Received:
04 April 2017

Revised:
31 January 2018

Accepted:
07 February 2018

<https://doi.org/10.1259/bjr.20170246>

Cite this article as:

Strudley CJ, Young KC, Warren LM. Mammography cancer detection: comparison of single 8MP and pair of 5MP reporting monitors. *Br J Radiol* 2018; **91**: 20170246.

THE ROLE OF IMAGING IN SCREENING SPECIAL FEATURE: FULL PAPER

Mammography cancer detection: comparison of single 8MP and pair of 5MP reporting monitors

¹CECILIA J STRUDLEY, MSc, ^{1,2}KENNETH C YOUNG, PhD and ¹LUCY M WARREN, PhD

¹National Co-ordinating Centre for the Physics of Mammography, Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK

²Department of Physics, University of Surrey, Guildford, UK

Address correspondence to: Dr Lucy M Warren

E-mail: lucy.warren@nhs.net

Objective: To compare breast cancer detection using a single 8MP display with using a standard pair of 5MP monitors.

Methods: An observer study was carried out in which mammograms were read using full field views only, and again with the additional use of magnified quadrant views. Seven observers read 300 cases, one view per breast, using each display type. Cases comprised 100 normal cases and 200 cases with cancers of subtle or very subtle appearance: 100 with malignant calcification clusters and 100 with non-calcified lesions. JAFROC software was used to analyse the results.

Results: When mammograms were viewed full field only, observers performed better ($p = 0.050$) in detecting malignant calcification clusters when using the pair of 5MP monitors compared with a single 8MP monitor.

This result became non-significant when results were generalised to a population of readers. Performance in detecting calcification clusters was improved by using quadrant view in addition to full field view when using either the pair of 5MP monitors or the 8MP monitor. There was no significant difference in detection of all types of cancer between the pair of 5MP monitors and the 8MP monitor when quadrant zoom was used.

Conclusion: Providing quadrant view is used in addition to full field view, there is no significant difference in cancer detection between the 8MP monitor and the pair of 5MP monitors.

Advances in knowledge: Effect of magnification on the detectability of subtle malignant calcification clusters in breast screening.

INTRODUCTION

It is current standard practice to use a pair of 5 megapixel (MP) monitors when reporting screening mammograms. Recent innovations in display technology have made available larger format monitors allowing display of a pair of mammograms on a single monitor. An example is an 8MP monitor, which approximates to a pair of 4MP displays on a single screen. This display format raises the question of whether there is a risk of missing cancers that may have been detected using a pair of 5MP monitors.

Several studies have compared mammography displays with different resolution^{1–4} and found no significant difference in cancer detection. In these studies, readers were allowed to use magnification. However, it may not be certain that, in a busy screening environment, readers of mammograms will always make full use of magnification. If not, there may be a reduction in cancer detection due to viewing mammograms full field at a lower magnification than that available with a 5MP display.

Two recent papers have described studies comparing performance of an 8MP monitor with a pair of 5MP monitors. A study by Krupinski⁵ looked at diagnostic accuracy with free use of zoom and also made a detailed study of eye movements and time to read images. With 6 observers reading 60 cases (20 with calcification clusters, 20 with masses and 20 normal), no significant difference was found in diagnostic accuracy between the two displays. A study by Yabuuchi et al⁶ which employed 6 observers to view 120 cases (30 cancers, 90 normal), found no significant difference in diagnostic accuracy.

The study presented in this paper was larger than those carried out previously, employing 7 observers to read 300 cases (200 cancers, 100 normal). Rather than allowing free use of zoom, observers were constrained to first assess cases using the full field view only, before using quadrant zoom.

Our aim was to detect any small difference in cancer detection between a single 8MP and a dual 5MP display when

viewing mammograms using only the full field view, and also when using quadrant zoom in addition to full field view.

METHODS AND MATERIALS

In this study, two types of EIZO monitor were compared (EIZO Corporation, Hakosan, Japan): an 8MP colour monitor, and a pair of 5MP monitors. Details of the monitors are given in [Table 1](#) and [Figure 1](#).

Physics tests

Physics quality control tests were carried out on the 8MP monitor and on the pair of 5MP monitors prior to the observer study, following National Health Service Breast Screening Programme (NHSBSP) guidelines⁷ for the testing of primary mammography displays.

Case selection

Cancer detection was investigated separately for calcification clusters and non-calcified lesions because the visibility of cancers with different radiological appearances may be affected differently by the two types of display. The number of cases with cancers required for this study was based on power estimations from two similar studies^{8,9} using jackknife alternative free-response receiver operating characteristic (JAFROC) software (JAFROC, v. 4.0, D.P. Chakraborty, www.devchakraborty.com). The prospective power analysis indicated that 80–100 cases containing subtle calcification clusters and 70–90 cases with subtle soft-tissue lesions were required for 80% power and an effect size of 0.01.

The cases in this study comprised 100 normal cases and 200 cases with biopsy proven cancers of subtle or very subtle appearance, of which 100 cases included calcification clusters and 100 cases included non-calcified lesions (masses, focal asymmetries and architectural distortions). The females comprising the normal cases had since returned for the next screening episode and were found to be still normal.

These 300 clinical cases, imaged using Hologic Selenia mammography systems, were selected from the pool of screening mammograms collected as part of the OPTIMAM Image Database.¹⁰ All the mammograms in the database were de-identified. Ethical approval for the establishment and use of the database was obtained from the NHS National Research Ethics Service.

The location (marked with a rectangle bounding box around the cancer) and conspicuity of each cancer in each image was judged by an experienced radiologist (who did not participate in the

study) and recorded in the image database. The radiologist categorised each cancer as obvious, subtle or very subtle according to their professional opinion. The diagnosis recorded for each cancer was confirmed by checking the pathology data records.

Observer study design

Seven observers (four radiologists and three advanced radiographer practitioners), with an average of 8.8 years of experience (range 1–24 years), currently reading an average of 6700 mammograms per annum (range 3200–12,000), were enrolled for the observer study.

In normal screening, four images are evaluated by the reader: a cranio-caudal (CC) view and a medio-lateral oblique (MLO) view of each breast. In this study, just two images were shown, one of each breast: either a pair of CC views or a pair of MLO views. This decision was made to maximise the number of cases in the study and to maximise any differences between the arms of the study. Assessing and marking just one view of both breasts decreases the time spent by observers on each case, enabling the inclusion of a sufficient number of cases to achieve the desired statistical power within the time constraints of the study. Also, use of just one view of both breasts allowed selection of the view in which the cancer was more subtle. Using more subtle cancers should allow for more sensitive detection of differences between arms in the study, since very obvious cancers may be still be visible with large changes in image quality.

For cases including a cancer, the view where the cancer was least obvious was selected, provided that it was judged to be visible by a radiologist who reviews and marks cancers in the OPTIMAM2 database. For normal cases the view to be used was randomly selected.

The study was conducted using MedXViewer software,¹¹ which allows the observer to mark suspected cancers on the images and record answers to questions about each mark.

The study was completed in ambient light conditions not exceeding 10 lux. NHSBSP Report 0604⁷ recommends an ambient light level not exceeding 10 lux. In NHSBSP Report 1411¹² ambient light conditions not exceeding 20 lux are recommended. Using a lower ambient light level may increase sensitivity,¹³ but increase fatigue. However, steps were taken to minimise the effects of fatigue in this study as described below. Prior to the main study a set of 15 cases containing more obvious cancers were used for training purposes to familiarise the observers with the software and to test the study setup.

Table 1. Details of the monitors used in this evaluation

	8MP monitor	Pair of 5MP monitors
Format	4096 pixels × 2160 pixels	Each: 2048 pixels × 2560 pixels
Pixel pitch	0.170 × 0.170 mm	0.165 × 0.165 mm
Viewable image size	697.9 × 368.0 mm	Each: 337.9 × 422.4 mm

Figure 1. Dimensions of the 8MP monitor (left) and the pair of 5MP monitors (right).

For the main study, cases were divided into five sets of 60 cases and each set was read by each observer on each monitor type, alternating between the two types of monitor. The sets were sequenced differently for each observer and the order in which cases within each set were presented was randomised for each observer. Restrictions were placed on the pace of the study such that there was a minimum of 2 weeks before observers saw any case for the second time, to reduce any risk of cases being remembered. Observers were also restricted to a maximum of two sets of cases per day to avoid fatigue.

Observers were first presented with two images (a pair of CC or a pair of MLO images of both breasts), displayed full field with quadrant view and magnification controls disabled, and asked to mark with a single point the centre of any possible cancers they could see, indicate their confidence level that this was a malignancy (1 to 100%) and to state whether they would recall on the basis of this lesion. The software then proceeded to the second stage of the assessment for that case, retaining any marks made in the first stage. Quadrant view (1:1 image pixel to monitor pixel) was enabled and observers were instructed to step through the four quadrant views as well as the full field view. Observers were asked to review any marks already made and to look for any further cancers. Options were available to change confidence levels and recall decisions regarding any existing marks, to delete marks and to add further marks. Marks and associated data for both stages were saved in the database.

Statistical analysis

Statistical analysis was performed on a per-lesion basis. Separate analysis was performed for calcification clusters and non-calcified lesions. This type of analysis has been performed previously in such similar studies.^{8,9,14}

Observer's marks were classified as either true positives or false positives. Each mark was compared to the rectangular box drawn around the cancers by the experienced radiologist. If the location fell within the boundary of a recorded cancer the mark was designated a true positive. The marks and rectangular ground truth were stored within the image database, allowing for automated comparison. Therefore, subsequent to the study, the marks made by the observers were displayed simultaneously with the rectangles indicating cancer boundaries, to ensure none were on/near the boundary.

The primary analysis for this work was JAFROC analysis, carried out using JAFROC analysis software (v. 4).¹⁵ The alternative

free response operating characteristic curve (AFROC) is a plot of the lesion localisation fraction (fraction of lesions correctly localised), against the false positive fraction (fraction of normal images with at least one false positive). Results are in terms of a figure of merit, which is the area under the AFROC curve. This figure of merit is the probability that a malignant lesion is rated as more suspicious than a false positive on a normal image. The JAFROC software includes significance testing using the Dorfman-Berbaum-Metz analysis of variance (ANOVA) technique,¹⁶ which generates *p* values relating to differences in figures of merit for pairs of modalities. A *p*-value of 0.05 or less indicates a confidence of 95% or more that there is a real difference between the two modalities. Two analyses were carried out: Analysis 1—random reader, random case (both cases and readers treated as random factors) and Analysis 2—fixed reader, random case (the results apply to the population of cases but only for the readers used in the study).

In addition to the JAFROC analysis, the maximum lesion localisation fraction was calculated for calcification clusters and non-calcified lesions termed the cancer detection rate. This is the fraction of calcification clusters and non-calcified lesions, which are detected, when including all ratings. The JAFROC analysis software (v. 4) was used for this analysis, including significance testing using ANOVA.

RESULTS

Physics tests

Compliance with standards was demonstrated but it was noted that it was more difficult to see the low contrast narrow (single pixel) line pairs on the 8MP monitor than on the pair of 5MP monitors. The minimum and maximum luminance was 0.41 to 471 cdm^{-2} for the 8MP display and, on average, 0.51 to 492 cdm^{-2} for the 5MP displays.

Effect of monitor type on cancer detection

The mean figures of merit for the seven readers were obtained for calcification clusters and for non-calcified lesions. These figures of merit are presented in [Tables 2 and 3](#), together with the differences between the two monitor types and the *p* values for these differences. *p* values resulting from two analyses are presented: Analysis 1 (random reader and random case) and Analysis 2 (fixed reader and random case).

The mean cancer detection rates for viewing calcification clusters and non-calcified lesions on each monitor type using full

Table 2. Mean FOM when cases are viewed full field, with *p* values calculated for Analysis 1 (random reader and random case) and Analysis 2 (fixed reader and random case)

	FOM (full field view)		Difference	<i>p</i> value	
	2 × 5MP	8MP		Analysis 1	Analysis 2
Calcifications	0.665	0.640	0.025	0.149	0.050 ^a
Non-calcified lesions	0.715	0.704	0.010	0.433	0.346

FOM, figures of merit.

^aResult indicating statistically significant difference between the two monitor types.

Table 3. Mean FOM when quadrant view is used in addition to full field, with p values calculated for Analysis 1 (random reader and random case) and Analysis 2 (fixed reader and random case)

	FOM (full field plus quadrant view)		Difference	p value	
	2 × 5MP	8MP		Analysis 1	Analysis 2
Calcifications	0.731	0.734	−0.003	0.880	0.841
Non-calcified lesions	0.732	0.719	0.014	0.307	0.218

FOM, figures of merit.

field view only and full field plus quadrant views, are shown in Figures 2 and 3. In agreement with the JAFROC figure of merit, there was no significant difference in cancer detection rate for non-calcified lesions with change in monitor type, both with and without the use of quadrant zoom.

For calcification clusters, there was no significant difference in performance between the 8MP and 5MP monitors when quadrant zoom was used. When quadrant zoom was not used to detect calcification clusters, cancer detection rate was significantly higher for the two 5MP monitors compared to a single 8MP monitor, when “Analysis 2” was performed (fixed readers). When “Analysis 1” was performed (readers generalised to the population) this difference was no longer significant.

Effect of using quadrant view in addition to full field view

An improvement in cancer detection was noted when quadrant view was used in addition to full field view for both the 5MP and 8MP monitors. This improvement was statistically significant for calcification clusters. Tables 4 and 5 show the differences in figure of merit between full field view and full field plus quadrant view for the pair of 5MP monitors and for the 8MP monitor, together with the p values for Analysis 1 and Analysis 2.

Figure 2. Mean cancer detection rates for calcification clusters viewed on each monitor type using full field view only or full field view plus quadrant views (error bars indicate two standard errors in the mean). An asterisk indicates a statistically significant difference.

DISCUSSION

Effect of monitor type on cancer detection

When mammograms were displayed using full field view only, the mean JAFROC figure of merit for the pair of 5MP monitors was significantly better ($p = 0.05$) when viewing calcification clusters and using Analysis 2 (fixed reader, random case). This difference became non-significant when using Analysis 1 (both readers and cases generalised to the population). There was no significant difference in detection of non-calcified lesions between monitors with both types of analysis. These findings were also found when investigating cancer detection rates.

When quadrant view was used in addition to full field view, there was no statistically significant difference in the JAFROC figure of merit between the two types of display, for both masses and calcification clusters. Our findings are consistent with those of other smaller studies^{5,6} where observers were allowed to compensate for the reduction in resolution of displays by using magnified views.

Unlike previous studies, our study controlled the use of zoom, allowing the comparison of cancer detection using an 8MP monitor or a pair of 5MP monitors when viewing images full field (no quadrant zoom). Our study also differed from these

Figure 3. Mean cancer detection rates for non-calcified lesions viewed on each monitor type using full field view only or full field view plus quadrant views (error bars indicate two standard errors in the mean). An asterisk indicates a statistically significant difference.

Table 4. Mean figures of merit for pair of 5MP monitors, with *p* values calculated for Analysis 1 (random reader and random case) and Analysis 2 (fixed reader and random case)

	FOM (2 × 5 MP)		Difference	<i>p</i> value	
	Full field view	Full field + quadrant view		Analysis 1	Analysis 2
Calcifications	0.665	0.731	0.067	<0.001 ^a	<0.001 ^a
Non-calcified lesions	0.715	0.732	0.018	0.187	0.108

FOM, figures of merit.

^aResult indicating statistically significant difference between the two monitor types.

prior studies, in that we carried out separate analysis of cases with calcification clusters and cases with non-calcified lesions, allowing us to investigate if different radiological features are affected differently by monitor type.

In the present study, observers included four radiologists and three advanced radiographer practitioners with a range of experiences and reading volume (1–24 years and 3200–18,000 per annum). Further analysis (not presented) showed that there was no significant difference in performance between radiologists and radiographers for all arms in the study ($p > 0.5$). We found no change in performance with experiences or reading volume with all arms ($p > 0.6$). In most observer studies the largest source of variation tends to be observer performance, resulting in wide error bars. This variation is taken into consideration when performing ANOVA to test for statistical significance.

The effect of using quadrant view in addition to full field view

Cancer detection rates when using quadrant view in addition to full field view were in all cases greater than when using full field view only, and the observed differences were greatest when viewing calcification clusters. For example, when using a pair of 5MP monitors to view calcification clusters, the use of quadrant view improved the detection rate from 50 to 65%. This difference was statistically significant ($p = 0.0002$).

JAFROC figures of merit were used to test the effect of using quadrant view in addition to full field view. The additional use of quadrant view gave statistically significant improvements in performance ($p < 0.001$) when viewing calcification clusters on either type of display, using either Analysis 1 or Analysis 2. Differences in performance when viewing non-calcified lesions were not statistically significant.

Limitations

Our study differed from normal clinical practice because pairs of images were used (one view, either CC or MLO, of both

breasts), no previous images were available and the image set was enriched with cancers. Use of two views per breast may have changed some readers' decisions, in particular regarding architectural distortions or focal asymmetries. However, since the same images were viewed on both monitor types (paired study design), this is unlikely to have affected the results of the study.

This study involved a greater number of cases than other studies investigating the effect of monitor type on cancer detection^{5,6} and the number of readers and cases is similar to other observer performance studies designed to test performance of breast screening technology.^{17–19} However, this study, in common with other observer performance studies, is inherently limited by the number of observers and cases available.

The study would have been designed differently if the primary aim were to measure the effect of using quadrant view in addition to full field view. As each case was assessed using full field view, immediately followed by additional use of the quadrant views, the two modalities were not independent and the validity of the statistical analysis in this regard is therefore questionable. However, this study does illustrate the considerable advantage of using magnified views in addition to full field views for the detection of calcification clusters.

CONCLUSIONS

When quadrant view was used in addition to full field view, there was no significant difference between using the pair of 5MP monitors and using the 8MP monitor to detect malignant calcifications or non-calcified lesions.

When only the full field view was used, observers performed better ($p = 0.050$) in detecting malignant calcifications when using the pair of 5MP monitors, for analysis with fixed-readers. When readers are generalised to the population, this difference becomes non-significant.

Table 5. Mean figures of merit for 8MP monitor, with *p* values calculated for Analysis 1 (random reader and random case) and Analysis 2 (fixed reader and random case)

	FOM (8MP)		Difference	<i>p</i> value	
	Full field view	Full field + quadrant view		Analysis 1	Analysis 2
Calcifications	0.640	0.734	0.094	<0.001 ^a	<0.001 ^a
Non-calcified lesions	0.704	0.719	0.014	0.277	0.188

FOM, figures of merit.

^aResult indicating statistically significant difference between the two monitor types.

There was no significant difference in cancer detection rate for non-calcified lesions between monitors for either analysis type, when only the full field view was used.

Performance in detecting calcifications was improved by using quadrant view in addition to full field view when using either the pair of 5MP monitors or the 8MP monitor.

ACKNOWLEDGEMENTS

The authors would like to thank the staff at the Jarvis Breast Screening Centre for their assistance in carrying out this study, and EIZO for the loan of the displays used.

FUNDING

This study was funded for the NHS Breast Screening Programme by Public Health England.

REFERENCES

- Chen Y, James JJ, Turnbull AE, Gale AG. The use of lower resolution viewing devices for mammographic interpretation: implications for education and training. *Eur Radiol* 2015; **25**: 3003–8. doi: <https://doi.org/10.1007/s00330-015-3718-z>
- Yamada T, Suzuki A, Uchiyama N, Ohuchi N, Takahashi S. Diagnostic performance of detecting breast cancer on computed radiographic (CR) mammograms: comparison of hard copy film, 3-megapixel liquid-crystal-display (LCD) monitor and 5-megapixel LCD monitor. *Eur Radiol* 2008; **18**: 2363–9. doi: <https://doi.org/10.1007/s00330-008-1016-8>
- Ong AH, Pitman AG, Tan SY, Gledhill S, Hennessy O, Lui B, et al. Comparison of 3MP medical-grade to 1MP office-grade LCD monitors in mammographic diagnostic and perceptual performance. *J Med Imaging Radiat Oncol* 2011; **55**: 153–62. doi: <https://doi.org/10.1111/j.1754-9485.2011.02245.x>
- Uematsu T, Kasami M. Soft-copy reading in digital mammography of mass: diagnostic performance of a 5-megapixel cathode ray tube monitor versus a 3-megapixel liquid crystal display monitor in a diagnostic setting. *Acta Radiol* 2008; **49**: 623–9. doi: <https://doi.org/10.1080/02841850802022993>
- Krupinski EA. Diagnostic accuracy and visual search efficiency: single 8 MP vs. Dual 5 MP displays. *J Digit Imaging* 2017; **30**: 144–7. doi: <https://doi.org/10.1007/s10278-016-9917-6>
- Yabuuchi H, Kawanami S, Kamitani T, Matsumura T, Yamasaki Y, Morishita J, et al. Detectability of BI-RADS category 3 or higher breast lesions and reading time on mammography: comparison between 5-MP and 8-MP LCD monitors. *Acta Radiol* 2017; **58**: 403–7. doi: <https://doi.org/10.1177/0284185116653279>
- Kulama E, Burch A, Castellano I, Lawinski CP, Marshall N, Young KC. *Commissioning and routine testing of full field digital mammography systems (NHSBSP equipment report 0604, version 3)*. Sheffield: NHS Cancer Screening Programmes; 2009.
- Warren LM, Given-Wilson RM, Wallis MG, Cooke J, Halling-Brown MD, Mackenzie A, et al. The effect of image processing on the detection of cancers in digital mammography. *AJR Am J Roentgenol* 2014; **203**: 387–93. doi: <https://doi.org/10.2214/AJR.13.11812>
- Warren LM, Halling-Brown MD, Looney PT, Dance DR, Wallis MG, Given-Wilson RM, et al. Image processing can cause some malignant soft-tissue lesions to be missed in digital mammography images. *Clin Radiol* 2017; **72**: 799.e1–799.e8. doi: <https://doi.org/10.1016/j.crad.2017.03.024>
- Halling-Brown MD, Looney PT, Patel MN, Warren LM, Mackenzie A, Young KC. Mammographic image database (MIDB) and associated web-enabled software for research. *LNCS* 2014; **8539**: 514–9.
- Looney PT, Young KC, Mackenzie A, Halling-Brown MD. *MedXViewer: an extensible web-enabled software package for medical imaging*. San Diego, CL, US: *Proceedings of SPIE*; 2014. pp. 9037:90371K.
- Baxter G, Jones V, Milnes V, Oduko J, Phillips V, Sellars S. *Routine quality control tests for full-field digital mammography systems, equipment report 1303*. 4th ed: Public Health England; 2013.
- Pollard BJ, Samei E, Chawla AS, Baker J, Ghate S, Kim C, et al. The influence of increased ambient lighting on mass detection in mammograms. *Acad Radiol* 2009; **16**: 299–304. doi: <https://doi.org/10.1016/j.acra.2008.08.017>
- Mackenzie A, Warren LM, Wallis MG, Cooke J, Given-Wilson RM, Dance DR, et al. Breast cancer detection rates using four different types of mammography detectors. *Eur Radiol* 2016; **26**: 874–83. doi: <https://doi.org/10.1007/s00330-015-3885-y>
- Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: modeling, analysis, and validation. *Med Phys* 2004; **31**: 2313–30. doi: <https://doi.org/10.1118/1.1769352>
- Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. *Invest Radiol* 1992; **27**: 723–31.
- Svahn TM, Chakraborty DP, Ikeda D, Zackrisson S, Do Y, Mattsson S, et al. Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy. *Br J Radiol* 2012; **85**: e1074–e1082. doi: <https://doi.org/10.1259/bjr/53282892>
- Zanca F, Jacobs J, Van Ongeval C, Claus F, Celis V, Geniets C, et al. Evaluation of clinical image processing algorithms used in digital mammography. *Med Phys* 2009; **36**: 765–75. doi: <https://doi.org/10.1118/1.3077121>
- Gennaro G, Toledano A, di Maggio C, Baldan E, Bezzon E, La Grassa M, et al. Digital breast tomosynthesis versus digital mammography: a clinical performance study. *Eur Radiol* 2010; **20**: 1545–53. doi: <https://doi.org/10.1007/s00330-009-1699-5>