poster.

**The Beth Whittaker Award** was won by Kieran Murphy from the Liverpool Heart and Chest Hospital for his poster, ‘Use of a blood pool contrast agent for MR vascular mapping in patients with cystic fibrosis’.

## A.2.5 UKRC Conference 2012

UKRC

POSTER PRESENTATIONS

Departmental protocol on imaging women with breast implants is based on guidance from the NHSBSP (2002) and RCR (2003) and should be followed. Despite this the protocol was not being adhered to and there was only a 10% increase in the use of the modified view even after training. As this is the case it could be concluded that there needs to be national guidelines on imaging women with augmented breasts.

### P-077 Six year longitudinal study of pressure force in screening mammography

*Claire Mercer; Peter Hogg; Katy Szczepura; Erika Denton; George McGill;*
*Royal Bolton Hospital NHS Foundation Trust; University of Salford; University of East Anglia, The Christie NHS Foundation Trust*

**Aims/Objectives:** Our previous research identified applied pressure in mammography is more heavily influenced by practitioners than clients. With pressure variability in mind, this retrospective longitudinal study (6 years) assessed 3 consecutive screening attendances, to determine how pressure varied within and between practitioners and clients.

**Content:** Consecutive screening mammography with retrospective selection of 500 clients, commencing at 50 years old. One centre using GE DMR+ analogue mammography machine. Recorded data included: practitioners, applied pressure, breast thickness, BI-RADS density, dose estimations. Exclusion criteria: previous breast surgery, previous /ongoing interventions including assessment, implants and volume change.

**Relevance /Impact:** To assess if pressure application has any dependence on the practitioner rather than the client.

**Outcomes:** pressure variations over 3 screens were noted for the same client. Amount of pressure applied highly dependent upon the practitioner. 3 practitioner compressor groups were demonstrated - high (mean 126N), intermediate (mean 89N) and low (mean 67N). Same client: when the same practitioner performed the 3 screens, pressure variation was low (-40N to +25N); when practitioners from different compressor groups performed 3 screens variations were higher (-20N to +100N). Retrospective dose analysis demonstrate mean reductions of 0.07mGy (MLO), 0.05mGy(CC) from an image taken by low compressors compared to an image taken by high compressors.

**Discussion:** The amount of pressure used seems highly dependent upon practitioner rather than client factors. Implications for radiation dose and image quality consistency. It may also affect client experience re-attendance over sequential attendances.

### P-078 The use of vacora biopsy in the NHSBSP

*Claire Mercer,*
*Royal Bolton Hospital NHS Foundation Trust*

**Aims/Objectives:** The aim of NHSBSP is to ensure accurate diagnosis at the earliest detectable stage whilst minimising the number of women for open biopsy for benign disease and maximising the number of women with cancer with a non-operative diagnosis of malignancy. As such, the requirement for diagnosis on first biopsy is high. The use of larger gauge needles and vacuum assistance for the assessment of suspicions lesions enables prompt diagnosis and can obliterate the requirement for open biopsy.

**Content:** Our service utilises various biopsy devices; namely the Mammotome®, Vacora® and Achieve® systems. The use of Vacora® is a relatively new introduction to our service. The device is used in conjunction with other biopsy products and is not considered to be a replacement. An introduction to the device and the training that is to be considered if introducing this technique to your service follows.

**Relevance/Impact:** This will aid to highlight good practice and identify training requirements for staff new to this technique.

Results:The system is intended for diagnostic sampling of breast tissue and is not used for therapeutic excision. It is excellent for calcification evaluation and is a simple procedure which is carried out during the assessment clinic appointment. It achieves highly accurate diagnostic results with the advantages over traditional 14g core biopsy being its ability to target vague diffuse areas. The patient's acceptance of the procedure remains high in our service.

**Discussion:** The Vacora® breast biopsy under stereotactic guidance is used in this department with increasing regularity and is now becoming the procedure of choice for first line investigation for an increasing number of breast lesions. It may also affect client experience re-attendance over sequential attendances as women find higher pressures more uncomfortable.

### P-079 Optimising paddle and detector pressures and footprints in mammography

*Alison Darlington; Peter Hogg; Katy Szczepura; Anthony Maxwell;*
*Pennine Acute NHS Trust; University of Salford; Royal Bolton Hospitals NHS Foundation Trust;*

**Introduction:** The breast is compressed during mammography between a fixed detector plate and a moveable compression plate. Ideally these should exert a uniform pressure on the breast. This study compares different detector positions relative to the breast to determine which give the most balanced surface pressures and contact areas (footprints).

**Method:** A breast phantom of similar compression characteristics to female breast was mounted on a rigid torso. Positioning (CC projection only) was in line with recommended practice. A flexible multi-sensor pressure mat was wrapped around the phantom so that breast/detector and breast/paddle pressure readings could be taken simultaneously. Readings were taken using Hologic Selenia and Selenia Dimensions mammography units, each with two different paddles, at 60N, 80N and 100N and at five vertical detector positions (-2cm, -1cm, 0, +1cm and +2cm) relative to the infra-mammary fold (IMF).

15 cases were imaged using both techniques (CR and FFD) between November 2010 and November 2011. Specimen radiographs were assessed by three breast radiology practitioners – a consultant breast radiologist, an advanced breast practitioner and a radiology SpR year 5. We scored the conspicuity of microcalcifications (both within the lesion and in the surrounding tissue) and lesion margins. Scores of 1, 2 or 3 indicated whether FFD images were of lower, equal or higher quality than CR images. A third assessment, collating the first two results, gave an overall appraisal of FFD versus CR. Images were interpreted on a mammography-quality PACS workstation with a single monitor in ideal lighting conditions and optimum windowing.

**Results:** FFD images were rated better than CR images in 76% of cases, and better or equal in 98% of cases. A particular strength of FFD is better conspicuity of microcalcifications within the excised lesion ($p<0.0001$, Fisher exact test). Our gold standard for WLE specimen radiography has now shifted from cabinet CR to FFD mammography-acquired DR.

## P-083 Breast density measurements in digital mammography: detector stability analysis

*Oliver Putt; Charlotte Kerrison; Jamie Sergeant; Tina Dunn; Susan Astley; Alan Hufton*
*Department of Physics and Astronomy, School of Cancer and Enabling Sciences, The University of Manchester; Nightingale Centre and Genesis Prevention Centre, Wythenshawe Hostpital;*

**Introduction:** For the longitudinal assessment of breast density detector stability is essential. We have thus analysed mean pixel value (MPV) per unit exposure (mAs) from five Full Field Digital Mammography (FFDM) systems over a period of 22 months.

**Methods:** Daily quality control (QC) data and servicing information were collected from five GE Senographe Essential FFDM units located at a static site and on mobile screening units between January 2010 and October 2011. The QC data were plotted for each machine as MPV/mAs against time. Linear fits were applied to the data to determine whether the value of MPV/mAs could be regarded as constant over an extended period.

**Results:** Periods of up to 6 months in which the MPV/mAs data were stable were identified. Consecutive periods of stability were separated by sudden changes in the mean value of MPV/mAs. By comparing the dates of step changes with servicing records every change can be accounted for. Changes were a result of events such as detector replacement or detector recalibration following routine servicing.

**Discussion:** Our results provide conclusive evidence that the GE Senographe Essential machines are stable over extended periods of time, with the mean value of MPV/mAs only changing in response to machine servicing. Stability allows longitudinal measurements of breast density to be made without the need to image a calibration object alongside the breast. The changes in MPV/mAs can thus be accounted for in the calibration data set.

## P-084 Minimising pressure variability in mammography – an exploratory calibration study

*Peter Hogg; Melanie Taylor; Claire Mercer; Erika Denton; Katy Szczepura;*
*University of Salford; North Manchester General Hospital; Bolton Royal Hospital; University of East Anglia*

Pressure variations in mammography exist between and within practitioners. Variation may affect client experience, radiation dose and image quality. This research reports on a calibration study to improve consistency.

Automatic readouts of breast thickness accuracy vary between mammography machines. Therefore one machine (Hologic Selenia), serving a symptomatic population, was selected for calibration. 250 randomly selected clients were invited to participate; 235 agreed and 940 compression datasets were recorded (comprising breast thickness, breast density and pressure). Pressure was increased from 50N stepping through 10N aliquots until the practitioner felt pressure was appropriate for imaging; at each pressure increment breast thickness was recorded.

Graphs were generated and equations derived; second order polynomial trendlines were applied to the data using least squares method. No difference existed between breast densities but a difference did exist between 'small paddle' and 'medium/large paddles'. Accordingly data was combined, with the Y axis representing average change in breast tissue thickness from 50N. 4 composite graphs were created. Small paddle: CC y=0.0944x2-3.4742x+15.968 ($R^2$=0.9809); MLO y=0.0944x2-3.4742x+15.968 ($R^2$=0.9809). Medium/large paddle: CC y=0.1313x2-4.4331x+19.21 ($R^2$=0.9984); MLO y=0.1323x2-4.575x+19.88 ($R^2$=0.9994). Graphs were colour coded into 3 segments - low, intermediate and high gradients (<-2 (amber); -1<>-2 (green); <-1 (red)). We propose 130/135N could be an appropiate termination pressure using this mammography machine.

Using client 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 so improving client experience and reducing potential variation in image quality.

## P-085 Results of a CT dose audit of new technology scanners

*Catherine Gascoigne; Arnold Rust; James Roberts;*
*Medical Physics Department, Cardiff & Vale UHB, University Hospital of Wales, Cardiff; Radiation Protection Department, Velindre Hospital, Whitchurch, Cardiff,*

A review of radiation doses delivered to patients undergoing X-ray CT examinations was undertaken for three newly installed scanners, in addition to one having undergone a major software upgrade and another having undergone optimisation of all thorax protocols. Three of the scanners featured the GE ASiR reconstruction algorithm, claimed by the manufacturer to reduce radiation

## A.2.6 UKRC Conference 2014

**Clinical: Breast**

**P-040 Mucinous carcinoma and fibroadenoma case study**

Claire Mercer; Valerie Reece
*University Hospital of South Manchester NHS Foundation Trust*

**Aims/objectives:** To investigate the possibility of the misdiagnosis of mucinous breast cancer for a common benign breast lesion eg., a fibroadenoma in the younger age group.

**Content:** 36 year old patient attended the symptomatic clinic with a palpable lump in the inner half of the right breast and a family history of breast cancer. It was initially thought to be a fibroadenoma but later confirmed by histology to be a mucinous carcinoma.

The case study includes images and reports of the following: mammogram, ultrasound, FNA, core biopsy, axilla ultrasound and pathology slide.

The patient was listed for WLE and sentinel node biopsy. The histology report demonstrated an invasive mucinous carcinoma grade 2. There were 0/1 lymph nodes with no lympho-vascular invasion. Low grade cribriform DCIS was also present.

**Relevance/impact:** There is a potential for misdiagnosis when two breast pathologies exhibit similar appearances on imaging and could have an effect on the correct outcome for the patient.

**Outcomes:** The MDT decision recommended adjuvant Radiotherapy and Endocrine therapy (Tamoxifen 20mg per day for 5 years). Local recurrence is a problem with mucinous carcinoma so good margins are required. The patient is awaiting radiotherapy and has been referred for egg preservation.

**Discussion:** The possibility of misdiagnosis can arise due to the fact that pure or nearly pure mucinous carcinoma accounts for no more than 2% of all breast cancers and it occurs more so in older women. This case study discusses other diagnostic differences between fibroadenoma and mucinous carcinoma.

**P-041 A mammography image set for observer training and assessment in BI-RADS density classification**

Claire Mercer; Peter Hogg; Judith Kelly; Rita Borgen; Sara Millington; Beverley Hilton; Patsy Whelehan; David Enion
*University Hospital of South Manchester NHS Foundation Trust; University of Salford; Countess of Chester Hospital NHS Foundation Trust; Burnley General Hospital; Countess of Chester Hospital NHS Foundation Trust; Burnley General Hospital; Medical Research Institute University of Dundee; Burnley General Hospital*

**Aims/objectives:** Breast density categorisation consistency is important when performing research where density is a relevant variable. Minimisation of inter and intra-operator variability is essential if findings are to be meaningful. This research aimed to validate a set of mammography images for visual breast density estimation to help achieve consistency in future research projects, and to determine observer performance (inter- and intra-observer agreement).

**Content:** A set of 50 film-screen mammograms was scored twice by each of eight observers, using the American College of Radiology BI-RADS (Breast Imaging Reporting and Data System) four-category density scale. Scoring agreement within and between observers was assessed.

**Relevance/impact:** This exercise has set a gold-standard score for the test set and enabled the observers' scoring consistency to be evaluated. This will facilitate rigour in future research where BIRADS mammographic density scores are relevant.

**Outcomes:** Six of eight observers achieved strong intra-observer agreement (Cohens' Kappa >0.81). Strong agreement between paired observers was demonstrated in 10 of 28 pairs on first scoring round, and 12 of 28 on second. No observers demonstrated a delta variance above 1. Fleiss' Kappa was used to evaluate concordance between all observers on first and second scoring rounds, with values of 0.64 and 0.56 respectively.

The CNR in Phantom 1 ranged from 3.00 - 9.68 with the highest value at a thickness reduction of 62%. The CNR in Phantom 2 ranged from 4.29 - 10.69 with the highest value at a thickness reduction of 49%.

A linear relationship was shown between thickness reduction and the area of the lesion.

**Conclusion:** For the deformable phantom, using 2AFC, lesion visibility increases as thickness reduces to a certain point beyond which lesion visibility deteriorates. Further research is necessary to understand why visibility deteriorates.

**P-044 A call for client consistency in compression**

Claire Mercer; Peter Hogg; Katy Szczepura; Judith Kelly; Rita Borgen; Erika Denton; Sara Millington; Beverley Hilton
*University Hospital of South Manchester NHS Foundation Trust; University of Salford; Countess of Chester Hospital NHS Foundation Trust; Burnley General Hospital; University of East Anglia and Norfolk & Norwich University Hospital; Countess of Chester Hospital NHS Foundation Trust; Burnley General Hospital*

**Aims/objectives:** The application of mammographic compression force is influenced by the practitioner which may affect client experience. This study establishes if practitioners vary in compression force application, and the resultant compressed breast thickness, at 3 NHS Breast Screening Service (NHSBSP) sites.

**Content:** Each site provided data from 3 consecutive screens for 500 clients and recorded: practitioner code, compression force(N), breast thickness(mm), BI-RADS® density. Exclusion criteria: breast surgery, previous/ongoing assessment, breast implants. 975 clients met the inclusion criteria: 2925 images. Variation of compression force(N) and breast thickness(mm) were analysed.

**Relevance/impact:** Demonstrated that practitioners vary in compression force and resultant compressed breast thickness applied at different NHSBSP sites.

**Outcomes:** Compression force varied significantly between sites. Site 1 had three varying practitioner compressor groups each significantly different to each other. Site 3 had a protocol for required minimal compression of 100N.

**Results:** Sites 1&2 demonstrated no significant difference in mean, 1st & 3rd quartile compression force and breast thickness values CC($p>0.5$), MLO($p>0.1$); with sites 1&3 and sites 2&3 demonstrating a significant difference($p<0.001$).

**Discussion:** The amount of compression force applied by practitioners and the resultant compressed breast thickness is not consistent across these 3 sites. Certain standardisation is found when guidance dictates minimum force in site 3. This may have a positive impact on image quality comparisons over time, radiation dose, potentially cancer detection. A large variation could negatively impact on patient experience; varying pain each attendance; potentially reducing rates of re-attendance and cancer detection. NHSBSP standards required to guide practitioners to ensure consistency in image quality and re-attendance over screens.

**P-045 The role of magnetic resonance image guided 2nd look ultrasound - effecting change in management for patients considered for breast conserving surgery**

Nikhil Rao; Praveen Varra; Muthyala Sreenivas
*University Hospitals Coventry & Warwickshire NHS Trust*

**Introduction:** Magnetic Resonance Imaging (MRI) guided second look ultrasound (US) is an established technique for detecting areas of suspicious breast tissue adjacent to a primary breast carcinoma and distinguishing solitary from multifocal disease. It has the advantage of being able to identify and sample these lesions. This has a key role in determining whether breast conserving or mastectomy is performed.

**Methods:** A retrospective study of 50 cases in which MRI guided 2nd look breast ultrasounds was carried out over the period of 30th December 2011 to 3rd of July 2013 (18 months, 240 total MRIs). 90% of 2nd look US were performed by single MRI reporting radiologist. Data was analysed for 29 cases from our institution with completed information. This included; correlation between MRI and US findings, histology results and whether patient management was impacted, in terms of solitary or multifocal disease and subsequent treatment.