Introduction

Preliminary clinical evaluation (PCE) was introduced as an abnormality flagging system to improve the communication of positive findings on extremity radiographs, replacing the red dot system, predominantly used in an Emergency Department (ED) setting (1). Previous studies have investigated various elements of the PCE system including the effects of focussed training (2), accuracy (3,4), impact on patient management (5), common errors (6) structure of the written comment (7), and the type of comment preferred (8). The benefits of the PCE system for extremity X-ray examinations are that it reduces interpretation errors (5,9) and increases confidence in decision making by referrers (5). Some studies have described the expansion of PCE in abdominal X-ray examinations (10) and Computed Tomography (CT) colonography (10,11). To date, there have been no studies specifically assessing the possibility of extending the scope of the PCE service to include other body parts that often present with traumatic mechanisms of injury, such as the chest. Recent research has indicated the inclusion of the chest X-ray (CXR) in a PCE system in at least one institution (6,12) though there has been no published assessment of performance showing how, or if, training can improve radiographers' abnormality detection and/or commenting accuracy.

The CXR is the most requested radiographic examination (13) and contributes a large portion of a general radiographer's daily workload. As previously suggested (5), the PCE will likely be most useful in the out-of-hours (OOH) setting in which urgent radiology reports may be provided by off-site reporting agencies or in the following days or weeks. Subsequently, a PCE comment for a CXR examination suspected to have traumatic findings has the potential to reduce mismanagement and expedite correct management and treatment.

This study aims to assess radiographers' ability to localise common traumatic pathologies and to accurately describe the pathology, in pre- and post-training conditions.

Method

This observer study was performed in a district general hospital in the West Midlands region of England, United Kingdom (UK). The Health Research Authority tool (14) did not classify this work as research but local institutional approval was provided. Participants were invited by poster advertisement and in-person discussions. Participants consented to allow their data to be used in this evaluation. Participants were required to localise an abnormality and describe the appearance, to generate a PCE.

Image bank formation

A 3-month survey revealed an average of 31±4.6 (range 26-35) CXR examinations performed in the OOH setting (8pm-8am) for ED patients. This was not considered to be an adequate caseload to provide a meaningful number of positive cases. Subsequently, the image bank caseload was increased to 58 based on the sample size estimations used for observer studies (15). This was considered adequate to keep the image bank at a practicable size to prevent observer fatigue. An earlier survey of trauma X-ray examinations in same centre indicated a 35% abnormality prevalence (5); this was applied to the current study to produce an image bank with 20 positive cases. The caseload was not enriched with positive cases beyond those identified in a prior survey, reducing the risk of overestimating performance (16). Participants were required to identify the precise anatomical location of a suspected abnormality; 10 cases contained bony pathology (rib, clavicle, or humerus fracture or dislocation) and 10 cases contained a pneumothorax. No clinical history was provided for the cases, but participants were informed that all positive cases presented following a traumatic incident and the diagnostic question consisted of "? Fracture/dislocation and/or ? Pneumothorax".

Scoring system

Participants responses were awarded a maximum of three points, based on abnormality recognition and descriptive accuracy; one point for correct physical localisation, one point for correctly describing what the abnormality is, and one point for correctly describing where it is. All image evaluations were completed on a 30" Barco Coronis Fusion 6MP MDCC-6530 LED colour monitor (Barco, Duluth, Georgia, USA). Localisation data were recorded with ROCView (17).

Participants were instructed to localise a pneumothorax by marking the image in a rib space avoiding bony anatomy (Image 1), and to localise a bony abnormality they were informed to accurately mark the bony anatomy. These points helped to determine whether participants have accurately localised the abnormality. Participants were advised to spend no more than two minutes appraising each image to closely replicate the time they would spend reviewing images in clinical practice. Participant localisations of pathology were used to determine scores in a 2x2 matrix (true positive, false positive, true negative, false negative), which allowed calculations of sensitivity, specificity, and accuracy. All participants completed a training dataset to ensure they would complete the task correctly. Participants typed their descriptive answers in to an electronic answer sheet and these were compared to a benchmark PCE prepared by a Consultant Radiographer and corroborated by an Advanced Practitioner Reporting Radiographer with 15 years' experience. The scoring system is defined in Table 1.

Score	0	0.5	1
Localisation	Incorrect, or no localisation(s)	Correct and incorrect localisation(s)	Correct localisation(s) only
PCE Score A	Incorrect type of abnormality	N/A	Correct type of abnormality
PCE Score B	Incorrectly describing location	N/A-	Correctly describing location

Table 1: The scoring system used to assess participants answers

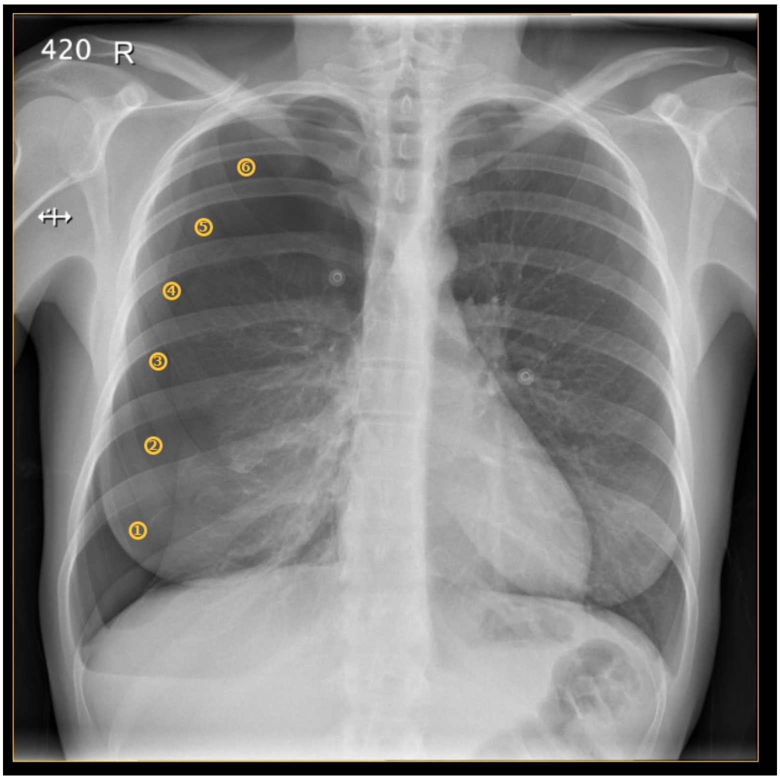


Image 1: Examples of how to mark a pneumothorax for analysis purposes. Participants only had to mark the image once.

Training intervention

Following their initial test, participants were provided with three pre-recorded online video tutorials lasting 30-45 minutes each, on consecutive weeks. The training sessions included an introduction to a systematic search strategy for recognising abnormalities in CXR examination, how to structure a PCE and practice cases to review at their own pace. The training was developed by a Consultant Radiographer. A minimum 6-week wash-out period between pre- and post-training tests was applied to reduce the potential of case memory.

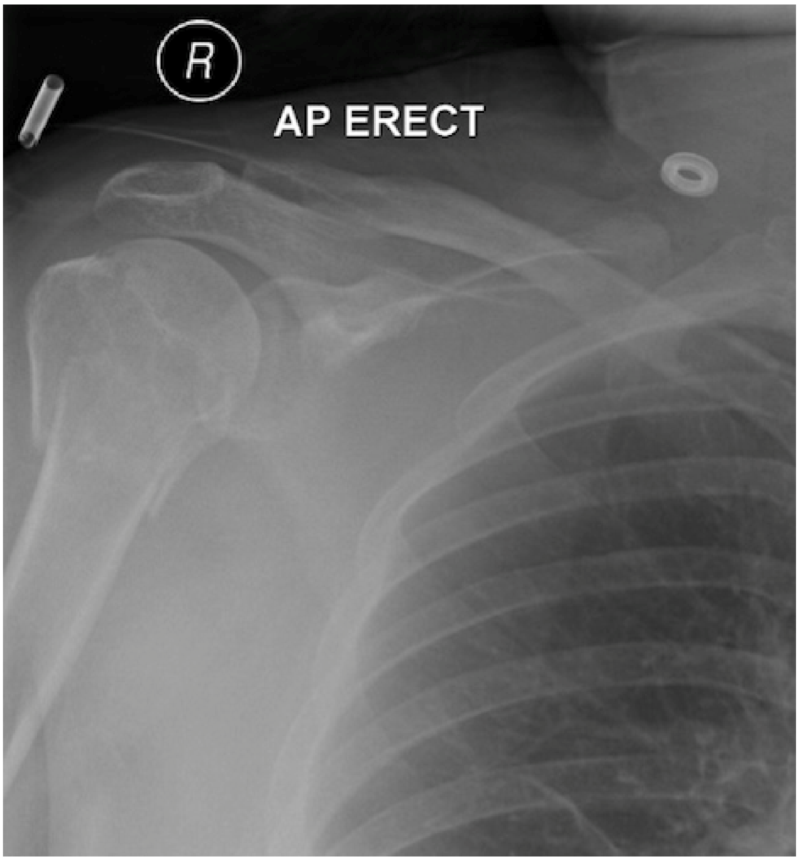
Results

Eleven participants were recruited but only nine completed the study. Two participants could not complete due to a combination of sickness absence and the data collection phase ending. Mean post registration experience was 4.1±5.2 years (range 1-17). The mean time between pre- and post-training evaluations was 84±14 days (range 59-104). Overall, pooled sensitivity remained consistent (78.9% to 78.8%) following training, whereas pooled specificity and accuracy showed moderate improvement, 79.0% to 89.9% and 78.9% to 86.0% respectively. Accuracy in localisation and PCE scores are also improved. Individually, n = 5 participants (55.5%) improved their sensitivity score on their post-training test, n = 6 (66.6%) improved their specificity, and n = 6 (66.6%) improved their accuracy. A summary of pooled and individual sensitivity, specificity and accuracy can be found in Table 2.

	User	A	B	C	D	E	F	G	H	I	Pooled
Pre	Sens. (%)	80.0	85.0	66.7	82.4	85.0	85.0	61.1	80.0	85.0	78.9
	Spec. (%)	68.4	84.2	82.5	51.2	81.6	100.0	75.0	78.9	89.5	79.0
	Acc. (%)	72.4	84.5	77.6	60.3	82.8	94.8	70.7	79.3	87.9	78.9
Post	Sens. (%)	65.0	75.0	88.9	65.0	65.0	90.0	85.7	85.0	90.0	78.8
	Spec. (%)	86.8	76.3	67.5	97.4	100.0	97.4	91.9	97.4	94.7	89.9
	Acc. (%)	79.3	75.9	74.1	86.2	87.9	94.8	89.7	93.1	93.1	86.0

Table 2: Pooled and individual sensitivity, specificity, and accuracy pre- and post-training intervention.

Most participants improved their mean score on the post-test (n = 6, 66.6%). Overall, participants performed better at correctly localising a pneumothorax compared to skeletal abnormalities, in the pre-test (78 vs 55) and the post-test (82 vs 56.5), respectively. There were 10 instances of 0.5 scoring for localisations in the pre-tests, whereby there was a correct localisation and an incorrect localisation, and one instance in the post-tests. The right ACJ dislocation had the greatest difference, + 6, in pooled post-test PCE scores compared to the pre-test. The right surgical neck/humeral head fracture had a negative difference, -5, in pooled post-test PCE scores compared to the pre-test. This case is illustrated in Image 2. A 33% drop was evident in the number of participants correctly describing what this abnormality was, and a 29% drop in the number of participants correctly describing where the abnormality was, in the post-tests. Further breakdown can be seen in tables 3 and 4.



Benchmark	Pre (score)	Post (score)	User
Right Surgical neck/humeral head fracture	Right neck of humerus fracture, no humeral head displacement from the Glenoid cavity (2)	no comment relevant to right humerus (0)	A
	Fracture of right humeral head (2)	no comment relevant to right humerus (0)	E

Image 2. Right surgical neck/humeral head fracture. This case had the greatest drop in pooled PCE scores from the pre-test to the post-test (-5). Examples of inconsistent performance contributing to reduction of PCE scores for the right surgical neck/humeral head fracture case are provided for 2 participants. For each participant they failed to localise the abnormality in the post-test and made no comment relevant to this anatomical area.

Examination groups (number of cases)	Pooled Correct Localisations		
	Pre	Post	TOTAL
Rib #s (3)	6.5	6.5	13
Clavicle # & ACJ (2)	6.5	9	15.5
Humeral head/neck/greater tuberosity #s (5)	42	41	83
Left pneumothorax (3)	23.5	26	49.5
Right pneumothorax (7)	54.5	56	110.5
Overall	**133**	**138.5**	**271.5**

Table 3. Pooled correct localisations for the pre- and post-tests categorised into examination groups. A more detailed table outlining individual performances for each positive case is presented in the supplemental material.

	Pre		**Post**		**Total**		
Abnormality Description	A	B	A	B	Pre	Post	Diff.
Right ACJ dislocation	1	1	4	4	2	8	6
Left Pneumothorax	7	5	8	8	12	16	4
Left 7th/8th rib #	4	3	6	4	7	10	3
Right Pneumothorax	4	4	5	6	8	11	3
Left Pneumothorax	8	7	9	9	15	18	3
Right 3rd rib #	2	0	2	3	2	5	3
Right surgical neck of humerus #	7	7	9	7	14	16	2
Right Pneumothorax	8	8	9	9	16	18	2
Right Pneumothorax	8	8	9	9	16	18	2
Left surgical neck of humerus #	9	5	9	7	14	16	2
Left distal clavicle #	4	5	5	5	9	10	1
Right Pneumothorax	9	8	9	9	17	18	1
Right Pneumothorax	9	8	9	9	17	18	1
Right Pneumothorax	9	8	9	9	17	18	1
Left Pneumothorax	9	8	9	9	17	18	1
Left surgical neck of humerus #	8	5	7	7	13	14	1
Right Pneumothorax	8	7	7	7	15	14	-1
Left 9th rib #	1	1	0	0	2	0	-2
Left humeral head/greater tuberosity #	9	4	6	4	13	10	-3
Right Surgical neck/humeral head #	9	7	6	5	16	11	-5
	133	**109**	**137**	**130**	**242**	**267**	**25**

Key - A = What (abnormality type); B = Where (anatomical location)

Table 4. Participants' Pooled PCE Scores

Discussion

This study suggests that radiographers can contribute to the identification of traumatic CXR findings and there may be some benefit of short and intensive recorded video tutorials to help radiographers develop skills in recognising and describing fractures and pneumothoraces on chest X-ray images. More than half of the participants showed improvement in sensitivity, specificity and/or accuracy scores, and their overall PCE.

Incorrect localisations were far less prevalent following training. This could be a result of the tutorial videos reinforcing the premise of the PCE to focus on the traumatic abnormality rather than being concerned about other appearances, such as heart size or consolidation. The greatest positive difference in PCE scores from the pre- to post-test was seen in the right ACJ dislocation (+6). This may illustrate the benefit of the training in reiterating the importance of comment structure and terminology. The decrease in PCE scores for the right surgical neck/humeral head fracture (-5) indicates a lack of consistency when using anatomical terminology and perception errors as outlined in the examples provided.

Participants performed better at recognising and describing pneumothoraces than the bony abnormalities. British Thoracic Society guidelines (18) state that the depth of a pneumothorax should be determined by the interpleural distance measured at the level of the hilum; small is classed as less than 2cm and large is greater than 2cm. The guidelines in America differ slightly in that depth is determined by the lung apex to cupola distance (19). Interestingly, in this study the pneumothorax cases (n = 3) that had the fewest correct localisations by participants in both tests were those that are classified as small using the British guidelines, and the case with the fewest correct localisations (pre-test, 4 out of 9 and post-test 5 out of 9) had a 1.61cm apex to cupola distance. This suggests that whilst radiographers in this sample can accurately locate large pneumothoraces, there may be need for further training and education to improve detection of apical pneumothoraces, specifically if the pneumothorax is small in volume. The presentation of a pneumothorax can vary from asymptomatic to life-threatening (20), and small pneumothoraces typically resolve with no treatment and only monitoring (21) but this does not detract from the importance of identifying a pneumothorax at the earliest opportunity. The impact of missing a pneumothorax can lead to failure to treat and can have wide-ranging outcomes for the patient (22).

Our results also suggest that bony abnormalities overlying the thorax such as the right 3rd rib fracture, or those that are more subtle like the left 9th rib fracture at the inferior margin of the image, may be more challenging for radiographers to recognise and subsequently would require additional attention with regards to further training sessions. A possible reason for this could be attributed to inattentional blindness (23) whereby observers were over-focussed on the task of looking at the lungs or the large humeral bones that they simply did not see the superimposed or subtler abnormalities. The cases with fractures of the ribs (n = 3) and distal clavicle (n = 1) and acromioclavicular joint dislocation (n = 1) returned the fewest correct localisations amongst the bony abnormalities, whereas those cases with fractures involving the head and surgical neck of humerus (n = 5) returned the most correct localisations. This suggests that further focused training may be required to reinforce observers' search patterns and to reiterate the importance of reviewing these areas when assessing the image.

Whilst our results suggest benefits of focussed training for recognising and describing abnormalities, it is accepted that radiographers may already be providing PCE comments for CXR in some institutions. However, chest examinations account for the biggest proportion of non-participation in a PCE system, accounting for 53% of all instances (12). Interestingly, the study by Alexander-Bates et al (6) also indicated that that traumatic chest examinations had the highest percentage of participants who were unsure or never provided a comment for a CXR examination, possibly due to reduced confidence, knowledge and/or understanding. Additionally, for those who did provide a comment on the CXR examinations, the overall sensitivity and specificity scores were 71% and 99%, respectively, showing excellent ability to recognise normal appearances but with room for improvement regarding describing CXR abnormalities. The study by Alexander-Bates et al (6) used a correct comment as the determinant of accuracy, whereas our study used a correct localisation; however, the sensitivity and specificity scores in our study, of 79% and 90%, respectively, were comparable. This suggests that there might be a need for additional training in CXR abnormality detection for radiographers participating in a PCE system, in which CXR examinations are within the scope of practice.

Previous PCE studies (2,24) using pre and post training image interpretation tests reported increases in both sensitivity, specificity, and accuracy, respectively, which contrasts with the findings in this study in which only specificity and accuracy increased. The study by Stevens & Thompson (2) utilised face-to-face teaching with 56 days between the pre- and post-tests. Williams et al (24) used the same method of delivery of teaching via recorded PowerPoint presentations as our study, with the addition of online content and textbook teaching, and had 112 days between tests, which is greater than the time frame in our study. It is possible that if there was additional supplementary teaching content in our study, to compliment the recorded PowerPoint tutorials, this could have helped to increase the sensitivity from pre- to post-test and provides an interesting consideration for future studies.

Limitations

A small sample size may limit the generalisability of these findings, but they do provide insight into the challenges of implementing a PCE system that incorporates CXR. The mean number of days between tests was longer than originally planned, this was caused by several participants and an author contracting COVID-19, and consequently most post-test dates had to be rescheduled. Due to the extended washout period, it is possible that those participants who had the longest break between the first test and viewing the training material, and the second test may not have retained the knowledge gained, and how this may have impacted on participants' performance and overall findings needs to be acknowledged. The use of pre-recorded tutorials prevented participants from having opportunity to immediately interact to seek clarification as would be the case with face-to-face teaching, and the impact of this should be considered alongside the results. This may also suggest that short, focussed training is not as effective as continued improvement in knowledge over time. However, experiential knowledge, by nature, is very difficult to measure.

It is also accepted that the direct nature of the task, where the participants knew that they were looking to establish no pathology, or either a fracture or pneumothorax may have influenced performance. In addition, co-existing pathology was not considered. However, this type of experimental control is valuable when assessing performance.

Conclusion

Radiographers can contribute to the identification of traumatic CXR findings sufficiently. Improvement in performance was evident in most participants' abnormality localisations and PCE scores, following the training intervention. The pooled results showed increases in specificity and accuracy. The study highlighted areas of CXR PCE provision that require further training, such as detecting superimposed or subtle abnormalities. Further investigation assessing the localisation and description of non-traumatic CXR pathologies is recommended to supplement the results presented here. This study provides additional support to the growing PCE knowledge base, and it is hoped these findings can support the development of future PCE systems.

References

1. The Society and College of Radiographers. Preliminary Clinical Evaluation and Clinical Reporting by Radiographers: Policy and Practice Guidance. 2013.

2. Stevens BJ, Thompson JD. The impact of focused training on abnormality detection and provision of accurate preliminary clinical evaluation in newly qualified radiographers. Radiography. 2017; Volume 24, Issue 1, 2018, Pages 47-51, https://doi.org/10.1016/j.radi.2017.08.007

3. Verrier W, Pittock LJ, Bodoceanu M, Piper K. Accuracy of radiographer preliminary clinical evaluation of skeletal trauma radiographs, in clinical practice at a district general hospital. Radiography. 2022 May;28(2):312-318. http://doi: 10.1016/j.radi.2021.12.010

4 Hardy M, Culpan G. Accident and emergency radiography: A comparison of radiographer commenting and "red dotting." Radiography. 2007;13(1):65–71. http://doi:0.1016/j.radi.2005.09.009

5. Stevens BJ, Thompson JD. The value of preliminary clinical evaluation for decision making in injuries of the hand and wrist. International Emergency Nursing [Internet]. 2019 Jul; Available from: https://linkinghub.elsevier.com/retrieve/pii/S1755599X19300515

6. Alexander-Bates I, Neep MJ, Davis B, Starkey D. An analysis of radiographer preliminary image evaluation – A focus on common false negatives. Journal of Medical Radiation Sciences. 2021;68(3). https://doi.org/10.1002/jmrs.466

7. Stevens BJ. An analysis of the structure and brevity of preliminary clinical evaluations describing traumatic abnormalities on extremity X-ray images. Radiography. 2020 Nov 1;26(4):302–7. https://doi.org/10.1016/j.radi.2020.02.010

8. Harcus JW, Stevens BJ. What information is required in a preliminary clinical evaluation? A service evaluation. Radiography. 2021, 27 (4), 1033-1037. https://doi.org/10.1016/j.radi.2021.04.001

9. Stevens BJ, Thompson JD. The impact of focused training on abnormality detection and provision of accurate preliminary clinical evaluation in newly qualified radiographers. Radiography [Internet]. 2017;5–9. Available from: http://dx.doi.org/10.1016/j.radi.2017.08.007

10. Bradbury C, Britton I, Lille K, Wright-White H. Abdominal radiograph preliminary clinical evaluation image test bank project. Radiography. 2019;25(3). http://doi:10.1016/j.radi.2019.02.006

11. Rimes SJ, Knapp KM, Meertens RM, Fox DL. Computed tomography colonography: Radiographer independent preliminary clinical evaluation for intraluminal pathology. Radiography. 2019; Nov;25(4):359-364. http://doi:10.1016/j.radi.2019.04.014

12. Neep MJ, Sci B, Rad M, Brown C, Sci B, Rad M, et al. Reducing risk in the emergency department: a 12-month prospective longitudinal study of radiographer preliminary image evaluations. J Med Radiat Sci. 2019 Sep;66(3):154-162. https://doi: 10.1002/jmrs.341

13. England NHS. Diagnostic Imaging Dataset Statistical Release. 2022. Available at: https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/

14. Health Research Authority. Is my study research? [Internet]. 2022. Available from: http://www.hra-decisiontools.org.uk/research/

15. Obuchowski N. Sample size tables for receiver operating characteristic studies. AJR Am J Roentgenol. 2000;175(3):603–8. http://doi. 10.2214/ajr.175.3.1750603

16. Hardy M, Flintham K, Snaith B, Lewis EF. The impact of image test bank construction on radiographic interpretation outcomes: A comparison study. Radiography [Internet]. 2016;22(2):166–70. Available from: http://dx.doi.org/10.1016/j.radi.2015.10.010

17. Thompson, J. D., Thompson, S., Hogg, P., Manning, D., and Szcepura K. ROCView : prototype software for data collection in jackknife alternative free-response receiver operating characteristic analysis. 2012;85(September):1320–6. http://doi: 10.1259/bjr/99497945

18. MacDuff A, Arnold A, Harvey J. Management of spontaneous pneumothorax: British Thoracic Society pleural disease guideline 2010. Thorax. 2010;65(SUPPL. 2). http://dx.doi.org/10.1136/thx.2010.136986

19. Bauman MH, Strange C, Heffner JE, Light R, Kirby TJ, Klein J, et al. Management of spontaneous pneumothorax: An American College of Chest Physicians Delphi Consensus Statement. Chest. 2001;119(2):590–602. http://doi. 10.1378/chest.119.2.590

20. Sharma A, Jindal P. Principles of diagnosis and management of traumatic pneumothorax. J Emerg Trauma Shock. 2008 Jan;1(1):34–41. https://doi:10.4103/0974-2700.41789

21. Zarogoulidis P, Kioumis I, Pitsiou G, Porpodis K, Lampaki S, Papaiwannou A, et al. Pneumothorax: from definition to diagnosis and treatment. J Thorac Dis [Internet]. 2014 Oct;6(Suppl 4):S372-6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25337391

22. Sajadi-Ernazarova KR, Martin J, Gupta N. Acute Pneumothorax Evaluation and Treatment. 2022. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. https://www.ncbi.nlm.nih.gov/books/NBK538316/

23. Drew T, Võ MLH, Wolfe JM. The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. Psychological Science. 2013;24(9):1848–53. http://doi:10.1177/0956797613479386

24. 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. https://doi.org/10.1002/jmrs.306