

# Development and validation of a novel measure of adverse patient positioning in mammography

\*Patsy Whelehan<sup>a,b,c</sup>, Maria Pampaka<sup>d</sup>, Jennifer Boyd<sup>a</sup>, Sarah Armstrong<sup>a</sup>, Andy Evans<sup>a,c</sup>, Gozde Ozakinci<sup>b</sup>

<sup>a</sup>University of Dundee and NHS Tayside, Mailbox 4, Ninewells Hospital & Medical School, Dundee, DD1 9SY, UK

<sup>b</sup>University of St Andrews, School of Medicine, North Haugh, St Andrews, KY16 9TF, UK

<sup>c</sup>NHS Tayside, Breast Screening & Imaging Department, Ninewells Hospital & Medical School, Dundee, DD1 9SY

<sup>d</sup>The University of Manchester, School of Social Sciences, Oxford Road, Manchester, M13 9PL, UK

\*Corresponding author; email address: [p.j.whelehan@dundee.ac.uk](mailto:p.j.whelehan@dundee.ac.uk) ; Telephone: +44 (0)1382 383292

Additional author email addresses:

Dr Ozakinci: [go10@st-andrews.ac.uk](mailto:go10@st-andrews.ac.uk)

Dr Pampaka: [maria.pampaka@manchester.ac.uk](mailto:maria.pampaka@manchester.ac.uk)

Professor Evans: [a.z.evans@dundee.ac.uk](mailto:a.z.evans@dundee.ac.uk)

## Abstract

### Purpose

The primary aim was to develop and validate a novel mammography positioning measure, specifically incorporating parameters which might relate to mammography pain. We then explored relationships between the new adverse positioning score and (1) pain; (2) patient and technique factors.

### Methods

A 15-item instrument incorporating positioning features with potential to relate to mammography pain was developed. Participants' mammograms ( $n=310$ ) were reviewed for presence of these features. Validity was investigated using the Rasch model. Scores produced by the resultant measure were investigated for associations with patients' pain scores and relevant patient and technique factors, using Pearson correlation, analysis of variance, and multiple linear regression.

### Results

Statistical indices within the Rasch measurement framework provided good evidence that the measure reflected a coherent construct of adverse positioning. Thus, the scores produced with the measurement instrument were valid for use in further statistical analysis. There is, however, scope for improvement of the measure's discriminatory properties.

Adverse positioning scores were higher for greater breast volumes ( $r=0.12$ ,  $p=.0391$ ) and body mass index (BMI) ( $r=0.13$ ,  $p=.0349$ ), and varied by mammographer ( $F(11,298) 2.38$ ,  $p=.0078$ ). The relationships with BMI and mammographer persisted in regression modelling. No relationship was found between adverse positioning and pain.

### Conclusions

Evidence from Rasch analysis suggests that this novel measure is valid for quantifying a coherent "adverse positioning" construct in mammography. Adverse positioning scores varied by mammographer and were related to higher patient BMI but not to mammography pain. The measure warrants expansion, further refinement, and testing in larger studies.

### Keywords

Rasch model, quality, patient experience, pain

### Study registration

ClinicalTrials.gov: [number redacted for peer review]

## Introduction

Experiences of mammography, notably pain from the examination procedure, have been shown to affect breast screening participation [1]. As mortality reductions achieved by screening depend upon participation rates, it is important to ensure that the mammographic examination is as acceptable as possible to women who wish to be screened [2].

Many factors have been linked to levels of pain and discomfort experienced during mammography. They include physical characteristics of the breast, patients' socio-demographic characteristics, psychological influences, and factors relating to quality of care, communication and information provision [3].

Recent research, including intervention development to reduce pain, has tended to focus on compression of the breast during mammography [4,5]. However, higher compression force is not consistently associated with higher pain levels [6].

To our knowledge, no research has been published which directly investigates the effect of mammographers' positioning technique on levels of pain and discomfort, despite longstanding expert opinion that positioning affects pain [7].

Positioning technique has been shown to vary by mammographer [8]. Positioning quality is assessed mainly by mammographers or radiologists examining the resultant images. However, such assessments are notoriously subject to observer variability and there is limited evidence of validity for the measures used [9]. Despite those limitations, research where mammographic image quality is a relevant variable has continued to employ traditional assessment methods, often without including any investigation of the measurement validity of the instrument [10]. Recent efforts have been made to establish greater consensus on appropriate parameters and their assessment, aiming to reduce observer variability in image scoring [11].

This study brings together two major challenges of mammography research and practice: the pain experienced by patients and the validity of positioning technique evaluation. The primary aim was to develop and validate a new instrument for measuring adverse positioning in mammography, focussing specifically on parameters hypothesised to have potential to influence the painfulness of the examination. The secondary aim was to explore relationships between the scores on our new adverse positioning measure and (1) pain during mammography; (2) patient and technical factors.

A further aim of this paper is to provide an introduction to the Rasch measurement framework for a medical imaging audience. The methodology will be unfamiliar to many radiography and radiology professionals but has potential to add value to various aspects of clinical imaging research.

## Methods

### Setting and participant recruitment

The study was approved by the [region blinded for peer-review] National Health Service Research Ethics Committee and the [institution blinded for peer-review] Teaching and Research Ethics Committee and was conducted in line with applicable governance frameworks.

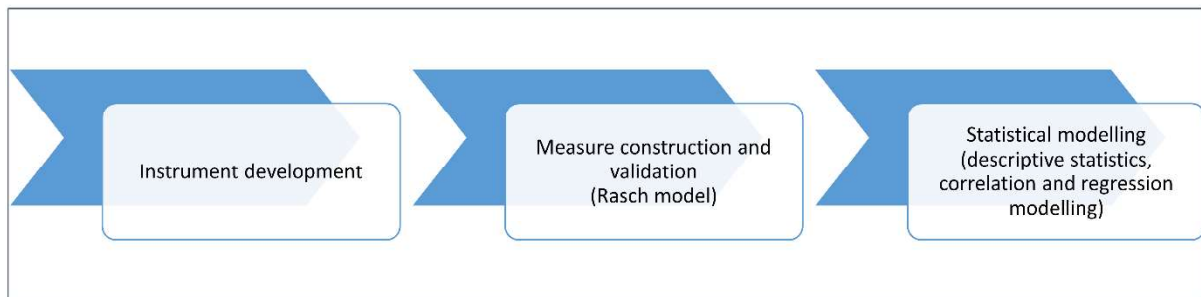
Patients were recruited from a single centre within a UK population-based breast screening service, in which eligible women are invited by letter to attend for mammography. Patients attending the hospital-based clinic for screening were eligible for the study, and were identified in advance through the screening appointment system.

1 Women due to attend clinics between 30 November 2016 and 12 July 2017, on dates when a  
2 researcher would be available, were sent study invitation letters prior to the appointment date. No  
3 additional inclusion criteria were applied but patients not able to read and write English could not be  
4 included.

5 An initial questionnaire was included with the study invitation letter. Women attending for screening  
6 could opt into the study by completing the questionnaire and bringing it to their appointment. On  
7 attendance, they completed additional questionnaires and were asked for written informed consent  
8 for their responses to be linked to their images and associated data.  
9

### 10 11 Overview of measurement and modelling framework

12 The study employed a methodological/analytical framework which is represented in Figure 1 and has  
13 been extensively applied in previous work [12]. The basis of the approach is that, according to  
14 existing research and/or theory, groups of “items” may capture an underlying construct (or “latent  
15 variable”). The Rasch model then enables the validation of the construct and the construction of  
16 valid measures to quantify participants’ positions on the construct. In this case, the theoretical  
17 construct was “adverse positioning in mammography”, with particular reference to pain.  
18  
19  
20



21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32 *Figure 1: A comprehensive measurement and modelling framework*

### 33 34 Instrument development for the adverse positioning measure

35 A list of items, consisting of features of mammographic positioning identifiable on the images, was  
36 postulated as having potential to increase risk of pain during the mammogram. For example, a  
37 trapped skin fold or the inclusion of extra muscle in the compressed field might be expected to add  
38 to the painful potential of the examination.  
39  
40

41 The initial list of variables was reviewed by six field experts, and a refined list of items was  
42 formulated (Table 1). It was hypothesised that these items together might form an underlying  
43 construct of mammography “adverse positioning”.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Item Number	(Short name)	Description
1	PecVisCC	Pectoralis major muscle is visible on either Cranio-Caudal (CC) projection image.
2	FoldsCC	Skin fold(s) visible overlying any part of either CC projection image
3	AirGapCC*	An air gap can be seen in association with any skin fold on either CC projection image.
4	ShoulderCC	Part of the patient's shoulder is visible on either CC projection image.
5	CentredHigh	On either Mediolateral Oblique (MLO) projection image, it is considered that the image receptor was positioned too high for the patient.
6	WidePec	On either MLO projection image, it is considered that too much of the pectoralis major muscle has been included in the field.
7	PecConcave	On either MLO projection image, the edge of the pectoralis major muscle has a concave contour.
8	PecConvex	On either MLO projection image, the edge of the pectoralis major muscle has a convex contour.
9	PecSigmoid	On either MLO projection image, the edge of the pectoralis major muscle has a sigmoid contour.
10	FoldsUpper	On either MLO projection, skin fold(s) visible over the upper part of the breast or overlying the pectoralis major muscle.
11	FoldsLower	On either MLO projection, skin fold(s) visible over the lower part of the breast, including at the infra-mammary angle.
12	AirGapMLO*	An air gap can be seen in association with any skin fold on either MLO projection image.
13	MuscleOther	On either MLO projection, a muscle other than pectoralis major is visible, e.g. pectoralis minor.
14	AnatOther	On a MLO projection, extraneous anatomy is visible, e.g. shoulder, arm or chin.
15	Blur	Movement unsharpness is seen on any image.

Table 1: Items within the proposed "adverse positioning" construct

\*This feature only applies in the presence of one or more skin folds, and is hypothesised to be an indicator of the "severity" of a skin fold.

Each participant's mammogram was reviewed by a single expert radiographer [initials blinded] for the presence or absence of the features listed in Table 1. Scoring was at mammogram level rather than individual image level, to facilitate comparison with a single overall pain score for each examination. A subset of the mammograms (n=50) was read by a second observer – a highly experienced breast radiologist [initials blinded] - to assess reproducibility of the classifications. Agreement between the readers was tested using Cohen's kappa statistic, which is suitable for two observers and binary response data.

## Analysis

Using the response data generated by examining the mammograms, we then attempted to validate the hypothetical "adverse positioning" construct by conducting a psychometric analysis within the Rasch measurement framework [13]. Rasch analysis, in its simplest form (Dichotomous Rasch Model used in this analysis) performed using Winsteps® software [14], mathematically models the probability of a positive response to a given item (e.g. question) in a given measurement instrument (e.g. questionnaire/test) as a function of the respondents' level on the underlying construct

1 measured. If the data adequately fit the Rasch model, valid measurement has been achieved. In the  
2 presence of adequate fit, the Rasch techniques then allow ordinal/categorical raw data (from a set  
3 of items/questions) to be converted to a continuous measure (representing the underlying  
4 construct), facilitating further statistical analysis using the scores that have been produced.

5  
6 Statistical indices comparing the observed data to the (Rasch) model enable assessments of various  
7 aspects of validity; in other words they describe the fit of the items to the model. For example, item  
8 fit statistics and principal components analysis of the residuals (dimensionality analysis) examine  
9 whether the items capture a single coherent construct. Item separation and reliability indices can  
10 verify the item hierarchy (i.e. spread of items in relation to the overall construct being measured),  
11 and the reproducibility of that ordering/hierarchy. Meanwhile, person separation and reliability  
12 indices test the ability of the measure to differentiate participants (participants' mammograms in  
13 this case), and the reproducibility of that differentiation across different samples of participants.  
14 Differential item functioning assessments are used to assess and establish measurement invariance  
15 when the instrument is to be used to compare groups of participants; i.e. they can help to identify  
16 sources of potential construct-irrelevant bias relating to participant characteristics.

17  
18 The application of the Rasch measurement framework in healthcare research, and its advantages  
19 compared to the well-known "classical test theory" approach to psychometric analysis, have been  
20 described in detail in a seminal monograph [15]. For the present study, the validation process and its  
21 results are described in detail in a companion article in *Data in Brief* [citation to Data in Brief article  
22 to be inserted].

23  
24 Following validation and transformation of the data to a continuous variable through Rasch  
25 procedures, one-way analysis of variance (ANOVA) and Pearson correlations were performed to test  
26 for relationships between the adverse positioning score and pain, as well as key patient, breast and  
27 technique variables with theoretical potential to be related to the score. Further exploration with  
28 linear regression modelling was conducted, with scores on the new measure of "adverse  
29 mammographic positioning" as the outcome and a theoretically-informed step-wise procedure  
30 guiding the selection of explanatory variables. Correlation and regression analyses were conducted  
31 using Stata (Stata Statistical Software: Release 14. College Station, TX: StataCorp LP). These analyses  
32 serve as a predictive validity check of the continuous score produced through the Rasch analysis, as  
33 well as being of interest per se.

34  
35 Self-reported height and weight from the patient questionnaires were used to compute body mass  
36 index [16]. Patient age, breast thickness and compression force were extracted from the Digital  
37 Imaging and Communications in Medicine (DICOM) headers of the mammographic images by  
38 VolparaDataManager® software (Volpara Health Technologies Ltd, Wellington, New Zealand),  
39 algorithm version 1.5.2. Breast volume and breast density (volumetric percentage of fibro-glandular  
40 tissue) were estimated by the Volpara algorithm.

41  
42 Pain during mammography was measured using established validated pain scales [17] within the  
43 patient questionnaires. There were two pilot phases within the study, the second taking place after  
44 initial testing and refinement of the questionnaires. For the purposes of this analysis, relevant data  
45 from both phases were compatible and were merged. However, the pain scale used in Pilot 1 was a  
46 0-10 numerical rating scale (NRS), whereas a four-point (None, Mild, Moderate, Severe) verbal rating  
47 scale (VRS) was used in Pilot 2, because there had been too many redundant categories with the  
48 NRS. Pain data were therefore combined and collapsed into three categories for analysis, based on  
49 the findings of Woo and colleagues regarding NRS cut-points for VRS values: 0-4 = none or mild; 5-6  
50 = moderate; 7-10 = severe [18].

# Results

## Sample description

Study recruitment is shown in Figure 2 and participant characteristics in Table 2.

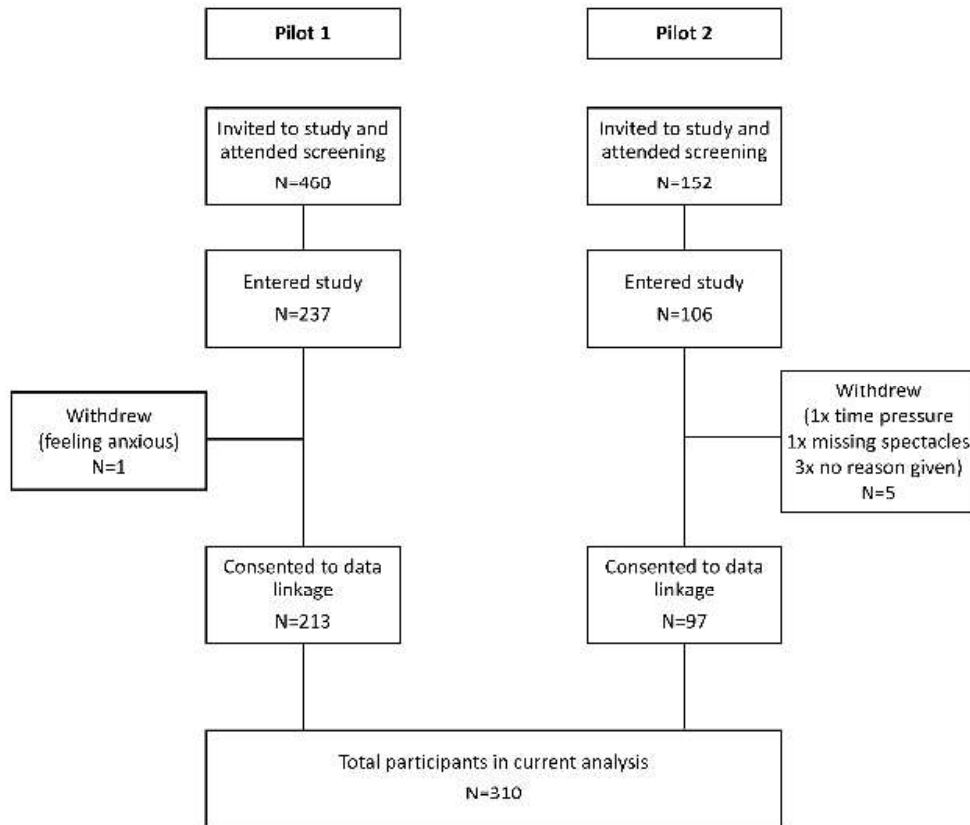


Figure 2: Participation flowchart

Variable	Categories	Frequencies	Percentages
<b>Highest Educational Level</b>	1 – Schooled to age 16	121	39%
	2 – Schooled post-16	19	6%
	3 – Completed further education/ vocational training	97	31%
	4 – Undergraduate Degree	36	12%
	5 - Postgraduate Degree	35	11%
	Missing	2	1%
<b>BMI category</b>	Underweight range (below 18.5)	-	-
	Healthy weight range (18.5-24.9)	87	28%
	Overweight range (25-29.9)	99	32%
	Obese range (30 and above)	90	29%
	Missing	34	11%
<b>Age</b>	50 to 59 years old	140	45%
	60 to 69 years old	138	45%
	70 and older	32	10%
<b>Parity</b>	Has not given birth to any children	52	17%
	Has given birth to at least one child	253	81.5%
	Missing	5	1.5%
<b>Previous breast cancer</b>	No	283	91%
	Yes	15	5%
	Missing	12	4%
<b>Previous breast surgery</b>	No	245	79%
	Yes	56	18%
	Missing	9	3%

Table 2: Participant characteristics

### Inter-observer agreement on mammogram feature ratings

In the subset of mammograms classified by a second reader (50 mammograms, 15 parameters each, totalling 750 observations), agreement between the two readers was 85%, Cohen’s kappa: 0.621 ( $p < 0.001$ ). This kappa statistic can be interpreted as “good” agreement, according to Altman’s specification [19]. For some of the individual features, there were insufficient observations of the feature being present in the subset to calculate kappa statistics but there did appear to be some variability in agreement levels by feature. For example, there was good agreement when identifying skin folds but moderate agreement in identifying whether the image receptor was positioned too high for the woman.

### Measure validation results

The results of validation of the new measure via the Rasch Model are provided in detail in our companion article in *Data in Brief* [citation to Data in Brief article to be inserted]. In summary, there was good evidence of unidimensionality (ability of the measure to capture a single, coherent construct). There was good evidence of the measure’s ability to produce a meaningful and reliable range of responses across items but the sensitivity of the measure to differences between the mammograms was suboptimal. The measure functioned reliably across different participant age groups but some items functioned differently according to patient BMI and mammographer. This differential functioning was considered likely to indicate genuine differences rather than bias, and was therefore investigated as part of further statistical analysis.



Figure 3 shows the mammograms' scores and items' positions on the scale produced via the Rasch procedures. On the left side of the figure, the logit scale (from -4 to 4) is the common measurement scale for both items and persons. In this case, "persons" are patients' mammograms. On the right-hand-side of the "histogram", the items that constitute the scale are presented, ranging from the most frequently observed at the bottom to the least frequent at the top. On the left of the map, the mammograms' distribution on the scale is presented. The higher the mammogram's place in that scale, the more adverse the positioning (as scored by the observer). It is the scores on this logit scale which are taken forward in further statistical analysis concerning substantive research questions.

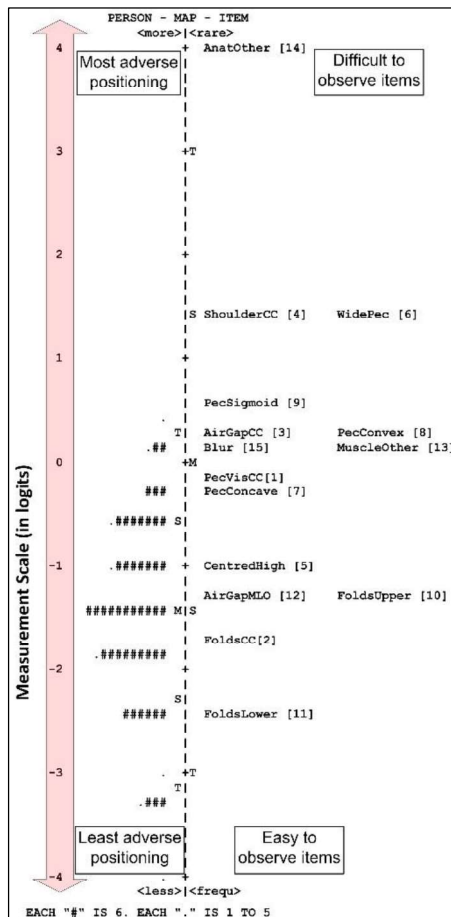


Figure 3: Person-item map

### Further Statistical Analysis using the new measure

We matched the logit scores to the rest of the dataset which included further information about the patients and their mammograms. We investigated correlations between adverse positioning and: (1) pain; (2) patient, breast and technique characteristics that may be expected to affect positioning. We particularly included patient BMI and (anonymised) mammographer identity, on account of the DIF identified during measure validation (see our companion article in *Data in Brief* [citation to Data in Brief article to be inserted]) and the theoretical likelihood of real effects on adverse positioning score.

Frequencies per pain category [18] were: 1 (none/mild): 233 (75%); 2 (moderate): 40 (13%); 3 (severe): 32 (10%); missing data: 5 (2%). Patient-reported mammography pain category was not associated with adverse positioning score according to one-way ANOVA ( $F(2,302) = 0.08, p=0.93$ ).

Pearson correlations between adverse positioning score and key patient and technique characteristics (Table 3) show that some but not all indicators of breast size were significantly associated with adverse positioning score. While compression pressure was associated with adverse positioning, compression force was not.

Positioning correlation with:	Pearson correlation	p-value (*sig)
<b>Breast characteristics:</b>		
Mean breast thickness	0.0235	.6801
Volpara mean breast volume	0.1174	*.0391
Volpara mean breast density	-0.1577	*.0055
<b>Technique characteristics:</b>		
Mean compression force	0.0286	.6166
Volpara mean compression pressure	-0.1587	*.0052
<b>Patients' characteristics:</b>		
BMI	0.1270	*.0349
Weight	0.1078	.0711
Height	-0.0222	.7037

Table 3: Correlations between adverse positioning score and key breast, technique and patient characteristics

Adverse positioning scores by mammographer are shown in Figure 4. Mean adverse positioning scores appear to differ by mammographer. The variation was statistically significant as determined by one-way ANOVA ( $F(11,298) 2.38, p=0.0078$ ).

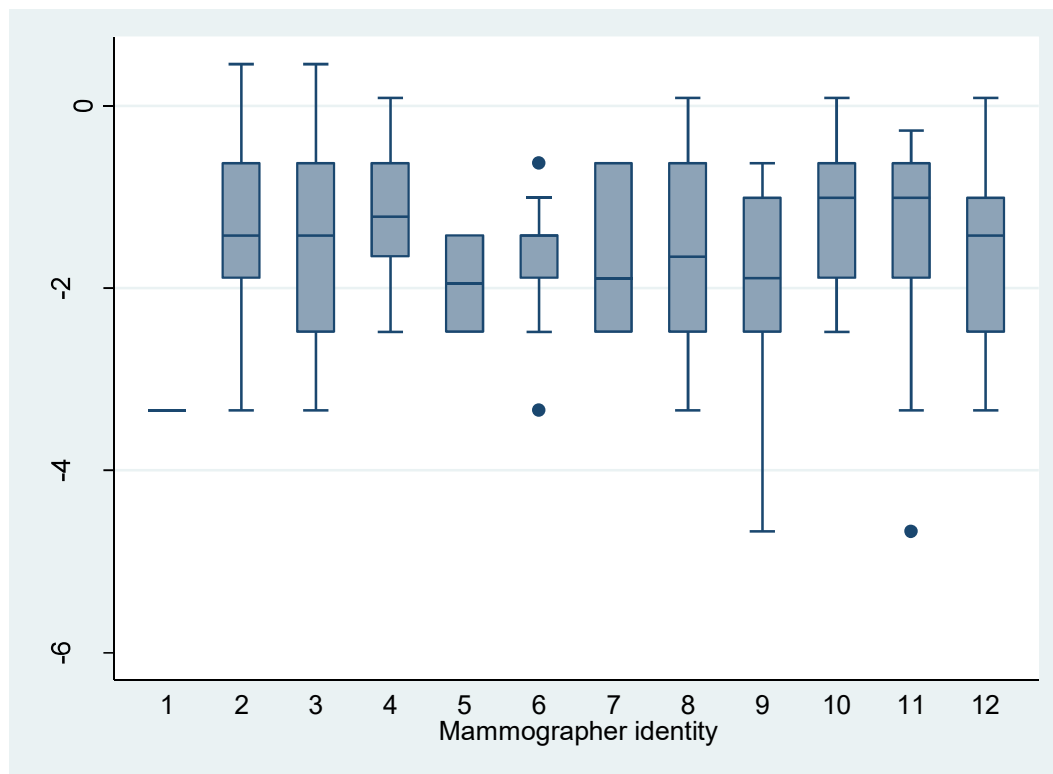


Figure 4: Box plot showing variance in adverse positioning score by mammographer (person performing the mammogram)

Results of multiple linear regression modelling (Table 4) show the variables that significantly relate to adverse positioning, after accounting for the rest of the variables. BMI is positively related (i.e. the greater the BMI, the higher the adverse positioning score), while breast thickness and breast pressure are negatively related. The positive coefficients for operators should be interpreted in comparison to the reference category (Operator 1).

The presented model was chosen due to theoretical reasons for the included variables and better model fit compared to other models, but still only explains a small proportion (14%) of the variance in adverse positioning score.

Variable	Coefficient	St. error	z	P>z
Constant	-1.62	1.054	-1.53	.127
Mean breast thickness	-0.02	0.007	-2.78	*.006
BMI	0.03	0.011	2.66	*.008
Mean compression force	-0.01	0.005	-1.49	.137
Volpara mean compression pressure	-0.05	0.019	-2.80	*.005
Operator (Ref: Operator_1)				
Operator 2	1.94	0.866	2.24	*.026
Operator 3	1.83	0.881	2.08	.039
Operator 4	2.41	0.925	2.61	*.010
Operator 5	1.25	1.047	1.19	.233
Operator 6	1.72	0.886	1.94	.053
Operator 7	1.33	0.996	1.33	.184
Operator 8	1.78	0.881	2.02	*.045
Operator 9	1.43	0.889	1.61	.109
Operator 10	1.85	0.896	2.06	*.040
Operator 11	1.97	0.868	2.27	*.024
Operator 12	1.47	0.871	1.69	.093
<b>Model Fit Information:</b>				
Number of observations	275			
F(15, 259)	2.75			
Probability > F (significance)	<0.001			
R <sup>2</sup>	0.14			
Adjusted R <sup>2</sup>	0.09			

Table 4: Results of regression modelling for outcome "adverse positioning score" (\*statistically significant result)

## Discussion

The primary aim of this work was to develop and validate a new measure of mammographic positioning with particular reference to features hypothesised to be linked to pain experienced by patients during mammography. A measure has been developed and evidence has been generated which supports its unidimensionality, i.e. ability to measure a coherent construct (latent variable). We have called the construct "adverse positioning". Although there is good evidence that the measure is valid, its discriminatory capacity could potentially be improved by further refinement.

Pearson correlations indicated possible relationships between the adverse positioning score and breast volume and BMI (positive correlations), as well as breast density and mean compression pressure (negative correlations). With higher breast density and higher compression pressure (force per unit area) expected to be related to lower BMI and smaller breast size [20], these findings together suggest that adverse positioning is more likely in patients with higher BMI and larger breasts. In regression analysis, the positive relationship between BMI and adverse positioning score

1 persisted whereas breast volume did not add to the model and breast thickness was negatively  
2 associated with adverse positioning. Higher compression pressure remained negatively related.  
3 Overall, these results could conceivably indicate that our measure specifically captures BMI-related  
4 adverse positioning.

5  
6 Our analysis indicated a relationship between mammographer and adverse positioning score, which  
7 is intuitively likely and therefore adds to the validity arguments for our measure. In regression  
8 modelling, relationships between positioning score and some of the individual mammographers  
9 were seen, while also accounting for breast size and BMI. These results indicate potential utility of  
10 the new measure for assessing and monitoring staff and service performance and identifying training  
11 needs.  
12

13  
14 Using ANOVA to compare scores on the new “adverse positioning” measure with patient-reported  
15 pain score categories did not reveal any significant differences. This may be a result of insufficient  
16 sample size, and low numbers in two of the three pain categories used for analysis. It may also be a  
17 result of the suboptimal discriminatory capacity of the new measure.  
18

19  
20 Although we are not aware of any previous published work attempting to quantify relationships  
21 between patient positioning and mammography pain, qualitative work has identified that  
22 mammographers believe positioning can affect experiences of pain [21,22]. However, pain is a  
23 complex phenomenon and it is possible that physical technique factors are less important than  
24 patient factors and practitioner-patient communication. Previous lack of evidence for compression  
25 force being a reliable predictor of mammography pain [6] also supports the possibility that physical  
26 technique factors do not predominate in determining mammography pain. As our ANOVA analysis  
27 did not reveal any significant differences between adverse positioning and pain, we have not  
28 explored this question by multiple regression modelling at this time. We plan a further manuscript  
29 reporting on a more extensive dataset from our study, including multiple co-variables with potential  
30 to affect pain.  
31  
32

33  
34 We can find only one previous published report of a study using psychometric methods to validate a  
35 measure of mammography image quality. Preliminary evidence of validity via a classical test theory  
36 (CTT) approach was presented [23]. We can find no examples of modern psychometric approaches,  
37 such as the Rasch model, being used to validate image quality evaluation measures in diagnostic  
38 images, although the Rasch approach has been applied in scoring radiological images for the  
39 presence of clinical conditions [24]. The appropriateness of using psychometric methods to test a  
40 measure derived from image features which are considered to reflect physical realities can be  
41 questioned. However, the fact that the features are conceptualised and observed by humans  
42 provides justification.  
43  
44

45  
46 Rasch analysis has several advantages over CTT, including insensitivity to missing data, non-reliance  
47 on simple summation of item scores, and the ability to produce scales at the interval level of  
48 measurement, facilitating parametric statistical analysis using the scores produced [25].  
49

50  
51 One of the limitations of our study concerns subjectivity in visual classification of image features, a  
52 well-recognised problem in mammographic image quality assessment [9]. We used only one  
53 observer to classify the images, which is acceptable for a preliminary investigation, but more  
54 observers and investigation of observer variability will be included in future work after further  
55 refinement of the measure. However, in the subset of mammograms that we subjected to a second  
56 read, agreement between two readers was good, comparing well with other relevant recent work  
57 where moderate agreement between experts was achieved [11].  
58  
59  
60  
61  
62  
63  
64  
65

1 Another limitation of our study is the sample size. With over 300 participants, numbers were  
2 adequate for investigating the measurement properties of a new measure but small in the context of  
3 high-volume breast screening. The analyses conducted using the new measure should therefore be  
4 considered as preliminary work and as part of the validation of the measure.

5  
6 Our measure is not a complete clinical image quality measure for mammography, having been  
7 purposely limited to parameters with theoretical potential for association with pain experienced  
8 during the examination. Indeed, at least one of the parameters – pectoral muscle visible on the CC  
9 projection – would normally be considered an indicator of good rather than poor quality. It was  
10 incorporated because inclusion of chest wall tissues in the compressed field may in theory increase  
11 examination painfulness. However, it did not show poor fit within the measure so we did not  
12 remove it at this stage.  
13  
14

15 Valid mammographic clinical image quality measures have long been recognised as both important  
16 and difficult to develop [9], and the lack of optimised measures still persists. Even one of the most  
17 notable recent studies on measuring mammographic clinical image quality, which relied heavily  
18 upon inter-rater agreement as evidence of validity, produced an instrument which the authors  
19 acknowledged still needed further evaluation and validation [11]. The continuing need for a robust  
20 and valid measure of observed clinical image quality in mammography is currently heightened by the  
21 emergence of computerised mammographic positioning assessment methods which themselves  
22 require validation against human concepts of image quality if they are to be trusted.  
23  
24  
25  
26

## 27 Summary, conclusions and future work

28  
29 In this first example of the Rasch model being applied to the evaluation of radiographic positioning,  
30 we have developed a measure with generally good measurement properties. The measure has  
31 enabled possible relationships to be highlighted between adverse positioning score and patient BMI,  
32 as well as showing potential to differentiate positioning performance among mammography  
33 practitioners.  
34  
35

36 No association between positioning features on mammography images and pain experienced during  
37 the examination was found. However, the relatively small study size and the suboptimal  
38 discriminatory power of our measure, as well as the fact that the pain scores in this particular  
39 sample of women were skewed to low levels of pain raise the possibility of a false negative finding.  
40 Therefore, a larger study after further refinement of the measure is warranted.  
41  
42

43 The largely good measurement properties of our new measure suggest that it merits expansion,  
44 refinement and further testing to evaluate its potential as a comprehensive measure of clinical  
45 image quality. Our approach has the potential to provide measurement reliability and validity where  
46 they have hitherto proved highly elusive.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Figure captions

Figure 1: A comprehensive measurement and modelling framework

Figure 2: Participation flowchart

Figure 3: Person-item map

Figure 4: Box plot showing variance in adverse positioning score by mammographer (person performing the mammogram)

## References

- [1] P. Whelehan, A. Evans, M. Wells, S. MacGillivray, The effect of mammography pain on repeat participation in breast cancer screening: A systematic review, *The Breast*. 22 (2013) 389–394. <https://doi.org/10.1016/j.breast.2013.03.003>.
- [2] K.K. Jacobsen, M. von Euler-Chelpin, Performance indicators for participation in organized mammography screening, *J. Public Health (Bangkok)*. 34 (2012) 272–278. <https://doi.org/10.1093/pubmed/fdr106>.
- [3] B. Davey, Pain during mammography: Possible risk factors and ways to alleviate pain, *Radiography*. 13 (2007) 229–234. <https://doi.org/https://doi.org/10.1016/j.radi.2006.03.001>.
- [4] C. Balleyguier, M. Cousin, A. Dunant, M. Attard, S. Delalogue, J. Arfi-Rouche, Patient-assisted compression helps for image quality reduction dose and improves patient experience in mammography, *Eur. J. Cancer*. 103 (2018) 137–142. <https://doi.org/10.1016/j.ejca.2018.08.009>.
- [5] C.R.L.P.N. Jeukens, T. van Dijk, C. Berben, J.E. Wildberger, M.B.I. Lobbes, Evaluation of pressure-controlled mammography compression paddles with respect to force-controlled compression paddles in clinical practice, *Eur. Radiol*. 29 (2019) 2545–2552. <https://doi.org/10.1007/s00330-018-5953-6>.
- [6] N. Moshina, S. Sagstad, S. Sebuødegård, G.G. Waade, E. Gran, J. Music, S. Hofvind, Breast compression and reported pain during mammographic screening, *Radiography*. 26 (2019) 133–139. <https://doi.org/10.1016/j.radi.2019.10.003>.
- [7] G.W. Eklund, Mammographic compression: Science or art?, *Radiology*. 181 (1991) 339–341. <https://doi.org/10.1148/radiology.181.2.1924767>.
- [8] M.-H. Guertin, I. Théberge, H.T.V. Zomahoun, M.-P. Dufresne, É. Pelletier, J. Brisson, Technologists' Characteristics and Quality of Positioning in Daily Practice in a Canadian Breast Cancer Screening Program., *Acad. Radiol*. 23 (2016) 1359–1366. <https://doi.org/10.1016/j.acra.2016.07.002>.
- [9] C. Moreira, K. Svoboda, A. Poulos, R. Taylor, A. Page, M. Rickard, Comparison of the validity and reliability of two image classification systems for the assessment of mammogram quality, *J. Med. Screen*. 12 (2005) 38–42.
- [10] S. Pal, D.M. Ikeda, R.A. Jesinger, L. Jake Mickelsen, C.A. Chen, D.B. Larson, Improving performance of mammographic breast positioning in an academic radiology practice, *Am. J. Roentgenol*. 210 (2018) 807–815. <https://doi.org/10.2214/AJR.17.18212>.
- [11] K. Taylor, D. Parashar, G. Bouverat, A. Poulos, R. Gullien, E. Stewart, R. Aarre, P. Crystal, M. Wallis, Mammographic image quality in relation to positioning of the breast: A multicentre international evaluation of the assessment systems currently used, to provide an evidence base for establishing a standardised method of assessment, *Radiography*. 23 (2017) 343–349. <https://doi.org/10.1016/j.radi.2017.03.004>.
- [12] M. Pampaka, D. Swain, S. Jones, J. Williams, M. Edwards, L. Wo, Validating constructs of learners' academic self-efficacy for measuring learning gain, *High. Educ. Pedagog*. 3 (2018) 118–144. <https://doi.org/10.1080/23752696.2018.1454264>.
- [13] E.W. Wolfe, E. V Smith Jr., Instrument development tools and activities for measure validation using rasch models: Part II - Validation activities, *J. Appl. Meas*. 8 (2007) 204–234.
- [14] J.M. Linacre, Winsteps® Rasch measurement computer programme, (2018). Winsteps.com

(accessed April 19, 2020).

- 1  
2 [15] J. Hobart, S. Cano, Improving the evaluation of therapeutic interventions in multiple sclerosis:  
3 the role of new psychometric methods, *Health Technol. Assess. (Rockv)*. 13 (2009).  
4 <https://doi.org/10.3310/hta13120>.  
5
- 6 [16] NHS, What is the body mass index (BMI)?, (2016). [www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/](http://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/) (accessed April 16, 2019).  
7  
8
- 9 [17] M.P. Jensen, P. Karoly, Self-Report Scales and Procedures for Assessing Pain in Adults, in: D.C.  
10 Turk, R. Melzack (Eds.), *Handb. Pain Assess.* (3rd Ed., The Guildford Press, New York, 2011:  
11 pp. 19–44.  
12
- 13 [18] A. Woo, B. Lechner, T. Fu, C.S. Wong, N. Chiu, H. Lam, N. Pulenzas, H. Soliman, C. DeAngelis,  
14 E. Chow, Cut points for mild, moderate, and severe pain among cancer and non-cancer  
15 patients: a literature review, *Ann Palliat Med.* 4 (2015) 176–183.  
16 <https://doi.org/10.3978/j.issn.2224-5820.2015.09.04>.  
17  
18
- 19 [19] D.G. Altman, *Practical Statistics for Medical Research*, Chapman and Hall, New York, 1990.  
20
- 21 [20] J.E. De Groot, M.J.M. Broeders, C.A. Grimbergen, G.J. Den Heeten, Pain-preventing strategies  
22 in mammography: An observational study of simultaneously recorded pain and breast  
23 mechanics throughout the entire breast compression cycle, *BMC Womens. Health.* 15 (2015).  
24 <https://doi.org/10.1186/s12905-015-0185-2>.  
25
- 26 [21] P. Whelehan, A. Evans, G. Ozakinci, Client and practitioner perspectives on the screening  
27 mammography experience, *Eur. J. Cancer Care (Engl)*. 26 (2017) e12580.  
28 <https://doi.org/10.1111/ecc.12580>.  
29
- 30 [22] H. Dumky, K. Leifland, K. Fridell, The Art of Mammography With Respect to Positioning and  
31 Compression—A Swedish Perspective., *J. Radiol. Nurs.* 37 (2018) 41–48.  
32 [http://libproxy.dundee.ac.uk/login?url=http://search.ebscohost.com/login.aspx?direct=true](http://libproxy.dundee.ac.uk/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=jlh&AN=128127024&site=ehost-live&scope=site)  
33 [&db=jlh&AN=128127024&site=ehost-live&scope=site](http://libproxy.dundee.ac.uk/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=jlh&AN=128127024&site=ehost-live&scope=site).  
34  
35
- 36 [23] C.E. Mercer, S. Cassidy, P. Hogg, Letter to the Editor, *Radiography.* 22 (2016) 101–102.  
37 <https://doi.org/10.1016/j.radi.2015.12.005>.  
38
- 39 [24] P.G. Conaghan, A. Tennant, C.G. Peterfy, T. Woodworth, R. Stevens, A. Guermazi, H. Genant,  
40 D.T. Felson, D. Hunter, Examining a whole-organ magnetic resonance imaging scoring system  
41 for osteoarthritis of the knee using Rasch analysis, *Osteoarthr. Cartil.* 14 (2006) 116–121.  
42 <https://doi.org/10.1016/j.joca.2006.03.011>.  
43
- 44 [25] J. Kean, E.F. Bisson, D.S. Brodke, J. Biber, P.H. Gross, An Introduction to Item Response Theory  
45 and Rasch Analysis: Application Using the Eating Assessment Tool (EAT-10), *Brain Impair.* 19  
46 (2018) 91–102. <https://doi.org/10.1017/BrImp.2017.31>.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



## Acknowledgements

We gratefully acknowledge the contributions to this work of the study participants and the administrative and radiographic staff in the screening clinics. We are also grateful to Dr Violet Warwick for database design, to Dr Kulsam Ali and Dr Rachel Nutt for assistance with data management, and to Dr Sarah Vinnicombe for comments on the draft manuscript.

## Funding sources

This work was supported by the School of Medicine Research Investment Fund at the University of St Andrews, and by charitable donations from Mrs Fiona Edwards (no grant numbers).

VolparaDataManager® software was supplied free-of-charge in exchange for expert advice over several years. The company was not involved in conducting the work and none of the authors has received any payment from the company.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65