

**Ann J Carrigan**<sup>1,2,4</sup>

BAppSc (MRS), GradDip  
Medical Sonography, BSc  
Psych (Hons)

**Patrick C Brennan**<sup>3</sup>

PhD

**Mariusz Pietrzyk**<sup>3</sup>

PhD

**Jillian Clarke**<sup>3</sup>

BAppSc (Hons),  
GradDip Medical  
Sonography, MHIthScEd

**Eugene Chekaluk**<sup>4</sup>

PhD

<sup>1</sup>Perception in Action  
Research Centre &  
Department of Cognitive  
Science  
Macquarie University  
Sydney  
New South Wales  
Australia

<sup>2</sup>ARC Centre of Excellence in  
Cognition & Its Disorders  
Macquarie University  
Sydney  
New South Wales  
Australia

<sup>3</sup>Medical Image Optimisa-  
tion and Perception Group  
(MIOPeG)  
Discipline of Medical  
Imaging and Radiation  
Sciences  
Faculty of Health Sciences  
University of Sydney  
Sydney  
New South Wales  
Australia

<sup>4</sup>Department of Psychology  
Macquarie University  
Sydney  
New South Wales  
Australia

Correspondence to email  
ann.carrigan@mq.edu.au

# A 'snapshot' of the visual search behaviours of medical sonographers

## Abstract

*Introduction:* Visual search is a task that humans perform in everyday life. Whether it involves looking for a pen on a desk or a mass in a mammogram, the cognitive and perceptual processes that underpin these tasks are identical. Radiologists are experts in visual search of medical images and studies on their visual search behaviours have revealed some interesting findings with regard to diagnostic errors. In Australia, within the modality of ultrasound, sonographers perform the diagnostic scan, select images and present to the radiologist for reporting. Therefore the visual task and potential for errors is similar to a radiologist. Our aim was to explore and understand the detection, localisation and eye-gaze behaviours of a group of qualified sonographers.

*Method:* We measured clinical performance and analysed diagnostic errors by presenting fifty sonographic breast images that varied on cancer present and degree of difficulty to a group of sonographers in their clinical workplace. For a sub-set of sonographers we obtained eye-tracking metrics such as time-to-first fixation, total visit duration and cumulative dwell time heat maps.

*Results:* The results indicate that the sonographers' clinical performance was high and the eye-tracking metrics showed diagnostic error types similar to those found in studies on radiologist visual search.

*Conclusion:* This study informs us about sonographer visual search patterns and highlights possible ways to improve diagnostic performance via targeted education.

*Keywords:* diagnostic errors, medical imaging, sonographers, visual search.

## Introduction

Visual search, the task of finding a target among distractors is commonplace in daily life. The processes involved in visual search have been studied extensively in laboratory settings and more recently in more real world settings on experts in a range of domains (e.g., airport luggage screening,<sup>1</sup> aeroplane piloting<sup>2</sup> and medical imaging<sup>3</sup>).

Radiologists are experts in visual search of medical images. They are required to visually search a cluttered medical image and then make a diagnostic decision based on abstract anatomical features. This task demands a high accuracy of performance. To reach a level of expertise in image interpretation, comprehensive training and practice is required.

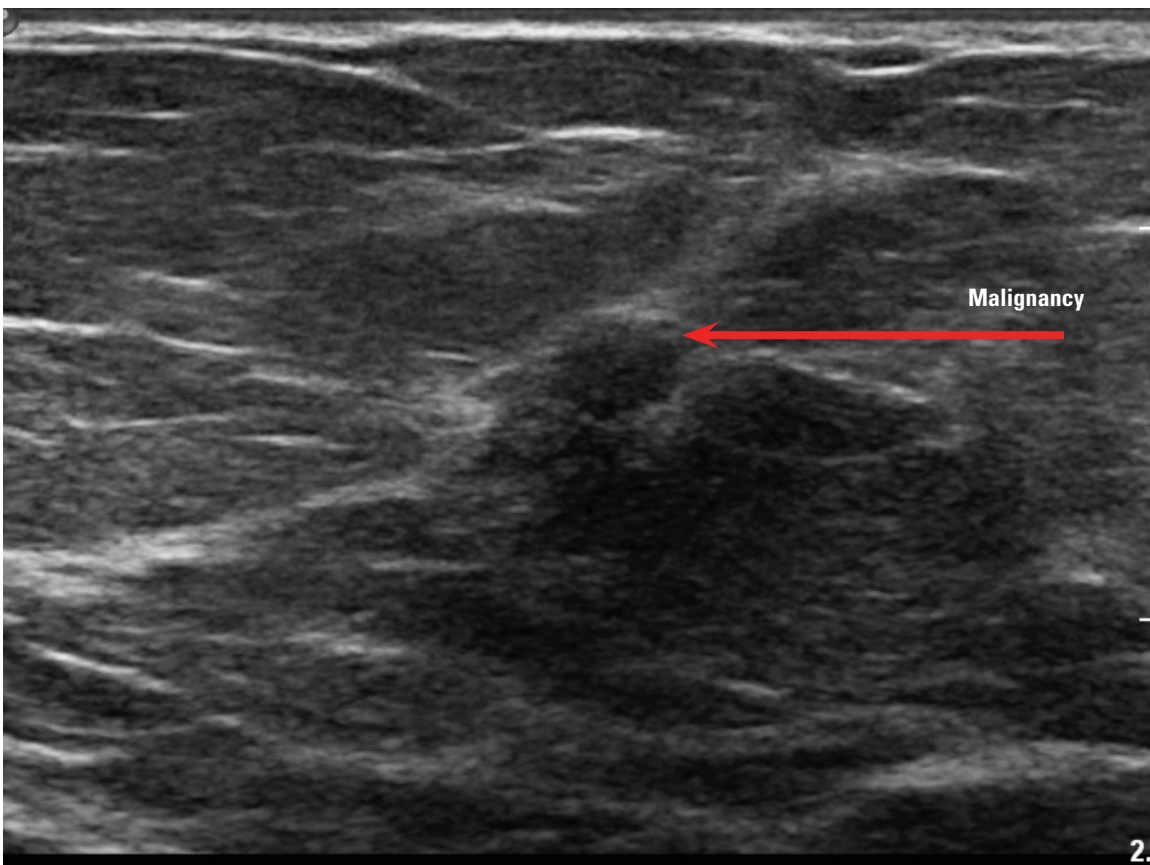
All tasks that rely on human visual search are prone to error. When errors occur within medicine, the societal cost is high and so it is important to study the processes that underpin this type of search. Medical imaging practices are trending toward higher volume and increasingly complex examinations, meaning human errors could potentially increase. It has been reported in some areas within radiology that there may be up to 30% miss error rates and equally high false alarm rates.<sup>4</sup>

In Australia, within the sub-specialty of ultrasound, medical sonographers perform the diagnostic scan, select images and present to the radiologist for reporting. This means that if the sonographer has not detected an abnormality, it cannot be diagnosed and reported. Among technologists, sonographers hold the unique responsibility of having to perform continuous visual and cognitive tasks, to make judgments and capture pathologies in real-time to ensure all relevant images are available for diagnosis by the radiologist. Therefore, it makes sense to study sonographers, as the underlying cognitive and perceptual processes involved in the visual task are identical to radiologists and the potential for errors are similar.

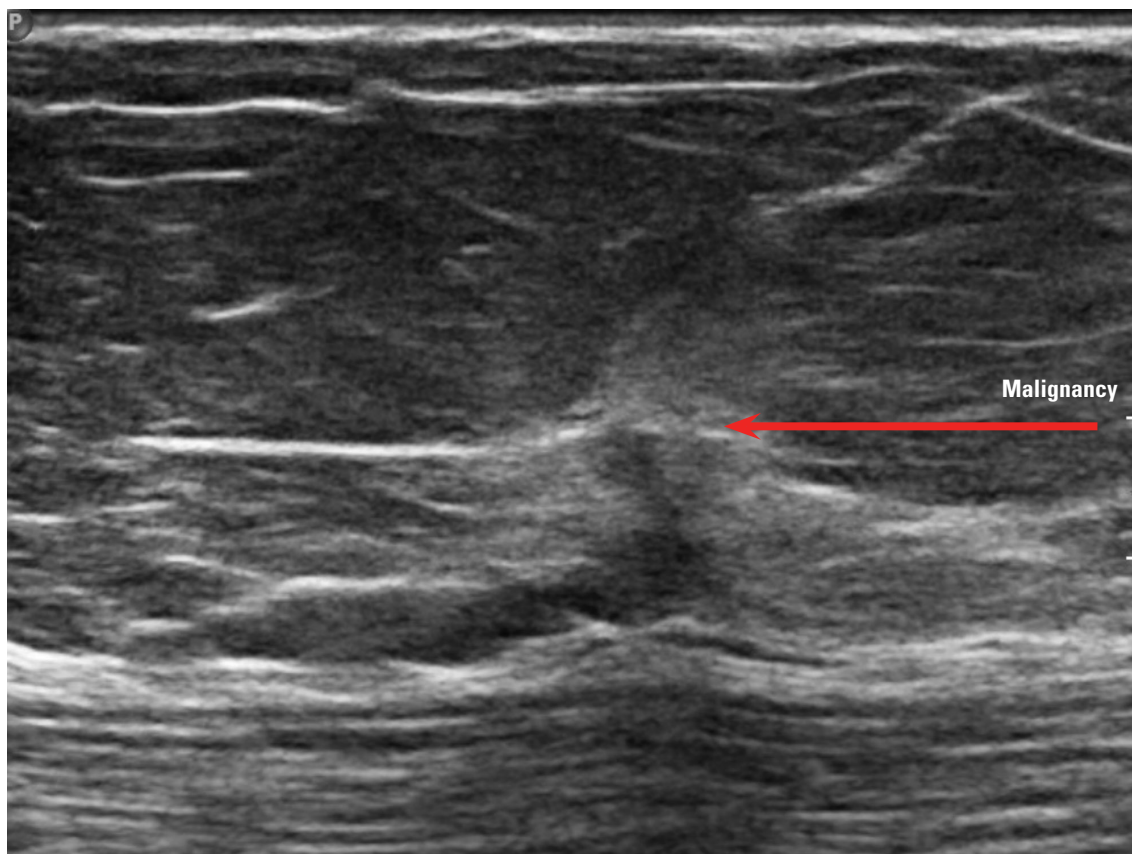
In regard to diagnostic accuracy, approximately 60% of all radiological diagnostic errors can be attributed to cognitive or perceptual errors.<sup>5</sup> A primary goal of image perception research in radiology has been to model the visual search strategies that radiologists use, in order to mitigate errors. The majority of studies have taken a perceptual approach, using eye tracking to quantify image perception.<sup>6</sup> Eye or gaze tracking is a method used to measure visual search behaviour, such as where an observer is focusing, how often and length of time.



**Figure 1:** Example of an easy case.



**Figure 2:** Example of a Medium case.



**Figure 3:** Example of a Difficult case.

Kundel, *et al.*<sup>7</sup> reported that medical imaging errors can be attributed to three types of errors that occur when searching an image for pathology. These errors were categorised into the following: visual search errors, where they never fixate the abnormality (30%); recognition errors, where the abnormality is fixated but may be camouflaged in the image and therefore no meaning is assigned to the area (25%); and decision errors, where the abnormality is fixated but actively dismissed as an abnormality (45%). While experience and specialisation, such as frequent exposure to certain pathologies over a number of years may reduce errors in identification of those pathologies, this likelihood of errors highlights the necessity to understand the clinical and image interaction and the psychophysical phenomena that exist.

While there have been studies on radiologist visual search and observer performance,<sup>8,9</sup> comparable studies of medical sonographers are not known to have been reported. It follows to study sonographers, given the clinical importance of their role in medical imaging. The goal of the current study was to explore and understand the visual search behaviours of medical sonographers and the diagnostic errors that may occur. This was achieved by studying clinical performance and eye-gaze behaviours of a group of Australian accredited medical sonographers in their workplace.

## Method

### Participants

Thirty accredited medical sonographers (26 females) with a mean age of 42 years, ranging from 27 to 62 years (SD =

9.84), were recruited on a voluntary basis via medical imaging departments throughout metropolitan and rural New South Wales. In Australia there are approximately four-and-a-half thousand accredited medical sonographers, of which 76% are female.<sup>10</sup> Demographic information included the number of years spent working as a medical sonographer (M = 14.24, SD = 8.47, range 1 to 34 years). The study was approved by the Macquarie University Human Research Ethics Committee (reference number: 5201300123).

### Stimuli

The test set consisted of fifty de-identified ultrasound images (50% contained a pathologically confirmed cancer) of the female breast which were collected prior to the study. These were imaged with a Philips iU22 (Philips Healthcare Solutions, Bothell, WA, USA), using a 17MHz linear transducer and were randomly presented on a MacBook Pro (resolution of 1280 x 800 pixels). Difficulty ratings of the images were independently assessed and categorised by a NSW BreastScreen radiologist (PS; > 25 years experience) and classed as easy (8), medium (9) or difficult (8) (Figure 3).

### Design and procedure

The study was conducted in each sonographer's usual work area 45 mins prior to their workday. We informed the sonographers that the images were of the female breast and that some contained a biopsy confirmed malignancy but no benign masses. We asked the sonographers to view the images on the laptop, make a binary classification (normal or cancer), then to "please mark on any malignancy" on a corresponding photocopied image on

**Table 1:** ROC, JAFROC, Sensitivity and Specificity: Mean, SD and 95% CI.

	M	SD	95% CI	
			Lower	Upper
ROC	.84	.07	.78	.91
JAFROC	.84	.07	.78	.90
Sensitivity	.75	.15	.69	.80
Specificity	.87	.10	.83	.91

Note. ROC = receiver operating characteristic; JAFROC = jackknifing free response operating characteristic; CI = confidence interval.

N = 29

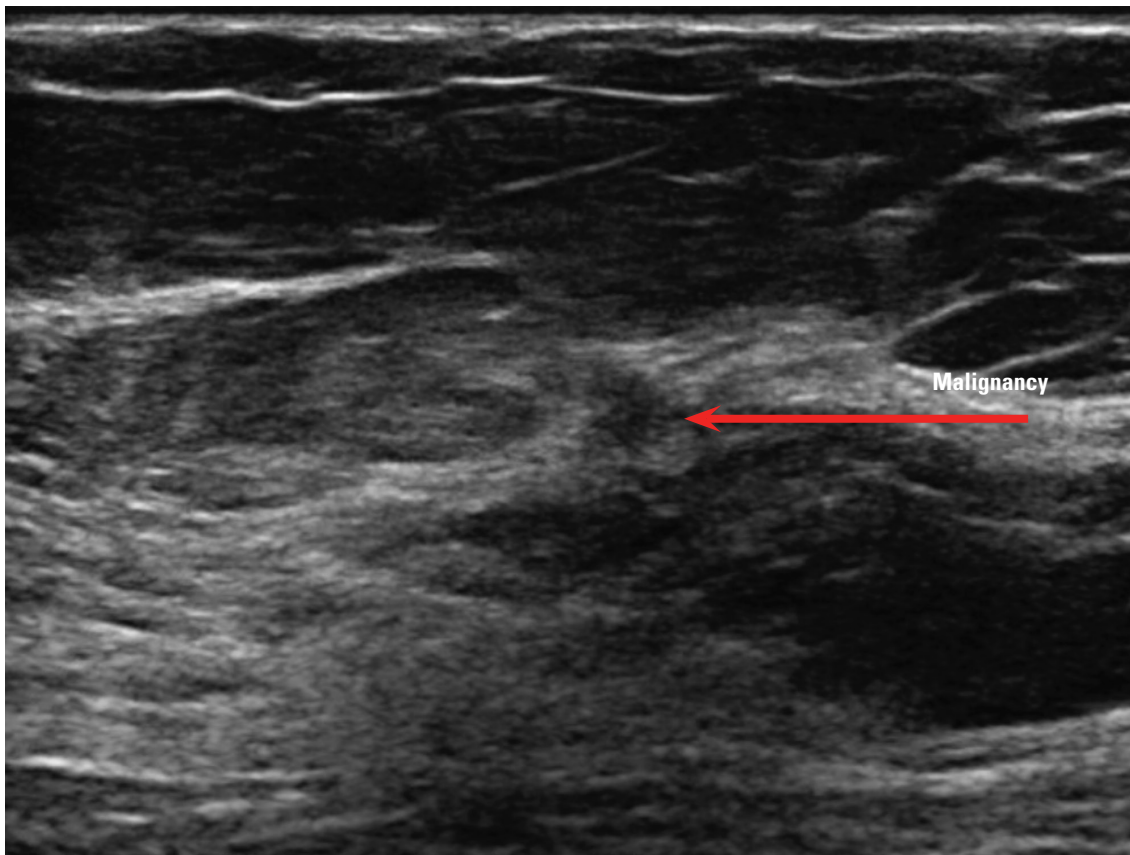
**Table 2:** Eye-tracking Metrics: Median Time-to-First-Fixation (TFF) (s).

	Seconds
Easy	2.18
Medium	2.77
Difficult	4.24

**Table 3:** Eye tracking Metrics: Median Total Visit Duration (TVD) (s).

Difficult 8	
AOI	Outside AOI
0.54	8.17

Note: AOI = area of interest. n = 4

**Figure 4:** Difficult 8.

A4 paper. They were instructed to “rate on a scale how confident you are that there is a malignant lesion” on a 6-point scale where 1 represented they were absolutely confident there is NO malignancy and 6 represented being absolutely confident there IS a malignancy. This confidence scale facilitates the calculation of a receiver operating characteristic (ROC), a function of observer sensitivity and specificity and has a range of 0–1, where 1 represents the ‘perfect’ observer. The jack-knifing free response operating characteristic method (JAFROC) extends upon ROC by taking lesion location into account. JAFROC has a range of 0-1 and it is defined as the probability that an obvious lesion seen on an abnormal image will be rated higher than the highest rated non-lesion on a normal image.<sup>11</sup> JAFROC has been used widely for the analysis of human observer, free-response data and validated in radiological populations.<sup>9,12,13</sup> For radiology observer performance studies, JAFROC statistical power has been reported as 0.8.<sup>14</sup>

The sonographers were allowed free viewing time and began the next image with a key press. They viewed the clinical images binocularly at a distance of approximately 57 cm and no feedback was provided during the study. At the completion of the tasks they commenced their normal workday.

The responses to the 25 cancer absent images were classified either as a true negative (TN) when confidence was marked less than four, or a false positive (FP) when confidence was marked greater than three. The cancer present images were recorded as a true positive (TP) when confidence marked greater than three, or a false negative (FN) when confidence marked less than four. A TP was scored if they marked within a region of acceptance surrounding the cancer (defined as the radius of the largest lesion = 1.5 cm). Sensitivity was calculated as the number of correct TP/25 on the cancer present images and specificity was calculated as the number of TN/25 on the cancer absent images.

## Apparatus

For seven participants, we recorded eye-tracking data during the clinical performance trials. The eye-tracking component of the study consisted of a single computer with a double screen LCD monitor: HP 2035, with a resolution of 1600 x 1200 pixels. The data was recorded using a remote eye tracking system, Tobii X50™ (Tobii Technology, Danderyd, Sweden).

The sonographers sat at a distance of 57 cm and following callibration we asked them to view the images and click in the centre of an area with a computer mouse to indicate a cancer. We then asked them to rate their confidence on a 1 to 6 scale using the computer keyboard. We measured: time-to-first-fixation (TFF), which represents the time in seconds taken from when the stimulus (image) was shown until the start of the first time the eye pauses and focuses on a lesion; total visit duration (TVD), total amount of time in seconds the eye pauses on a lesion and cumulative dwell time (CDT). Due to the quality of the recordings, eye-tracking metrics were extracted and analysed for four participants.

## Results

### Clinical Performance

Table 1 shows the results for twenty-nine participants (data for one sonographer was excluded due to error) and shows that their clinical performance was high (Table 1).

### Eye-tracking recordings

The metrics for median TFF (Table 2) shows that as difficulty of the cancer increased, the sonographers took longer time to fixate on the cancer.

A second metric, TVD (seconds) for Difficult 8 (D8) is presented in Table 3.

The sonographers fixated longer outside as compared with inside the area of interest (AOI). For lesion D8, none of the sonographers ( $n = 4$ ) indicated a true positive.

The cumulative dwell time (CDT) heat maps, where red indicates longer dwell times, provide insight into visual behaviour and the possible causes of sonographer diagnostic errors. The cancer absent image that had the highest number of false positives recorded was Normal 11 (N11), with cancer incorrectly indicated by 34% of the sonographers. The eye tracking pattern is recorded on the CDT heat map (Figures 5 and 6).

The cancer present image with the highest number of false negatives overall was Difficult 7 (D7), for which 28 out of 29 failed to detect a cancer. This lesion had the pathological diagnosis of an 8 mm ductal carcinoma in-situ (DCIS). The CDT heat map for this image indicates that although the sonographers fixated on the lesion (red in colour) only one indicated a true positive (Figures 7 and 8).

## Discussion

Using established observer performance measures, these results indicate that for the sample of sonographers their clinical performance as measured by ROC, JAFROC, sensitivity and specificity was high. In comparison, studies on breast radiologists have a reported JAFROC score of .79.<sup>12</sup> There may be a range of reasons for this outcome. The current sample were predominantly experienced and comprised of accredited medical sonographers

trained in breast sonography. This training has facilitated fine-tuning of their perceptual and cognitive processes to detect breast cancer. Another possible reason is the central location of the majority of the cancers within the image. The sonographers may have implicitly learned location and set up a location expectancy for responses over the course of the study. The images were obtained retrospective to the study design and in their practice sonographers are trained to optimise a pathology (e.g., centre and enlarge) for the radiologist to report. A future study using a further bank of images, where lesions are imaged across a range of locations, may be beneficial.

With regard to the eye-tracking metrics, the patterns relating to three images discussed below suggest that the sonographers' errors were predominantly cognitive or decision-related. Although the sample size was modest ( $n = 4$ ), the eye-tracking data in our study is concordant with previous studies on radiologists' diagnostic errors.<sup>15</sup>

The cancer absent image N11 scored the highest number of false positives. The CDT heat map shows that the sonographers were distracted by, and fixated for, longer periods of time on the darker area rather than the surrounding tissue. This 'mass-like' area represents a fat lobule that is a normal variant found among the glandular tissue of the female breast. Ongoing education about the sonographic appearances of the normal breast with varying features such as fat lobules may help sonographer specificity.

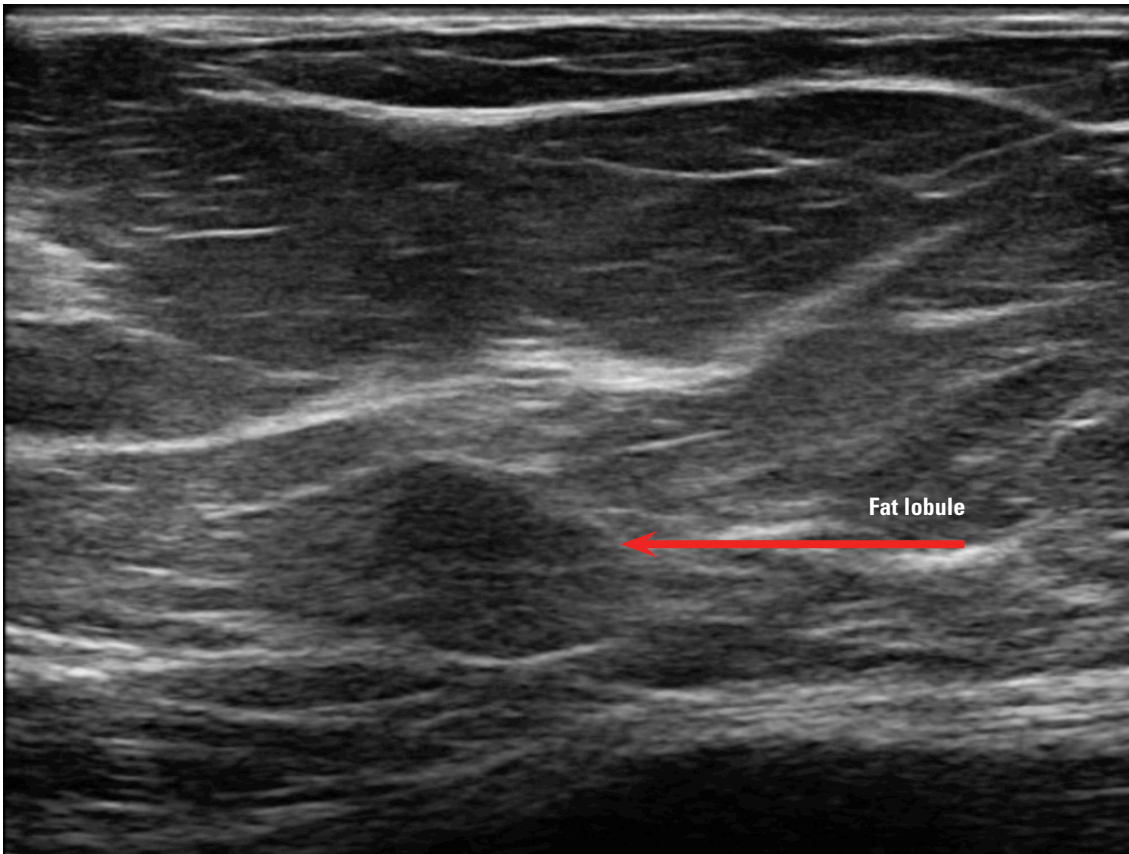
Cancer present image D8 shows more time was spent visually fixating outside the AOI (8.2 seconds), than in the AOI (0.5 seconds) and no sonographers decided this was a TP ( $n = 4$ ). For this sub-set of participants the eye-tracking patterns show that the sonographers were gazing around the cancer looking outside the AOI longer than the cancer and could indicate errors in visual search (perceptual error).

The cancer present image D7 (DCIS), scored the highest number of false negatives across all participants ( $n = 29$ ). The CDT heat map demonstrates total fixation times for four sonographers on D7 and shows that the sonographers dwelled on the lesion but most failed to report it as cancerous, despite knowing there were no benign masses in the test set. These findings are indicative of decision errors, where the sonographers conducted an adequate visual search, recognised an area disrupting normal breast anatomy, and drawing on prior knowledge and experience still made an incorrect decision.

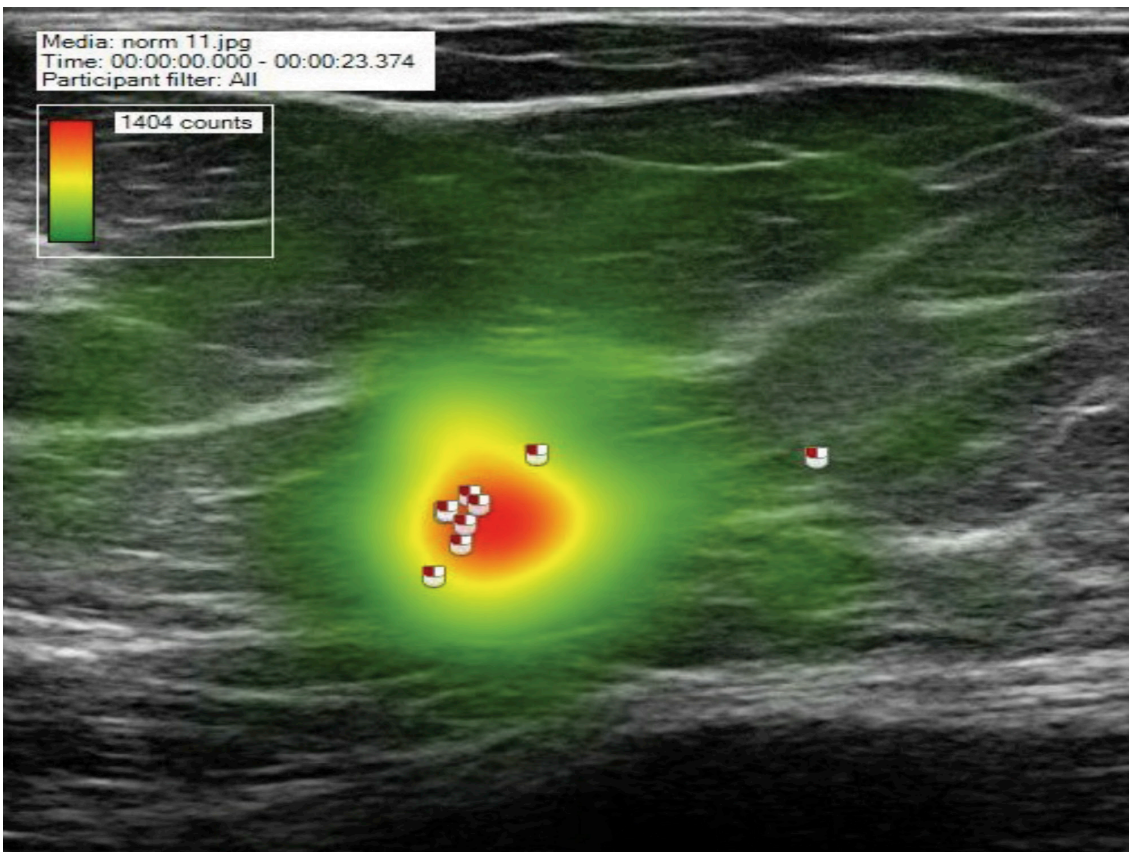
Breast cancer has low prevalence (0.3%),<sup>16</sup> and in low prevalence conditions, miss error rates are known to be high.<sup>17</sup> Ultrasound is not the primary diagnostic tool for DCIS which are commonly detected as micro-calcifications on a mammogram. Given this, it may be that sonographers do not as easily detect such cancers, as they do not expect to find them during the course of an ultrasound, without prior indication.

Sonographer education includes topics such as: breast anatomy, scanning techniques and the sonographic appearances of pathology. Lesion criteria include border integrity, echotexture and dimensions. Specifically, sonographers are familiar at identifying whether the dimensions in the antero-posterior projection are higher than the width of the transverse plane of identified lesions (taller than wider).<sup>18</sup>

DCIS arises in the terminal ductolobular units (TDLU) and



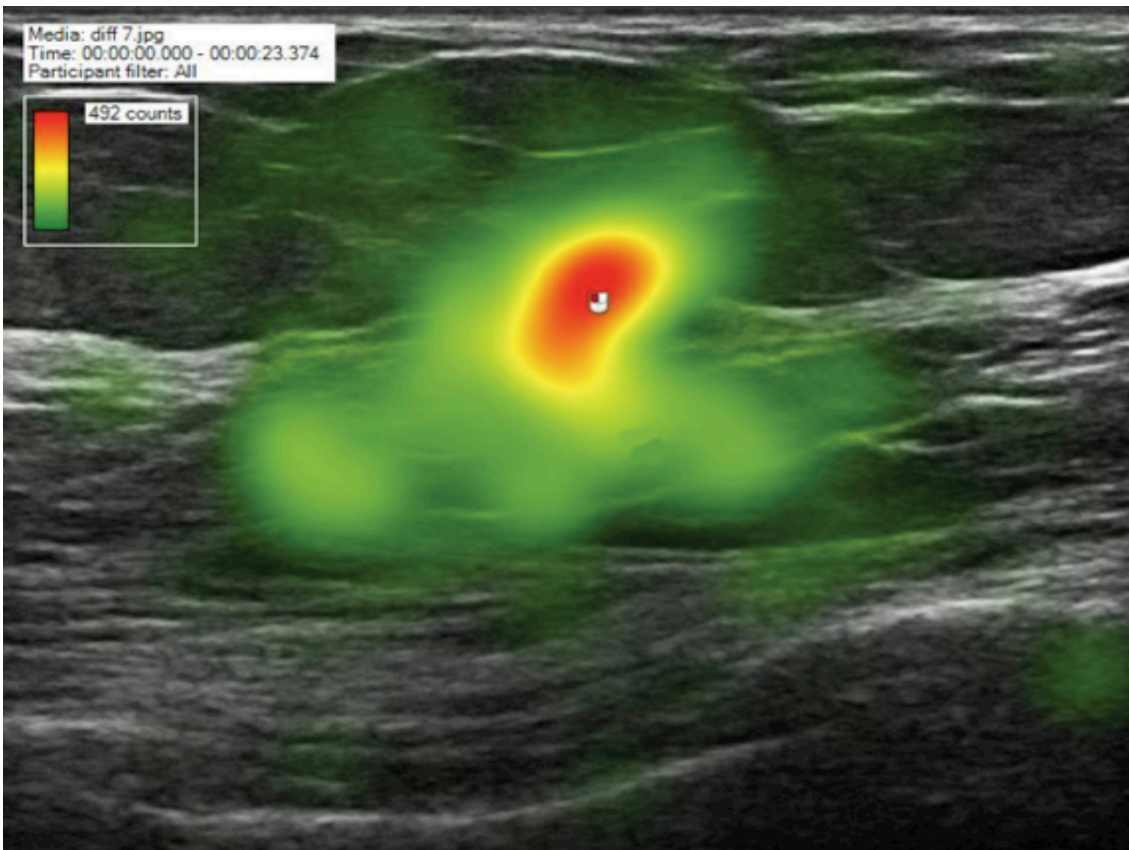
**Figure 5:** Normal Breast 11: Grey Scale.



**Figure 6:** Normal Breast 11: Cumulative Dwell Time heat map  
NB: Marks indicate a 'yes' response.



**Figure 7:** Difficult 7:  
Grey Scale.



**Figure 8:** Difficult 7:  
Cumulative Dwell Time  
heat map.

grows in the terminal duct proximally, parallel to the horizontal orientated TDLU. At no point in development will these be taller than wider so this diagnostic criterion does not apply for DCIS.<sup>18</sup>

Looking to D7, this image does not appear to possess any of the 'typical' sonographic features for a malignancy. However, on closer inspection certain features and 'soft signs' for a DCIS can be identified as the TDLU is located in the horizontal plane and the microlobulations of the lesion can be seen extending into the duct. Increasing sonographer sensitivity for the soft signs of DCIS is likely to increase false-positives and decrease positive predictive value.<sup>18</sup> Nevertheless, sonographic detection, localisation and obtaining the dimensions of a DCIS, along with mammographic correlation can improve patient care (e.g., ultrasound guided biopsy and hook wire placement is considered to provide higher patient comfort and does not require ionising radiation). Therefore, ongoing sonographer education which includes understanding the pathogenesis of DCIS and breast scanning techniques (in the radial and anti-radial planes in addition to longitudinal and transverse planes) is vital in order to detect ductal pathology.

These results also highlight an area for possible targeted education. It may be beneficial to provide retraining with feedback for sonographers, perhaps during a morning viewing session prior to their start of work. This has been shown to be beneficial for airport security workers searching for weapons in luggage.<sup>17</sup> A training task such as this, could not only apply to ultrasound, but also any diagnostic modality that involves visual search.

## Conclusion

The current study presents the first known data collected on medical sonographers and has provided a 'snapshot' of sonographer clinical performance and visual search behaviours. The rich data obtained from the eye-tracking recordings enabled the investigation of the cognitive and perceptual errors that sonographers are at risk of. The study has highlighted relevant and clinically important findings, which is crucial for sonographer education. It has provided the groundwork for future research on sonographers who are an integral component of medical imaging departments. Learning about and improving observer performance is the common goal among medical perception researchers, with the overall aim being to reduce errors, advance diagnostic accuracy and provide positive patient outcomes.

## Acknowledgements

The authors would like to thank the sonographers who gave up their time to participate in this research and the radiology practices and hospitals for allowing us to conduct the study on their premises. Also to Warren Reed for his help with study design.

This paper is a component of an empirical thesis that was submitted in partial fulfilment of the requirements for the degree of Bachelor of Science, Psychology (Honours), Macquarie University, 2013.

## References

- 1 Wolfe JM, Horowitz TS, Kenner NM. Rare items often missed in visual searches. *Nature* 2005; 435: 439–40.
- 2 Lopez N, Previc FH, Fischer J, Heitz RP, Engle RW. Effects of sleep deprivation on cognitive performance by United States Air Force pilots. *J Appl Res Mem Cogn* 2012; 1: 27–33.
- 3 Drew T, Vo ML, Wolfe JM. The invisible gorilla strikes again: Sustained inattentive blindness in expert observers. *Psychol Sci* 2013; 20: 1–6.
- 4 Berlin L. Errors of omission. *AJR Am J Roentgenol* 2005; 185: 1416–21.
- 5 Brem R, Baum J, Lechner M, Kaplan S, Souders S, Naul G, *et al.* Improvement in sensitivity of screening mammography with computer-aided detection: A multi-institutional trial. *AJR Am J Roentgenol* 2003; 181: 687–93.
- 6 Nodine CF, Mello-Thoms C. The role of expertise in radiologic image interpretation. In E. Samei & E. Krupinski (Eds.), *The handbook of medical image perception and techniques* (pp. 139–156). Cambridge: Cambridge University Press; 2010.
- 7 Kundel H, Nodine C, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 1978; 13: 175–81.
- 8 Nodine C, Kundel H. Using eye movements to study visual search and to improve tumor detection. *Radiographics* 1987; 7: 1241–50.
- 9 Krupinski E, Reiner BI. Real-time occupational stress and fatigue measurement in medical imaging practice. *J Digit Imaging* 2012; 25: 319–24.
- 10 Australian Sonographer Accreditation Registry Ltd. 2013.
- 11 Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: Modeling, analysis and validation. *Med Phys* 2004; 31: 2313–30.
- 12 Rawashdeh M, Lee W, Bourne R, Ryan E, Pietrzyk M, Reed WM, *et al.* Markers of good performance in mammography depend on number of annual readings. *Radiology* 2013; 269: 61–67.
- 13 Reed WM, Ryan JT, McEntee MF, Evanoff MG, Brennan PC. The effect of abnormality-prevalence expectation on expert observer performance and visual search. *Radiology* 2011; 258: 938–43.
- 14 Chakraborty DP. Analysis of location specific observer performance data: validated extensions of the jackknife free-response (JAFROC) method. *Acad Radiol* 2006; 13: 1187–93.
- 15 Krupinski EA. Current perspectives in medical image perception. *Atten Percept Psychophys* 2010; 72: 1205–17.
- 16 Gur D, Sumkin JH, Rockette HE, Ganott M, Hakim C, Hardesty L, *et al.* Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *JNCI Journal of the National Cancer Institute* 2004; 96: 185–90.
- 17 Wolfe JM, Horowitz TS, Van Wert MJ, Kenner NM, Place SS, Kibbi N. Low target prevalence is a stubborn source of errors in visual search tasks. *J Exp Psychol Gen* 2007; 136: 623–38.
- 18 Stavros AT. *Breast Ultrasound*. Philadelphia, PA: Lippincott, Williams & Wilkins; 2004.