

Received:
23 March 2016

Revised:
16 January 2017

Accepted:
25 January 2017

<https://doi.org/10.1259/bjr.20160271>

Cite this article as:

Ma WK, Borgen R, Kelly J, Millington S, Hilton B, Aspin R, et al. Blurred digital mammography images: an analysis of technical recall and observer detection performance. *Br J Radiol* 2017; **90**: 20160271.

FULL PAPER

Blurred digital mammography images: an analysis of technical recall and observer detection performance

¹WANG KEI MA, MSc, ²RITA BORGEN, DCR, MSc, ³JUDITH KELLY, DCR, MSc, ³SARA MILLINGTON, BSc (Hons), PGD, ²BEVERLEY HILTON, BSc (Hons), ⁴ROB ASPIN, PhD, ^{5,6}CARLA LANÇA, PhD and ^{1,7}PETER HOGG, FCR

¹Department of Radiography, University of Salford, Salford, UK

²East Lancashire Breast Screening Unit, Burnley General Hospital, Burnley, UK

³Department of Radiography, Countess of Chester Hospital NHS Foundation Trust, Chester, UK

⁴Department of Computer Science and Software Engineering, University of Salford, Salford, UK

⁵Department of Sciences and Rehabilitation Technologies, Lisbon School of Health Technology, Lisbon, Portugal

⁶Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Lisbon, Portugal

⁷Department of Radiography, Karolinska Institute, Stockholm, Sweden

Address correspondence to: Mr Wang Kei Ma

E-mail: carnby2000@gmail.com

Objective: Blurred images in full-field digital mammography are a problem in the UK Breast Screening Programme. Technical recalls may be due to blurring not being seen on lower resolution monitors used for review. This study assesses the visual detection of blurring on a 2.3-MP monitor and a 5-MP report grade monitor and proposes an observer standard for the visual detection of blurring on a 5-MP reporting grade monitor.

Methods: 28 observers assessed 120 images for blurring; 20 images had no blurring present, whereas 100 images had blurring imposed through mathematical simulation at 0.2, 0.4, 0.6, 0.8 and 1.0 mm levels of motion. Technical recall rate for both monitors and angular size at each level of motion were calculated. χ^2 tests were used to test

whether significant differences in blurring detection existed between 2.3- and 5-MP monitors.

Results: The technical recall rate for 2.3- and 5-MP monitors are 20.3% and 9.1%, respectively. The angular size for 0.2- to 1-mm motion varied from 55 to 275 arcs. The minimum amount of motion for visual detection of blurring in this study is 0.4 mm. For 0.2-mm simulated motion, there was no significant difference [χ^2 (1, $N=1095$) = 1.61, $p=0.20$] in blurring detection between the 2.3- and 5-MP monitors.

Conclusion: According to this study, monitors ≤ 2.3 MP are not suitable for technical review of full-field digital mammography images for the detection of blur.

Advances in knowledge: This research proposes the first observer standard for the visual detection of blurring.

INTRODUCTION

Image blurring due to motion unsharpness in full-field digital mammography (FFDM) is a widely recognized problem in the UK and various explanations exist about how it occurs.^{1,2} One explanation is the breast/paddle movement whilst the exposure is being made.¹⁻⁴ Other factors such as inadequate compression and patient movement together with long exposures may also cause blurring.⁵

Blurring has the potential to increase false-negative results as it may obscure small or low-density microcalcification cancers and larger lesions particularly in dense breast tissue. Technical repeat due to blurring increases client radiation dose and overall examination time and can raise client anxiety. Technical recall is necessary if blurring is not seen at the attendance time, and it could add further to client and family anxiety,⁶ as unlike a technical repeat taken

at the time of the initial examination, the female will have to wait several days for repeat imaging.

Little has been published about blurred mammography images. In 2000, Seddon et al⁵ reported that >90% of their screening mammogram technical recalls were due to blurred images. More recently, blurred images were found to be a major source of technical recall in Manchester, UK.⁷ In an unpublished audit, in one of our breast screening units, we found that 0.86% (40 out of 4650 FFDM examinations) of clients were recalled due to image blur; this contributed to almost one-third (29%) of the 3% maximum permissible recall rate in the National Health Breast Screening Programme.⁸ For some of these images, the blurring could only be detected when they were displayed on 5-MP reporting grade monitors at the time of reporting. In many instances, blurring was missed when

the images were checked for technical accuracy at the time of imaging. We believe this discrepancy could be due to the lower quality non-diagnostic quality monitors used in clinical rooms coupled with variable and also generally brighter ambient lighting when compared with reporting rooms. Interestingly, a good deal of research emphasis has been placed on the evaluation of reporting grade monitors and the environment in which they sit,^{9–11} but surprisingly, little has been placed on the evaluation of technical review monitors used within mammography imaging rooms or radiography imaging rooms generally. In the context of breast screening, only one study in 2016 by Kinnear and Mercer¹² was found which reported the ability of observers to visually detect image blurring in FFDM images on 5- and 1-MP monitors; the lower resolution monitor resulted in a lower visual detection rate for blurred images. Kinnear and Mercer's study represents an important first step, and our study builds on this in various ways. First, our study has a much larger group of observers thereby enabling interobserver differences to be considered; second, simulation of blurring is used in which the exact amount of blurring is known; third, image selection went through a rigorous and carefully documented evidence-based approach; finally, the images were displayed in a room where the ambient lighting was controlled and standardized.

Aside from monitor resolution, it is possible that observer ability to visually identify blur will also affect technical recall rates (TCs). Currently, no performance data exist on observer ability to detect blur. However, early work by Ma et al³ suggested that 0.4 mm of simulated blur can be visually detected on 5-MP reporting grade monitors. Limitations of Ma et al's study relate to the low number of observers used and the observers being experienced image readers who are not representative of the practitioners who undertake mammography imaging.

Our study has two aims: to investigate whether there is a difference in the visual detection of blurring between a 2.3-MP technical review monitor and a 5-MP reporting grade monitor; to propose an observer standard for the visual detection of blurring on reporting grade 5-MP monitors.

METHODS AND MATERIALS

Mammography images were acquired in 2014 on a Selenia® Dimensions® FFDM unit (Hologic®, Bedford, MA) which has a 24 × 29-cm amorphous silicon thin-film transistor image receptor with 70- μm pixel size and spatial resolution of 7.1 lp mm⁻¹ within the UK Breast Screening Programme.¹³ Two experienced image readers independently reviewed a number of images using published quality criteria¹⁴ to identify 20 normal and artefact-free FFDM images. These were comprised of craniocaudal and mediolateral oblique images. Mathematical simulation software³ with a soft-edge mask was used to simulate the effect of motion in the 20 images. Soft-edge mask simulates motion by applying a mathematical algorithm known as convolution function based on a Gaussian distributed pixel under simulated motion.^{15,16} Motion blurring was added to the images by accumulating the pixel intensity of randomized micro steps within 1.5-mm motion boundary.³ The soft-edge mask method was chosen because it best represents the physical process that caused the blurring effect.

Simulated blurring was imposed to the 20 artefact-free FFDM images from 0.2 to 1.0 mm at 0.2-mm increments. 120 images were available for use—100 with 5 levels of simulated motion and 20 with no blur. Figures 1 and 2 show examples of FFDM images with and without simulated blur imposed.

The 120 images were deidentified, randomized and displayed at full screen size on a 24-inch 2.3-MP monitor (Multisync 243wm; NEC, Itasca, IL) with 0.27-mm pixel pitch and 1920 × 1200 display resolution; and a 21.3-inch 5-MP monitor (Dome E5; NDS, San Jose, CA) with 0.17-mm pixel pitch and 2560 × 2048 display resolution. Both monitors were calibrated to the digital imaging and communications in medicine greyscale standard display function.¹⁷ Dimmed ambient lighting (<10 lux) was used for both monitor viewing sessions, being consistent with that employed in normal image-reading conditions.¹⁴ Images were displayed using MediViewer (Schaefer Systemtechnik, Petersaurach, Germany). No interpolation method was used to map image pixels onto the display pixels. Observers were blinded to the type of monitor used as both monitors have similar dimensions and appearance; and information about the

Figure 1. Full-field digital mammography image with no blur.

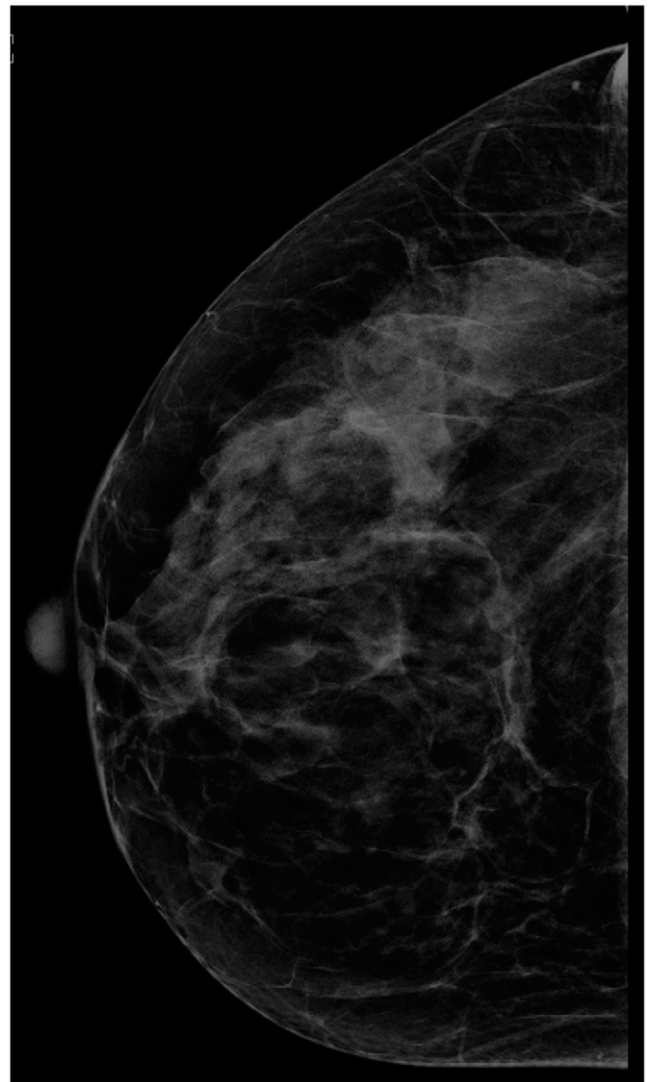
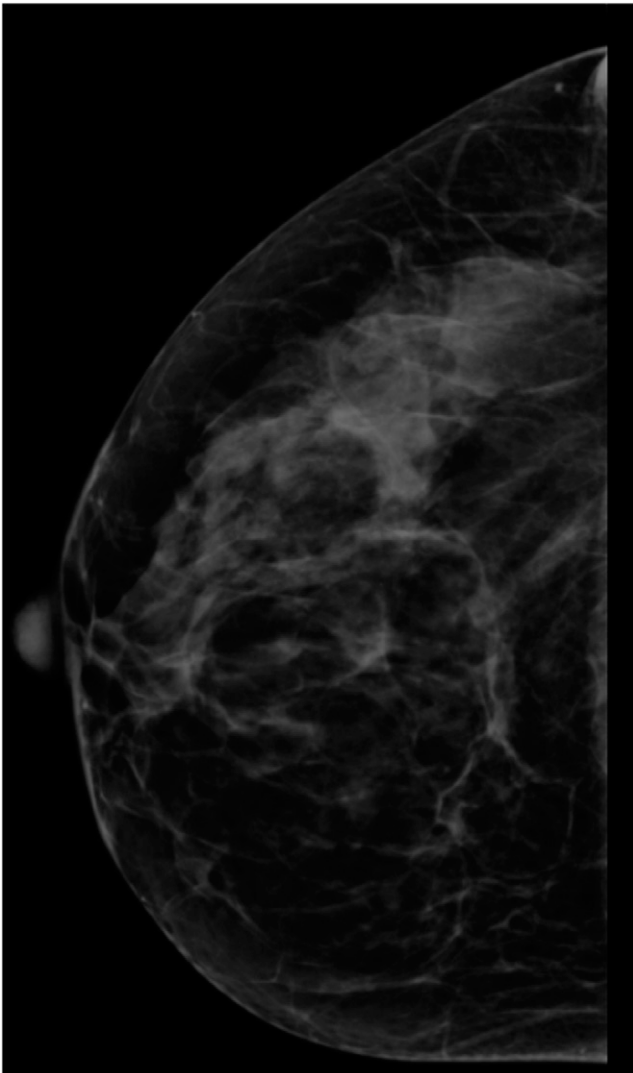


Figure 2. Full-field digital mammography image with 1-mm simulated blur.



monitor was not displayed anywhere. Images were viewed on a blinded basis by 28 observers without knowing the amount of blurring. Window width and level was set to values agreed by consensus between two experienced FFDM image readers prior to the observers commencing the study; width and levels were set to give image appearances similar to those seen in routine practice.

In clinical practice, the distance between the monitor and observer's eye is not standardized or controlled. This is because observers constantly change the distance between their eye and the monitor when viewing images. Our study allows this variation of distance to be preserved by positioning the chair such that the observers' eye to monitor distance would not exceed 75 cm. A viewing distance of 75 cm was chosen because it is within the viewing range (64–89 cm) which maintains the extraocular muscles in a more relaxed state and minimizes eye strain.¹⁸ However, we did not control or measure the distance from eyes to monitors as this was not the focus of our study. Two calculations on angular size were performed, one at 30 cm

and one at 75 cm, as these are likely to be the extremes of distance that observers might view images.

Angular size is a measurement that describes how large an object appears from a given point of view, defining the distance between the two ends of an object. The capacity to identify blurring depends on the potentialities of the human visual system. To identify the minimum amount of blurring that can be detected by the observer, the angular size for each level of motion was calculated with the equation shown below:¹⁹

$$\text{Angular size in degree} = 57.3 \times \text{physical size} / \text{viewing distance}$$

where physical size is the level of motion in millimetres.

26 radiographers qualified in mammography imaging and two radiologists (observers) from 2 breast screening centres in the North West of England (UK) were invited to review the 120 images on the 2.3-MP technical review monitor and the 5-MP reporting grade monitor. None of the observers reported visual pathologies, and image evaluation was conducted with optical correction if glasses had been prescribed previously. The observers were approached individually and asked if they would be willing to participate; of those who agreed, they were provided with written information about the research before conducting it. This study was classified as service evaluation in both breast-screening centres; Clinical Audit Department permission was granted formally on this basis from both hospitals. Anonymity was provided by one co-ordinating staff member within each centre assigning a unique code to each observer; only the observer and coordinating staff member knew the code. Feedback was given only on an individual basis to each observer. Observers' age varied from 26 to 59 years (mean = 44.5, standard deviation = 8.3 years). Mammography experience varied from 0.4 to 25 years (mean and median experiences were 9.9 and 10 years, respectively; standard deviation = 4.9 years; interquartile range = 7.5 years).

The observers were not permitted to magnify the images or adjust the window width and level. Image manipulation was not permitted due to the need to tightly control the viewing conditions to exclude sources of error.^{20–22} If the observers were allowed to manipulate images based on their personal preferences in display, then the study could be comparing the ability of the observers to manipulate images as well as detect blurring on the two monitors.

For each image, the observers had to indicate whether blurring was present or not; this was a binary decision (yes = 1, no = 0). As in Mucci et al's study,²³ Fleiss' kappa analysis was carried out to determine the interobserver variability.²⁴ To minimize fatigue, image review sessions did not exceed 30 min,²⁵ and each monitor took approximately 1 h to complete, therefore four viewing sessions were required (approximately 2 h per observer was needed) to review the images on the 2.3- and 5-MP monitors. Owing to clinical demands, data collection had to be conducted over an 8-month period. Experimental conditions and observer training for the experiments were overseen and controlled/standardized by two members of staff—one in each clinical

centre. Also, all observers underwent a training exercise to help them identify blurred and non-blurred images. This exercise was conducted by an experienced image reader using a 5-MP reporting grade monitor; for this exercise, clinical FFDM images were drawn from each of the two screening programmes to train the observers. These images contained blurred and non-blurred examples.

Blurring detection rate (BD) at each level of motion for 2.3- and 5-MP monitors was calculated. The equation for BD is shown below.

$$BD = N_i / N_b$$

where N_i is the number of blurred mammograms identified by the observers and N_b is the number of blurred mammograms.

χ^2 test was used to determine whether significant differences in BD existed between the 2.3- and 5-MP monitors. The influence of the level of motion, monitor resolution, observers' experience and age on blurring detection was modelled in a logistical regression model.

TC at each level of motion for 2.3- and 5-MP monitors was calculated according to the National Health Breast Screening Programme recommendations.²⁶ In this study, the number of mammograms required to repeat (N_r) was estimated by the number of blurred mammograms missed by the observers (N_m) which is equal to the difference between the number of blurred mammograms (N_b) and the number of blurred mammograms identified by the observers (N_i).

The equation for TC is shown below:

$$\begin{aligned} TC &= N_r / N_t \\ &= N_m / N_t \\ &= (N_b - N_i) / N_t \end{aligned}$$

where N_r is the number of mammograms required to repeat; N_t is the total number of mammograms taken; N_b is the number of blurred mammograms; and N_i is the number of blurred mammograms identified by the observers.

The upper quartile for the BD on the 5-MP monitor was calculated to develop the observer standard for the visual detection of blurring. The upper quartile was used to set the minimum standard for blur detection because it represents the highest 25% of the data. If the BD is at the 75th percentile, it means 75% of the observers would perform the same as or less than this level and 25% would perform better than this level.

RESULTS

The average BD for the 2.3- and 5-MP monitors is shown in Figure 3. All the non-motion images were identified correctly. As can be seen in Figure 3, the BD increases with simulated motion and monitor resolution. The 5-MP monitor has a higher average BD than the 2.3-MP monitor.

χ^2 test revealed that there was no significant difference in blurring detection between the 2.3- and 5-MP monitors for 0.2 mm motion, [χ^2 (1, $N = 1095$) = 1.61, $p = 0.20$]. While there were significant differences in blurring detection between 2.3- and 5-MP monitors for 0.4 mm [χ^2 (1, $N = 1095$) = 17.50, $p < 0.001$], 0.6 mm [χ^2 (1, $N = 1095$) = 44.44, $p < 0.001$], 0.8 mm [χ^2 (1, $N = 1095$) = 75.26, $p < 0.001$] and 1 mm [χ^2 (1, $N = 1095$) = 108.32, $p < 0.001$] motion.

Figure 3. Blurring detection rate against the level of motion; the error bars represent the standard errors.

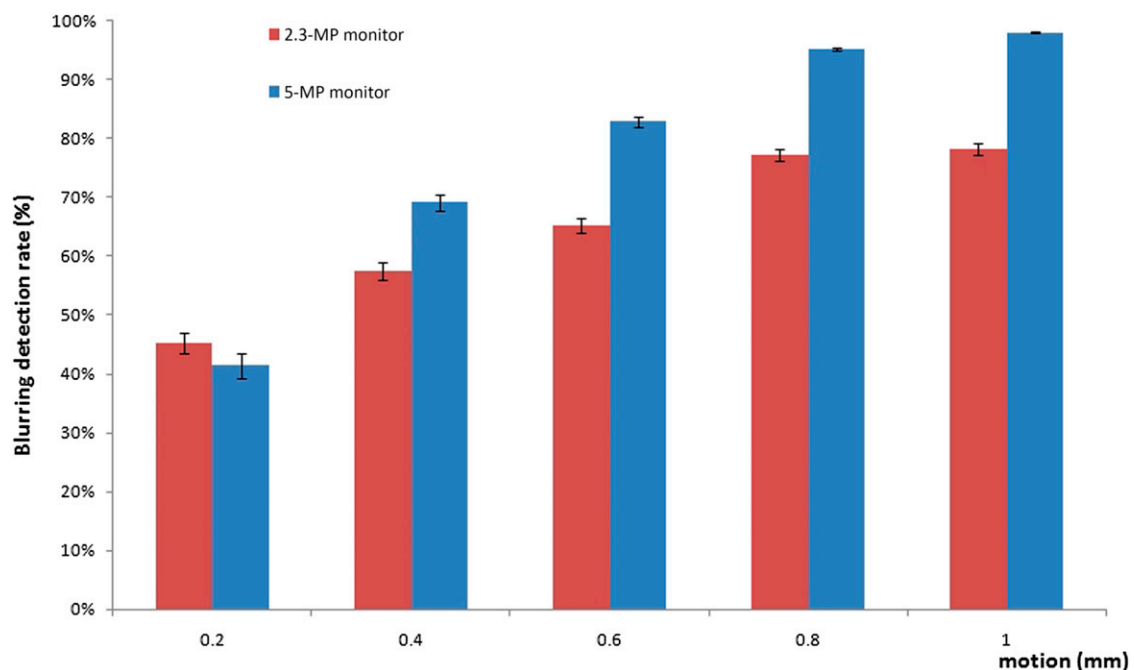


Table 1. Angular size for the different levels of motion

Level of motion (mm)	Angular size (°)		Angular size (arcsec)	
	30 cm	75 cm	30 cm	75 cm
0.2	0.0382	0.01528	138	55
0.4	0.0764	0.03056	275	110
0.6	0.1146	0.04584	413	165
0.8	0.1528	0.06112	550	220
1	0.1910	0.07640	688	275

Fleiss' kappa for 5- and 2.3-MP monitors is 0.48 and 0.11, respectively, and the mean kappa is 0.26. A kappa of 1 indicates perfect agreement, whereas a kappa of 0 indicates agreement equal to chance.²⁴

Cohen's *d* was used to indicate the effect size for factors in the logistical regression model. The Cohen's *d* values for the level of motion, monitor resolution, observers' experience and age are 0.38, 0.35, 0.09 and 0.05, respectively. Cohen's *d* of 0.2 can be considered as "small" effect, around 0.5 as "medium" effect and >0.8 as "large" effect.²⁷ Therefore, the Cohen's *d* value indicated that in this study, observers' experience and age are not good predictors for blurring detection.

The angular size for each level of motion for viewing distances of 30 and 75 cm is summarized in Table 1. As can be seen, the angular size increases with the level of motion, and it is bigger when the observers are closer to the monitor (30 cm). Individuals with 20/20 vision have the ability to recognize a pixel if the angular size is ≥ 60 arcsec.

The angular size for 0.2 mm motion at 75 cm is 55 arcsec which is smaller than the threshold and such a small change cannot be identified by the human eye.^{28,29} With this in mind, we propose the minimum amount of motion required for visual detection of blurring in this study is 0.4 mm.

The TCs for 2.3- and 5-MP monitors were calculated and summarized in Table 2. As can be seen in Table 2, the TC decreased with the level of motion and monitor resolution. The TC for the 2.3-MP monitor varies from 3.6% to 7.1%, and for the 5-MP monitor, it varies from 0.3% to 5.1%. The 2.3-MP monitor has a higher overall TC (20.3%) than the 5-MP monitor (9.1%). For example, at 1 mm motion, the recall rate for 2.3- and 5-MP monitors is 3.6% and 0.3%, respectively, which means for 1000 clients, the number of recall would be 36 and 3, respectively.

The upper quartile for the BDs on the 5-MP monitor are summarized in Table 3. The observer standard for the minimum

standard of blurring detection at 0.4, 0.6, 0.8 and 1.00 mm levels of motion is 96%, 100%, 100% and 100%, respectively.

DISCUSSION

The results from the monitor comparison study confirm that a monitor with lower resolution (e.g. 2.3 MP) would likely have a poorer visual detection rate for FFDM image blurring than a higher resolution reporting grade monitor (5 MP). The number of blurred images missed by the observers (N_m) for the lower resolution monitor is higher than the number in the higher resolution monitor, which leads to a higher TC for the lower resolution monitor.

In clinical practice, as some technical review monitors have resolutions as low as 1 MP,¹² we can confidently propose that such monitors would have even poorer blurred image visual detection rates than the one used in our study (2.3 MP). Further work is needed to determine the minimum specifications of a technical review monitor for use in imaging rooms for which TCs could be suitably low for clinical purposes. It is worth noting that our data suggest that there is a 55% reduction in the TC if a 5-MP reporting grade monitor is used for checking images in the clinical rooms. This would reduce the need for additional time slots for appointments as well as the cost of the administrative overhead for booking the appointments. Also, it would minimize client/client family anxiety and costs for the reattendance.

Resolution acuity refers to the smallest amount of spatial detail necessary to distinguish a difference between patterns or features in a visible target.²⁸ Individuals with 20/20 vision have the ability to recognize a minimal angle of resolution subtended by the components of the stimulus, which has an angular size of 60 arcsec.^{28,29} At 0.2 mm of simulated blurring, there is no significant difference [χ^2 (1, $N = 1095$) = 1.61, $p = 0.20$] in blurring detection between the 2.3- and 5.0-MP monitors. One of the possible explanations is that the human visual system is not able to resolve this level of detail at a distance of 75 cm as the angular size is <60 arcsec.

Table 2. Technical recall rate (TC) for 2.3- and 5-MP monitors

Level of motion (mm)	0.4	0.6	0.8	1	Total
TC for 2.3-MP monitor	7.1%	5.8%	3.8%	3.6%	20.3%
TC for 5-MP monitor	5.1%	2.9%	0.8%	0.3%	9.1%

Table 3. Observer standard for the minimum standard of blurring detection for 5-MP monitor

Level of motion (mm)	0.4	0.6	0.8	1
Upper quartile (75th percentile)	96%	100%	100%	100%

Angular size calculations demonstrate that blur of 0.2 mm motion is not possible to identify if the viewing distance is increased to 75 cm, independently of the monitor used. The impact of the visual system on diagnostic decision-making is not well understood. However, it is known that visual acuity and accommodation accuracy get worse at the end of a long radiology workday.^{30,31} Variance in the viewing distance combined with visual fatigue and a low-resolution monitor can be a potential risk factor for missing the detection of blur on 2.3-MP monitors.

The selection of the motion levels was related to the early work by Ma et al.³ Detection performance between the limits of 30 and 75 cm was not tested for 0.3 mm. According with our current calculations of angular size for 0.3 mm of motion, it could be argued that if 0.3 mm of blurring had been used, the blurring should be identifiable by the observers at 75 cm (82 arcsec). This warrants further research to determine threshold values for detection of blurring at different distances from the monitor.

Fleiss' kappa for 2.3-MP monitors is much lower than that for the 5-MP monitor which suggests that using the lower resolution monitor to see blurring is more difficult compared with the higher resolution monitor.

On the other hand, the mean kappa in our study is 0.26 which indicates poor agreement between observers.²⁴ In observer studies, it is very rare to achieve perfect agreement, and a range of cognitive, visual and environmental factors can be used to explain this. Also, anecdotally we know that some people find the task of differentiating blurred from non-blurred images very difficult, so this could be another explanation for poor agreement. One conclusion from this could be that observers who performed less well could need additional training. This poor level of agreement raises questions about the blur detection abilities between observers which is the second aim of this study. In view of this, the observer standard developed in our study could be used to help inform the development of competence assessment standards of observers in training programmes and in routine practice.

Intraobserver variation and interobserver variation across professional disciplines was not included into this study. As observers only viewed each image once, it is not possible to calculate the intraobserver variation. For interobserver variation

across professional disciplines, the sample size for radiologists is too small ($n = 2$) to conduct meaningful analysis. Further research is therefore warranted for intra- and interobserver variability for different professional groups.

One of the limitations of our study is the use of motion simulation as this may not fully represent real blurring. For instance, the mathematical simulation used in our study blurs the whole image, whereas real mammography image blurring may fully or partly affect the image. An updated version of our mathematical simulation has the ability to introduce regional blurring. Using this updated version, further studies could be carried out to investigate the effect of regional blurring on observer and monitor BDs. Aside from proposing an extension to our study using regional blurring, it could be valuable to conduct a study using real blurred FFDM images. However, it should be noted that for real blurring, it would be hard to control and identify the exact amount of blurring in the images.

Another limitation of our study is that the normal mammography screening environment might not be fully recreated in our study. For example, practitioners working in imaging rooms often do not work in levels of subdued light consistent with common reporting conditions, and they probably do not have the same amount of time as image readers to scrutinize the image. Further studies could be carried out to investigate the effect of lighting and image viewing time on BD for technical review monitors.

Finally, we did not take into account observers' previous activities. For example, visual fatigue may occur if a radiologist or radiographer finished a reporting session and then immediately took part in the study. Further studies could be carried out to investigate the effect of visual fatigue on BDs and also other factors, as indicated earlier, which can impact upon observer performance.

CONCLUSION

According to our study monitors ≤ 2.3 MP are not suitable for technical review of FFDM images for the detection of blur. The minimum amount of motion required for visual detection of blurring in our study is 0.4 mm, and the observer standard for blur detection at 0.4, 0.6, 0.8 and 1 mm levels of simulated blurring are 96%, 100%, 100% and 100% on a 5-MP monitor.

REFERENCES

1. Ma WK, Brett D, Howard D, Kelly J, Millington S, Hogg P. Extra patient movement during mammographic imaging: an experimental study. *Br J Radiol* 2014; **87**: 20140241. doi: <https://doi.org/10.1259/bjr.20140241>
2. Ma WK, McEntee MF, Mercer C, 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.20150715>
3. 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>
4. Ma WK, Hogg P, Kelly J, Millington S. A method to investigate image blurring due to mammography machine compression paddle movement. *Radiography* 2015; **21**: 36–41. doi: <https://doi.org/10.1016/j.radi.2014.06.004>
 5. Seddon D, Schofield KA, Waite CA. Investigation into possible causes of blurring in mammograms. *Breast Cancer Res* 2000; **2** (Suppl. 2): A64. doi: <https://doi.org/10.1186/bcr253>
 6. Hogg P, Kelly J, Claire E, eds. *Digital mammography: a holistic approach*. 1st edn. London, UK: Springer; 2015.
 7. O'Rourke J, Mercer CE, Starr L. Programme evaluation: technical recall and image blur within a breast screening service. *Symposium Mammographicum 2014*. Bournemouth, UK: British Institute of Radiology; 2014. July [cited 2017 January 16]. Available from: <http://www.birpublications.org/doi/pdf/10.1259/conf-symp.2014>
 8. NHS Cancer Screening Programmes. Consolidated guidance on standards for the NHS Breast Screening Programme. NHSBSP Publication No 60 (v. 2). Sheffield, UK; 2005.
 9. Shiraishi J, Abe H, Ichikawa K, Schmidt RA, Doi K. Observer study for evaluating potential utility of a super-high-resolution LCD in the detection of clustered microcalcifications on digital mammograms. *J Digit Imaging* 2010; **23**: 161–9. doi: <https://doi.org/10.1007/s10278-009-9192-x>
 10. Schueller G, Schueller-Weidekamm C, Pinker K, Memarsadeghi M, Weber M, Helbich TH. Comparison of 5-megapixel cathode ray tube monitors and 5-megapixel liquid crystal monitors for soft-copy reading in full-field digital mammography. *Eur J Radiol* 2010; **76**: 68–72. doi: <https://doi.org/10.1016/j.ejrad.2009.04.070>
 11. Kamitani T, Yabuuchi H, Matsuo Y, Setoguchi T, Sakai S, Okafuji T, et al. Diagnostic performance in differentiation of breast lesion on digital mammograms: comparison among hard-copy film, 3-megapixel LCD monitor, and 5-megapixel LCD monitor. *Clin Imaging* 2011; **35**: 341–5. doi: <https://doi.org/10.1016/j.clinimag.2010.08.015>
 12. Kinnear L, Mercer CE. A study to compare the detection of visual blurring in 1 MP and 5 MP monitors within mammography clinical practice. *Imaging and Therapy Practice*; 2016. pp. 23–28. ISSN 1360-5518. Available from: <http://content.yudu.com/Library/A3yIot/ImagingampTherapyPra/resources/23.htm>
 13. Selenia Dimensions Mammography System datasheet; 2016 [updated 2016 January 13; cited 2017 January 16]. Available from: http://www.hologic.com/sites/default/files/DS-05534_rev002_SeleniaDimensions13-Jan2016.pdf
 14. Perry N, Broeders M, Wolf C, Törnberg S, Holland R, Karsa L. *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th edn. Luxembourg: European Communities; 2006.
 15. Young SS, Driggers GR, Jacobs LE. *Signal processing and performance analysis for imaging systems*. 1st edn. New York, NY: Artech House; 2008.
 16. Dougherty G. *Digital image processing for medical applications*. 1st edn. Cambridge, UK: Cambridge University Press; 2009.
 17. National Electrical Manufacturers Association (NEMA). *Digital Imaging and Communications in Medicine (DICOM) Part 14: Grayscale Standard Display Function*; 2011.
 18. Krupinski EA, Flynn MJ. *IT reference guide for the practicing radiologist*. Reston, VA: American College of Radiology; 2013. [cited 2017 January 16]. Available from: <http://www.acr.org/~media/ACR/Documents/PDF/Advocacy/IT%20Reference%20Guide/IT%20Ref%20Guide%20Displays.pdf>
 19. Legge GE, Bigelow CA. Does print size matter for reading? A review of findings from vision science and typography. *J Vis* 2011; **11**: 8. doi: <https://doi.org/10.1167/11.5.8>
 20. Thompson JD, Chakraborty DP, Szczepura K, Tootell AK, Manning JD, Hogg P. Effect of reconstruction methods and X-ray tube current–time product on nodule detection in an anthropomorphic thorax phantom: a crossed-modality JAFROC observer study. *Med Phys* 2016; **43**: 1265–74. doi: [https://doi.org/0094-2405/2016/43\(3\)/1265/10](https://doi.org/0094-2405/2016/43(3)/1265/10)
 21. Thompson JD, Hogg P, Manning DJ, Szczepura K, Chakraborty DP. A free-response evaluation determining value in the computed tomography attenuation correction image for revealing pulmonary incidental findings: a phantom study. *Acad Radiol* 2014; **21**: 538–45. doi: <https://doi.org/10.1016/j.acra.2014.01.003>
 22. Thompson JD, Hogg P, Higham S, Manning D. Accurate localization of incidental findings on the computed tomography attenuation correction image: the influence of tube current variation. *Nucl Med Commun* 2013; **34**: 180–4. doi: <https://doi.org/10.1097/MNM.0b013e32835c0984>
 23. Mucci B, Murray H, Downie A, Osborne K. Interrater variation in scoring radiological discrepancies. *Br J Radiol* 2013; **86**: 20130245. doi: <https://doi.org/10.1259/bjr.20130245>
 24. Gwet KL. *Handbook of inter-rater reliability*. 4th edn. Gaithersburg, MD: Advanced Analytics; 2014.
 25. Anshel JR. Visual ergonomics in the workplace. *AAOHN J* 2007; **55**: 414–20. doi: <https://doi.org/10.1177/216507990705501004>
 26. NHS Cancer Screening Programmes. *Collecting, Monitoring and Reporting Repeat Examinations. NHSBSP good practice guide No 4 (v. 2)*. Sheffield, UK; 2006.
 27. Geher G, Hall S. *Straight forward statistics: understanding the tools of research*. 1st edn. London, UK: Oxford University Press; 2014.
 28. Glaser JS. *Neuro-ophthalmology*. 3rd edn. Philadelphia, PA: Lippincott Williams & Wilkins; 1999.
 29. Healey CG, Sawant A. On the limits of resolution and visual angle in visualization. *ACM T Appl Percept* 2012; **9**: 4. doi: <https://doi.org/10.1145/2355598.2355603>
 30. Safdar NM, Siddiqui KM, Qureshi F, Mirza MK, Knight N, Nagy P, et al. Vision and quality in the digital imaging environment: how much does the visual acuity of radiologists vary at an intermediate distance? *AJR Am J Roentgenol* 2009; **192**: W335–40. doi: <https://doi.org/10.2214/AJR.07.3515>
 31. Krupinski EA, Berbaum KS, Caldwell RT, Scharz KM, Kim J. Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol* 2010; **7**: 698–704. doi: <https://doi.org/10.1016/j.jacr.2010.03.004>