BJR

Received: 16 November 2016 | Revised: 2 May 2017 | Accepted: 9 May 2017

https://doi.org/10.1259/bjr.20160871

Cite this article as:
Abdullah AK, Kelly J, Thompson JD, Mercer CE, Aspin R, Hogg P. The impact of simulated motion blur on lesion detection performance in full-field digital mammography. *Br J Radiol* 2017; **90**: 20160871.

FULL PAPER

The impact of simulated motion blur on lesion detection performance in full-field digital mammography

[1,2]AHMED K ABDULLAH, MSc, BSc, [4]JUDITH KELLY, MSc, DCR, [2]JOHN D THOMPSON, MSc, PhD, [2]CLAIRE E MERCER, MSc, PhD, [3]ROB ASPIN, MSc, PhD and [2,5]PETER HOGG, MPhil, FCR

[1]University of Diyala, Baqubah, Diyala, Iraq
[2]Directorate of Radiography, University of Salford, Greater Manchester, UK
[3]School of Computing, Science and Engineering, University of Salford, Greater Manchester, UK
[4]Chester Breast Imaging Unit, Countess of Chester Hospital NHS Foundation Trust, Chester, UK
[5]Karolinska Institute, Stockholm, Sweden

Address correspondence to: Mr Ahmed K Abdullah
E-mail: *a.k.abdullah@edu.salford.ac.uk*

Objective: Motion blur is a known phenomenon in full-field digital mammography, but the impact on lesion detection is unknown. This is the first study to investigate detection performance with varying magnitudes of simulated motion blur.

Methods: 7 observers (15 ± 5 years' reporting experience) evaluated 248 cases (62 containing malignant masses, 62 containing malignant microcalcifications and 124 normal cases) for 3 conditions: no blurring (0 mm) and 2 magnitudes of simulated blurring (0.7 and 1.5 mm). Abnormal cases were biopsy proven. Mathematical simulation was used to provide a pixel shift in order to simulate motion blur. A free-response observer study was conducted to compare lesion detection performance for the three conditions. The equally weighted jackknife alternative free-response receiver operating characteristic was used as the figure of merit. Test alpha was set at 0.05 to control probability of Type I error.

Results: The equally weighted jackknife alternative free-response receiver operating characteristic analysis found a statistically significant difference in lesion detection performance for both masses [$F(2,22) = 6.01$, $p = 0.0084$] and microcalcifications [$F(2,49) = 23.14$, $p < 0.0001$]. The figures of merit reduced as the magnitude of simulated blurring increased. Statistical differences were found between some of the pairs investigated for the detection of masses (0.0 *vs* 0.7 and 0.0 *vs* 1.5 mm) and all pairs for microcalcifications (0.0 *vs* 0.7, 0.0 *vs* 1.5 and 0.7 *vs* 1.5 mm). No difference was detected between 0.7 and 1.5 mm for masses.

Conclusion: The mathematical simulation of motion blur caused a statistically significant reduction in lesion detection performance. These false-negative decisions could have implications for clinical practice.

Advances in knowledge: This research demonstrates for the first time that motion blur has a negative and statistically significant impact on lesion detection performance in digital mammography.

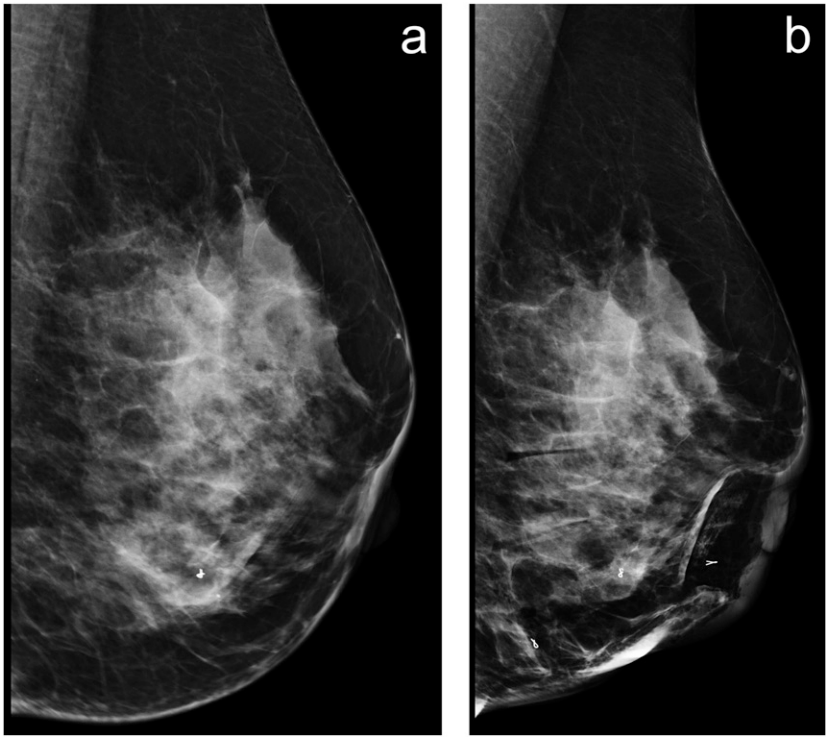
INTRODUCTION

Full-field digital mammography (FFDM) is the current standard imaging technique for the early detection of breast cancer,[1–3] and high-quality, artefact-free, diagnostic images are crucial to the accuracy of this process. Unwanted motion during the image acquisition phase and subsequent image blurring are unfortunate consequences in some FFDM images.[4] It is thought that this could lead to a reduction in diagnostic performance. Figure 1 illustrates a typical example of motion blur in mammography.

The causes of image blur can be patient based (*e.g.* breast and/or chest wall motion) or technology based (*e.g.* paddle movement).[5,6] This can lead to distortion of the image in one or more directions.[7] Chest wall motion could be due to respiration,[8] but we hypothesize that breast motion could be more complex and could be the outcome of a combination of paddle movement, thixotropic behaviour and blood being forced away from the breast due to applied compression force. Thixotropic behaviour[9] can be defined as a time-dependent reduction of viscosity and modulus induced by deformation when mechanical loading changes breast volume and results in the motion of fixed structures (glandular and adipose tissues).

Compression paddle motion has been reported to occur during the "clamping" phase,[10] and it has been hypothesized that this may cause image blur.[11] Recent research identified paddle motion to be present in a number FFDM machines during the clamping phase, with estimates of motion being as high as 1.7 mm.[12,13] Further reports

Figure 1. (a, b) Blurred image (a), the internal breast anatomical structures show no clearly defined edges or borders but appear unfocussed and a single metallic marker within the breast resembles two (one superimposed on another), as a result of motion occurring during image acquisition. (b) No significant blurring as the breast structures are much sharper and focused, and three metallic markers within the breast have clearly defined borders.



suggest that the visual impact of simulated image blurring can be detected from 0.4 mm of movement.[14]

Anecdotal evidence within the National Health Service Breast Screening Programme (NHSBSP) suggests that image blurring may require images to be repeated, thus increasing patient radiation dose, anxiety and service costs. The paucity of literature on this topic suggests that this technical issue continues to be under-reported. Some studies[15,16] have calculated the repeat and technical recall rates with direct reference to image blurring. Several studies report image blurring to be a dominating factor in overall recall rates,[18,19] causing up to 90% of all recalls.[19] Results from another screening service found that 0.86% of all screening candidates were recalled due to image blurring;[16] a high proportion of the 3% maximum NHSBSP permissible rate for repeated images.[17]

Recent research[12] suggests blurring is visible at submillimetre levels, but presently, we do not know the impact of blurring on breast cancer detection. Consequently, our current study seeks to understand whether blurring has an impact on cancer detection performance in FFDM. Our approach uses novel software to perform a pixel shift simulation of motion to introduce blurring to clinical FFDM images.

METHODS AND MATERIALS

Case selection

Ethical approval for this study was granted by The Nightingale Breast Cancer Prevention Centre at University Hospitals of South Manchester NHS Foundation Trust, UK (HSCR15-107) and with consent from The University of Salford, Manchester, UK. This was a retrospective study of breast screening images drawn from the PROCAS database.[20] Initially, 150 cases containing micro-calcifications, 150 cases containing masses and 150 normal cases were made available. These were reviewed visually to identify a range of Breast Imaging-Reporting and Data System density grades and to ensure that the cases did not contain blurring. Cases were chosen from a bank of 300 to ensure a representative distribution of breast density (A = 10%, B = 40%, C = 40% and D = 10%) while also excluding cases where the pathology was too obvious, to control difficulty and also excluding cases that contained artefacts other than blurring. The FFDM system used for image acquisition was the GE Seno Essential, General Electric Medical Systems, Milwaukee, WI. This FFDM unit has a $23 \times 19.2\text{-cm}^2$ field-of-view alpha-silicon flat panels coupled with a thallium doped caesium iodide scintillator image receptor with 100-μm pixel size. This system was operating within the NHSBSP quality assurance guidelines.[21] All images from clients deemed to be mammographically normal had gone through a subsequent breast screening cycle (3 years) to confirm that no cancer was present. Images demonstrating either malignant microcalcifications or masses were biopsy-proven cancers. A mammography image reader with 17 years' mammography-reporting experience re-confirmed the location (s) of masses and microcalcifications in all images. This acted as the truth for the observer study.

In total, 248 cases (124 normal; 62 containing microcalcifications; 62 containing masses) were evaluated by the observers at 0 mm

(no blurring) and two magnitudes of simulated blurring, 0.7 and 1.5 mm. The sample size was guided by tables provided by Obuchowski.[28] Free-response data were analysed separately for microcalcifications and masses. All images were assessed visually by an experienced mammography advanced practitioner to exclude any images which may have contained blur.

Simulated blurring and image display

A mathematical model[22,23] was used to simulate motion in the FFDM images.[13] Simulated motion blur was applied using a convolution mask that provided a three-standard-deviation distribution of blur over the desired blur radius. The three-standard-deviation range is consistent with the application of a Gaussian blur mask, typically used to generate generic blur effects (equivalent to a semi-transparent film being placed over an image). However, the Gaussian distribution profile did not match the characteristics of a typical blur effect. To determine an appropriate blur distribution function, a simulation of image pixel motion, under elastic restitution, was made. This allowed an individual pixel to be displaced by a random vector (within the range of the blur effect) and the pixel contribution to the overall image sampled by substeps, as the pixel returned to the central position. Sampling of the motion pixel was enacted as a pixel-sized Gaussian distribution within a supersampled image frame to allow for fractional motion within each substep. Repeated iterations of this process enabled a representative distribution profile to be generated that showed a sharper central peak, more rapid initial distribution decay and longer continuation than a traditional Gaussian function. Multiple applications of the simulation were made to define an average distribution function. To ensure that the intensity window of the pixel values remained the same after blurring, the pre-blurring minimum and maximum pixel intensities were corrected post blurring through intensity scale and shift.

An initial face validity check with eight mammography practitioners suggested the visual appearance of simulated blur was comparable with real blur. Subsequently, 5 mammography practitioners who had been trained to identify image blur were presented with 20 real and 20 simulated (10 at 0.7 mm and 10 at 1.5 mm) blurred images in a randomized and anonymized fashion. The images were displayed on a 5-MP monitor calibrated to the digital imaging and communications in medicine greyscale standard; ambient lighting was set below 10 lux.[24] For images containing simulated blurring, the average incorrect rate was 34% (standard deviation = 13.8); for real blur, the average incorrect rate was 34% (standard deviation = 20). The incorrect rate refers to the proportion of images incorrectly identified as either real blur or simulated blur. On this basis, we propose the visual appearance of simulated blur to be comparable to that of a real blur.

In accordance with observations made by Ma et al,[12,13] two levels of simulated blurring were used in our study and the images were evaluated under three conditions—without blurring (0 mm) and with two magnitudes of simulated blurring (0.7 and 1.5 mm). Ma et al concluded that the extent of paddle motion, through the acquisition of mammographic images, could be as much as 1.5 mm in the vertical plane. Ma et al illustrated that image blurring at 0.7 mm is the minimum amount of simulated breast movement required for visual detection of soft-edge mask estimation of blurring, as used in this study.

For the free-response receiver operating characteristic (FROC) study, images were displayed on a 5-MP reporting-grade monitor calibrated to the digital imaging and communications in medicine greyscale display function standard.[24] Ambient room lighting was set to below 10 lux. ROCView,[25] which provides zoom up to 100%, was used to provide a randomized order of cases for each observer in each evaluation and to record the observer data from the free-response study.

Observer performance study

Seven observers (15 ± 5 years' clinical reporting experience in mammography) evaluated image sets containing malignant masses, microcalcifications and normal cases for the three conditions. All observers participate in the NHSBSP-approved biannual external audit which evaluates their performance for difficult cases specifically selected by expert radiologists.[26,27] It was agreed that local directors of screening would be notified of any outliers regarding poor performance; however, no outliers were identified and this was not required.

All observers were provided with relevant training prior to beginning the free-response study. Observers were shown 15 images, not used in the main study, comprising of 5 normal images, 5 images containing masses and 5 images containing microcalcifications. This introduced the observers to the task and familiarized them with creating mark-rating pairs[28] (localization and confidence score) using mouse clicks and a slider-bar confidence scale. Observers were instructed to move the slider bar (Scales 1–10) further to the right for an increasing suspicion of malignancy. Observer ratings were then displayed alongside the case. All observers were advised of the importance of localizing the centre of each lesion, as all localizations are compared with a reference map (truth) and determined as lesion localization or non-lesion localization by an acceptance radius emanating from the centre of each lesion/cluster. A minimum period of 2 weeks was imposed between image evaluations to reduce the influence of case memory. Each observer completed the evaluations (0, 0.7 and 1.5 mm) in a different order to reduce the dependence of evaluation order on the overall figure of merit (FOM).

Statistical analysis

Free-response data were analysed primarily using the equally weighted jackknife Alternative Free-Response Receiver Operating Characteristic (wJAFROC) FOM. This represents the empirical probability that a lesion localization is rated higher than a non-lesion localization on normal cases.[29,30] Data analysis was performed using RJafroc,[31] where we also used alternative FOMs to provide us with values of sensitivity (FOM = HrSe) and specificity (FOM = HrSp). Test alpha was set at 0.05 to control probability of Type I error.

Separate analyses were performed for the detection of microcalcifications and masses. For each analysis, an acceptance radius was used based on the maximum size of the mass or spread of an individual cluster of microcalcifications. The acceptance radius

was set at 42 pixels (11 mm) for masses and 50 pixels (13 mm) for microcalcifications.

The wJAFROC FOM was calculated to reward correct localizations and to penalize errors. It provides a single value summarizing performance which can be compared statistically. For instance, comparing two magnitudes of simulated blurring, one calculates a FOM for each method, and a statistical test is performed to identify the difference between two FOMs; if the difference is large enough to be different in consideration of the pre-test value of alpha (0.05), then there is a statistical difference if the result of the overall *F*-test is also significant.[32] We report the result of the overall *F*-test, *p*-values for FOM pairs and the observer-averaged FOM and 95% confidence interval (CI) for each magnitude of simulated blurring.

RESULTS

Free-response data were collected for the detection of malignant microcalcifications and masses for three conditions: (i) no simulated blurring (0 mm) and for two magnitudes of simulated blurring (ii) 0.7 mm and (iii) 1.5 mm. A statistically significant difference was found for the detection of masses [$F(2,21) = 6.01$, $p = 0.0084$] and for the detection of microcalcifications [$F(2,49) = 23.14$, $p < 0.0001$]. For both analyses, a significant difference was observed between 0 and 0.7 mm and between 0 and 1.5 mm of simulated blurring, and also between 0.7 and 1.5 mm for microcalcifications. No significant difference was detected between 0.7 and 1.5 mm for masses. RJafroc was also used to calculate observer-averaged sensitivity (FOM = HrSe) and specificity (FOM = HrSp) as the FOM for all conditions for microcalcifications and masses (Tables 1 and 2).

Two cases (Figure 2a,b) illustrate the impact of simulated blurring on the visual task. Figure 2a demonstrates a spiculated mass of irregular shape, low density and indefinite borders. The percentage of observers detecting this abnormality reduced from 100% (7/7) to 71% (5/7) for 0.7 mm of simulated blurring and to 71% (5/7) for 1.5 mm of simulated blurring. Figure 2b illustrates a case with a single cluster of granular microcalcifications of varying shape and density in the outer half of the right breast representing ductal carcinoma *in situ*. Again, this case saw a reduction in the number of observers detecting this cluster, from 100% (7/7) to 43% (3/7) for 0.7 mm of simulated blurring and to 29% (2/7) for 1.5 mm of simulated blurring.

Microcalcifications

For microcalcifications, the observer averaged wJAFROC FOM and 95% CIs are displayed in Table 1. Differences between FOM pairs (magnitudes of simulated blurring) are displayed in Figure 3a with the *p*-values to indicate significance. For a difference in FOMs to be declared significant, the 95% CI of the FOM pair must not include zero, in addition to the result of the overall *F*-test being significant. The observer-averaged wAFROC curves for microcalcifications are displayed in Figure 4a.

When sensitivity (HrSe) was used as the FOM, a significant difference was found between all pairs of magnitudes of simulated blurring [$F(2,18) = 10.48$, $p = 0.0010$]. This implies that the false-negative rate was increasing significantly as the magnitude of simulated motion blur was increased. When specificity (HrSp) was used as the FOM, there was no significant difference between magnitudes of simulated blurring [$F(2,13) = 0.21$, $p = 0.8110$]. This reveals that the false-positive rate did not increase significantly with image blurring.

Masses

For masses, the observer-averaged wJAFROC FOM and 95% CIs are displayed in Table 2. The differences between the FOM pairs (magnitudes of simulated blurring) are displayed in Figure 3b with the *p*-values to indicate significance. The observer-averaged wAFROC curves for masses are displayed in Figure 4b.

When sensitivity (HrSe) was used as the FOM, there was no significant difference between magnitudes of simulated blurring [$F(2,16) = 0.43$, $p = 0.6575$]. This implies that the false-negative rate was not changing significantly as a result of simulated motion blur. When specificity (HrSp) was used as the FOM, again there was no significant difference between magnitudes of simulated blurring [$F(2,12) = 1.31$, $p = 0.3043$].

DISCUSSION

This study has investigated the impact of computer-simulated motion by means of shifting accumulated pixel points to blur the resultant image. We have found simulated motion blur to have a significant effect on observer performance, with performance becoming statistically worse for the detection of microcalcifications, as simulated blurring was increased from 0 to 0.7 mm, and then on to 1.5 mm. For masses, a statistical difference in detection performance was also observed when blurring was applied to the images at 0.7 mm. However, in this instance, observer performance did not become incrementally worse when the higher magnitude of blurring (1.5 mm) was applied. To be clear, there was no significant difference in detection performance between images blurred with a magnitude of 0.7 mm and those with 1.5 mm, for cases containing masses. This was not the case for microcalcifications, where detection

Table 1. The equally weighted jackknife alternative free-response receiver operating characteristic figure of merit (wJAFROC FOM) and 95% confidence interval (CI), sensitivity and specificity for each magnitude of simulated blurring for the detection of microcalcifications

Magnitude of simulated blurring (mm)	wJAFROC FOM (95% CI)	Sensitivity (%)	Specificity (%)
0	0.899 (0.859, 0.939)	97.9	84.8
0.7	0.813 (0.757, 0.870)	86.4	84.3
1.5	0.746 (0.679, 0.812)	76.5	86.6

Table 2. The equally weighted jackknife alternative free-response receiver operating characteristic figure of merit (wJAFROC FOM) and 95% confidence interval (CI), sensitivity and specificity for each magnitude of simulated blurring for the detection of masses

Magnitude of simulated blurring (mm)	wJAFROC FOM (95% CI)	Sensitivity (%)	Specificity (%)
0	0.905 (0.859, 0.952)	92.3	82.7
0.7	0.869 (0.814, 0.924)	91.9	73.3
1.5	0.862 (0.810, 0.915)	90.5	77.6

performance became statistically worse as the magnitude of blurring was increased. The previous work[13] has suggested that motion blur is visible at 0.7 mm for a soft-edged blur, and our work seems to confirm this. This could have implications for practice as it could mean that when blur is observed in an image, repeat imaging should be considered because, in clinical work, one would simply not know how much blurring is present and what impact it has.

An example is provided in Figure 2b. Here, 7/7 observers detected the lesion in Figure 2b when there was no simulated blurring (0 mm); this decreased to 3/7 observers at 0.7 mm and only 2/7 detected the lesion at 1.5 mm. This is a typical example of the reduction in detection performance, and this trend was observed over a large number of cases containing microcalcifications.

Conversely, we found that mass lesions that have higher contrast with their background, and/or have defined borders (oval or round), do not cause difficulties for detection in the presence of simulated blurring. This means that motion blur has less impact on higher contrast and well-defined masses.

For microcalcifications, we can be less predictive of the impact of simulated blurring on different presentations, other than to say that the impact is greater (higher level of significance) than for masses. The variation in presentation of microcalcifications may be a factor in detection performance and the influence of motion blur, but we have been unable to establish any trend.

There are many factors related to the appearance of breast lesions within FFDM images that can affect lesion detection performance: location within the breast; lesion size, shape and contrast; and the texture and complexity of the surrounding tissue. Lesions located within fibroglandular regions of high-density breast or those complicated by overlapping anatomical structures are more challenging to detect. Small calcification clusters with indefinite edges are considered the most difficult lesions to identify due to size and poor contrast. Lesion shape can be used as a predictor of malignancy,[33] therefore it is important that this can be adequately characterized.

Figure 2. (a, b) Zoomed areas of full-field digital mammography images at 0, 0.7 and 1.5 mm of simulated blurring. (a) A spiculated mass of irregular shape with indefinite borders. (b) A single cluster of granular microcalcifications with different shapes, densities and sizes. Although the mass becomes increasingly difficult to visualize, the microcalcifications are no longer visible with 1.5 mm of simulated blur.

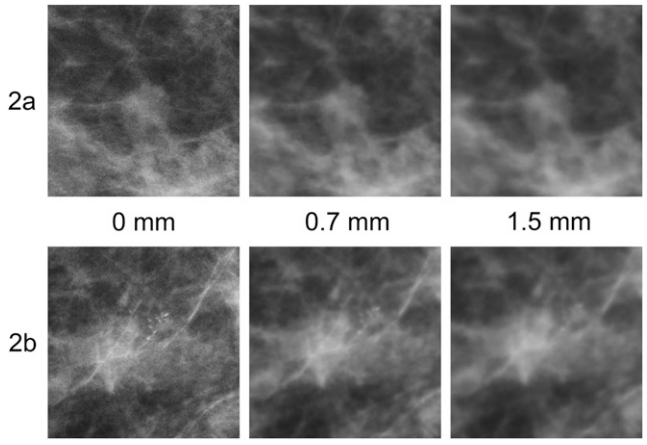


Figure 3. (a, b) The magnitude difference for all pairs of simulated blurring for microcalcifications (a) and for masses (b). For a difference between pairs of the figures of merit (FOMs) to be declared significant, the result of the overall *F*-test must be significant, and the 95% confidence interval (CI) of the pair must not include zero. Statistical differences are evident between all pairs except between 0.7 and 1.5 mm for masses. wJAFROC, equally weighted jackknife alternative free-response receiver operating characteristic.

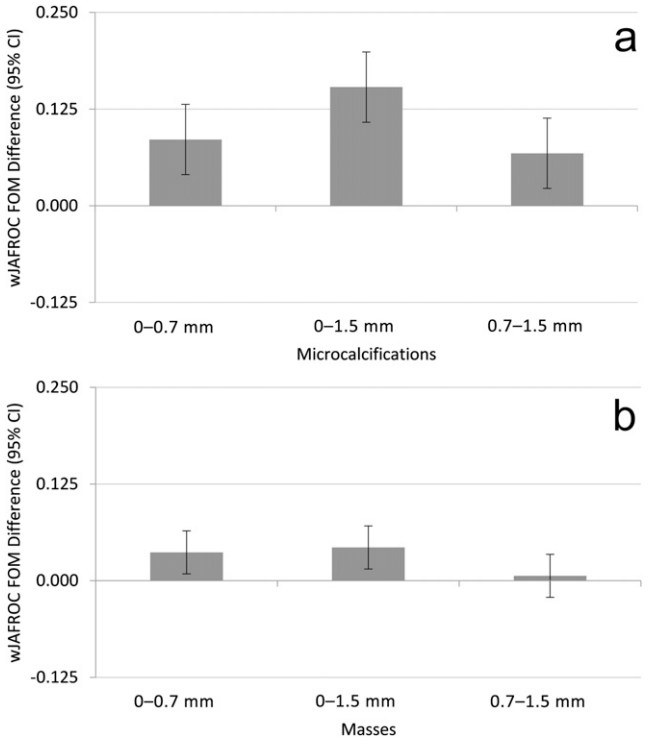
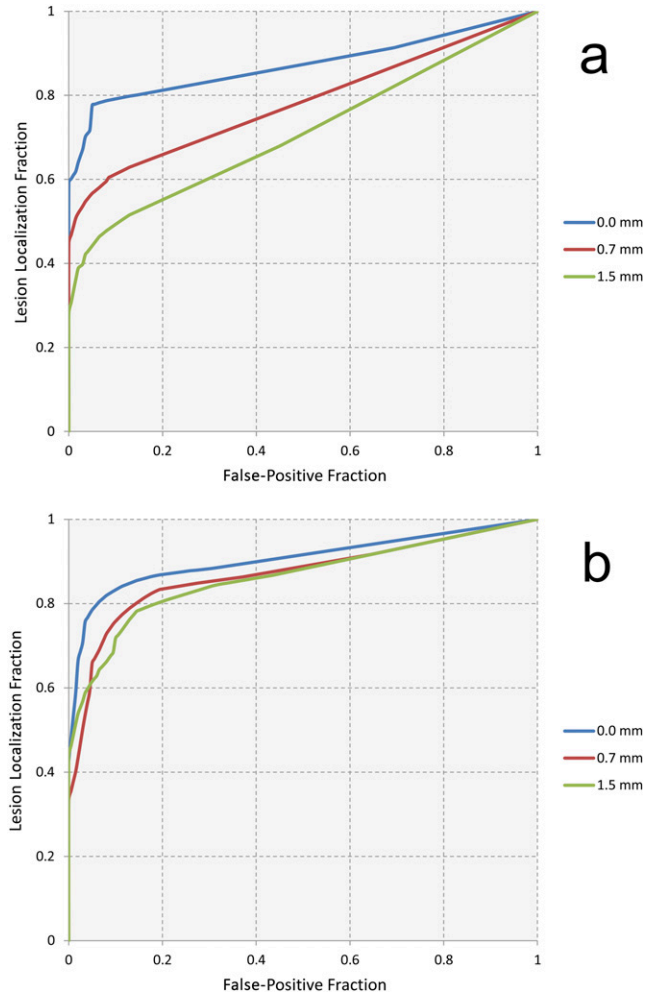


Figure 4. (a, b) The wAFROC curves for all magnitudes of simulated image blurring for microcalcifications (a) and masses (b).



We also analysed the observer data using sensitivity and specificity as FOMs to obtain a better understanding of the impact of simulated motion blur. For both microcalcifications and masses, there was a reduction in sensitivity as the magnitude of simulated image blurring was increased. For masses, this was not statistically significant, and the values in Table 2 demonstrate that the false-negative rate changed little as blurring increased. For microcalcifications, this was not the case, and there was a statistically significant reduction in sensitivity (Table 1), suggesting that the increase in motion blur caused the smaller lesions to become visually imperceptible. Figure 2b provides a typical example of this. The change in specificity was not significantly different for masses or microcalcifications.

There are some limitations to our work. In clinical mammography, the operator does not know what magnitude of motion blur they are inspecting, therefore it could be suggested that it was superfluous to investigate two different magnitudes of motion blur. However, we know from previous work[12] that image blurring is visible at about the level of 0.7 mm for the soft-edge mask used in this simulation, therefore it is of interest to understand if this caused a reduction in observer performance; if it did not, we needed to understand whether a higher magnitude of motion blur did cause an effect. Of course, the image blurring in this study is a simulation, and it has a global effect on the image. In clinical mammography, the motion blur may be global or regional, and for regional blurring, we are not able to predict the impact of this on lesion detection performance from our current work. Additionally, image noise may be blurred by our mathematical simulation, whereas real movement blur would not affect quantum mottle. To overcome a potential smoothing effect on quantum mottle brought about by mathematical simulation, it may be possible to adapt our method by adding noise back into the newly created blurred image. Despite this, our method gives us a certain level of control on motion blur that could not be achieved with blurred images from a clinical setting. With respect to the power of the study, it should also be noted that the prevalence of disease in our study is much higher than would be expected in a screening population, but this is difficult to overcome in observer studies.

A further limitation of the blurring process is also worth raising. The blurring process is enacted as a convolution mask that, in effect, spreads each pixel, redistributing its intensity into the neighbouring pixels based on a function and mask size determined by the modelling of the pixel motion as a random vector path parameterized by the characteristics of breast tissue (generalized) elastic coefficient, required duration and required displacement. The latter two factors act as input to the simulation to determine the magnitude of the blur effect. This creates a controllable blur mask for convolution that has a distribution curve reflective of the intensity spread within a collimated light (energy) propagation system reflective of the X-ray system used. Without modelling actual motion within the breast, it is not possible to determine direction of motion at a specific locality within the breast, therefore this is an approximation to the blur effect that is uniform for the entire image region. Localization is possible but requires each source image to have a specific region of blur defined, and in this case, motion is assumed to be radial and the mask application adjusted accordingly on a per pixel basis from the centre of the defined region, with maximum motion at the centre, reducing to zero motion at the perimeter of the region. Given the large number of source images processed for this study and the requirement for a consistent blur effect on all the generated image sets, regional blurring was not used in this study. This is a limitation in that the blur effect is indicative of the blur that would be present within a "real" patient image in terms of magnitude and effect but does not replicate the directional nature of the blur that would occur for a real image.

CONCLUSION

Simulated motion blur has a statistically significant and negative impact on lesion detection performance for the detection of malignant microcalcifications and masses in FFDM imaging. In view of this, caution should be exercised when making decisions about the acceptability of images that appear to contain blur as false-negative decisions could be reached.

REFERENCES

1. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM. Full-field digital *versus* screen-film mammography: comparison within the UK breast screening program and systematic review of published data. *Radiology* 2009; **251**: 347–58. doi: https://doi.org/10.1148/radiol.2512081235
2. Hambly NM, McNicholas MM, Phelan N, Hargaden GC, O'Doherty A, Flanagan FL. Comparison of digital mammography and screen-film mammography in breast cancer screening: a review in the Irish Breast Screening Program. *AJR Am J Roentgenol* 2009; **193**: 1010–18. doi: https://doi.org/10.2214/AJR.08.2157
3. Knox M, O'Brien A, Szabó E, Smith CS, Fenlon HM, McNicholas MM, et al. Impact of full field digital mammography on the classification and mammographic characteristics of interval breast cancers. *Eur J Radiol* 2015; **84**: 1056–61. doi: https://doi.org/10.1016/j.ejrad.2015.03.007
4. Rosen EL, Baker JA, Soo MS. Malignant lesions initially subjected to short-term mammographic follow-up. *Radiology* 2002; **223**: 221–8. doi: https://doi.org/10.1148/radiol.2231011355
5. Choi JJ, Kim SH, Kang BJ. Mammographic artifacts on full-field digital mammography. *J Digit Imaging* 2014; **27**: 231–6. doi: https://doi.org/10.1007/s10278-013-9641-4
6. Massanes F, Brankov JG. Motion perception in medical imaging. *SPIE Med Imaging* 2011; **7966**: 1–10. doi: https://doi.org/10.1117/12.878417
7. Cao G, Zhao Y, Ni R. Edge-based blur metric for tamper detection. *J Inform Hiding Multimed Signal Process* 2010; **1**: 20–7.
8. Boas FE, Fleischmann D. CT artifacts: causes and reduction techniques. *Imaging Med* 2012; **4**: 229–40. doi: https://doi.org/10.2217/iim.12.13
9. Geerligs M, Peters GW, Ackermans PA, Oomens CW. Does subcutaneous adipose tissue behave as an (anti-)thixotropic material? *J Mech Behav Biomed Mater* 2010; **43**: 1153–9. doi: https://doi.org/10.1016/j.jmbbm.2012.05.014
10. Ma WK, McEntee MF, Mercer CE, Kelly J, Millington S, Hogg P. Analysis of motion during the breast clamping phase of mammography. *Br J Radiol* 2016; **89**: 20150715. doi: https://doi.org/10.1259/bjr.2015071511
11. Kelly J, Hogg P, Millington S, Sanderud A, Willcock C. *Paddle motion analysis – preliminary research*. In: University of Salford; 2011. Manchester, UK: United Kingdom Radiological Congress (UKRC); 2012. pp. 25–7.
12. Ma WK, Hogg P, Kelly J, Millington S. A method to investigate image blurring due to mammography machine compression paddle movement. *Radiography* 2014: 4–9. doi: https://doi.org/10.1016/j.radi.2014.06.004
13. Ma WK, Aspin R, Kelly J, Millington S, Hogg P. What is the minimum amount of simulated breast movement required for visual detection of blurring? An exploratory investigation. *Br J Radiol* 2015; **88**: 20150126. doi: https://doi.org/10.1259/bjr.20150126
14. Ma W, Hilton B, Borgen R. Comparative analysis of visual blurring detection in mammography images—clinical room monitor *versus* reporting grade monitor: initial results. *Eur Radiol* 2016: 1–14. doi: https://doi.org/10.1594/ecr2016/C-0316
15. Bisset L. Clinical audit of technical recall data for blur following the introduction of the breath hold technique in breast screening LE. *Breast Cancer Res* 2014; **15**: 1–15.
16. O'Rourke J, Mercer C, Starr L. Programme evaluation: technical recall and image blur within a breast screening service. In: Symposium. Mammographicum 2014 Meet. Abstract; 2014. p. 4.
17. NHS Breast Screening Program Radiologists Quality Assurance guidelines for breast cancer screening radiology NHSBSP Publication No. 59, March; 2011.
18. Berry-Smith J, Lonsdale C. Image blur: back to basics. *Breast Cancer Res* 2000; **2**: A48. doi: https://doi.org/10.1186/bcr237
19. Seddon D, Schofield K, Waite C. Investigation into possible causes of blurring in mammograms. *Breast Cancer Res* 2000; **2**: A64. doi: https://doi.org/10.1186/bcr253
20. A study looking at breast cancer risk during screening (PROCAS). Cancer research UK. Available from: http://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-study-looking-breast-cancer-risk-screening-procas#undefined
21. NHSBSP Publication No 63; 2006. Quality assurance guidelines for mammography including radiographic quality control. Available from: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp63.pdf
22. Young SS, Driggers RG, Jacobs EL. *Signal processing and performance analysis for imaging systems*. Norwood, MA: Artech House; 2008.
23. Dougherty G. *Digital image processing for medical applications*. New York, NY: Cambridge University Press; 2009.
24. The Royal College of Radiologists. Picture archiving and communication systems (PACS) and guidelines on diagnostic display devices; 2012.
25. Thompson J, Hogg P, Thompson S. ROCView: prototype software for data collection in jackknife alternative free-response receiver operating characteristic analysis. *Br J Radiol* 2012; **85**: 1320–6. doi: https://doi.org/10.1259/bjr/99497945
26. Wivell G, Denton ER, Eve CB. Can radiographers read screening mammograms? *Clin Radiol* 2003; **58**: 63–7. doi: https://doi.org/10.1053/crad.2002.1087
27. van den Biggelaar FJ, Nelemans PJ, Flobbe K. Performance of radiographers in mammogram interpretation: a systematic review. *Breast* 2008; **17**: 85–90. doi: https://doi.org/10.1016/j.breast.2007.07.035
28. Obuchowski N. Sample size tables for receiver operating characteristic studies. *AJR Am J Roentgenol* 2000; **175**: 603–8.
29. Zanca F, Hillis SL, Claus F. Correlation of free-response and receiver-operating-characteristic area-under-the-curve estimates: results from independently conducted FROC/ROC studies in mammography. *Med Phys* 2012; **39**: 5917–29. doi: https://doi.org/10.1118/1.4747262
30. Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: modeling, analysis, and validation. *Med Phys* 2004; **31**: 2313. doi: https://doi.org/10.1118/1.1769352
31. Chakraborty DP, Zhai X. RJafroc version 4.2; 2015. Available from: http://www.devchakraborty.com/RJafroc.php
32. Chakraborty DP. New developments in observer performance methodology in medical imaging. *Semin Nucl Med* 2011; **41**: 401–18. doi: https://doi.org/10.1053/j.semnuclmed.2011.07.001
33. Evans A, Ellis I, Pinder S, Wilson R. Breast calcification a diagnostic manual. London, UK: Greenwich Medical Media; 2002.