

Evaluation of radiographers' mammography screen-reading accuracy in Australia

Josephine C. Debono, AssocDip (DR), MAppSc,¹ Ann E. Poulos, PhD, BAHons, DipEd, DipRad,² Nehmat Houssami, MBBS (Hons), MPH, MEd, FASBP, FAFPHM, PhD,³ Robin M. Turner, MBiostat, PhD, BSc (Hons),⁴ & John Boyages, MBBS (Hons), FRANZCR, PhD,⁵

¹Westmead Breast Cancer Institute, Westmead, New South Wales, Australia

²Discipline of Medical Radiation Sciences, Faculty of Health Sciences, University of Sydney, Lidcombe, New South Wales, Australia

³Screening and Test Evaluation Program, School of Public Health (A27), Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia

⁴School of Public Health and Community Medicine, University of New South Wales, Sydney, New South Wales, Australia

⁵Macquarie University Cancer Institute, Macquarie University Hospital, Australian School of Advanced Medicine, Macquarie University, Sydney, New South Wales, Australia

Keywords

Accuracy, image interpretation, mammogram, radiographer, screen-reading

Correspondence

Josephine C. Debono, Westmead Breast Cancer Institute, PO Box 143, Westmead, NSW 2145, Australia. Tel: +61 402 232 511; Fax: + 61 2 9845 8491; E-mails: josephine.debono@bci.org.au, jodebono61@gmail.com

Funding Information

No funding information provided.

Received: 30 August 2013; Revised: 15 May 2014; Accepted: 1 June 2014

J Med Radiat Sci 62 (2015) 15–22

doi: 10.1002/jmrs.59

Abstract

Introduction: This study aimed to evaluate the accuracy of radiographers' screen-reading mammograms. Currently, radiologist workforce shortages may be compromising the BreastScreen Australia screening program goal to detect early breast cancer. The solution to a similar problem in the United Kingdom has successfully encouraged radiographers to take on the role as one of two screen-readers. Prior to consideration of this strategy in Australia, educational and experiential differences between radiographers in the United Kingdom and Australia emphasise the need for an investigation of Australian radiographers' screen-reading accuracy. **Methods:** Ten radiographers employed by the Westmead Breast Cancer Institute with a range of radiographic (median = 28 years), mammographic (median = 13 years) and BreastScreen (median = 8 years) experience were recruited to blindly and independently screen-read an image test set of 500 mammograms, without formal training. The radiographers indicated the presence of an abnormality using BI-RADS[®]. Accuracy was determined by comparison with the gold standard of known outcomes of pathology results, interval matching and client 6-year follow-up. **Results:** Individual sensitivity and specificity levels ranged between 76.0% and 92.0%, and 74.8% and 96.2% respectively. Pooled screen-reader accuracy across the radiographers estimated sensitivity as 82.2% and specificity as 89.5%. Areas under the reading operating characteristic curve ranged between 0.842 and 0.923. **Conclusions:** This sample of radiographers in an Australian setting have adequate accuracy levels when screen-reading mammograms. It is expected that with formal screen-reading training, accuracy levels will improve, and with support, radiographers have the potential to be one of the two screen-readers in the BreastScreen Australia program, contributing to timeliness and improved program outcomes.

Introduction

The goal of the BreastScreen Australia (BSA) program is the reduction of breast cancer mortality and morbidity through early detection and treatment. This goal may be compromised due to radiologist workforce shortages

contributing to delays in women receiving their screening results.¹ The number of women undergoing assessment, within the national standard of 28 days from their screening mammogram has reduced significantly between 1996 and 2005.² In 2010 there were 70 radiologists providing services per million of the Australian population,

while international levels average 100 radiologists per million of the population. Furthermore, the projected need of 1000 additional radiologists by 2021 to reach international average levels is considered unachievable.³ These delays are anticipated to increase as a result of the number of women reaching age of eligibility. There will be more women in the target age group, relative to the health and skilled labour needed for screening.² The problem is anticipated to be further compounded by the recent increase in the breast screening target age from the current 50–69 years to 50–74 years of age.²

In the United Kingdom, as a solution to similar breast screening program delays, following support on many levels, radiographers have been trained and employed alongside radiologists as second screen-readers,^{4,5} to ensure mammograms were accurately and efficiently double screen-read.^{4,6} Evidence suggests that the accuracy of radiographer screen-readers in the United Kingdom is acceptable for practice. Previous screen-reading studies have reported acceptable radiographer screen-reading accuracy levels, when compared to an appropriate gold standard of known pathology and a minimum 1-year follow-up.^{5,7–9} Further evidence reports that the addition of radiographers as screen-readers increased, rather than decreased, cancer detection rates.^{10–12}

This same strategy could potentially be applied within Australian and other contexts, as a solution to radiologist shortages. In fact the BSA Evaluation Report suggests that increasing the number of available non-radiologist screen-readers may address the increasing radiologist shortages that contribute to increasing delays.² However, Australia may differ from its U.K. counterpart due to many contributing factors including prevailing workplace attitudes, differing radiographer education and training and potential remote working conditions. These factors may contribute to variations in radiographer screen-reading accuracy. Therefore, prior to radiographers taking on the role as one of two screen-readers in any setting, it is critical to evaluate the accuracy of these radiographers as screen-readers.

Not only is there a need for an Australian study but also the need for a quality study. In their systematic review, van den Biggelaar, Nelemans, and Flobbe identified very few well-designed studies evaluating the accuracy of radiographers screen-reading.¹³ The design of this study aims to improve on previous studies by evaluating accuracy by comparison with a robust gold standard, while applying rigorous study design characteristics, including BI-RADS[®] and extensive data analysis, such as reader operating characteristic (ROC) curve analysis. The aim of this study is to evaluate the accuracy of radiographers' screen-reading mammograms in an Australian setting.

Materials and Methods

Ethics approval was obtained through Sydney West Area Health Service Human Research Ethics Committee and University of Sydney Human Research Ethics Committee.

Screen-reader recruitment

All 20 radiographers employed in both a screening (BreastScreen) and diagnostic (symptomatic) capacity by the Westmead Breast Cancer Institute (BCI) were invited to participate in this study. Ten radiographers consented to participate with no further selection criteria applied, and were aged between 27 and 64 years. They ranged in radiographic experience between 7 and 47 years (median = 28 years, interquartile range [IQR] = 30 years); mammographic experience between 3 and 27 years (median = 13 years, IQR = 20 years) and BreastScreen experience between 3 and 17 years (median = 8 years, IQR = 10 years) and were representative of BSA radiographers. They had not received any formalised screen-reading training and were required to screen-read an image test set of 500 mammogram examinations including prior images as appropriate. They were not informed of cancer prevalence so that expectation bias was minimised.

Image test set

The film-screen (analogue) image test set was obtained from screening mammograms previously read during routine screening from the BreastScreen NSW Sydney West 2004 database. All interval cancers were excluded, to ensure that the cancers included in the screening set were visible on mammography. Images were stratified into three groups: normal mammograms; mammograms recalled and assessed as benign; and mammograms recalled and confirmed as cancers on the basis of histology. A systematic selection from each stratum was undertaken to obtain a sample of 500 mammogram examinations: 400 normal mammograms (80%); 50 benign lesions (10%) and 50 cancers (10%); each consisting of cranio-caudal (CC) and medio-lateral oblique (MLO) views.

The method of representative image test set compilation is presented in Figure 1. There was a representation of most typical lesion types and no limitations were applied to levels of difficulty. The distribution of benign and malignant lesion proportions is presented in Table 1. By enriching the image test set with malignant lesions it was possible to calculate accuracy levels without the impractical time-consuming task of screen-reading several thousand consecutive population mammograms to produce a sufficiently

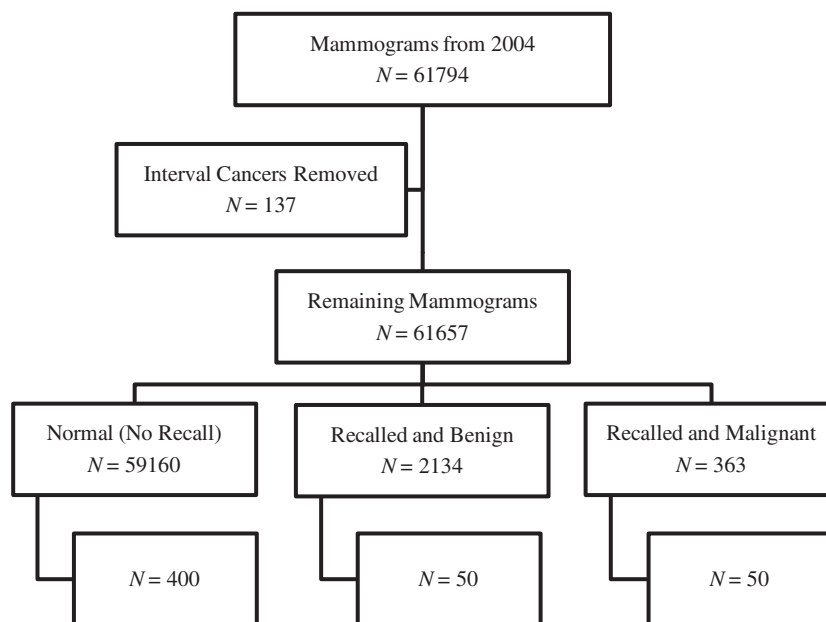


Figure 1. Method of image test set compilation.

Table 1. Distribution of benign and malignant lesion proportions.

Recalled lesion	Benign lesion		Malignant lesion	
	lesion	% Benign	lesion	% Malignant
Calcifications	2	4	10	20
Discrete mass with/without calcifications	12	24	17	34
Stellate lesion	0	0	17	34
Architectural distortion	7	14	2	4
Non-specific density	29	58	4	8
Total	50	100	50	100

powered study to calculate meaningful accuracy levels. The original mammograms were double screen-read (triple screen-read in the case of discrepancies) by various combinations of 16 radiologists employed as first or second (or third) screen-readers.

The test set was randomly sorted, and then separated into 10 batches consisting of between 30 and 55 mammograms for screen-reading by the radiographers, in order to minimise potential fatigue levels.

The test set contained three unilateral images due to women having undergone mastectomy surgery of one of their breasts giving 997 individual right or left breast images.

Procedure

Images were placed for viewing (hung) on one of three Mammoviewer 810TM image viewers (Diversified

Diagnostics, Inc., Houston, TX) in a quiet reporting room under the identical optimal viewing conditions and using the same standardised hanging protocol as originally viewed, with previous comparison images hung below the current 500 mammogram examinations.

Screen-reading process

The radiographers performed screen-reading over a series of sessions without any pre-specified time limits. Screen-reading was blinded to and independent of both the original radiologists report and other radiographers' screen-reads. Screen-reading varied between daytime, evenings and weekends, and the radiographers individually chose to read between partial or multiple batches of between 20 and 155 mammograms in each screen-reading session, depending on personal, time and travelling constraints. No participants withdrew over the screen-reading period from April 2010 to May 2011.

Radiographers read the screening images using a paper standardised reporting form with a classification scale, based on the modified BI-RADS[®] (Reston, VA) classification lexicon¹⁴ used in BreastScreen NSW, Sydney West, as presented in Table 2. This was the same protocol used by radiologist readers in 2004. The radiographers then marked any potential lesion visualised on paper on a breast diagram in the MLO and CC projections of each breast. The radiographers were told that the circling of a '3, 4 or 5' would be considered as 'recall to assessment', identical to normal screen-reading practice.

Table 2. BreastScreen Australia (BSA) recommendation requirement and BI-RADS® classification.

BSA recommendation requirement	Modified BI-RADS® lexicon of mammogram classification
1-Normal	'1' (no lesion) '2' (benign lesion)
2-Suspicious	'3' (probably benign) '4' (probably malignant) '5' (malignant)

Data analysis

The gold standard for all analyses was pathology results, interval matching and client 6-year follow-up. Accuracy for each radiographer was evaluated, and then accuracy was pooled across all radiographers. Individual radiographer accuracy levels were evaluated when screen-reading two images (CC and MLO). Both the right and left image of each mammogram (except for the three mastectomy clients, where only one breast was imaged) were included in the accuracy evaluation.

Individual accuracy

The radiographer classifications of '3, 4 or 5' were defined as test positive and '1 or 2' test negative, identically to routine screen-reading practice. If a lesion was drawn, it was checked that it corresponded with the location of the mammogram lesion. Side-specific analysis was undertaken. In the six examinations where there was more than one lesion in a breast, the lesion scoring a higher level of suspicion was used for analysis. These radiographer classifications were then compared to the gold standard classification of cancer or normal/benign to obtain the number of true and false positives (TP, FP) and true and false negatives (TN, FN). We estimated the sensitivity (true-positive rate) and specificity (true-

negative rate) together with 95% confidence intervals (95% CI) for each radiographer. The empiric ROC curve was estimated for each radiographer.

Pooled accuracy

A bivariate model^{15,16} to account for within-screen-reader correlation was used to estimate sensitivity and specificity pooled across all radiographers. This model only accounts for the within-radiographer correlation and not the within-woman correlation. To adjust pooled results for within-woman correlation, an inflation factor, based on a model of specificity alone that accounts for this correlation, was applied to the standard errors prior to the calculation of the CIs. We used separate models for each possible threshold of test positivity to estimate an approximate ROC curve. All analyses were conducted in SAS Version 9.2 (SAS Institute, Inc., Cary, NC, 2008) program and Stata Version 11.2 (Statacorp. 2009. Stata Statistical Software: Release 11, College Station, TX).

Results

Individual screen-reader accuracy

Sensitivity and specificity

Table 3 presents values for each radiographer of TP, FN, TN, FP, sensitivity, specificity and areas under the characteristic (AUC) curve (including 95% CI). Sensitivity was between 76.0% and 92.0%, and specificity was between 74.8% and 96.2%. AUC values were between 0.842 and 0.923.

Reader operating characteristic curve

The accuracy of the radiographers across varying thresholds of test positivity are shown in Figure 2. The

Table 3. Individual radiographer accuracy with BI-RADS® classification of '3, 4 or 5' considered as positive.

Reader	True positive	False negative	True negative	False positive	Sensitivity (95% CI)	Specificity (95% CI)	Area under curve (95% CI)
1	41	9	887	60	82.0 (71.2, 92.8)	93.7 (92.1, 95.2)	0.903 (0.848, 0.958)
2	38	12	911	36	76.0 (64.0, 88.0)	96.2 (95.0, 97.4)	0.899 (0.845, 0.953)
3	42	8	860	87	84.0 (73.7, 94.3)	90.8 (89.0, 92.7)	0.890 (0.835, 0.945)
4	38	12	881	66	76.0 (64.0, 88.0)	93.0 (91.4, 94.7)	0.887 (0.827, 0.947)
5	43	7	786	161	86.0 (76.3, 95.7)	83.0 (80.6, 85.4)	0.902 (0.849, 0.956)
6	40	10	862	85	80.0 (68.8, 91.2)	91.0 (89.2, 92.8)	0.896 (0.839, 0.954)
7	44	6	815	132	88.0 (78.9, 97.1)	86.1 (83.9, 88.3)	0.912 (0.862, 0.961)
8	41	9	831	116	82.0 (71.2, 92.8)	87.8 (85.7, 89.8)	0.881 (0.825, 0.936)
9	38	12	818	129	76.0 (64.0, 88.0)	86.4 (84.2, 88.6)	0.842 (0.776, 0.909)
10	46	4	708	239	92.0 (84.4, 99.6)	74.8 (72.0, 77.5)	0.923 (0.881, 0.966)

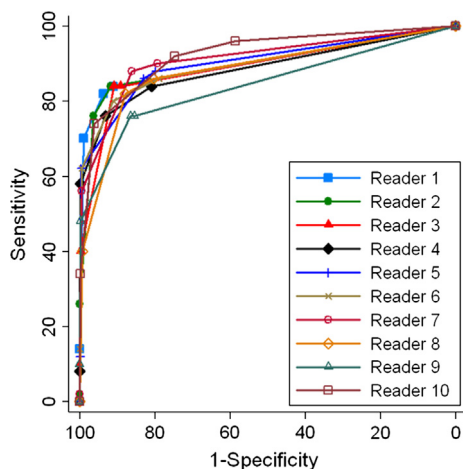


Figure 2. Reader operating characteristic (ROC) curve for each radiographer.

closer the curve to the top left-hand corner of the figure, the more accurate the reader.

Pooled screen-reader accuracy

Pooled sensitivity and specificity are presented in Table 4 for each of the separate models for the varying thresholds of test positivity.

Figure 3 shows the implied ROC curve from the estimated pooled sensitivity and specificity at the various thresholds. For the most commonly used threshold of '3, 4 or 5', we have also presented the original radiographers accuracy as red dots and the confidence region for the pooled accuracy shown in red circle. At this threshold the pooled sensitivity was 82.2% (77.1, 86.3) and specificity 89.5% (83.8, 93.3).

Discussion

Overall we found that accuracy varied across screen-readers, with sensitivity levels of 76.0–92.0% (median = 82, IQR = 10) and specificity levels of 74.8–96.2% (median = 89.3, IQR = 6.9), comparing favourably with other reported accuracy levels of radiographers and radiologists.^{5–12,17–21} These results suggest radiographers as screen-readers have reasonable levels of sensitivity and specificity when compared to the applied gold standard.

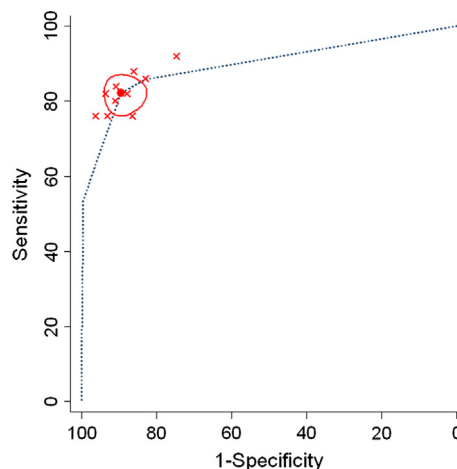


Figure 3. Pooled estimates of sensitivity and specificity for threshold of screen-reading of BI-RADS® classification of '3, 4 and 5' versus '1 and 2' with estimated reader operating characteristic (ROC) curve.

This support of ability provides further evidence in addition to previous international accuracy studies, of the ability of radiographers to read screening mammograms.

The radiographers screen-reading in this study are likely to be broadly representative of the BSA radiographer population based on their demographic and BreastScreen experience characteristics. Individual motivation and skill levels, however, may vary and while we cannot absolutely say that the study findings will generalise to the entire BSA radiographer workforce, the number of participating readers enhances the generalisability of our (pooled) estimates to all BSA radiographers.

Previous studies reported radiographer sensitivity levels of 61.0–91.42% and specificity levels of 45.0–99.1%, as presented in Table 5.^{5,7,8,10,17–20} When compared to previous studies, the accuracy values found in this study demonstrate higher minimum values of both sensitivity and specificity. This may indicate that a majority of the radiographers in this study had greater abnormality detection ability in comparison to radiographers in several previous studies, or that the difference can be explained by study differences. It is difficult to make comparisons between studies as there is such variation in methods employed. Differences in rigour of individual studies may have contributed such as variations in size, setting, composition, cancer prevalence and lesion

Table 4. Pooled sensitivity and specificity (including 95% CI) at each threshold.

Positive threshold	≥BI-RADS® '1, 2, 3, 4 or 5'	≥BI-RADS® '2, 3, 4 or 5'	≥BI-RADS® '3, 4 or 5'	≥BI-RADS® '4 or 5'	≥BI-RADS® '5'	>BI-RADS® '5'
Pooled sensitivity	100	85.8 (80.5, 89.8)	82.2 (77.1, 86.3)	53.8 (41.1, 65.9)	8.1 (3.4, 18.3)	0
Pooled specificity	0	82.9 (73.9, 89.2)	89.5 (83.8, 93.3)	99.5 (98.7, 99.8)	100.0 (99.8, 100.0)	100

Table 5. Radiographer sample size and accuracy of previous studies.

Study	Number of radiographers	Sensitivity	Specificity
Haiart and Henderson ⁸	1	80%	78%
Bassett <i>et al.</i> ⁷			
Inst. 1	8	90%	75%
Inst. 2		84%	64%
Pauli <i>et al.</i> ⁵	7	73%	86%
Pauli <i>et al.</i> ²⁰	7	83%	80%
Tonita <i>et al.</i> ⁹	3	–	–
Wivell <i>et al.</i> ⁶	3	–	–
Sumkin <i>et al.</i> ²¹	33	–	–
Holt ¹⁷	5	91.42%	87.62%
Duijm <i>et al.</i> ¹⁰	21	61.5%	99.1%
Duijm <i>et al.</i> ¹¹	21	–	–
Duijm <i>et al.</i> ¹²	21	–	–
Moran and Warren-Forward ¹⁹	11	61–89%	45–97%
Holt and Pollard ¹⁸	12	77.1–82.6%	75.4–79.6%
This study, 2013	10	82.2%	89.5%

proportion of the image set, combined with screen-reader experience, may potentially have influenced accuracy levels. The readers in our study may indeed have been more experienced than some readers in previous studies, but we cannot ascertain whether this accounts for the different findings in our study. In addition, the gold standard applied and the screen-reading process itself, as well as analysis of accuracy could account for accuracy differences.

A strength that can be attributed to the design of this study is the use of a modified BI-RADS[®] classification lexicon which facilitates ROC analysis. Pooled analysis was carried out to maximise the ability to infer the study sample results to the population.^{15,16} Pooled sensitivity and specificity levels were reported in four previous studies.^{7,10,12,17} Our approximate pooled ROC curve (see Fig. 3) represented further analysis that has not been undertaken in previous radiographer accuracy studies. ROC curves are used extensively in radiologist accuracy studies to analyse performance²² and this omission is a deficiency of previous radiographer accuracy studies, potentially due to not having a continuous test measure such as the BI-RADS[®] classification lexicon. This work therefore maximises the use of established analytic methods to estimate accuracy through ROC analysis, demonstrating high-accuracy levels relative to previous international radiographer accuracy studies.

It can be expected that these accuracy levels would increase with screen-reading education and training, as has been reported in previous studies.^{7,20} The accuracy identified from this sample of Australian radiographers is encouraging, particularly considering the results which are

comparable to trained screen-readers from other studies.^{5–8,10–12,20} This study therefore provides a baseline to measure any potential post-training accuracy differences. Several studies have recommended the necessity of training radiographers to improve accuracy prior to taking on the role as second screen-readers.^{5–7,17,20,21} Interestingly, two of those studies^{17,21} report comparable accuracy to radiologists, by radiographers without formalised screen-reading training, and may be explained through differences in experience of participants. Due to the small number of images read by some of the individual radiologists, we did not compare the accuracy of radiographers and radiologists, however, this was not an aim of this study.

In screening practice, both sensitivity and specificity are relevant due to the importance of maximising the potential for an accurate breast cancer diagnosis, while minimising anxiety caused by unnecessary recall to assessment. In typical screen-reading practice, mammograms are independently double screen-read to improve detection^{23–25} and increase overall accuracy levels by 10–15%.²⁵ In potential future radiographer/radiologist double screen-reading practice, the overall accuracy levels can be expected to increase above those reported in this study.

The representative image test set, though artificially enriched with lesions, was representative of 2004 screens, the last year that had been matched for interval cancers at the time of test set compilation. This approach increased study quality by ensuring known outcomes (and hence correct classification of cases) and by including lesions representative of screen-detected cancers. Enriching the image test set with lesions allowed a sufficiently powered study to calculate meaningful accuracy levels. Although enriched test sets may lead to expectation bias due to radiographers expecting to find more cancers than are likely in consecutive population screen-reading, this is an established method in observer studies. Pauli *et al.* and Scott *et al.* have undertaken studies that report correlation between image test set results and consecutive screen-reading practice.^{20,26} Furthermore, Gur *et al.* have reported consecutive screen-reading accuracy being significantly better than image test set results.²⁷ Of note, we designed the study with attention to minimising the likelihood of expectation bias by ensuring that readers were not informed of the cancer prevalence or lesion proportion within the image set, therefore, this bias does not constitute a major limitation of our work.

It could be argued that a potential limitation of the study is that the method used to compile the image set did not allow us to calculate radiologists' accuracy. While such a comparison would have provided some information it would not have been representative of the screen-readings performed by the radiologists (because it

included a small subset of the screens from the study timeframe) and this was not a primary aim of our work and therefore does not undermine the integrity of the study. The test set sample limitations include the lack of allowance for within-woman correlation, potentially leading to confidence intervals that are minimally underestimated, however, appropriate adjustment was made.

Australia predominantly uses digital technology; however, it is likely that these findings on readers' accuracy using film-screen mammograms will transfer to screen-reading of digital mammograms. Previous studies have reported overall similar accuracy levels when comparing film-screen images and digital images.²⁸

The results of this study provide evidence of the ability of radiographers to identify abnormalities in screening mammograms in an Australian setting. Future work is necessary to determine whether radiographers can be accepted as screen-readers in Australia. The BSA National Accreditation Standards (NAS)²⁹ state that 'For both medical and legal acceptance of the BreastScreen Australia program, it is necessary that at least one reader be a radiologist', and 'if the need arises...specifically trained non-radiologist readers could be employed' (2008, p. 43). While maintaining this important standard, and to maintain the efficiency and accuracy of the BSA program, a radiographer could take on the role of one of the two screen-readers.

For radiographers to be accepted as non-radiologist readers, support is needed from relevant bodies within Australia. Currently, there are both political and legislative challenges to radiographers taking on the role of screen-reading. This is a very complex arena comprising a multitude of competing arguments which deserve in-depth analysis outside the scope of this paper. Suffice to say much has been written on this topic³⁰ and much has been discussed and reported.³¹

Conclusion

This reported study has provided evidence that radiographers in an Australian setting demonstrate adequate screen-reading accuracy. While radiographers have the potential to be one of the two screen-readers within the BreastScreen Australia program, the current Australian context requires political and legislative change for this to take place. It may yet be some time before radiographers can contribute in this way to the timeliness and accuracy of program outcomes.

Acknowledgements

Acknowledgement goes to the Westmead Breast Cancer Institute for enabling the undertaking of this research.

Thanks to the members of staff who generously contributed their skills and time.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Commonwealth of Australia Department of Health and Ageing. BreastScreen Australia Evaluation Review of Infrastructure and Capacity. Screening Monograph No.6/2009.
2. Commonwealth of Australia Department of Health and Ageing. BreastScreen Australia Evaluation. Evaluation Final Report. Screening Monograph No.1/2009.
3. Royal Australian New Zealand College of Radiologists (RANZCR). RANZCR Radiology Workforce Report: Australia 2010. Available from: http://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&sqi=2&ved=0CCEQFjAA&url=http%3A%2F%2Fwww.ranzcr.edu.au%2Fcomponent%2Fdocman%2Fdoc_download%2F718-2010-radiology-workforce-survey-report&ei=hJNEUPHRPM6fiAfBpYCwBg&usg=AFQjCNGCKV04VqGwTgE0GsbVCwMkexVIHA&sig2=BiA0VbROi-SKcmbtiuJXlw (accessed 2 December 2013).
4. Bennett R, Sellars S, Blanks R, Moss S. An observational study to evaluate the performance of units using two radiographers to read screening mammograms. *Clin Radiol* 2011; **67**: 114–21.
5. Pauli R, Hammond S, Cooke J, Ansell J. Comparison of radiographer/radiologist double film reading with single reading in breast cancer screening. *J Med Screen* 1996; **3**: 18–22.
6. Wivell G, Denton E, Eve CB, Inglis JC, Harvey I. Can radiographers read screening mammograms? *Clin Radiol* 2003; **58**: 63–7.
7. Bassett L, Hollatz-Brown A, Bastani R, Pearce J, Hirji K, Chen L. Effects of a program to train radiologic technologists to identify abnormalities on mammograms. *Radiology* 1995; **194**: 189–92.
8. Haiart DC, Henderson J. A comparison of interpretation of screening mammograms by a radiographer, a doctor and a radiologist. *Br J Clin Pract* 1991; **45**: 43–5.
9. Tonita J, Hillis J, Lim C. Medical radiologic technologist review: effects on a population-based breast cancer screening program. *Radiology* 1999; **211**: 529–33.
10. Duijm L, Groenewoud J, Fracheboud J, de Koning H. Additional double reading of screening mammograms by radiologic technologists: impact on screening performance parameters. *J Natl Cancer Inst* 2007; **99**: 1162–70.
11. Duijm L, Groenewoud J, Fracheboud J, van Ineveld B, Roumen R, de Koning H. Introduction of additional

- double reading of mammograms by radiographers: effects on a biennial screening programme outcome. *Eur J Cancer* 2008; **44**: 1223–8.
12. Duijm L, Louwman M, Groenewoud J, van de Poll-Franse L, Fracheboud J, Coebergh J. Inter-observer variability in mammography screening and effect of type and number of readers on screening outcome. *Br J Cancer* 2009; **100**: 901–7.
 13. van den Biggelaar FJHM, Nelemans PJ, Flobbe K. Performance of radiographers in mammogram interpretation: a systematic review. *Breast* 2008; **17**: 87–92.
 14. American College of Radiology. BI-RADS-Mammography, 4th Edn. Available from: http://www.acr.org/SecondaryMainMenuCategories/quality_safety/BIRADSAAtlas/BIRADSAAtlasexcerptedtext/BIRADSMammographyFourthEdition.aspx (accessed 4 February 2010).
 15. Reitsma J, Rutjes A, Scholten R, Bossuyt P, Zwinerman A. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982–90.
 16. van Houwelingen H, Zwinderman K, Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993; **12**: 2273–84.
 17. Holt J. Evaluating radiological technologists ability to detect abnormalities in film-screen mammographic images: a decision analysis pilot project. *Can J Med Radiat Technol* 2006; **37**: 24–9.
 18. Holt J, Pollard K. Radiographers' ability to perceive and classify abnormalities on mammographic images-results of a pilot project. *Radiographer* 2010; **57**: 8–14.
 19. Moran S, Warren-Forward H. A retrospective pilot study of the performance of mammographers in interpreting screening mammograms. *Radiographer* 2010; **57**: 12–9.
 20. 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; **69**: 10–4.
 21. Sumkin JH, Klamon HM, Graham M, et al. Prescreening mammography by technologists: a preliminary assessment. *AJR Am J Roentgenol* 2003; **180**: 253–6.
 22. Goddard C, Gilbert F, Needham G, Deans H. Routine receiver operating characteristic analysis in mammography as a measure of radiologists' performance. *Br J Radiol* 1998; **71**: 1012–7.
 23. Anderson E, Muir B, Walsh J, Kirkpatrick A. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994; **49**: 248–51.
 24. Beam C, Sullivan D, Layde P. Effect of human variability on independent double reading in screening mammography. *Acad Radiol* 1996; **3**: 891–7.
 25. Thurfsjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. [see comment]. *Radiology* 1994; **191**: 241–4.
 26. Scott H, Gale A. How much is enough? Factors affecting the optimal interpretation of breast screening mammograms. *SPIE Proceedings: Medical Imaging Conference 2007: Image Perception, Observer Perception, and Technology Assessment*. 2007;8(6515, 65150F).
 27. Gur D, Bandos A, Cohen C, et al. The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. *Radiology* 2008; **249**: 47–53.
 28. Hendrick R, Cole E, Pisano E, et al. Accuracy of soft-copy digital mammography versus that of screen-film mammography according to digital manufacturer: ACRIN DMIST retrospective multireader study. *Radiology* 2008; **247**: 38–48.
 29. BreastScreen Australia. 2008. BreastScreen Australia National Accreditation Guidelines. BreastScreen Australia Quality Improvement Program [Internet]: [43 p.]. Available from: [http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/br-accreditation/\\$File/standards.pdf](http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/br-accreditation/$File/standards.pdf) (accessed 2 December 2013).
 30. Smith T, Baird M. Radiographers' role in radiological reporting: a model to support future demand. *Med J Aust* 2007; **186**: 629–31.
 31. Australian Institute of Radiography (AIR). A Model of Advanced Practice in Diagnostic Imaging and Radiation Therapy in Australia. Melbourne. 2009. Available from: http://www.air.asn.au/cms_files/09_AdvancedPractice/APWG_Final_Report_260609.pdf (accessed 2 December 2013).