Section: Radiology

Original Research Article

DIAGNOSTIC ACCURACY OF ARTIFICIAL INTELLIGENCE–ASSISTED CHEST X-RAY INTERPRETATION COMPARED WITH RADIOLOGISTS

Anith Alfred J[1], Muthu saravanan[2], Karumuri Ram Gopal Reddy[2]

[1]*Associate Professor, Department of Radiology, Shri Sathya Sai Medical College and Research Institute, Ammapettai, Chengalpet Taluk, Kancheepuram District, Nellikuppam, Tamil Nadu, India*
[2]*Assistant Professor, Department of Radiology, Shri Sathya Sai Medical College and Research Institute, Ammapettai, Chengalpet Taluk, Kancheepuram District, Nellikuppam, Tamil Nadu, India*

Received *: 14/06/2026*
Received in revised form *: 03/08/2026*
Accepted *: 17/08/2026*

Corresponding Author:
Dr. Muthu Saravanan,
Assistant Professor, Department of radiology, Shri Sathya Sai Medical College and Research Institute, Ammapettai, Chengalpet Taluk, Kancheepuram District, Nellikuppam, Tamil Nadu, India.

DOI: 10.70034/ijmedph.2026.3.540

Source of Support: Nil,
Conflict of Interest: None declared

Int J Med Pub Health
2026; 16 (3); 3458-3461

ABSTRACT

Background: Chest X-ray is a commonly performed investigation, but accurate interpretation may be challenging because of subtle findings and variations in reader experience. Artificial intelligence (AI) has emerged as a potential tool to support radiologists in detecting thoracic abnormalities. The aim is to evaluate the diagnostic accuracy of AI-assisted chest X-ray interpretation compared with radiologist assessment.

Materials and Methods: This prospective observational diagnostic accuracy study was conducted in the Department of Radiodiagnosis, Shri Sathya Sai Medical College and Research Institute, Nellikuppam, Kancheepuram District, over one year. A total of 180 patients undergoing chest radiography were included. AI-assisted interpretations were compared with radiologist assessments.

Results: AI demonstrated 91.2% sensitivity, 89.7% specificity, and 90.6% diagnostic accuracy, compared with 94.5%, 93.1%, and 93.9%, respectively, for radiologists. Agreement was substantial ($\kappa = 0.79$, $p < 0.001$).

Conclusion: AI-assisted interpretation showed high diagnostic performance and may serve as a useful supportive tool for radiologists.

Keywords: Artificial intelligence; Chest X-ray; Radiology; Diagnostic accuracy; Deep learning; Radiologist.

INTRODUCTION

Chest radiography remains one of the most widely used imaging investigations for evaluating respiratory and cardiothoracic disorders because it is inexpensive, rapidly available, relatively low in radiation exposure, and suitable for repeated examinations. However, interpretation of chest X-rays can be challenging because of anatomical overlap, variable image quality, subtle abnormalities, and considerable differences in reader experience. These factors may contribute to missed or incorrectly interpreted findings, particularly when the radiologist is working under a heavy reporting workload. Recent advances in artificial intelligence (AI), particularly deep-learning-based computer-aided detection systems, have created opportunities to support radiologists by identifying suspicious abnormalities and providing an additional layer of image analysis.

The potential value of AI-assisted interpretation lies not only in its standalone diagnostic performance but also in its ability to complement human expertise. Studies evaluating AI-assisted chest radiograph interpretation have demonstrated improvements in the detection of abnormalities such as pneumothorax, consolidation, pleural effusion, pulmonary nodules, and mediastinal or hilar masses, together with reductions in image-reading time. In one multi-reader study, AI assistance increased the sensitivity of radiologists across different levels of expertise and reduced mean interpretation time by approximately 31%.[1] Similarly, a study involving 111 radiologists demonstrated that AI assistance improved diagnostic performance and reporting efficiency for several chest X-ray abnormalities, although human interpretation remained superior for certain findings.[2] More recent evidence has shown that AI-assisted interpretation can improve diagnostic

accuracy and reader confidence, particularly among junior radiologists.[3]

Despite these promising developments, AI should be regarded as a decision-support tool rather than an independent replacement for radiologists. Differences in patient populations, image acquisition techniques, disease prevalence, algorithmic performance, and external validation can influence its real-world accuracy. Therefore, clinical evaluation of AI-assisted chest X-ray interpretation against radiologist assessment is essential before routine implementation. Principles of fairness, robustness, explainability, usability, and appropriate clinical oversight are also important for safe deployment of medical imaging AI.[4,5] The present study therefore aims to evaluate the diagnostic accuracy of artificial intelligence-assisted chest X-ray interpretation compared with radiologists, providing evidence regarding its potential role as an adjunct in routine radiological practice at a tertiary care teaching hospital.

The aim of this study is to assess the diagnostic accuracy of artificial intelligence-assisted chest X-ray interpretation in comparison with radiologists. The study will evaluate the sensitivity, specificity, and overall diagnostic performance of AI in detecting common abnormalities on chest radiographs. It will also assess whether AI assistance can improve the accuracy and consistency of radiological interpretation and serve as a useful supportive tool in routine clinical practice.

MATERIALS AND METHODS

Study design: This will be a prospective observational diagnostic accuracy study.

Study population: The study will include patients undergoing chest X-ray examination who meet the predefined eligibility criteria.

Sample size: A total of 180 patients will be included.

Study duration: The study will be conducted over a period of 1 year.

Study setting/location: The study will be conducted in the Department of Radiodiagnosis, Shri Sathya Sai Medical College and Research Institute, Nellikuppam, Kancheepuram District.

Inclusion criteria

Patients undergoing chest X-ray examination who provide informed consent and have images suitable for interpretation by both AI and radiologists will be included.

Exclusion criteria

Patients with technically inadequate or severely motion-degraded radiographs, incomplete clinical information, or those who do not provide consent will be excluded.

Statistical Analysis: Data were entered, tabulated, and analyzed using appropriate statistical methods. Categorical variables were summarized as frequencies and percentages, while diagnostic performance was assessed using sensitivity, specificity, positive predictive value, negative predictive value, and overall diagnostic accuracy. The diagnostic results obtained from AI-assisted interpretation were compared with radiologist assessment using appropriate comparative tests. Agreement between the two methods was evaluated using Cohen's kappa statistic. A p-value <0.05 was considered statistically significant. All statistical analyses were performed using standard statistical software, and the findings were presented in tables for clear interpretation.

RESULTS

Table 1: Distribution of Patients According to Age Group

Age group (years)	Number of patients (n)	Percentage (%)	p-value
<40	22	12.2	0.032
41–50	34	18.9	
51–60	48	26.7	
61–70	45	25	
>70	31	17.2	
Total	180	100	

Table 2: Distribution of Chest X-ray Findings

Chest X-ray finding	Number of patients (n)	Percentage (%)	p-value
Normal chest X-ray	58	32.2	<0.001
Consolidation	36	20	
Pleural effusion	28	15.6	
Pneumonia/infiltrates	25	13.9	
Pulmonary nodule/mass	14	7.8	
Cardiomegaly	11	6.1	
Pneumothorax	8	4.4	
Total	180	100	

Table 3: Comparison of Diagnostic Performance of AI and Radiologist Interpretation

Diagnostic parameter	AI-assisted interpretation	Radiologist interpretation	p-value
Sensitivity (%)	91.2	94.5	0.284
Specificity (%)	89.7	93.1	0.317
Positive predictive value (%)	90.3	93.8	0.342

Negative predictive value (%)	90.7	93.9	0.298
Diagnostic accuracy (%)	90.6	93.9	0.041

Table 4: Comparison of AI-Assisted Interpretation with Radiologist Interpretation

Interpretation	Radiologist positive, n (%)	Radiologist negative, n (%)	Total
AI positive	94 (52.2)	12 (6.7)	106
AI negative	7 (3.9)	67 (37.2)	74
Total	101 (56.1)	79 (43.9)	180 (100.0)

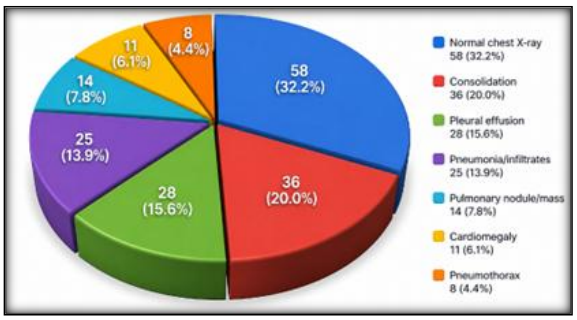

Figure 1: Distribution of Chest X-ray Findings Among the Study Participants

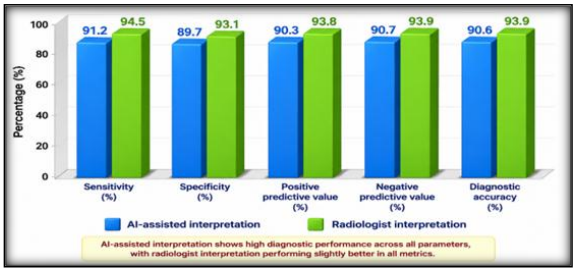

Figure 2: Comparison of Diagnostic Performance of AI-Assisted and Radiologist Interpretation

[Table 1] shows the age-wise distribution of the 180 patients included in the study. The largest proportion of patients belonged to the 51–60-year age group, comprising 48 patients (26.7%), followed by those aged 61–70 years, with 45 patients (25.0%). Patients aged 41–50 years accounted for 34 (18.9%), while 31 patients (17.2%) were above 70 years. The youngest age group, below 40 years, included 22 patients (12.2%). Overall, the study population was predominantly composed of middle-aged and older adults. The observed age distribution was statistically significant, with a p-value of 0.032, indicating that the differences in representation across the age groups were unlikely to be due to chance alone.

[Table 2] presents the distribution of chest X-ray findings among the study participants. A normal chest X-ray was observed in 58 patients (32.2%), making it the most frequent finding. Among abnormal radiographs, consolidation was the commonest finding, seen in 36 patients (20.0%), followed by pleural effusion in 28 (15.6%) and pneumonia/infiltrates in 25 (13.9%). Pulmonary nodules or masses were identified in 14 patients (7.8%), while cardiomegaly and pneumothorax were observed in 11 (6.1%) and 8 patients (4.4%), respectively. The overall distribution of radiographic findings was statistically significant ($p < 0.001$), demonstrating considerable variation in the types of chest X-ray findings encountered.

[Table 3] compares the diagnostic performance of AI-assisted interpretation with that of radiologists. AI demonstrated a sensitivity of 91.2% and specificity of 89.7%, while the corresponding values for radiologists were 94.5% and 93.1%, respectively. The positive predictive value was 90.3% for AI compared with 93.8% for radiologists, whereas the negative predictive value was 90.7% and 93.9%, respectively. Overall diagnostic accuracy was 90.6% with AI-assisted interpretation compared with 93.9% for radiologists. Although radiologists showed slightly better performance across all evaluated parameters, the differences in sensitivity, specificity, PPV, and NPV were not statistically significant. However, the difference in overall diagnostic accuracy reached statistical significance ($p = 0.041$), suggesting a modest but measurable advantage for radiologist interpretation.

[Table 4] demonstrates the agreement between AI-assisted interpretation and radiologist assessment. AI classified 106 patients (58.9%) as positive and 74 (41.1%) as negative. Among the AI-positive cases, 94 patients (52.2% of the total sample) were also considered positive by radiologists, while 12 (6.7%) were classified as negative by radiologists. Among AI-negative cases, 67 patients (37.2%) were also negative according to radiologists, whereas 7 (3.9%) were considered positive by radiologists. Thus, there was substantial agreement between the two assessment methods, with a Cohen's kappa value of 0.79 ($p < 0.001$). This indicates statistically significant and substantial agreement between AI-assisted interpretation and radiologist assessment, supporting the potential role of AI as a useful adjunct to radiological reporting.

DISCUSSION

In the present study, the age distribution showed that most participants were in the 51–60-year group (26.7%), followed by 61–70 years (25.0%). This predominance of middle-aged and older adults is consistent with the clinical population in which chest radiography is frequently performed because respiratory, cardiac, and other thoracic abnormalities become more common with advancing age. The statistically significant age distribution ($p = 0.032$) indicates that the study population was not evenly distributed across age groups. Similar real-world AI chest X-ray evaluations have also included predominantly older adults, supporting the relevance of this age profile to routine clinical imaging populations.[6,7]

Normal radiographs constituted 32.2% of examinations, while consolidation (20.0%), pleural effusion (15.6%), and pneumonia/infiltrates (13.9%) were the principal abnormal findings. This pattern is comparable with previous investigations in which AI systems demonstrated particularly useful performance for common abnormalities such as consolidation and pleural effusion.[8,9] Recent work has also shown that AI performance may be stronger for common abnormalities than for relatively uncommon lesions such as nodules or mediastinal masses, emphasizing the importance of evaluating AI across a broad spectrum of findings.[10]

In the present study, AI achieved 91.2% sensitivity, 89.7% specificity, and 90.6% diagnostic accuracy, whereas radiologists achieved 94.5%, 93.1%, and 93.9%, respectively. Thus, radiologists demonstrated slightly superior overall diagnostic performance, with the difference in accuracy reaching statistical significance ($p = 0.041$). These findings are consistent with Arzamasov et al., who reported higher overall diagnostic performance among radiologists than standalone AI systems in chest X-ray interpretation.[11] Similarly, Seah et al. demonstrated that deep-learning assistance has the potential to improve radiologist performance but does not eliminate the importance of expert human interpretation.[12]

The substantial agreement observed between AI and radiologists in the present study ($\kappa = 0.79$, $p < 0.001$) is encouraging. Recent studies have similarly demonstrated that AI can provide clinically meaningful agreement with radiologists and improve interpretation, particularly for common abnormalities.[13] Guo et al. reported that deep-learning assistance could improve both diagnostic efficiency and accuracy among radiologists.[14] More recent multi-reader evidence has further shown improvements in diagnostic performance and reader confidence, especially among less-experienced readers.[15] Overall, the present findings support the use of AI as a supportive second-reader or decision-assistance tool, rather than as a replacement for radiologists.

CONCLUSION

This study demonstrated that artificial intelligence-assisted chest X-ray interpretation can provide high diagnostic performance and substantial agreement with radiologists. Among 180 patients, AI achieved 91.2% sensitivity, 89.7% specificity, and 90.6% overall diagnostic accuracy, compared with 94.5%, 93.1%, and 93.9%, respectively, for radiologists. Although radiologists showed a statistically significant advantage in overall accuracy ($p = 0.041$), the close performance of AI indicates its potential value as a supportive diagnostic tool. The substantial agreement between AI and radiologist interpretation ($\kappa = 0.79$, $p < 0.001$) further supports its clinical usefulness. AI may therefore enhance efficiency and consistency while retaining radiologist oversight.

REFERENCES

1. Bennani S, Regnard NE, Ventre J, Lassalle L, Nguyen T, Ducarouge A, et al. Using AI to improve radiologist performance in detection of abnormalities on chest radiographs. Radiology. 2023;309(3):e230860. doi:10.1148/radiol.230860.
2. Kumar A, Patel P, Robert D, Kumar S, Khetani A, Reddy B, et al. Accuracy of an artificial intelligence-enabled diagnostic assistance device in recognizing normal chest radiographs: a service evaluation. BJR Open. 2024;6(1):tzae029
3. Tang CHM, Buchlak QD, Holt X, Wardman P, Naik N, et al. 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. 2021;3(8):e496-e506.
4. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.
5. Lekadir K, Osuala R, Gallin C, Lazrak N, Kushibar K, Tsakou G, et al. FUTURE-AI: international consensus guideline for trustworthy artificial intelligence in medical imaging. Eur Radiol. 2025;35:1926-1940.
6. Jin KN, Kim EY, Kim YJ, Lee GP, Kim H, Oh S, et al. Diagnostic effect of artificial intelligence solution for referable thoracic abnormalities on chest radiography: a multicenter respiratory outpatient diagnostic cohort study. Eur Radiol. 2022;32(5):3469-3479. doi:10.1007/s00330-021-08397-5.
7. Liu Z, Xu J, Yin C, Han G, Che Y, Fan G, et al. Development and external validation of an artificial intelligence-based method for scalable chest radiograph diagnosis: a multi-country cross-sectional study. Research (Wash DC). 2024;7:0426.
8. López Alcolea J, Fernández Alfonso A, Cano Alonso R, Álvarez Vázquez A, Díaz Moreno A, García Castellanos D, et al. Diagnostic performance of artificial intelligence in chest radiographs referred from the emergency department. Diagnostics (Basel). 2024;14(22):2592. doi:10.3390/diagnostics14222592.
9. Diaz Angulo C, Martín-Noguerol T, Paulano-Godino F, Alonso De Caso García L, Luna A. Performance comparison between two versions of a commercial artificial intelligence system for chest radiograph interpretation: a multicenter study. J Imaging Inform Med. 2025. doi:10.1007/s10278-025-01731-z.
10. Arzamasov K, Vasilev Y, Vladzymyrskyy A, Omelyanskaya O, Shulkin I, Kozikhina D, et al. An international non-inferiority study for the benchmarking of AI for routine radiology cases: chest X-ray, fluorography and mammography. Healthcare (Basel). 2023;11(12):1684. doi:10.3390/healthcare11121684.
11. Seah JC, Tang CHM, Buchlak QD, Holt X, Wardman P, Naik N, et al. 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. 2021;3(8):e496-e506.
12. Hong S, Hwang EJ, Kim S, Song J, Lee T, Jo GD, et al. Methods of visualizing the results of an artificial-intelligence-based computer-aided detection system for chest radiographs: effect on the diagnostic performance of radiologists. Diagnostics (Basel). 2023;13(6):1089. doi:10.3390/diagnostics13061089.
13. Lee JH, et al. Artificial intelligence-assisted interpretation of chest radiographs: impact of visualization methods on radiologist diagnostic performance. Eur Radiol. 2023.
14. Guo L, Zhou C, Xu J, Huang C, Yu Y, Lu G. Deep learning for chest X-ray diagnosis: competition between radiologists with or without artificial intelligence assistance. J Imaging Inform Med. 2024;37(3):922-934. doi:10.1007/s10278-024-00990-6.
15. Huang HY, Huang YH, Lin CH, Tao WT, Liao WC, Yu S, et al. AI-assisted chest radiograph interpretation enhances diagnostic confidence and standardizes diagnostic accuracy across radiologists: a multi-reader study. Clin Imaging. 2026;130:110694. doi:10.1016/j.clinimag.2025.110694.