The Radiographer 2010; 57 (2): 8–14

Radiographers' ability to perceive and classify abnormalities on mammographic images – results of a pilot project

JJ Holt[1] and K Pollard[2]

[1]Statewide Medical Imaging Support Service – Queensland Health, Locked Bag 66, Coorparoo DC, Queensland 4151, Australia.
[2]Dentistry and Health Sciences, Charles Sturt University, Wagga Wagga, New South Wales 2678, Australia.
Correspondence Judith_holt@health.qld.gov.au

Abstract *Purpose:* The literature indicates that image interpretation as a developing role for radiographers is becoming more popular overseas. However, little is known about the ability of Australian radiographers to interpret radiographic images. The aim of this study is to provide pilot data to determine how well radiographers, without formal image interpretation training, can read mammograms. *Methods:* Twelve radiographers employed by the BreastScreen Queensland screening service were divided into two equal groups based on years of mammographic imaging experience. Each participant was asked to interpret a purposive image set containing 60 two-view primary screening bilateral mammographic images. The participants were asked to determine if an abnormality was perceived, indicate the location on a breast diagram and then classify the image set according to a 5 point scale; no specific finding, benign finding, probably benign finding, probably malignant finding, highly suggestive of malignancy. Sensitivity and specificity rates for each group was calculated and compared. *Results:* Compared to the "Gold Standard" outcomes, the sensitivity of Group A in detecting abnormalities which had resulted in a positive screening result was 82.6% as compared to Group B at 77.5%. Specificity demonstrated Group B higher than A: 79.6% compared to 75.4 %. These results were minimally lower than those found in overseas studies. *Conclusion:* The results indicate that further work in this area is justified to support the training and role development of radiographers in the formal reporting role of screening mammograms.

Keywords: categorise, mammography, perception, radiographer reporting.

Introduction

Worldwide**,** breast cancer is an important health problem affecting one in 11 women[1] over the age of 75 regardless of race.[2] If you are a woman, breast cancer is the cancer you are most likely to get.[3]

To help combat this disease, breast screening services were established in Western Europe,[4–7] in the 1980s in an attempt to reduce mortality by detecting early stage breast cancer. Following the lead of Europe, BreastScreen Australia commenced a national breast cancer screening program in 1991 by offering asymptomatic women aged 50–69 years, a free biennial screening mammogram.[8] With this program came an increased demand for radiological services; radiographers were needed to take the mammographic images and readers to interpret them. In 1997 and 1998, once the program was firmly established, over 1.25 million Australian women were screened by BreastScreen Australia[8] with over 1.6 million screens in 2004–2005.[9] In order to achieve a high cancer detection rate, BreastScreen Australia's policy states, "All mammographic images will be read and reported independently by two or more readers, at least one whom shall be a radiologist".[10]

Any non-concordant reports between the two readers are reviewed and combined into a single recommendation. Given the number of screens annually and the need for a minimum of one radiologist reading each screen, a large number of radiologists are required by BreastScreen Services. According to the 2006 RANZAR Workforce Survey,[11] between 2004 and 2006 a 10.3% increase in the number of Medicare funded imaging services was only partially met by a 6.2% increase in the number of "billing radiologists". This radiologist shortage is particularly serious for breast screening services as fewer newly trained radiologists are opting for a career in mammography.[12,13,14] With fewer radiologists available to interpret images, meeting the needs and expectations of the public will become more difficult. One solution is that training be provided to radiographers to enable them to perform certain radiological tasks such as image interpretation.[15] In the past 15 years, role development for radiographers in the realm of image interpretation has become widely accepted in both the UK[15,16] and the USA.[17] The radiographer's ability to successfully report on radiographic images in the UK and the USA is well documented in the literature;[18,19,20] however, little is known about the ability of the Australian radiographer to interpret images; specifically those skills related to mammography. Only one Australian[21] study has investigated the ability of radiographers to interpret mammographic screening images, while another[22] reports on implementing the Red Dot system to provide provisional diagnosis.

Radiographers' mammography experience may differ between screening and diagnostic centres. To be eligible to participate in the BreastScreen Australia program, women must be considered "well-women", that is, asymptomatic of any breast disease signs and symptoms.[9] All women are considered to be "well" until the results of their screening indicate otherwise. In the BreastScreen Australia program radiographers routinely carry out an average of 25 two-view bilateral screening mammograms per day and are typically the first to recognise an abnormality/anomaly on the radiographic image. In fixed site clinics, radiographers have the opportunity to view mammographic images with one another, informally consulting and theorising on the potential outcomes of the presented images. Feedback from peers is important as it reinforces learning and could be viewed as an elementary training mechanism in image interpretation. However, radiographers working on a mobile breast screening van may not know the outcome of screen reading unless they also work in the assessment clinic to which women are recalled for diagnostic work-up investigation. In the clinic, the learning experience for the radiographer may be further consolidated by viewing what the screen reader perceives as abnormal. This is in contrast with a diagnostic mammography examination which caters for women with suspected breast abnormalities presenting with a sign or symptom of breast disease. A radiographer in the diagnostic setting images fewer clients daily but is usually aware of examination results due to the nature of the reader consultation process. Under the direction of the radiologist, the radiographer is able to focus on the abnormality to obtain specific images to investigate the nature of the abnormality. The goal of diagnostic mammography is to pin-point whether the breast complaints are radiologically visible.[23,24] and an underlying abnormality beneath a palpable lump may reinforce pattern recognition to the radiographer.

Without formalised training in image interpretation, how well can radiographers specialising in breast imaging perceive findings on the mammographic images? If they do perceive an abnormality how accurate are they in interpreting what the findings are?

If research demonstrates that Australian radiographers are as capable as their UK peers[18] in mammographic lesion detection, Australian radiographers would have evidence that they are able to take on the role as a "screening mammogram reader".

Aims

The aim of this study was to investigate the potential of radiographers as image readers in screening mammography. The objective was to determine if radiographers without formal training can firstly perceive and secondly classify an abnormality.

Materials and methods

Written permission to access BreastScreen Queensland patient records was given by BreastScreen Queensland. BreastScreen Queensland Research Governance and the Charles Sturt University Ethics in Human Research Committee provided ethics approval. Participation in the study indicated informed consent.

Participants

The study was non-experimental[25] and using the snowball effect[26] an email containing a cover letter and "invitation to participate" was sent to the BreastScreen Queensland State Radiographer. This information was then dispatched electronically to those BreastScreen Queensland chief radiographers within a one hour drive from Brisbane, requesting that it be further directed to radiographers interested in participating in this study. A total of 12 radiographers from a potential pool of approximately 59 full-time, part-time and casual staff[27] agreed to participate. A short demographics questionnaire preceded the study and in order to avoid identification of individuals, the participants' reading results were divided into two groups of six based on screening and diagnostic experience. Those with the most years of combined (diagnostic and screening) mammographic experience comprised Group A and the least combined experience comprised Group B.

Procedure

In order to evaluate the participants' image interpretation skills rather than their ability to detect change over time, a purposive sample of 60, two-view bilateral primary screening mammographic images were selected, ensuring the selected examinations fulfilled the requirements set out for this study. The samples were selected retrospectively from pre-2000 screening mammograms at one accredited BreastScreen Queensland clinic. This time period was specified to ensure that original images which were to be culled could be used for the study. The Gold Standard was established by selecting images where both readers were concordant in their perception and classification of any abnormalities. In addition, follow-up of three consecutive bi-annual screenings ensured the Gold Standard was true and that no interval cancers had been detected on follow-up.

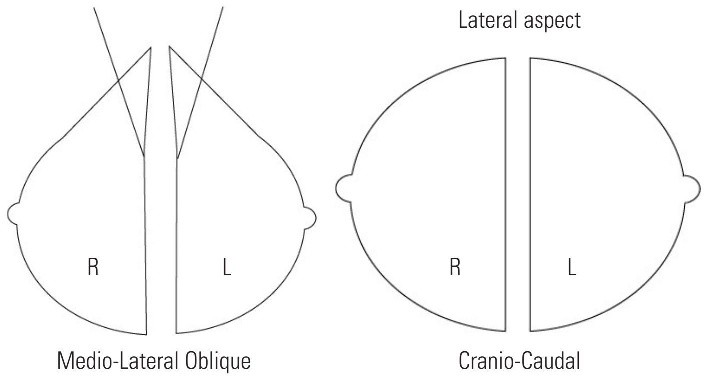

Figure 1: Pictorial outline of the Left and Right Breast in the Medio-Lateral Oblique (MLO) and Cranio-Caudal (CC) projection.[28]

Thirty-six image sets that met the selection criteria for a negative screening result were randomly selected from the pre-2000 archives. This was an attempt to ensure the normal sample was non-biased and represented a variety of breast tissue densities which might be found in the target population[26]. Twenty-four image sets that had a positive screening result (those that required additional imaging) were randomly interspersed with 36 image sets. All cases were made anonymous by applying black tape on both sides of the film over the clients' name, numbering the sets 1 to 50 with stickers and sealing with sticky tape. Only the researcher knew the status of each image set and patient confidentiality was maintained. The participants were unaware of the proportion of positive and negative cases. A standardised image set of 60 was chosen as a manageable sample size for this study to minimise the time burden for participants. In addition, a reporting session could feasibly comprise images for 50 to 100 cases, thus 60 image sets would mimic "normal" reporting conditions.

To reduce bias and to minimise the potential for variation, the radiographers reported the images independently of one another and under the same viewing conditions, such as ambient lighting, noise distractions and differences in intensity of the light boxes. A similar report form used for the Gold Standard was used by the participants. This form provided a pictorial outline of both breasts in the crainocaudal (CC) and medio-lateral oblique (MLO) view on which to mark the position of the abnormality. The radiographers were instructed to place a mark on the appropriate breast diagram to indicate any or all perceived abnormalities on each image set (Figure 1).[28]

Based on their observation of the mammogram, each radiographer then classified the perception of their findings on a five point scale (Table 1). This scoring system reflects their degree of certainty of the presence of malignant disease; with an increased score indicating increased certainty. A score of 3 means there is a 50:50 chance of malignancy in the readers' opinion. This system mimics the Breast Imaging Reporting and Data System (BI-RADS) classification scale[29] and is currently utilised for reporting mammograms in the BreastScreen Australia program.

Table 1

1 No Specific Finding
2 Benign finding
3 Probably benign finding
4 Probably malignant finding
5 Highly suggestive of malignancy

The BI-RADS system was established to give clarity to reports because it is common for those reporting any radiographs to vary in their interpretation of a finding[30].

Perception of an abnormality

The criteria to establish if the radiographer was able to perceive the same abnormality as the Gold Standard depended on determining if the radiographer saw an abnormality in the same breast; either the left or right. It was then determined if the abnormality was seen in either the CC or MLO view or both. If the radiographer stated an abnormality was in the MLO view, to further verify its position, an imaginary line which extended perpendicular from the nipple to the chest wall would determine if the abnormality was in the superior or inferior aspect of the breast. Any mark placed centrally on this line was deemed to be in the superior aspect of the breast. To substantiate this position on the CC view, an imaginary line running perpendicular from the nipple to the chest wall divided the breast into lateral and medial halves. Any mark placed centrally on this line was deemed to be in the lateral aspect of the breast. In a normal reading situation, not all abnormalities are seen in both views; therefore, a finding was scored as a true perception only if it agreed with the specified location of the selected abnormality as seen by the Gold Standard. Visualising one abnormality in one region on one of the images was considered a perception for that image set.

Classification of an abnormality

After detection, it is then necessary to decide whether the abnormality is real and significant by classifying it as a positive or negative screening. The radiographers had to judge each mammogram set according to a confidence scale as in Table 1 then each decision was compared to the Gold Standard.

With this scale, by choosing *probably benign finding, probably malignant finding* or *highly suggestive of malignancy* the radiographer was declaring a "positive screening" result, which would lead to a RECALL and the necessity for additional imaging. Choosing no *specific finding* or *benign finding* indicated a "negative screening" result and hence NO RECALL would be required.

Data analysis

The True Positive (TP), False Negative (FN), True Negative (TN) and False Positive (FP) rate (classification) for each participant was calculated and scrutinised to determine whether the radiographer was identifying the same lesion as the Gold Standard (perception). These results were combined within the participants' designated group to give the overall response rate for that group. Sensitivity, specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were calculated and compared between the two groups.

Results

Sample of participants

Of the 12 participating radiographers, 11 were over 40 years of age of which three were over the age of 60 years. Years of experience in radiography ranged from nine to 44 years (mean 23.58 years).

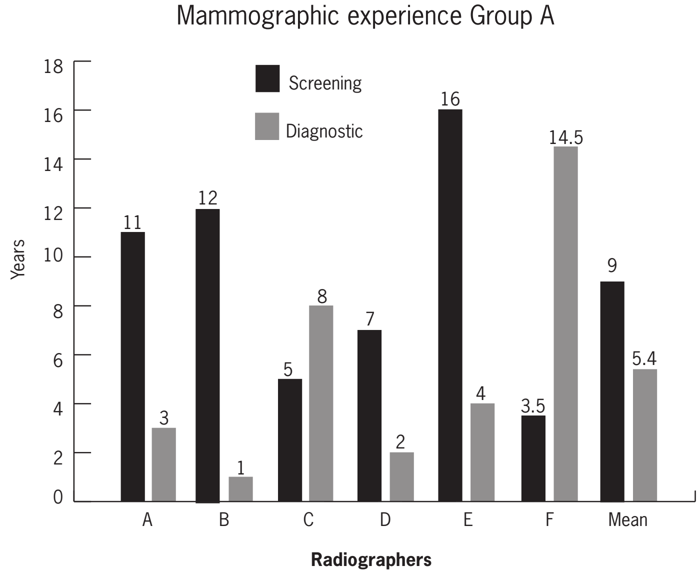

Figure 2a: Mammographic experience for all Radiographers in Group A.

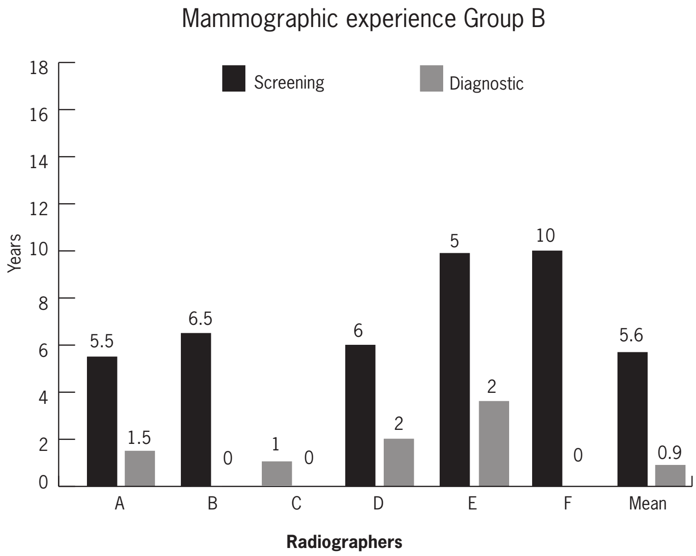

Figure 2b: Mammographic experience for all Radiographers in Group B.

The breakdown of experience in mammography for Group A and Group B is demonstrated in Figures 2a and 2b; illustrating that Group A is much more experienced overall than Group B in screening and diagnostic mammography.

Image interpretation education obtained from conference attendance confirmed a lack of expertise as 83% (10/12) of the participants indicated they had less than 10 hours of image interpretation education and 16% (2/12) specified they received more than 10 but less than 20 hours of training. Radiography qualifications of half the radiographers were from overseas with the remainder trained in Australia; only one possessed a Bachelors degree. The mean time to read the set of 60 images was 58 minutes (range 42–78). None of the radiographers indicated that they had spent any time working as a screen reader (radiologist or non-radiologist who interprets mammographic images), but 11 of the radiographers felt they should be able to work as a screen reader after undergoing training; only one radiographer felt that radiographers should not be screen readers.

A TP image set was established by the Gold Standard having perceived at least one abnormality on the two-view bilateral mammogram. With 24 TP images sets interspersed within the reading sample and six radiographers in each group, a potential existed for

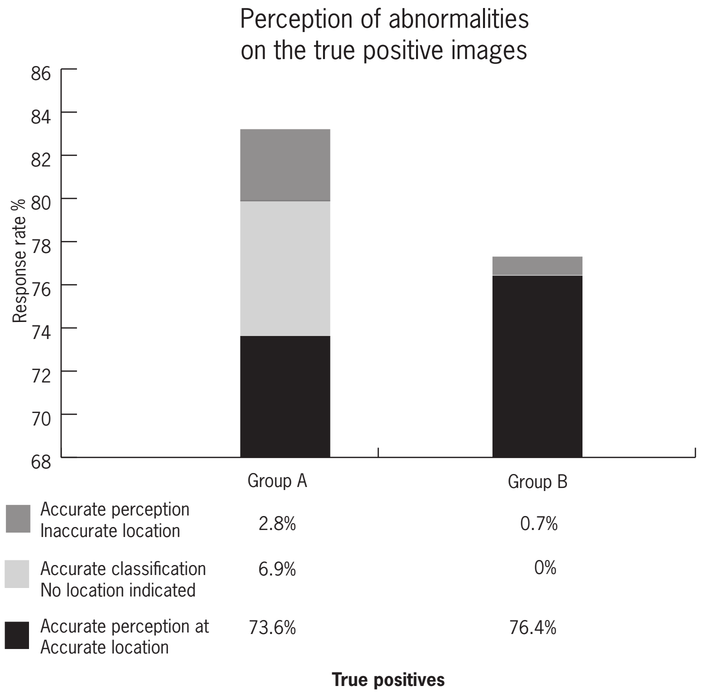

Figure 3: Images deemed True Positive by Group A and B on the True Positive images.

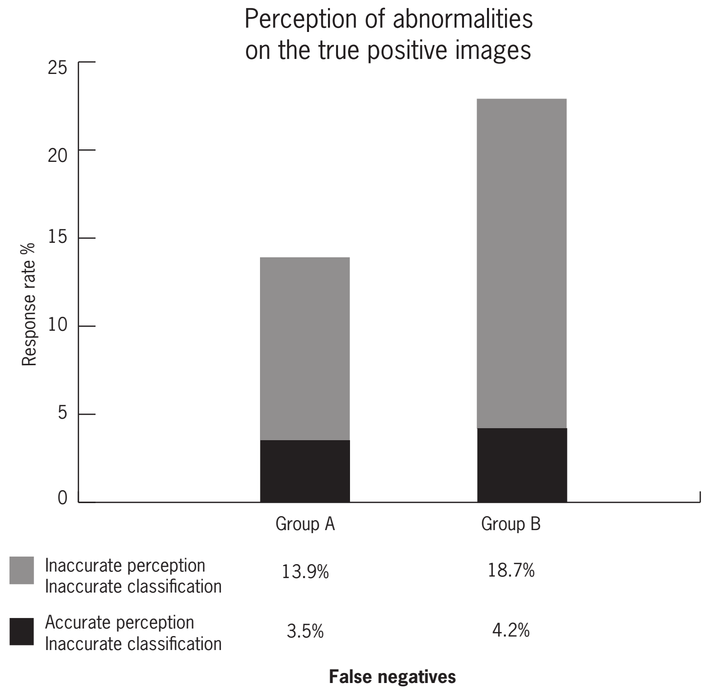

Figure 4: Images deemed False Negative by Group A and B on the True Positive images.

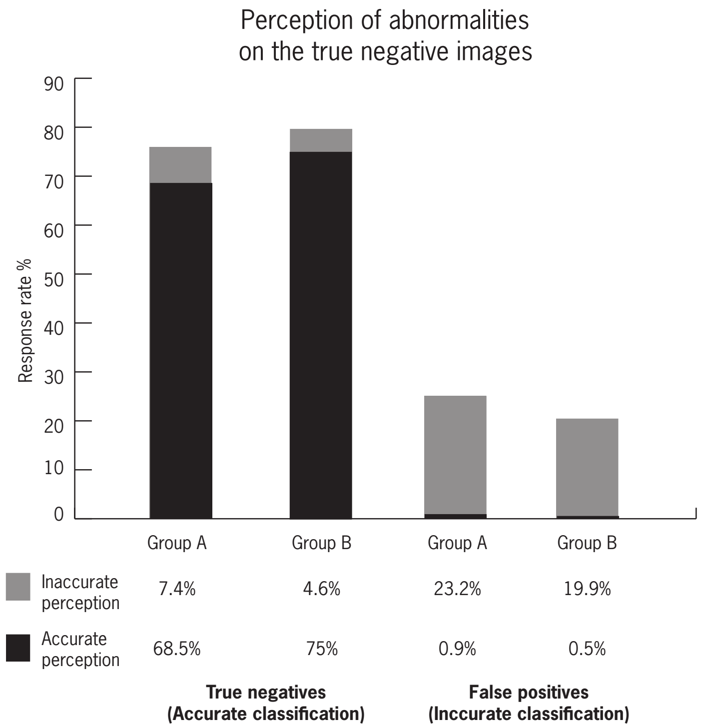

Figure 5: Perception of Abnormalities for Group A and B on the True Negative images.

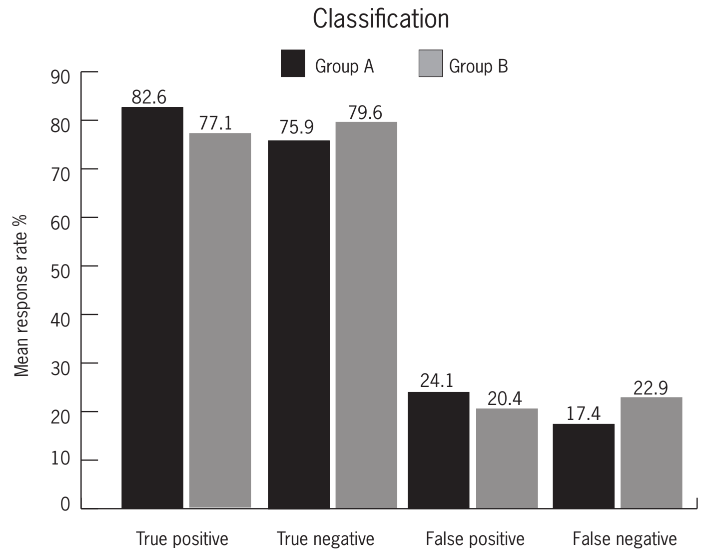

Figure 6: Mean response rate in percentage for comparing True Positive, True Negative, False Positive, and False Negative between Group A and Group B.

144 correct perceptions for each group.

The images deemed as TP by both groups on the TP images are compared in Figure 3. Group A was only slightly better (82.6%) than Group B (77.1%) in agreeing with the Gold Standard that the images had an abnormality. However, Group B perceived the correct location of the abnormality more often than Group A (76.4% versus 73.6%). Both groups indicated they perceived abnormalities on the images which was not in location concordance with the Gold Standard; incorrect region of the breast or in the unilateral breast.

Figure 4 compares the results of the images deemed as FN by both groups on the TP images.

Group A called 25/144 (17.4%) of the TP images as FN. Five (3.5%) lesions were accurately perceived but incorrectly classified, whereas 20 errors (13.9%) were attributed to poor perception. Six (4.2%) of Group B's errors were due to accurate perceptions incorrectly classified. On 27 (18.7%) occasions, Group B reported they did not see an abnormality and consequently classified the image incorrectly.

Thirty-six negative screening image sets were considered to be TN by the Gold Standard; no abnormality present. Therefore, a potential existed for each group of six radiographers to NOT perceive an abnormality in 216 instances.

Figure 5 demonstrates Group B is slightly better than Group A at identifying TN images and although Group A saw a (non-

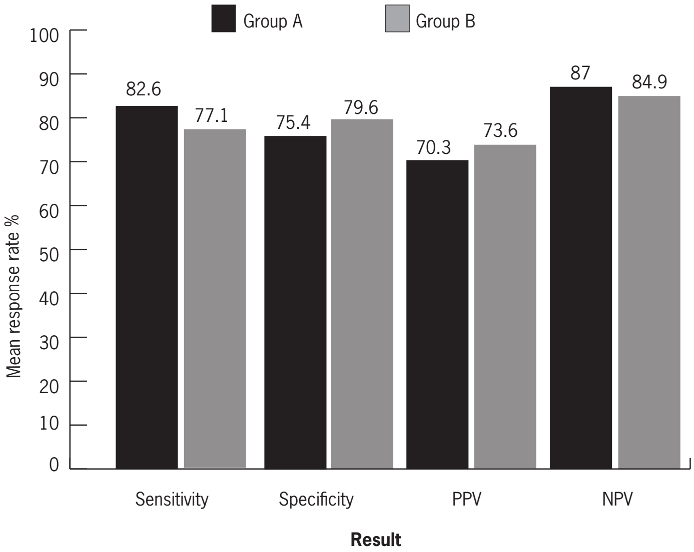

Figure 7: Comparison of Sensitivity, Specificity, Positive Predictive Value (PPV), and Negative Predictive Value (NPV) between Group A and Group B.

existent) abnormality on the images more frequently than Group B they were still able to correctly classify their perception as “benign”. Regarding the FPs, Group A believed that 2 (0.9%) had no abnormality (accurate perception, inaccurate classification) but in the remaining 50 images, stated there was an abnormality and were subsequently classified as significant. Of these 50 circumstances, 45 were categorised as being a *probably benign finding* (#3); in a normal reading situation this would require agreement between the readers. Group B indicated a FP finding on 44 occasions (20.4%); one (0.5%) did not have an abnormality (accurate perception, inaccurate classification) but 43 (19.9%) opinions resulted in a positive screening as an abnormality was (incorrectly) perceived.

Determining the TP, TN, FP and FN rate is a good indication as to how well the radiographers could accurately categorise their findings even if the radiographers did not always perceive the same abnormality as the Gold Standard. Figure 6 displays the direct classification comparison of the TP, TN, FP, FN rates for each group.

As this project was a pilot, involving only 12 radiographers reporting on 60 images, the sample size was deemed too small for a receiver operating characteristic (ROC) curve to be drawn since insufficient data can lead to unreliable statistical power.[31]

The sensitivity/specificity method of performance assessment is a popular and sound means of comparing observers if the true disease-state of the test population is known. Sensitivity is defined as the percentage of cases correctly identified as a positive screening out of all of the cases known to be a positive screen.[32] Specificity is defined as the percentage of cases correctly classified as a negative screening out of all of the cases known to be a negative screen.[32] Sensitivity and specificity were used to compare the overall report of the two groups with that of the Gold Standard; this method of analysis measured the ability of the radiographers to make a diagnosis based on the recognition or failure to identify an abnormality on the radiograph.[33]

For each radiographer in both groups, the number of TP, FP, TN and FN classifications were determined and the true positive fractions (TPF) and False Positive Fractions (FPF) were calculated, i.e. sensitivity and specificity, respectively as well as the Positive Predictive Value (PPV) and Negative Predictive Value (NPV). The PPV is the ability to say there is an abnormality when one is present, whereas the NPV is relates to the ability to express that no abnormality is present when one is not present.[26] The radiographers’ results were consolidated so that groups were compared rather than individual radiographers.

Of the 60 image sets, 40% (24/60) were categorised by the Gold Standard as requiring a “recall” for further imaging (Abnormality detected/positive screening result). Figure 7 shows a sensitivity percentage of 82.6 and specificity of 75.4% for Group A, calculating to an overall accuracy of 79.0%. Group B fell slightly short with a sensitivity of 77.1% and specificity of 79.6%; overall accuracy of 78.3%. On remaining 36 image sets, “no recall” for further imaging was required by the Gold Standard (i.e. TN). Group A recorded a specificity of 75.4% scoring slightly less than Group B at 79.6%. An approximate 3% difference separated the two groups when it came to the PPV and NPV with both groups more accurate at identifying a negative screening image as negative.

Discussion

This study was undertaken to determine how well radiographers could perceive abnormalities on mammographic images and then how well they could classify the abnormalities they detected.

It might be expected that Group A, being the more experienced group in both screening and diagnostic mammography, would perform much better than Group B, however the overall results did not denote a substantial statistical difference between the two groups to perceive and classify abnormalities. These overall results are similar to studies done in the UK[18,22,34,35] and within range of the Australian study[21].

Group A had an overall perception accuracy rate of 79% compared to 78.3% of Group B with Group A being more precise than Group B when it came to identifying images that were TP (82.6% versus 77.1%). This may be a consequence of the experience obtained in diagnostic imaging as the radiographers had the opportunity to become familiar with the appearance of screen positive abnormalities. However, it should be noted that Group A had an accurate classification on the TP images 6.9% of the time but had neglected to actually mark a location on the diagram. These sensitivity rates are similar to those published in a study by Haiart and Henderson[18] who report a sensitivity of 80% and specificity of 78% for non-trained radiographers. Group B was well within this range with a specificity of 79.6%, followed closely by Group A’s 75.4%. This is also within 5% of Sumkin, *et al.’s* [34] study where the radiographers classifying screening mammograms agreed with the reader 82% of the time. As this is a pilot study using a small sample size, only limited conclusions can be made regarding radiographers’ ability to read mammographic images but these results suggest that both groups of radiographers are accurate when reporting mammographic images and comparability between both groups and the Gold Standard is evident.

Results by Tudor, *et al.*[35] in their plain radiography study reported a sensitivity rate of 83% along with a high specificity rate of 73% with which Group A compares, where as Group B is less sensitive but more specific than the results of Tudor, *et al.*[34] Hall, *et al.*[22] show that radiographers who had limited image interpretation training were capable of an 85% accuracy rate in determining abnormal images.

However, Brealey, *et al.*[36] stress that an overall accuracy rate of less than 80% is not impressive and these results indicate that only 78.3% and 79.0% of the radiographs are being correctly reported when compared to the Gold Standard. Brealey[37] further suggests that FPs might involve further investigations for the client possibly leading to increased stress, increased risks due to radiation exposure as well as an increased financial cost to the provider. Yet when one considers that these radiographers have not had any formal training when reporting on images, it can be surmised that these rates could increase following an organised training program[38] leading to less FPs.

With a high NPV, it is clear the radiographers were confident in determining negative images. If images were pre-read by specifically trained radiographers and deemed to be negative (i.e. no specific finding) then set aside for a one radiologist read, not only could increased job satisfaction be achieved for the radiographers but the radiologist screen readers would have less images overall to read and hence more time to attend to their own higher level duties.

The use of a standardised set, although unrepresentative of a normal screen reading situation, does permit the inclusion of a greater number of abnormalities other than that which might have been found in the regular screening population. The literature reports between three and eight cancers per 1000 mammograms[39] so this does make for a more difficult set of images to interpret. In addition, a form of context bias existed as the images were not representative of the true screening population in that there was a large portion of positive screening images in this study.[40]

An expectancy bias[41] may also be responsible for the radiographer's high FP rate. The radiographers understood the purpose of the study and wanted to perform well and perhaps in order to prevent missing any pathology may have called any unsure diagnosis as abnormal (*probably benign findings*).

The lack of reading experience and confidence may be the reason that both groups attributed greater than 86% of the FPs as being a *probably benign finding* (going to a third read and there is a 50/50 chance of the woman being called back). But this is not an unusual occurrence as it could be noted that the Gold Standard is not always confident when classifying lesions. In this image set, on the TP images, the Gold Standard called *probably benign finding* in 18/24 (75%) instances. One must remember that these images were randomly selected from the pre-2000 screening archives so that there was no context or expectancy bias on the part of the Gold Standard.

A further limitation on this study was evidence of image quality degradation due to fixer retention. This likely had occurred since interpretation by the Gold Standard and possibly contributed to misinterpretation of images by the radiographers. A further study would need to investigate the possibility of using digital images or a long term study which would use the same images that were currently being read by the screen readers.

Conclusion

This pilot study aimed to assess whether there was any significant differences in perception and classification of abnormalities between the Gold Standard and the radiographer. As a pilot, this was also an opportunity to determine if the methodology provided the information to answer the question and to identify problems and ways of addressing them for a larger study.

Accuracy of radiographic interpretation was the main outcome being considered in this study. The agreement between the recommendations of the Gold Standard and those of radiographers was high, especially when considering that the radiographers did not receive any special training to perform the required classification task. Despite these encouraging results, using radiographers as pre-readers in the screening process would clearly require that they be specifically trained for this purpose. Both their detection sensitivity and specificity must increase above the baseline levels reported here. Even if radiographers can be trained to detect suspected abnormalities, characterising the abnormalities as benign or malignant with a high degree of accuracy requires a substantial amount of training.

This pilot has highlighted problems in study design which would need to be addressed in a larger study. Utilising images that are currently being reported on versus images from the archives may give a future study more accurate interpretation results for the radiographers. A reduction in expectancy bias (and hence fewer FPs) may be achieved by engaging the radiographer in a normal screen reading environment (i.e. reading as a screen reader). The results suggest that further work in this area is justified to support the training and role development of radiographers in the formal reporting role of screening mammograms.

The authors
Judith J Holt MAppSc, ABIC, PID
K Pollard ADipMR, MHEd, Med(ResMeth)

Acknowledgements

Thank you to BreastScreen Queensland for allowing us to conduct this study. To the radiographers who participated and to those would have liked to participate. Special thanks to Doctor Jane Braizer for supervising our work while at BreastScreen Queensland.

References

1 Cancer Council SA – Breast Cancer. Available online at: http://www.cancersa.org.au/aspx/breast_cancer.aspx [verified May 2010].

2 Nielsen BB, Khoury EI, Miaskowski C, Templeton B, Spitzer M, Rudisch M, *et al*. Mobile mammography: the technologist's role. *Radiol Technol* 1992; 63: 316–23.

3 Baum M, Saunders C, Meredith S. Breast cancer: a guide for every woman. London: Oxford Medical Publications; 1994.

4 Frisell J, Glas U, Hellstrom L, Somell A. Randomized mammographic screening for breast cancer in Stockholm. *Breast Cancer Res Treat* 1986; 8: 45–54.

5 Gastrin G, Miller AB, To T, Aronson KJ, Wall C, Hakama M, *et al*. Incidence and mortality from breast cancer in the Mama program for breast screening in Finland 1973-1986. *Cancer* 1994; 73 (8): 2168–74.

6 Advisory Committee on Breast Cancer Screening. Screening for breast cancer in England: past and future. NHSBSP Publication No. 61 2006. Available online at: http://jms.rsmjournals.com/cgi/reprint/13/2/59.pdf [verified May 2010].

7 Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *Br Med J* 2004; 328: 921–4.

8 National Advisory Committee for the Early Detection of Breast Cancer. Commonwealth Department of Health, Housing and Community Services, Canberra BreastScreen Australia Achievement Report 1998. Available online at: http://www.aihw.gov.au/publications/can/bsa97-8/bsa97-8-c00.pdf. [verified May 2010].

9 Statistics and research. Breast Cancer Statistics in Australia 2008. Available online at: http://www.aihw.gov.au/publications/can/bsamr04-05/bsamr04-05.pdf [verified May 2010].

10 BreastScreen Australia Program – National Policy. 2004 Available online at: http://www.health.gov.au/internet/screening/publishing.nsf/Content/national-policy [verified May 2010].

11 Royal Australian and New Zealand College of Radiologists (RANZCR) Australian Radiology Workforce Report. 2006 Available online at: http://www.ranzcr.edu.au/documents/list.cfm?lngCat=69&ophilelngEditState=3&ophileintBase=9&ophileLibrary=29&ophilesrch=&ophilesrchName=All%20records [verified May 2010].

12 Jones N. Australian Radiology Workforce Report 2000. RANZCR 2000. Available online at: http://www.doh.gov.za/mts/docs/aus-rad-workforce.pdf [verified May 2010].

13 Bassett LW, Monsees BS, Smith RA, Wang L, Hooshi P, Farria D, *et al*. Survey

of radiology residents: breast imagining training attitudes. *Radiology* 2003; 227: 862–9.

14 Wivell G, Denton ERE, Eve CB, Inglis JC, Harvey I. Can radiographers read screening mammograms*? Clin Radiol* 2003; 58: 63–7.

15 Chapman AH. Changing work patterns. *Lancet* 1997; 350: 581–3.

16 Svasti-Salee D, Flanigan JJ, Conry BG, Wetton CW, Akhurst S. The reliability of radiographer based interpretation in acute reporting of V/Q scans: a prospective assessment. *Nucl Med Commun* 2004; 25 (4): 408.

17 American registry of radiologic technologists (ARRT*)*. 2006 Available online at: http://www.arrt.org/index.html?content=radasst/raintro.htm [verified May 2010].

18 Haiart DC, Henderson J. A comparison of interpretation of screening mammograms by a radiographer, a doctor and a radiologist: results and implications. *Br J Clin Pract* 1991; 45 (1): 43–5.

19 Clinical reporting in radiography 2008. Available online at: http://www.canterbury.ac.uk/health/allied-health-professions/research/ClinicalReportingInRadiography.aspx [verified May 2010].

20 Titcomb DR, Law RL, Longstaff AJ, Slack N, Dixon AR. Radiographers in the reporting of double contrast enemas in colorectal disease. *Colorectal Dis* 2002; 4 (1): 50.

21 Moran S, Warren-Forward H. A retrospective pilot study of the performance of mammographers in interpreting screening mammograms. *The Radiographer* 2010; 57 (1): 12–9.

22 Hall R, Kleeman SJ, Egan I. The red dot system: the outback experience. *The Radiographer* 1999; 46 (2): 83–9.

23 Imaginis – The women's diagnostic health resource. Diagnostic mammography 2007. Available online at: http://www.imaginis.com/breast-health/diagnostic-mammography-2 [verified May 2010].

24 BreastScreen Australia. Policy on screening versus diagnostic mammography 2006. Available online at: http://www.health.gov.au/internet/screening/publishing.nsf/Content/br-policy-diagnostic [verified May 2010].

25 Gambling T, Brown P, Hogg P. This is not the end, nor is it the beginning-but it is the end of the beginning – getting to grips with the research process. *Radiography* 2004; 92: 161–7.

26 Minichiello V, Sullivan G, Greenwood K, Axford R. Handbook of research methods for nursing and health sciences, 2nd ed. Australia: Pearson Education; 2004.

27 Allied Health & Oral Health Clinical Education and Training Unit Business Case of the Medical Radiation Professional. June 2009.

28 Breast Diagram from BreastScreen Queensland Assessment Performa. BreastScreen Queensland.

29 Stephan P. BIRADS Breast Imaging Reporting and Data System 2008. Available online at: http://breastcancer.about.com/od/diagnosis/a/birads.htm [verified May 2010].

30 Robinson PJA. Radiology's Achilles' heel: error and variation in the interpretation of the Roentgen image. *Br J Radiol* 1997; 70: 1085–98.

31 Haslam M, Manning D. Quantifying diagnostic performance: the log odds method as a practical option. *Radiography* 2006; 12 (1): 236–43.

32 Esserman L, Cowley H, Eberle C, Kirkpatrick A, Chang S, Berbaum K, *et al*. Improving the accuracy of mammography: volume and outcome relationships. *J Natl Cancer Inst* 2002; 94 (5): 369–75.

33 Anthony D. Understanding advanced statistics – a guide for nurses and health care researchers. London: Churchill Livingstone; 1999.

34 Sumkin J, Klaman H, Graham M, Ruskauff T, Gennari R, King J, *et al*. Prescreening mammography by technologists: a preliminary assessment. *Am J Radiol* 2003; 80: 253–6.

35 Tudor GR, Finlay D, Taub N. An assessment of interobserver agreement and accuracy when reporting plain radiographs. *Clin Radiol* 1997; 52: 235–8.

36 Brealey S, King D, Warnock N. An assessment of different healthcare professionals' attitudes towards radiographers' reporting A&E films. *Radiography* 2002; 8 (1): 2–34.

37 Brealey S. Quality assurance in radiographic reporting: a proposed framework. *Radiography* 2001; 7: 263–70.

38 Pauli R, Hammond S, Cooke J, Ansell J. Radiographers as film readers in screening mammography: an assessment of competence under test and screening conditions. *Br J Radiol* 1996; 9: 10–14.

39 Sickles EA, Ominsky SH, Solitto RA, Galvin HB, Monticciolo DL. Medical audit of a rapid- throughput mammography screening practice: methodology and results of 27,114 examinations. *Radiology* 1990; 175: 323–7.

40 Egglin TKP, Feinstein AR. Context bias: a problem in diagnostic radiology. *J Am Med Assoc* 1996; 276: 1752–5.

41 Gordon PB, Borugian MJ, Warren Burhenne LJ. A true screening environment for review of interval breast cancers: Pilot study to reduce Bias. *Radiology* 2007; 245 (2): 411–5.