

Radiologist reporting productivity benefits from AI-assisted triage of CXR studies in clinical practice

Sajith Karunasena^{1, 2}, Michael Milne², David Rosewarne³, Catherine Jones^{1, 2}

1. I-MED Radiology Network, Brisbane, Australia. 2. Annalise.AI, Sydney, Australia. 3. Royal Wolverhampton NHS Trust, Wolverhampton, UK.



Artificial intelligence (AI) has significant potential to advance clinical radiology through improvements to workflow efficiency. To date, AI applied to radiology has typically provided assistance for diagnostic image interpretation to radiologists. This has led to improved accuracy in detection of imaging findings, inherently reducing error rates with regards to missed findings (Aggarwal et al. 2021). In addition to diagnostic assistance, there is potential for AI to augment clinical workflow. With recent studies indicating that AI assistance improves radiologist reporting time and interobserver agreement (Lim et al. 2022).

An alternative aspect where AI has the potential to improve productivity is worklist triage, whereby an AI device categorises studies on the radiologists' worklist as either 'remarkable' or 'unremarkable' based on the presence of select radiologic findings. This has the potential to improve radiologist productivity through the ability of radiologists to quickly identify studies suspected of containing remarkable radiological findings. Previous studies have indicated that AI-assisted triage can reduce reporting time and report turnaround time for radiologic examinations (Nam et al. 2021), which may lead to earlier clinical management and improved patient outcomes. This is particularly important in large healthcare systems where backlogs of unreported studies can lead to significant delays.

Whilst there have been numerous studies validating the performance of AI assist devices, in particular their accuracy and effectiveness at augmenting radiologist accuracy (Seah et al. 2021), the ability of these devices to effectively provide improvements to productivity remains under-studied. This investigation evaluates an alternative use-case of an AI diagnostic assist device capable of detecting findings on chest x-ray (CXR) studies. Specifically, can this device provide AI-assisted worklist triage, based on the presence of remarkable radiological findings. This investigation will focus on the effect of triage on radiologist reporting productivity in a real-world teleradiology setting. The AI device used in this study was Annalise Enterprise CXR, which has been validated for the detection of 124 findings on frontal and lateral projection CXRs (Seah et al. 2021) and has been successfully implemented into real-world reporting environments (Jones et al. 2021).

Can AI-assisted worklist triage used to identify and prioritise studies containing clinically significant radiological findings impact reporting times and report turnaround times for radiologists, when compared to non-AI-assisted triage in an exploratory real-world study.

Of the 124 findings detectable by Annalise Enterprise CXR, 88 were considered clinically significant by a thoracic radiologist. The specific findings considered significant are defined in Table 1. For this investigation, the AI device categorised CXR studies as either remarkable or unremarkable, with any CXR containing one or more clinically significant findings being considered a remarkable.

Table 1 – List of 124 findings detected by Annalise Enterprise, organised into unremarkable and remarkable finding categories

Remarkable Findings											
Acute clavicle fracture	Breast implant	Cardiac valve prosthesis	Diffuse fibrotic volume loss	Electronic cardiac devices	In position Endotracheal tube (ETT)	Lung collapse	Calcified mass (>5mm)	Rib resection	Solitary lung mass	Suboptimal gastric band	
Acute humerus fracture	Bronchiectasis	Cavitating mass with content	Diffuse interstitial	Focal airspace opacity	In position Nasogastric tube (NGT)	Mastectomy	Pleural mass	Scapular fracture	Solitary lung nodule	Suboptimal Nasogastric tube (NGT)	
Acute rib fracture	Bullae diffuse	Cavitating mass(es)	Diffuse lower airspace opacity	Gastric band	In position pulmonary arterial catheter (PAC)	Multifocal Air Space Opacity	Pneumomediastinum	Scapular lesion	Spine lesion	Suboptimal pulmonary arterial catheter (PAC)	
Airway stent	Bullae lower	Chronic rib fracture	Diffuse nodular	Hiatus hernia	Inferior mediastinal mass	Oesophageal stent	Post resection volume loss	Segmental collapse	Spine wedge fracture	Superior mediastinal mass	
Aortic stent	Bullae upper	Clavicle lesion	milary lesions	Hilar lymphadenopathy	Intercostal drain	Multiple masses or nodules	Pulmonary artery enlargement	Shoulder arthritis	Subcutaneous emphysema	Tension pneumothorax	
Atelectasis	Calcified granuloma (<5mm)	Coronary stent	Diffuse pleural thickening	Humeral lesion	Internal foreign body	Osteopaenia	Pulmonary congestion	Shoulder dislocation	Subdiaphragmatic gas	Upper zone fibrotic volume loss	
Basal predominant interstitial	Calcified hilar lymphadenopathy	Diaphragmatic elevation	Diffuse upper airspace opacity	Hyperinflation	Loculated effusion	Peribronchovascular cuffing	Reduced lung markings	Simple effusion	Suboptimal central line (CVC)	Widened aortic contour	
Biliary stent	Calcified pleural plaques	Diffuse airspace opacity	Upper predominant interstitial	In position Central Line (CVC)	Lower zone fibrotic volume loss	Perihilar airspace opacity	Rib lesion	Simple pneumothorax	Suboptimal endotracheal tube (ETT)	Widened cardiac silhouette	
Unremarkable Findings											
Abdominal clips	Calcified neck nodes	Chronic humerus fracture	Distended bowel	Lung sutures	Nipple shadow	Pectus carinatum	Rib fixation	Shoulder fixation	Spine arthritis	Underexposed	
Aortic arch calcification	Cervical flexion	Clavicle fixation	Gallstones	Mediastinal clips	Overexposed	Pectus excavatum	Rotator cuff anchor	Shoulder replacement	Sternotomy wires	Underinflation	
Axillary clips	Chest incompletely imaged	Diaphragmatic eventration	Image obscured	Neck clips	Patient rotation	Pericardial fat pad	Scoliosis	Spinal fixation	Tracheal deviation	Unfolded aorta	
Calcified axillary nodes	Chronic clavicle fracture	Diffuse spinal osteophytes	Kyphosis								

METHODS

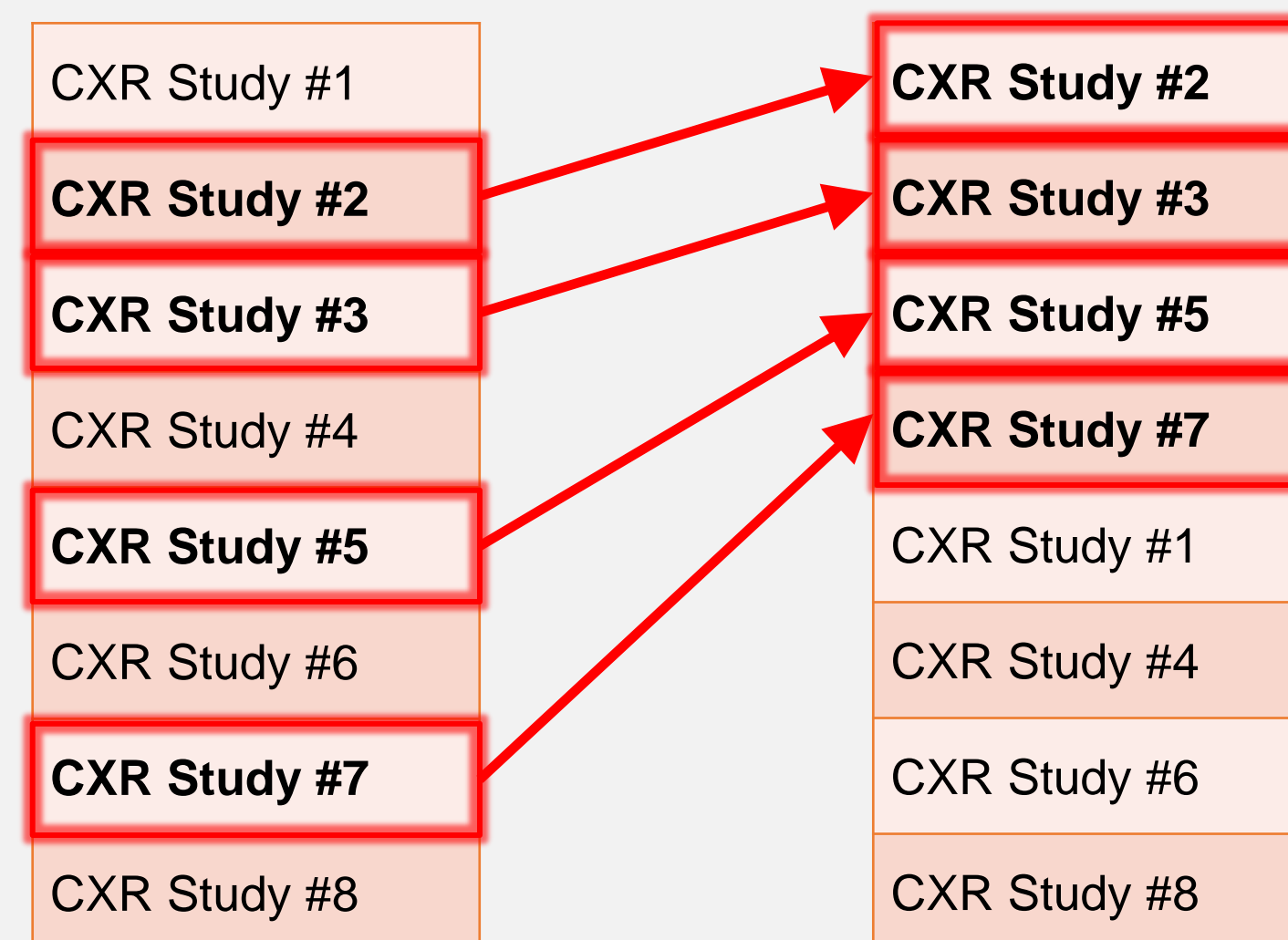
Four radiologists participated in the study. These included one senior thoracic radiologist (>10 years experience), two general radiologists (3-5years experience), and one junior (<1 year experience).

Each radiologist undertook two separate reporting sessions, one reading a normal 'unsorted' reporting worklist of CXR cases and the other reading a 'sorted' worklist of studies triaged by the AI device. Each worklist contained a list of 50 CXRs, which included 25 remarkable and 25 unremarkable studies. In the unsorted worklist, the CXR studies were reported in a first-in first-out order, with radiologists reporting all studies in one sitting. In a subsequent reading session, the radiologist reported a worklist of CXRs, where CXRs indicated by the AI to contain remarkable findings were presented first and the AI prediction of remarkable/unremarkable was visible to the radiologist. The worklist sorting process is shown in figure 1.

Typical 'unsorted' first-in first-out worklist

CXR Study #1
CXR Study #2
CXR Study #3
CXR Study #4
CXR Study #5
CXR Study #6
CXR Study #7
CXR Study #8

AI remarkable prioritised worklist



AI processing for clinically significant findings

Figure 1 – An example worklist describing the AI-assisted triage prioritisation system used in this investigation

Radiologist reporting time was measured per case, as the time from report initiation to finalisation. Average report turnaround times were also recorded for remarkable studies, defined as the time between the start of the reporting session and the completion of each CXR report. Radiologist agreement with the AI device indication of remarkable/unremarkable was recorded for each study, with overall results presented as a percentage of cases where the radiologist did not agree with the AI designation of remarkable/unremarkable. t-tests were conducted to compare the group averages.

RESULTS

There was a consistent trend towards reduced reporting time for the AI-triaged worklists compared to the unsorted worklists. Average reporting time across all CXRs decreased from 55.8 to 48.5 seconds, while the reporting time for remarkable CXR studies decreased from 75.5 seconds to 66.8 seconds and reporting time for unremarkable CXR studies decreased from 35.5 seconds to 30.3 seconds (Figure 2), with none of the differences being found to be statistically significant (all CXRs: 0.50, remarkable: p 0.55, unremarkable: p 0.26).

The report turnaround time for remarkable CXR studies was also reduced. The average report turnaround time for remarkable CXR studies for all radiologists decreased from 1450.5 to 889.3 seconds (Figure 2) and this decrease was statistically significant (p 0.015). Discrepancies between AI device predictions and the radiologist were reported in 6.5% of cases (13/200 studies).

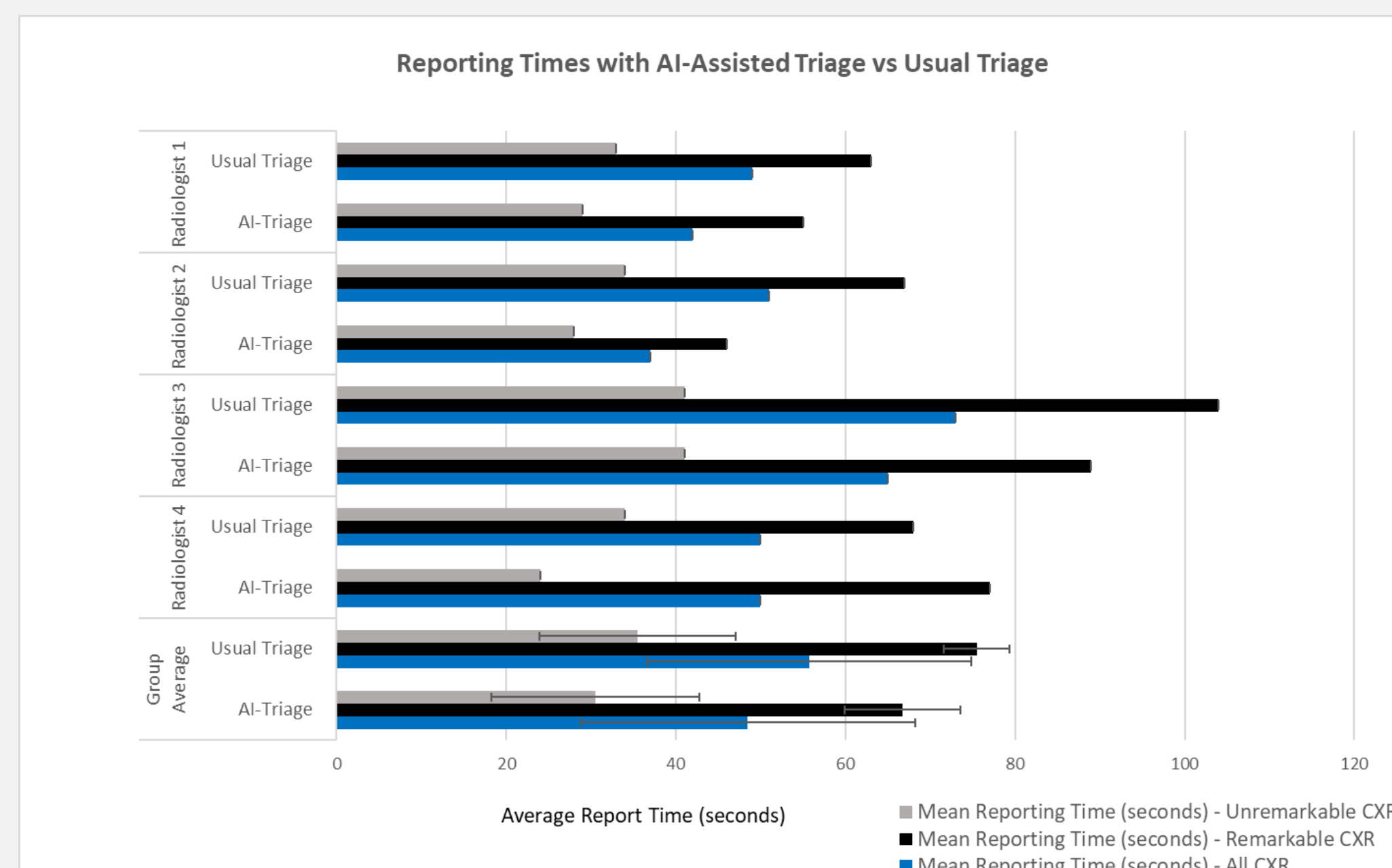


Figure 2 – Graph illustrating the impact of AI-assisted triage on radiologist reporting time. Results displayed as mean, error bars represent standard deviation (SD).

RESULTS

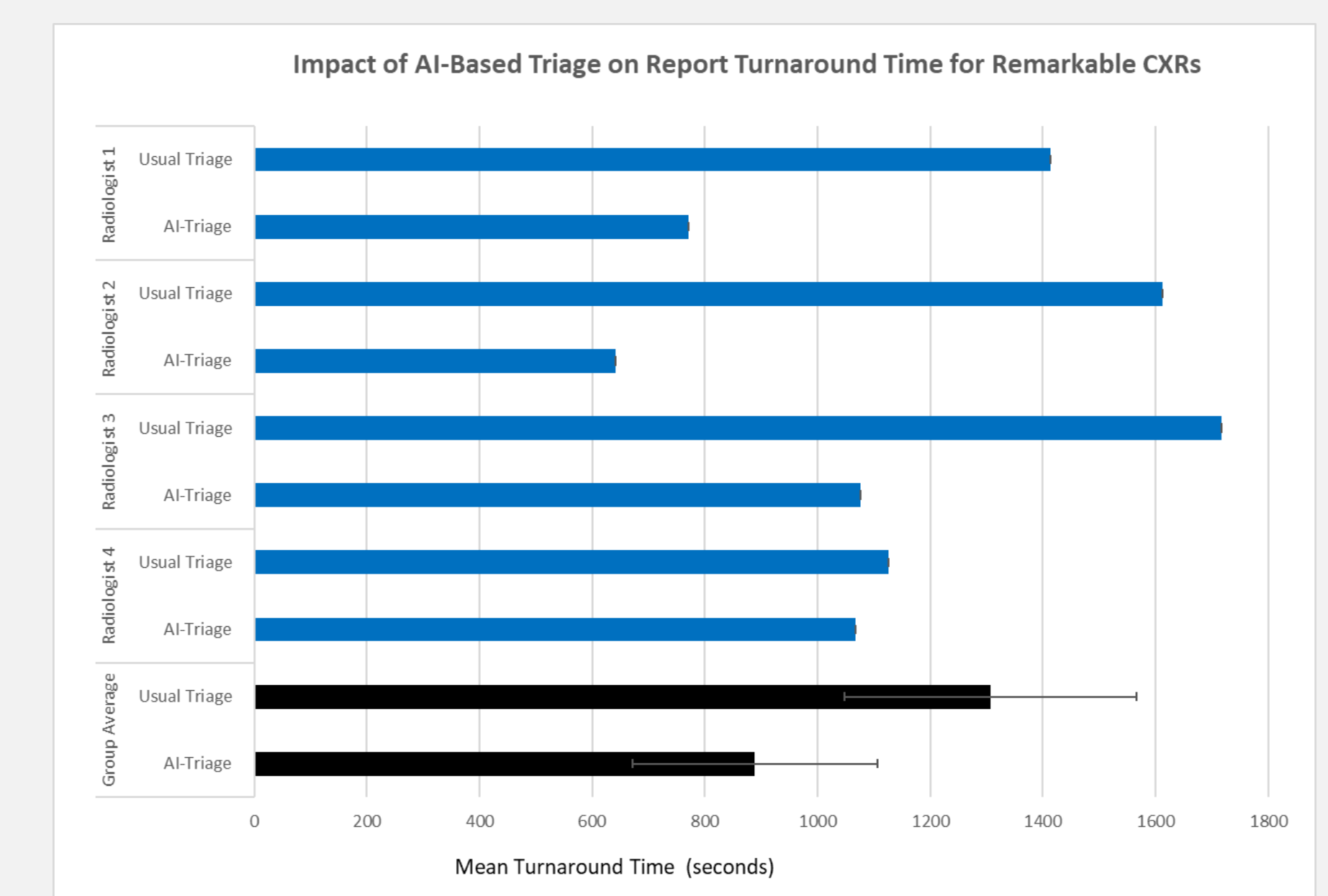


Figure 2 – Graph illustrating the impact of AI-assisted triage on radiologist report turnaround time. Results displayed as mean, error bars represent SD.

DISCUSSION

Results illustrate that AI-assisted triage of a reporting worklist by a commercially available AI device successfully reduced the reporting time across both remarkable and unremarkable CXR studies and reduced the report turnaround time for CXR studies containing clinically significant findings. In addition to this, discrepancies between radiologists and the AI device were low.

Three of four radiologists showed a reduction in reporting time across all CXR study groups. The remaining radiologist (radiologist 4) showed an equal overall reporting time between sorted and unsorted reporting sessions and demonstrating an increased reporting time when reporting remarkable studies with AI-assistance. Additionally, the combined average report turnaround time for remarkable studies was significantly reduced across all radiologists. The smallest time reduction was observed with radiologist 4, which was observed to likely result from a predominance of remarkable studies presented early in the unsorted list, a random variation.

This study showed high variability in reporting time and report turnaround times between radiologists. This may be caused by varying experience levels of participating radiologists and/or varying difficulty of cases in each radiologist's worklist. Both these factors would be mitigated by a larger study with more radiologists and higher number of cases. It was also noted that different reporting styles impact on reporting time, particularly around the use of reporting templates versus dictated reports in the setting of unremarkable studies.

It is important to ensure that AI device accuracy is high and there are systems in place to avoid extended waiting times. While not reported here, it is expected that the report turnaround time for unremarkable will be increased in the AI-assisted group. It is important to take this into consideration as any CXR studies that contain remarkable findings that are not detected by the AI device will take longer to report, potentially leading to worse patient outcomes.

CONCLUSION

Results from this investigation indicate that AI devices to assist CXR interpretation and triage could have a significant role in optimising efficiency and reducing report turnaround times for studies containing clinically significant findings. Even in this relatively small study, there was significant improvement in turnaround time for reports of remarkable studies when an AI-based triage system was in place.

A larger study with more radiologists and higher number of cases is planned.

REFERENCES

- Aggarwal, R., V. Sounderajah, G. Martin, D. S. W. Ting, A. Karthikesalingam, D. King, H. Ashrafian, and A. Darzi. 2021. "Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-Analysis." *Npj Digital Medicine* 4 (1).
- Jones, Catherine M., Luke Danaher, Michael R. Milne, Cyril Tang, Jarrel Seah, Luke Oakden-Rayner, Andrew Johnson, Quinlan D. Buchlak, and Nazanin Esmaili. 2021. "Assessment of the Effect of a Comprehensive Chest Radiograph Deep Learning Model on Radiologist Reports and Patient Outcomes: A Real-World Observational Study." *BMJ Open* 11 (12): e052902.
- Lim, Desmond Shi Wei, Andrew Makmur, Lei Zhu, Wenqiao Zhang, Amanda J. L. Cheng, David Soon Yiew Sia, Sterling Ellis Eide, et al. 2022. "Improved Productivity Using Deep Learning-Assisted Reporting for Lumbar Spine MRI." *Radiology*, June, 220076.
- Nam, Ju Gang, Minchul Kim, Jongchan Park, Eui Jin Hwang, Jong Hyuk Lee, Jung Hee Hong, Jin Mo Goo, and Chang Min Park. 2021. "Development and Validation of a Deep Learning Algorithm Detecting 10 Common Abnormalities on Chest Radiographs." *The European Respiratory Journal: Official Journal of the European Society for Clinical Respiratory Physiology* 57 (5). <https://doi.org/10.1183/13993003.03061-2020>.
- Seah, Jarrel C. Y., Cyril H. M. Tang, Quinlan D. Buchlak, Xavier G. Holt, Jeffrey B. Wardman, Anuar Aimoldin, Nazanin Esmaili, et al. 2021. "Effect of a Comprehensive Deep-Learning Model on the Accuracy of Chest x-Ray Interpretation by Radiologists: A Retrospective, Multireader Multicase Study." *Lancet Digit Health* 3 (8): e496-506.