The effect of a digital training tool to aid chest image interpretation: Hybridising eye tracking technology and a decision support tool

Keywords; chest, image interpretation, digital training tool, education

Introduction

Plain radiograph interpretation has been employed for over a century to detect and localise pathologies within the body[1]. If these pathologies were not recognised and treated, the consequences to the individual could be fatal or life inhibiting[2-4]. Limiting the errors made in detecting such diseases and pathologies can reduce time delays to patient care and improve patient outcomes. The previously mentioned can only be achieved through improving the ability of those interpreting the images to enhance the imaging service provided or indeed by enhancing the capabilities of artificially intelligent systems [5]. Various systems and devices have been tested for their effectiveness in the training of advanced healthcare practitioners and medical staff[6-13]. Eye tracking has been used to help understand the process of image interpretation and secondly to assess and provide feedback/training on the interpretation process[14,15]. Eye tracking feedback provided to participants has been shown to have a positive effect on interpretive performance[14,15].

Checklists have proven to be a valuable resource for image reporting within healthcare settings[16]. Submission of a digital checklist, formulated by Wang et al.[17], automatically generated a standardised image report, ensuring all aspects of an image were considered within the image interpretation process. Likewise, a search strategy reinforced by an interactive checklist to be used for image interpretation, is a method that can ensure all aspects of the image have been inspected to detect abnormal features [18]. The aforementioned study findings provided the rationale for the development and content of the chest image interpretation digital training tool by our research team[19,20]. The tool formed by the research team incorporates a checklist and search strategy, encouraging the user to search the image systematically to form a diagnosis. The tool generates an image report from the responses provided by the user. This study was carried out to test the effect of this digital training tool on participant performance.

Methodology

Ethical approval was obtained from the Nursing and Health Research Ethics Filter Committee in Ulster University. Written informed consent was received from participants before undertaking the study.

Study Design

A study was carried out over a nine month period with reporting radiographers (RRs). Participants completed an initial assessment at recruitment and were asked to re-attend nine months later for a follow-up assessment. The intervention group were given unlimited access to the training tool during the nine month period. The control group did not have access to the tool during this time.

During the assessments, each participant's diagnosis was voice recorded. The participants stated their confidence level on a scale between one and ten with each diagnosis (where 10 = very confident). In addition, a questionnaire on radiographers' clinical experience was completed. Participants completed the assessment of 20 images using eye tracking technology to enable eye gaze metrics to be collected during the image interpretation session.

Participant groups

A convenience sample was used to recruit radiographers who were registered on a postgraduate chest image interpretation programme. Participants were allocated to the control or intervention group depending on their postgraduate programme start date, to limit sharing of the intervention between participants. Participants were not all in the same cohort, they were recruited at the beginning of the postgraduate programme, the postgraduate programme was delivered by one institution at different points in the year. Another group of participants recruited were radiographers trained to report on images of the musculoskeletal system, but with no formal training in chest image interpretation. These participants were randomly allocated to a control group or intervention group.

Digital training tool

A tool was developed to include; A) a search strategy approach and B) an educational tool[17].

A). Search strategy training tool

A combination of open and closed questions, diagrams and guidelines pointed out areas on the image that the observer should interpret to complete the checklists within the pro-forma.

The search strategy comprised six sections, which focused on different anatomy, pathologies and artefacts which may be present within the image; (1) general image considerations, (2) tubes/lines/devices, (3) bony thorax, soft tissues, (4) diaphragm/heart/mediastinum, (5) lung zones and (6) lung shadows.

B). Educational programme

The educational tool consisted of videos comprised of radiologist/reporting radiographer eye gazes and eye movement paths recorded during chest image interpretation and collected whilst the expert used the search strategy training tool. Participant eye gaze behaviours were recorded.

Use of the tool

The tool was to be used when practicing chest image interpretation. It was envisaged that prolonged use of the online checklist when practicing image interpretation would lead to the search strategy becoming second nature to the image interpreter. It was anticipated that the search strategy would then form the structure of their chest image interpretation method. All intervention participants were encouraged to use the tool frequently.

Images

Postero-anterior (PA) and Antero-posterior (AP) chest images were sourced from an image repository previously used in research[21]. There was a prevalence of abnormal images of 45% in the initial assessment, 45% in the training tool and 50% in the follow up assessment, these reflect proportions used in previous studies[14, 21, 22]. Each of the images within the test bank were interpreted by either two or three consultant radiologists and a consultant reporting radiographer and the diagnosis was consensually agreed. Images included a range of chest pathologies including but not limited to; lung nodules, atelectasis, consolidation and pneumothorax. Similar pathologies were included in the initial and follow up assessment to ensure both interpretation tasks were a similar level of difficulty but without the possibility of memory influencing the interpretations.

Equipment

The Tobii Studio X60 eye tracker and the Tobii studio software© were utilised for data collection and for computing eye gaze metrics[23]. The remote non-intrusive eye tracker collected the data without interference to the participant's interpretation. The eye tracker was positioned inferior to the high resolution (1440px x 900px) 24" LCD monitor that displayed the images and angled upwards (approximately 30° cranially) to align with the participant's gaze.

Data analysis

Descriptive statistics included the mean, median and standard deviation. Analysis of Variance (ANOVA) and Analysis of Covariance (ANCOVA) as parametric testing for statistical significance were used.

Eye tracking data collected

The following eye gaze metrics were computed:

- Fixation duration: Measure of the sum of the duration for all fixations within a defined area of interest (AOI). Areas of interest were identified on the eye tracking software as the Area of Pathology(ies) within an abnormal image.
- Fixation count: Measure of the number of times the participant fixated on an AOI.
- Time to first fixation: Measure of how long it took before a test participant fixated on an AOI
- Visit duration: Measure of the duration of all visits within an AOI.
- Visit count: Measure of the number of visits within an AOI.

Heat maps were extracted using the eye tracking technology software. All images were overlaid with a heat map. Green areas represent areas of low fixation counts and red areas represent areas of high fixation counts.

Quality scores

Quality scores were awarded to participants for mentioning additional information regarding the quality of each image. Additional information could include; positioning errors (for example, a lordotic appearance), areas of abnormality, artefacts, lines/tubes/devices, normal variants or image features (for example raised diaphragm, enlarged heart etc.) which may look abnormal but were not the pathological abnormality associated with the chest image. A series of possible quality scores were determined following a discussion with the research team. Participants were then scored using these agreed criteria. Quality scores were awarded to participants to appreciate relevant information which did not contribute to the participant scores. The quality score highlighted how broadly the participant interacted with the image to demonstrate how image quality may impact on subsequent interpretation**.**

Eye tracking sampling quality

Eye tracking sampling quality was extracted and analysed for the participants in this study. No significant difference was noted in the initial and follow up eye tracking sampling quality.

Results

Demographics

35 radiographers took part in this study and a total of 1400 images were interpreted. The control group were qualified as radiographers for longer and had greater/more experience interpreting images as radiographers (Table 1). However, both the control and intervention group had similar experience in reporting images of approximately 6-7 years (Table 1). There were no significant differences between groups in terms of demographic data.

Scores

Diagnostic scores improved to a greater extent within the intervention group compared to the control group along with confidence in this diagnosis. True positive (TP) scores increased for intervention group (2.47 to 4.40) ($p<0.05$) and False Positive (FP) scores decreased for intervention group ($p<0.05$), from 5.87 to 3.27. In addition, True Negative (TN) scores increased, from 5.13 to 6.73 for the intervention group ($p<0.05$). Quality scores decreased for the control and the intervention group at the follow up testing ($p<0.05$), from 13.2 to 10.3 for the control group and from 13.2 to 9.67 for the intervention group (Table 2). Confidence increased significantly for the intervention group only ($P<0.05$) (Table 2).

Further subgroup analysis revealed that the improvements in performance and confidence observed were evident for the intervention groups of both MSK and CXR radiographers compared to the control groups.

Subgroup analysis of the MSK RR participants revealed that TP, TN scores increased significantly for the intervention group only ($P<0.05$). This was a TN mean difference of 11.14% for the control group and a TN mean difference of 21.36% for the intervention group (Table 3). Confidence in their diagnosis increased by 6.7% ($P<0.05$) in the intervention group only (Table 3).

In the CXR RR subgroup analysis, the intervention group demonstrated fewer FP values (19.27% decrease) and a greater number of TP values (9.11% increase) following the intervention period ($p<0.05$) which was not apparent in the control group (Table 4). Both groups were more confident in the diagnoses they provided at the end of the study, this was more apparent in the intervention group, with an increase of 16.3%, compared to the 3.4% rise observed within the control group, however these changes were not significant ($P>0.05$) (Table 4).

Eye tracking data

Fixation count and visit duration on the entire image decreased for the control group following the intervention period. Interpretation time for this group also decreased significantly (Table 2).

The MSK intervention group spent longer fixating, took longer to fixate, fixated a greater number of times and visited the area of pathology for longer or more often than the MSK control group, but these differences were not significant (Table 5).

For CXR RRs, mean fixation duration and mean fixation count significantly decreased for both groups following the intervention period ($p<0.05$) (Table 6). Decision time amongst the CXR reporting radiographers increased for the intervention group participants ($p<0.05$) (Table 6).

Heat maps

Figure 7(a) and Figure 7(b) are the same normal image as interpreted by a representative participant from the intervention group and control group (post intervention period) respectively. Figure 7(c) and Figure 7(d) are the same abnormal image as interpreted by a representative participant from the intervention group and control group respectively. Both the control and intervention group participants demonstrate areas of high fixation counts in the right middle zone, where the lung nodule is present.

Discussion

The improvement in performance observed in this study would suggest that the training tool developed from this research has the potential to greatly aid and improve chest image interpretation for reporting clinicians. These improvements in performance were similar to those identified by other studies employing educational strategies[21,24].

The increased TP and TN scores seen within the MSK RR intervention group may be a reflection of the value of using the tool given that this change was not seen in the MSK control group[6]. The improvement in performance mimics that observed following chest image interpretation training courses[15]. Therefore, engaging with this training tool may provide a less expensive and quicker means of educating radiographers in chest image interpretation roles[14]. Being a quick and cost effective method to aid chest image interpretation training and skills, the importance of such strategies are particularly relevant during the recent COVID-19 pandemic[25].The improvements in performance and confidence were most evident in the intervention group of both MSK radiographers.

Implementation of the tool may have led to the longer interpretation times and increased number of eye gaze metrics observed in the intervention group as they tried to follow the search strategy. Longer interpretation times were also reported in McLaughlin et al.[26] by the experienced reporting radiographer group which searched the entire image for pathology.

An increase in decision time was also reported by Litchfield et al.[22] where the type of training provided determined whether the decision time changed. In this study, images were already reported on and the patient had received their diagnosis, therefore patient treatment or care was not influenced by the participant's decisions. However, the utility function feature, whereby participants may take longer to decide on a diagnosis when there is a consequence of an incorrect diagnosis, may have influenced participant decision time[27,28].

Quality scores tended to decrease in the follow up period, especially for intervention groups. Perhaps participants did not raise issues of concern in the image because they were more confident in identifying abnormalities, therefore any image quality issues did not impact on the image interpretation process. This would reflect the increase in confidence seen also.

All participants were quicker during the follow up assessments compared to the initial testing period. Increased familiarity with the task at the follow up period may have contributed to this. In addition, different images were used in the follow up to prevent a memory effect through prior image recognition which otherwise may have contributed to this reduction in time[29]. Therefore, the training tool has the potential to improve interpretation performance however the time implications of its use on interpretive speed must also be taken into consideration. It is opined that over time, as participants become accustomed to using the search strategy and tool in practice, this slight delay may reduce and become obsolete.

In general, the high fixation count areas tended to be more widespread for the intervention group participants in the follow up testing period, this was also reflected in the heat maps. This is likely to be an effect of the training tool which aims to encourage users to search the entire image to exclude pathologies. It is opined that the questions asked in the search strategy and the answer detail required for the questions contributed to the extra attention given to the areas of pathology by the intervention group and therefore resulted in the additional time taken to reach a diagnosis. The systematic approach would be an important counter to satisfaction of search where multiple pathologies exist and may be missed[18]. The training tool emphasised a checklist approach and systematic search of the image. There was an increase in eye tracking data at critical areas identified in the training tool in the post-intervention cohort, demonstrating that when users are made aware of these search strategies and review areas, they adopt their own search pattern to include aspects of the training tool approach. This is a useful educational tool to encourage a more systematic approach in image interpretation.

Limitations

The methodology reflects studies which have been completed previously[21], therefore increasing the reliability of the methodology used within this study. However, a relatively small sample size was used and therefore the results must be used with caution.

Two participants were unable to complete the follow up data collections due to restricted schedules. The participant data was excluded from analysis.

Participant engagement with the training tool is not reported here and this may have varied greatly across participants, impacting study results. Further investigation exploring extent of engagement and subsequent effect on outcomes should be considered.

The monitor used would be inferior to clinical reporting workstations, however the monitor was taken to data collection sites and allowed standardisation across participants for their viewing environment.

Conclusion

Within the current study, the implementation of the digital training tool provided improvements in participant performance. The introduction of the tool may cause longer decision times but could lead to an overall cost effective and clinically safe service in reporting radiography. The use of pathology and search check-boxes, present in the tool, are of benefit to those who interpret images. Decision times are likely be reduced with use of the tool over a prolonged time period as staff become familiar with it and it is iteratively refined and optimized by the individual. The reported tool currently has the potential to ensure satisfaction of search errors are reduced and full image interrogation is achieved, to cost effectively aid chest image interpretation training.

Research can be expanded in this area; the use of the tool can be investigated in other professions and the tool can be developed to be used in the interpretation of images of different anatomical areas.

References

1 Thom SE. Does advanced practice in radiography benefit the healthcare system? A literature review. Radiography. 2018 Feb 1;24(1):84-9.

2 Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. The Ulster medical journal. 2012 Jan;81(1):3.

3 Krupinski EA. The importance of perception research in medical imaging. Radiation Medicine 2000; 18 (6), 329-334.

4 Krupinski EA, Tillack AA, Richter L, Henderson JT, Bhattacharyya AK, Scott KM, Graham AR, Descour MR, Davis JR. and Weinstein RS, Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. Human pathology 2006; 37 (12), 1543-1556.

5 McConnell J, Devaney C, Gordon M, Goodwin M, Strahan R. and Baird M. The impact of a pilot education programme on Queensland radiographer abnormality description of adult appendicular musculo-skeletal trauma. Radiography 2012; 18 (3), 184-190.

6 Auffermann WF, Henry T.S, Little BP, Tigges S. and Tridandapani S. Simulation for Teaching and Assessment of Nodule Perception on Chest Radiography in Nonradiology Health Care Trainees. Journal of the American College of Radiology 2015; 12 (11), 1215-1222.

7 Semakula Katende NS, Andronikou S. and Lucas, S. Digital platform for improving non-radiologists' and radiologists' interpretation of chest radiographs for suspected tuberculosis - a

method for supporting task-shifting in developing countries. Pediatric radiology 2016; 46 (10), 1384-1391.

8 Murphy A, Ekpo E, Steffens T, Neep MJ. Radiographic image interpretation by Australian radiographers: a systematic review. Journal of medical radiation sciences. 2019 Dec;66(4):269-83.

9 Williams I, Baird M, Pearce B, Schneider M. Improvement of radiographer commenting accuracy of the appendicular skeleton following a short course in plain radiography image interpretation: a pilot study. Journal of medical radiation sciences. 2019 Mar;66(1):14-9.

10 Neep MJ, Steffens T, Eastgate P, McPhail SM. Evaluating the effectiveness of intensive versus non-intensive image interpretation education for radiographers: a randomised controlled trial. Journal of medical radiation sciences. 2019 Mar;66(1):5-13.

11 Van Der Gijp A, Ravesloot CJ, Jarodzka H, Van der Schaaf MF, Van der Schaaf IC, van Schaik JP, Ten Cate TJ. How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology. Advances in Health Sciences Education. 2017 Aug;22(3):765-87.

12 Subesinghe M, Goldstone AR, Patel CN, Chowdhury FU. and Scarsbrook AF. Design and implementation of a web-based PET-CT reporting assessment and e-portfolio tool. Clinical Radiology 2015; 70(2), 123-127.

13 Wright C. and Reeves P. Image interpretation performance: A longitudinal study from novice to professional. Radiography 2017; 23 (1), e1-e7.

14 Donovan T, Manning DJ. and Crawford T, Performance changes in lung nodule detection following perceptual feedback of eye movements. In: Medical Imaging. International Society for Optics and Photonics, 2008; 691703-691703-9.

15 Kundel HL, Nodine CF. and Krupinski EA. Computer-displayed eye position as a visual aid to pulmonary nodule interpretation. Investigative radiology, 1990; 25 (8), 890-896.

16 Kramer HS. and Drews FA. Checking the lists: A systematic review of electronic checklist use in health care. Journal of Biomedical Informatics 2016; Available from: http://www.sciencedirect.com/science/article/pii/S1532046416301186

17 Wang KC, Jeanmenne A, Weber GM, Thawait S. and Carrino JA, An online evidence-based decision support system for distinguishing benign from malignant vertebral compression fractures by magnetic resonance imaging feature analysis. Journal of Digital Imaging, 2011; 24 (3), 507-515.

18 Kok EM, Jarodzka H, de Bruin ABH, BinAmir HAN, Robben SGF. and van Merriënboer JJG, Systematic viewing in radiology: seeing more, missing less? Advances in Health Sciences Education, 2015; 189-205.

19 McLaughlin L, McConnell J, McFadden, S, Bond R. and Hughes C. Methods employed for chest radiograph interpretation education for radiographers: A systematic review of the literature. Radiography 2017; 350-357. Available from: https://www.sciencedirect.com/science/article/pii/S1078817417301281

20 Blinded

21 Woznitza N, Piper K, Burke S, Patel K, Amin S, Grayson K. and Bothamley, G. Adult chest radiograph reporting by radiographers: Preliminary data from an in-house audit programme. Radiography, 2014; 20 (3), 223-229.

22 Litchfield D, Ball LJ, Donovan T, Manning DJ and Crawford T. Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. Journal of Experimental Psychology Applied, 2010; 16 (3), 251-262.

23 Tobii Pro, Envision human behaviour, [online]. Available: http://www.tobiipro.com/ ,2016.

24 Sonnex E, Tasker A. and Coulden R. The role of preliminary interpretation of chest radiographs by radiographers in the management of acute medical problems within a cardiothoracic centre. The British Journal of Radiology, 2001; 74(879), 230-233.

25 Wong HY, Lam HY, Fong AH, Leung ST, Chin TW, Lo CS, Lui MM, Lee JC, Chiu KW, Chung T, Lee EY. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. Radiology. 2020 Mar 27:201160.

26 McLaughlin L, Bond R, Hughes C, McConnell J. and McFadden S. Computing eye gaze metrics for the automatic assessment of radiographer performance during X-ray image interpretation. The International Journal of Medical Informatics, 2017; 11-21. Available from: http://www.sciencedirect.com/science/article/pii/S1386505617300540.

27 McConnell J, Eyres R, Nightingale J. Interpreting trauma radiographs. Blackwell; 2005 Aug 19.

28 Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. In Handbook of the fundamentals of financial decision making: Part I 2013 (pp. 99-127).

29 Ryan, J.T., Haygood, T.M., Yamal, J., Evanoff, M., O'Sullivan, P., McEntee, M. and Brennan, P.C. The "memory effect" for repeated radiologic observations. American Journal of Roentgenology, 2011; 197 (6), W985-W991.

Tables and Figures

Table 1: Descriptive statistics of reporting radiographers in the control and intervention group

	Control (n=20)	**Intervention (n=15)**
Qualified (years)	22.18 ± 10.28	14.97 ± 7.22
Experience interpreting images (years)	21.68 ± 9.65	13.90 ± 6.48
Experience reporting images (years)	6.70 ± 5.77	6.03 ± 5.41
Sex	Female 55% Male 45%	Female 86.7% Male 13.3%

*Mean ± standard deviation. *significant difference within group from pre to post intervention ~significant difference between groups post intervention (P<0.05)*

Table 2: Scores, confidence and eye tracking data of AOP for intervention and control groups

	Control pre (n=20)	Control post (n=20)	Intervention pre (n=15)	Intervention post (n=15)
False Positives	4.20 ± 1.70# (38.18%)	3.20 ± 1.96 (32.0%)	5.87 ± 1.96# (53.36%)	3.27 ± 1.62* (32.7%)
False Negatives	5.95 ± 1.73 (66.11%)	5.60 ± 1.47 (56.0%)	6.53 ± 1.25 (72.56%)	6.00 ± 1.25 (60.0%)
True Positives	3.05 ± 1.73 (33.89%)	4.40 ± 1.47* (44.0%)	2.47 ± 1.25 (27.44%)	4.40 ± 1.47* (44.0%)
True Negatives	6.80 ± 1.70# (61.82%)	4.40 ± 1.47 (44.0%)	5.13 ± 1.96# (46.64%)	6.73 ± 1.62* (67.30%)
Quality scores	13.20 ± 2.80 (44.0%)	10.30 ± 2.89* (46.82%)	13.20 ± 2.96 (44.0%)	9.67 ± 2.29* (43.95%)
Confidence	6.17 ± 1.23 (61.7%)	6.43 ± 1.20 (64.3%)	5.32 ± 1.70 (53.2%)	6.31 ± 1.07* (63.1%)
Mean fixation duration (secs)	8.88 ± 5.95	7.93 ± 4.70	7.02 ± 4.65	7.95 ± 5.62
Mean time to first fixate (secs)	7.86 ± 5.03	8.94 ± 8.07	13.34 ± 15.50	11.50 ± 15.91
Mean fixation count (n)	27.82 ± 15.97	27.40 ± 14.49	23.30 ± 13.36	31.76 ± 20.29
Mean visit duration (secs)	10.10 ± 6.21	9.81 ± 5.00	8.49 ± 4.41	10.72 ± 6.13
Mean visit count (n)	11.78 ± 5.91	9.09 ± 4.85	11.04 ± 6.36	11.44 ± 7.14
Mean decision time (secs)	66.19 ± 17.00	41.75 ± 27.65*	68.55 ± 22.33	54.51 ± 40.02

*Data is presented as mean ± standard deviation #significant difference between groups at baseline *significant difference within group from pre to post intervention ~significant difference between groups post intervention (P<0.05)*

Table 3: Scores and confidence of reporting radiographers trained in MSK image interpretation

	Control pre (n=13)	Control post (n=13)	Intervention pre (n=10)	Intervention post (n=10)
FP	4.69 ± 1.80 (42.64%)	3.00 ± 1.16* (30.0%)	6.20 ± 2.30 (56.36%)	3.50 ± 1.79* (35.0%)
FN	6.31 ± 1.80 (70.11%)	6.08 ± 1.32 (60.8%)	6.60 ± 1.43 (73.33%)	5.90 ± 1.52 (59.0%)
TP	2.69 ± 1.80 (29.89%)	3.92 ± 1.32 (39.2%)	2.40 ± 1.43 (26.67%)	4.10 ± 1.52* (41.0%)
TN	6.31 ± 1.80 (57.36%)	6.85 ± 1.35 (68.5%)	4.80 ± 2.30 (43.64%)	6.50 ± 1.78* (65.0%)
Quality scores	12.31 ± 2.32 (41.03%)	9.31 ± 2.69* (42.32%)	12.10 ± 2.28 (40.33%)	8.60 ± 1.65* (39.09%)
Confidence	5.63 ± 1.02 (56.3%)	5.84 ± 1.06 (58.4%)	5.60 ± 1.61 (56.0%)	6.27 ± 1.13* (62.7%)

*Data is presented as mean ± standard deviation *significant difference within group*

Table 4: Scores and confidence of reporting radiographers' training in chest image interpretation pre and post study

	Control pre (n=7)	**Control post (n=7)**	**Intervention pre (n=5)**	**Intervention post (n=5)**
FP	3.29 ± 1.11 (29.91%)	3.57 ± 3.05 (35.7%)	5.20 ± 0.84 (47.27%)	2.80 ± 1.30* (28.0%)
FN	5.29 ± 1.50# (58.78%)	4.71 ± 1.38~ (47.1%)	6.40 ± 0.89# (71.11%)	6.20 ± 0.48~ (62.0%)
TP	3.71 ±1.50# (41.22%)	5.29 ± 1.38~ (52.9%)	2.60 ± 0.89# (28.89%)	3.80 ± 0.48*~ (38.0%)
TN	7.71 ± 1.11# (70.09%)	6.43 ± 3.05 (64.3%)	5.80 ± 0.84# (52.73%)	7.20 ± 1.30 (72.0%)
Quality scores	14.86 ± 3.02 (49.53%)	12.14 ± 2.41* (55.18%)	15.4 ± 3.13 (51.33%)	11.80 ± 1.92 (53.64%)
Confidence	7.18 ± 0.95 # (71.8%)	7.52 ± 0.42 # (75.2%)	4.78 ± 1.91 # (47.8%)	6.41 ± 1.05 # (64.1%)

*Data is presented as mean ± standard deviation *significant difference within group (P<0.05) # significant difference between groups at baseline ~significant difference between groups post intervention (P<0.05)*

Table 5: Eye tracking data for the Areas of Pathology (AOP) of reporting radiographers trained in MSK image interpretation

	Control pre (n=13)	Control post (n=13)	Intervention pre (n=10)	Intervention post (n=10)
Mean fixation duration (secs)	7.37 ± 4.00	8.25 ± 5.39	6.57 ± 5.00	9.34 ± 4.76
Mean time to first fixate (secs)	7.62 ± 4.91	7.75 ± 6.86	13.36 ± 18.82	11.53 ± 18.88
Mean fixation count (n)	23.76 ± 12.25	28.65 ±19.66	22.59 ± 14.17	35.66 ± 18.87
Mean visit duration (secs)	8.51 ± 4.11	10.08 ± 5.83	8.19 ± 4.78	11.44 ± 5.82
Mean visit count (secs)	10.64 ± 5.33	8.76 ± 5.37	10.48 ± 6.65	12.60 ± 6.83
Mean decision time (secs)	70.17 ± 16.37	55.87 ± 24.04	68.97 ± 25.39	72.63 ± 37.16

Data is presented as mean ± standard deviation.

Table 6: Eye tracking data for the Areas of Pathology (AOP) of reporting radiographers training in chest image interpretation

	Control pre (n=7)	**Control post (n=7)**	**Intervention pre (n=5)**	**Intervention post (n=5)**
Mean fixation duration (secs)	11.69 ± 8.12	7.33 ± 3.36	7.93 ± 4.24	5.16 ± 6.70
Mean time to first fixate (secs)	8.32 ± 5.61	11.15 ± 10.17	13.28 ± 6.62	11.44 ± 9.14
Mean fixation count (n)	35.38 ± 20.13	25.08 ± 8.85	24.74 ± 13.02	23.95 ± 22.91
Mean visit duration (secs)	13.06 ± 8.53	9.29 ± 3.11	9.09 ± 4.01	9.30 ± 7.17
Mean visit count (n)	13.89 ± 6.77	9.70 ± 4.03	12.17 ± 6.32	9.12 ± 7.95
Mean decision time (secs)	58.80 ± 16.77	65.81 ± 28.09	67.72 ± 17.12	76.71 ± 18.86*

*Data is presented as mean ± standard deviation. All values are means ± SD. *significant difference within group (P<0.05)*

Figure 7 Integral heat maps showing the density of fixations for participants completing image interpretation of PA chest images.