European Journal of Radiology 86 (2017) 289–295

Contents lists available at ScienceDirect

European Journal of Radiology

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

Pressure-standardised mammography does not affect visibility, contrast and sharpness of stable lesions

J.E. de Groot [a,*], I.G.M. Hopman [b], M.G.J.T.B. van Lier [a], W. Branderhorst [a], C.A. Grimbergen [c], G.J. den Heeten [c]

[a] *Sigmascreening, Meibergdreef 45, 1105 BA, Amsterdam, The Netherlands*
[b] *University of Amsterdam, The Netherlands*
[c] *Academic Medical Center, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form
14 September 2016
Accepted 27 November 2016

Keywords:
Mammography
Compression
Standardisation
Lesion
Visibility

ABSTRACT

Introduction: A recent technological development allows pressure-standardised mammography by personalizing the compression force to the breast size and firmness. The technique has been shown to reduce pain and compression variability between consecutive exams, but also results in a slightly thicker compressed breast during exposure. This raises the question whether visibility, contrast and sharpness of lesions are affected?

Methods: Four experienced radiologists compared 188 stable lesions and structures including (clusters of) calcifications, (oil) cysts and lymph nodes that were visible in mammograms obtained in 2009 with a pain-tolerance limited 18 daN target force compression protocol, and in 2014/2015 obtained with a 10 kPa (75 mmHg) pressure-standardised compression protocol. Observers were blinded for all DICOM metadata and rated which of the randomly ordered, side by side presented images had better lesion visibility, contrast and sharpness, or whether they saw no difference. They also indicated which overall image they preferred, if any, and whether the non-preferred image was still adequate. Statistical non-inferiority is concluded when the lower limit of the 95% confidence interval of the 4-rater averaged 'new protocol better' proportions exceed the non-inferiority limit of 0.463.

Results: In 2014/2015, the compressions were significantly milder, with on average 17% (mediolateral oblique) to 29% (craniocaudal) lower forces. Breasts remained on average 2.4% (1.4 mm) thicker. Dose was significantly lower (6.5%), which is explained by glandular atrophy. The 95% confidence interval lower limits are 0.479 for visibility, 0.473 for contrast, 0.488 for sharpness and 0.486 for preference, all exceeding the non-inferiority limit. Of the 60 non-preferred mammograms, multiple observers found only five to be inadequate: 4 obtained with the force protocol and 1 with the pressure protocol.

Conclusion: Pain-reduced mammography with 10 kPa pressure-standardised compression has non-inferior visibility, contrast and sharpness for stable lesions compared to pain-tolerance limited 18 daN target force compression.

1. Introduction

Pain reduction in x-ray mammography has long been called for [1] but various attempts either had a negative side-effect [2–5] or no significant pain reduction [6]. A recent technological development [7] allows pain-reduced mammography by personalizing the used compression force to the individual breast size and firmness. To achieve this, the paddle measures both the force (in decanewton, daN) and the breast contact area (in square decimetre, dm^2) and calculates in real-time the contact pressure (in kilopascal, 1 kPa = 1 daN/1 dm^2). For the first time, this enables automatic real-time adjustment of the compression force to the size and firmness of each breast: personalized compression.

Pressure standardisation reduces compression variability between exams [8]. Mercer et al. found that the compression force applied by 14 technicians working in the same mammography screening unit in the United Kingdom is highly dependent on the technician rather than on the client [9,10]. This leads to differences

* Corresponding author.
E-mail addresses: jerry.degroot@sigmascreening.com (J.E. de Groot), ingrid-hopman@hetnet.nl (I.G.M. Hopman), monique.vanlier@sigmascreening.com (M.G.J.T.B. van Lier), w.branderhorst@amc.uva.nl (W. Branderhorst), c.a.grimbergen@amc.nl (C.A. Grimbergen), denheeten@gmail.com (G.J. den Heeten).

http://dx.doi.org/10.1016/j.ejrad.2016.11.030
0720-048X/

in applied pressure, physiological conditions, pain experiences and image quality consistency in women with the same breast size and/or firmness. The personalized compression paddle gives technicians the information needed to aim for the same target pressure each time.

A breast cancer screening performance study [28], and clinical validation studies [7,8] suggest standardising the compression to a pressure of 10 kPa (75 mmHg). This would correspond better to normal breast physiology (between venous and diastolic blood pressure) than stopping the compression at any target/minimum force, or continuing 'until the skin is taut', as suggested by some protocols [11]. One study compared 10 kPa pressure-standardised compressions with contralateral 14 daN force-standardised compressions for 433 women in a Dutch breast cancer screening unit. Results showed a significant average pain score reduction of 10% (mediolateral oblique, MLO) to 17% (craniocaudal, CC) on a 0–10 Numerical Rating Scale (NRS), and a significant reduction of the proportion of women who report severe pain scores (NRS $\geq$ 7): 27% (MLO) to 32% (CC).

However, that study also showed that 10 kPa pressure-standardised compressions leave the breasts on average 4.2% (MLO) to 6.3% (CC) thicker (less compressed). One could hypothesize that this affects the average glandular dose and the visibility, contrast and sharpness of lesions in a negative way. On the other hand, digital mammography devices nowadays automatically optimize the contrast-to-dose ratio and image exposure, and they perform digital image processing. Perhaps a small thickness increase is not visible to the human eye, in which case unnecessary pain from too much compression could be avoided. In this study, we hypothesize that the visibility, contrast and sharpness of stable lesions and structures in pressure-standardised images is not inferior to those in force-standardised images.

2. Methods

The Medical Ethical Committee of the Institution where this study was carried out, waived the need for informed consent for this retrospective observer study. Previous research has been conducted within the same general population (symptomatic mammography patients) [12–14]. Four radiologists compared 188 stable lesions and structures that were present in two rounds of regular clinical mammograms consisting of one CC and one MLO image per breast. The mammograms in 2009 were obtained using the gold standard compression protocol that could be phrased as: "apply a force of 18 daN to all breasts, or as much as the client tolerates". The mammograms in 2014/2015 were obtained using the 10 kPa pressure-standardised compression protocol. The same mammography device was used to obtain all RAW images (Senographe Essential, GE Healthcare, Buc, France), and the same image-processing algorithm was used for presentation (Premium View STD, GE Healthcare, Buc, France).

2.1. Selection of case pairs

Within our hospital, 530 women had a mammogram both in 2009 and in 2014/2015. Selection of case pairs was performed in three steps. First, all mammograms were processed by a Computer-Aided Detection (CAD) program (Transpara, version 1.0.0, ScreenPoint Medical, Nijmegen, The Netherlands), which identified possible lesions and annotated these with a level of suspected malignancy on a numerical scale. Second, after blinding and randomizing the dataset, one researcher (author IH) manually selected all lesions that were present in both images and had almost equal (within 5% difference) levels of suspected malignancy as indicated by the Transpara CAD program. Because of the 5-year time gap, all case pairs actually were non-suspect, benign stable lesions or structures. Large calcifications were added because the CAD program does not annotate these. This step yielded 310 case pairs. Third, an experienced radiologist (author GdH) piloted the intended observer-task, i.e. side-by-side comparison of the randomly ordered image pairs without any meta-data annotation. Thus blinded for chronological order, GdH excluded case pairs for which one or both of the images had insufficient breast positioning, cases that were not actual lesions but incomparable structures annotated by the CAD program, and image pairs from which an obvious chronological order could be recognized, e.g. due to surgery, scar formation or lesion involution/growth. This third step resulted in 188 case pairs.

2.2. Study sample

The 188 selected case pairs consisted of 45 well-defined normal structures (e.g. retromammilar/axillar glandular tissue), 40 (clusters of) calcifications, 38 (clusters of) lymph nodes, 22 (oil) cysts, 14 scar tissue structures, 7 fibroadenomas, 7 ill-defined masses, 6 skin abnormalities, 5 reproducible overprojections (e.g. Cooper's ligaments), 3 fibrotic structures and 1 hamartoma. These case pairs were present in a set of 158 mammogram pairs (multiple cases were present in several mammograms). The women's ages range from 29 to 72 (mean 53.1) at the force-protocol mammogram in 2009, and they were on average 5.3 years older at the pressure-protocol mammogram in 2014/2015.

To study the differences between the force- and pressure-protocol for this sample, a comparison (mean $\pm$ standard deviation and paired *t*-tests) was made for the compressed breast thicknesses and compression forces taken from the DICOM meta-data, as well as for the estimated contact areas, pressures, mean glandular doses, volumetric breast densities and breast volumes obtained from the RAW images with dedicated software (Volpara Analytics, version 1.5.2, Volpara Solutions, Wellington, New Zealand).

2.3. Observer scoring

Four radiologist observers with 12–30 (mean 21) years of mammography experience independently compared the visibility, contrast and sharpness of 188 stable lesions and structures. The mammogram pairs were shown on two dedicated 5 megapixel screens (BARCO, Kortrijk, Belgium) side-by-side in random left-right placement and without any annotation: all DICOM meta-data was hidden from view. To indicate to the observers which lesions and structures to score, the selected case pairs were annotated on a separate laptop. The observers were asked to indicate for each case pair whether they considered the visibility, contrast and sharpness better on the left or the right BARCO screen, or whether they saw no difference. The observers also indicated whether they preferred the overall quality of either image and, if this was the case, whether they still considered the non-preferred image adequate for diagnosis. To reduce order bias, the first two observers worked through the list from beginning to end and the other two started in the middle and wrapped around. Before starting, all four observers read and signed the same study information which described the aforementioned case selection procedure, stated that the image pairs differed in the amount of applied force/pressure during compression, and explained the four scoring parameters as shown in Table 1. The gap of five years between the mammograms was on purpose not mentioned to prevent learning bias based on possible visible clues of chronological order. One researcher (author IH) assisted all observers. This study did not include a consensus reading.