Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Artificial intelligence-generated preliminary reports from a radiologist's perspective - A systematic review.

Dr. Priyadarshini Subramani[1], Dr. Karthik Shunmugavelu[2*]
[1] *Assistant Professor, Department of Radiology, PSP Medical College Hospital and Research Institute, Tambaram, Kanchipuram main road, Oragadam, Panruti Kanchipuram district, Tamilnadu 631604 India*
[2] *Assistant Professor, Department of Dentistry / PSP Medical College Hospital and Research Institute Tambaram Kanchipuram main road Oragadam Panruti Kanchipuram district Tamilnadu 631604 India*

Abstract

Background
Artificial intelligence (AI) is increasingly integrated into radiology to generate preliminary reports. While AI promises to improve efficiency and standardization, its impact on diagnostic accuracy and radiologist workflow requires further investigation.
Objective: To systematically review evidence on radiologist interaction with AI-generated radiology reports, evaluating effects on reporting efficiency, diagnostic accuracy, report quality, radiologist confidence, and patient comprehension.

Methods
A systematic literature search of PubMed, MEDLINE, Embase, Scopus, and Web of Science was conducted for studies published between January 2016 and March 2025. Eligible studies included original research and reviews assessing AI-generated preliminary or structured radiology reports with radiologist involvement. Data were extracted on study characteristics, imaging modalities, AI system type, and key outcomes. Narrative synthesis was performed due to heterogeneity in study designs.

Results
Eight studies met the inclusion criteria, encompassing imaging modalities such as chest radiographs, MRI, CT, and spine imaging. Across studies, AI-assisted reporting consistently improved efficiency, with reductions in reporting time ranging from 15% to over 40%. Diagnostic accuracy and report quality were generally maintained, although variability was noted in complex or abnormal cases. Radiologists' confidence and experience were enhanced when interacting with AI-generated reports. Structured reporting and patient-friendly summaries further improved standardization and patient comprehension. Limitations included occasional AI misinterpretations, ethical and regulatory concerns, and the continued need for human oversight, particularly for complex imaging findings.

Conclusions
AI-generated radiology reports offer significant benefits in efficiency, standardization, and patient-centered communication while preserving diagnostic accuracy when combined with expert radiologist review. Human–AI collaboration emerges as the most effective model.

Future research
Future research should focus on optimizing AI performance for complex cases, validating patient-centered reporting, and establishing robust clinical and ethical frameworks for AI integration.

Keywords: *Artificial intelligence, Radiology, preliminary report, Structured report, Human–AI collaboration, Workflow efficiency, Diagnostic accuracy*
Submitted: *September 20, 2022* **Accepted:** *November 21, 2022* **Published:** *December 30, 2022*

Corresponding author: *Dr. Karthik Shunmugavelu**
Email: ***drkarthiks1981@gmail.com*** *https://orcid.org/0000-0001-7562-8802*
Senior Resident, Department of Dentistry, PSP Medical College Hospital and Research Institute, Tambaram Kanchipuram main road, Oragadam, Panruti Kanchipuram district, Tamilnadu 631604 India

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Introduction

Artificial intelligence (AI) is increasingly transforming medical imaging by augmenting radiologists' diagnostic capabilities and streamlining workflows. Among its many applications, AI-generated radiology reports ranging from preliminary draft interpretations to structured and patient-centered summaries have garnered significant attention for their potential to enhance efficiency, standardization, and communication [1]. Rapid advances in machine learning, deep learning, and transformer-based language models have enabled AI systems to not only detect abnormalities but also generate coherent textual reports, extract key findings from imaging data, and facilitate structured documentation [2].

Despite these technological advances, integration of AI into clinical practice raises critical questions regarding diagnostic accuracy, workflow impact, radiologist acceptance, and patient-centered communication [3]. While AI may improve efficiency for routine cases, its performance in complex or abnormal findings remains variable, highlighting the importance of human oversight. Additionally, the ethical, regulatory, and data privacy considerations surrounding AI-generated medical documentation further complicate clinical adoption [4].

This systematic review aims to synthesize current evidence on radiologists' interaction with AI-generated preliminary and structured reports. Specifically, it examines how AI integration affects reporting efficiency, diagnostic accuracy, report quality, radiologist confidence, and patient comprehension. By analysing findings across diverse imaging modalities, study designs, and AI architectures, this review seeks to identify patterns, highlight limitations, and provide insights into best practices for safe and effective human AI collaboration in radiology.

Methodology

Study design

A systematic review methodology was employed to comprehensively evaluate the current literature on AI-generated radiology reports and radiologist interaction. The review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological rigor and transparency.

Search Strategy

A structured search of PubMed, MEDLINE, Embase, Scopus, and Web of Science was conducted for studies published from January 2016 to March 2025. Search terms combined keywords and controlled vocabulary related to radiology, artificial intelligence, generative AI, large language models, preliminary reports, structured reporting, and radiologist workflow. Boolean operators and truncation were used to maximize sensitivity while minimizing irrelevant results. Additionally, reference lists of included studies were screened to identify further relevant publications.

Inclusion criteria

Studies were included if they met the following criteria:

- Original research or systematic/scoping reviews evaluating AI-generated radiology reports.
- Involvement of radiologists interacting with AI-generated preliminary or structured reports.
- Assessment of outcomes such as reporting efficiency, diagnostic accuracy, report quality, radiologist confidence, or patient comprehension.
- Published in English in peer-reviewed journals.

Exclusion criteria included

- Studies focused solely on AI-based abnormality detection without report generation.
- Non-human studies or conference abstracts without full-text availability.
- Reports not involving radiologist interaction or clinical workflow evaluation.

Risk of bias assessment

To assess the quality of the included studies, two reviewers independently evaluated the risk of bias using appropriate tools based on study design. For randomized and non-randomized original studies, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was applied. For systematic reviews, the AMSTAR-2 (A Measurement Tool to Assess Systematic Reviews) checklist was used. Disagreements were resolved through consensus or consultation with a third reviewer. The results of these assessments were considered when interpreting the findings and drawing conclusions.

Data extraction and synthesis

Two independent reviewers screened titles, abstracts, and full texts for eligibility. Discrepancies were resolved through discussion or consultation with a third reviewer.

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Data extraction captured the following domains: study characteristics (authors, year, country, design), imaging modality, AI system type, sample size, key outcomes (efficiency, accuracy, quality, confidence, patient comprehension), and reported limitations.

Effect measures

For the synthesis and presentation of results, effect measures were selected based on the nature of the outcome data reported in the included studies. For continuous outcomes such as reporting time and radiologist confidence scores, the primary effect measures were mean differences or percentage changes from baseline. For dichotomous outcomes, such as diagnostic accuracy and report acceptance rates, effect measures included proportions, percentages, and, where available, measures of agreement (e.g., RADPEER scores). Due to the heterogeneity of the included studies, a narrative synthesis was prioritized; therefore, pooled effect estimates (e.g., risk ratios) are not reported.

Certainty assessment

The certainty (or confidence) in the cumulative evidence for each key outcome was assessed using a narrative approach based on the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework. Factors considered included the methodological quality (risk of bias) of the included studies, consistency of findings across different study designs and imaging modalities, precision of the reported effect estimates, and relevance of the evidence to the review question. Due to the descriptive nature of the synthesis, a formal GRADE profile was not constructed; however, these factors informed the interpretation of the strength of the evidence presented in the discussion.
Findings were synthesized narratively, emphasizing patterns, consistencies, and divergences across studies. Quantitative outcomes were reported where available, and qualitative assessments were incorporated to evaluate radiologist perceptions and patient-centered impacts. Due to heterogeneity in study design, imaging modalities, and AI platforms, meta-analysis was not feasible; instead, a descriptive synthesis was used to present a coherent and comprehensive overview of the evidence.

RESULTS

Study selection

The systematic literature search of PubMed, MEDLINE, Embase, Scopus, and Web of Science yielded a total of 347 records. After the removal of 89 duplicate records, 258 unique titles and abstracts were screened for relevance. Of these, 236 records were excluded based on title and abstract review as they did not meet the preliminary inclusion criteria (e.g., focused solely on AI-based abnormality detection without report generation, were conference abstracts without full text, or did not involve radiologist interaction). This left 22 full-text articles to be assessed for eligibility.

Following full-text review, 14 studies were excluded for various reasons. A full-text article by Hwang et al. (2023) [15] was excluded because, although it evaluated AI-assisted interpretation of chest radiographs, it focused on diagnostic performance for acute respiratory symptoms rather than the generation or use of preliminary textual reports. Similarly, a study by Lee et al. (2020) [20] was excluded as it assessed a deep learning algorithm for lung cancer detection without integration into a report generation workflow. Another study by Archer et al. (2022) [3] was excluded because it evaluated AI-generated hip measurements but did not involve the generation of a full radiology report or direct radiologist interaction with a preliminary text output.
Ultimately, eight studies [5-11, 23] met all inclusion criteria and were included in the narrative synthesis. The study selection process is illustrated in Figure 1 (PRISMA flow diagram to be inserted here). The characteristics of the included studies are summarized in Table 1 [5–11].

Study characteristics

The eight included studies comprised a mix of original research and review articles, reflecting the evolving evidence base in this field. Study designs included one longitudinal multireader study [5], one prospective cohort study [9], one comparative study [6], one technical implementation study [10], one primary study on patient-centered reporting [11], one systematic review [7], and one scoping review protocol [8]. The studies were conducted across various countries, including the United States, South Korea, and multinational collaborations. Sample sizes for original studies ranged from 100 complex imaging cases [6] to 23,960 radiographs [9] and 685 spine MRI examinations [11]. Imaging modalities covered included chest radiographs [5, 9], CT [6, 10], MRI (knee, lumbar spine) [6, 11], and general plain radiography [9]. Efficiency and workflow outcomes reported across the included studies are presented in Table 2 [5,6,9,11].

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Risk of bias assessment

The risk of bias among the included studies varied depending on study design. The longitudinal study by Hong et al. (2026) [5] demonstrated a low risk of bias due to its prospective design, large sample size (756 radiographs), and clear outcome measures; however, being a single-center study, it had moderate applicability concerns. The prospective cohort study by Huang et al. (2025) [9] was assessed as having a low risk of bias due to its large sample size (23,960 radiographs) and matched-control design. The comparative study by Rajmohamed et al. (2025) [6] showed a moderate risk of bias due to its single-center design and lack of blinding, though objective efficiency metrics strengthened its validity.
The systematic review by Sacoransky et al. (2024) [7] was assessed using AMSTAR-2 and demonstrated high quality with a comprehensive search strategy and clear inclusion criteria, though it was limited by the heterogeneity of the primary studies included. The scoping review protocol by Feng et al. (2025) [8] was assessed as having a low risk of bias due to its methodological rigor, though, as a protocol, outcomes are not yet available. The technical implementation study by Mehdiratta et al. (2025) [10] had a low risk of bias for its specific objective of technical integration, though generalizability to clinical outcomes is limited. The patient-centered reporting study by Park et al. (2024) [11] demonstrated a moderate risk of bias due to the subjective nature of comprehension assessment, though objective metrics for hallucinations strengthened its validity.

Results of syntheses

The narrative synthesis was organized around four key outcome domains: efficiency, diagnostic accuracy, report standardization, and patient-centered communication. For each synthesis, the characteristics and risk of bias of contributing studies were considered.
Efficiency and Workflow Integration: Across five studies reporting efficiency outcomes [5, 6, 8, 9, 11], AI-assisted reporting consistently improved reporting times. Hong et al. (2026) [5] (low risk of bias) demonstrated a reduction in mean reading time from 25.8 to 19.3 seconds across sequential batches of chest radiographs. Huang et al. (2025) [9] (low risk of bias) reported a 15.5% decrease in reporting time (189.2 to 159.8 seconds) for plain radiographs. Rajmohamed et al. (2025) [6] (moderate risk of bias) observed a reduction from 6.1 to 3.43 minutes for complex imaging studies. Park et al. (2024) [11] (moderate risk of bias) reported improved workflow efficiency in spine MRI reporting.
Diagnostic Accuracy and Report Quality: Four studies evaluated diagnostic accuracy or report quality [5, 6, 9, 11]. Hong et al. (2026) [5] (low risk of bias) found stable RADPEER agreement and quality scores for normal chest radiographs but significant variability for abnormal cases ($P < .05$). Huang et al. (2025) [9] (low risk of bias) reported no significant difference in peer-reviewed accuracy ($P = .41$) or textual quality ($P = .06$) between AI-assisted and standard reports. Rajmohamed et al. (2025) [6] (moderate risk of bias) noted occasional anatomical misinterpretations that were easily correctable. Park et al. (2024) [11] (moderate risk of bias) reported artificial hallucinations in 1.12% of cases and potentially harmful translations in 7.40%. Diagnostic accuracy and report quality findings are summarized in Table 3 [5,6,9,11].

Report standardization: Two studies addressed report standardization [7, 10]. Sacoransky et al. (2024) [7] (low risk of bias systematic review) synthesized evidence that transformer-based models effectively convert unstructured reports into structured formats. Mehdiratta et al. (2025) [10] (low risk of bias for technical objective) successfully demonstrated integration of AI-derived measurements into reports using Common Data Elements (CDEs). Outcomes related to structured reporting, quantitative data integration, and patient-centered communication are shown in Table 4 [7,10,11].
Patient-Centered Communication: One primary study [11] (moderate risk of bias) specifically evaluated patient-centered outcomes, demonstrating that patient-friendly report formats achieved significantly higher understanding scores (4.69 ± 0.48) compared with conventional reports (2.71 ± 0.73) ($p < 0.001$).

Table 1. Characteristics of included studies (n = 8)

Author (Year)	Country	Study Design	Imaging Modality	Sample Size	AI System Type
Hong et al., 2026	USA	Longitudinal multireader	Chest radiographs	756	Multimodal radiology-specific AI
Rajmohamed et al., 2025	UK	Comparative study	MRI, CT	100	Semi-automated AI reporting platform
Huang et al., 2025	USA	Prospective cohort	Plain radiographs	23,960	Generative AI draft reporting

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Park et al., 2024	South Korea	Primary study	Spine MRI	685	Generative AI language model
Mehdiratta et al., 2025	USA	Technical implementation	Abdominal CT	Not a clinical cohort	AI segmentation + CDE integration
Sacoransky et al., 2024	Multinational	Systematic review	Multiple	8 included studies	Transformer-based LLMs
Feng et al., 2025	Switzerland	Scoping review protocol	Multiple	NA	Vision–language models

Table 2. Efficiency and workflow outcomes

Author	Reporting Time Reduction	Statistical Significance	Additional Workflow Findings
Hong et al., 2026	25.8 → 19.3 sec	P < .001	Progressive learning effect
Huang et al., 2025	189.2 → 159.8 sec (15.5%)	P = .02	Maintained quality
Rajmohamed et al., 2025	6.1 → 3.43 min	P < .0001	Improved radiologist confidence
Park et al., 2024	Improved efficiency	Not specified	Faster structured output

Table 3. Diagnostic accuracy and report quality

Author	Accuracy Outcome	Statistical Findings	Key Observations
Hong et al., 2026	Stable RADPEER (normal cases)	P < .05 for abnormal variability	Lower reliability in complex cases
Huang et al., 2025	No difference vs standard	P = .41 (accuracy), P = .06 (quality)	Comparable clinical performance
Rajmohamed et al., 2025	Accuracy improved (3.81 → 4.65)	P < .0001	Minor anatomical misinterpretations
Park et al., 2024	Hallucinations 1.12%	—	7.40% potentially harmful translations

Table 4. Standardization and patient-centered reporting

Author	Structured Reporting	Quantitative Integration	Patient Comprehension
Sacoransky et al., 2024	Effective structured conversion	Yes	Not evaluated
Mehdiratta et al., 2025	CDE-based structured integration	Yes (DICOM-SR)	Not evaluated
Park et al., 2024	Structured summaries	Yes	2.71 → 4.69 (P < .001)

Discussion

This systematic review synthesizes evidence regarding radiologists' interaction with AI-generated preliminary and structured reports, highlighting both the promise and the challenges of AI integration into clinical radiology workflows. Across the studies reviewed, several consistent themes emerge, including improvements in reporting efficiency, maintenance or enhancement of diagnostic accuracy, variability in performance for complex cases, and the potential to improve patient-centered communication. A nuanced examination of these findings provides insight into the broader implications of AI-assisted radiology practice.

Efficiency and workflow integration

A recurrent finding across multiple studies is the significant reduction in reporting time associated with AI-assisted workflows. Hong et al. (2026) demonstrated that

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

radiologists interpreting 756 chest radiographs with AI-generated preliminary reports reduced their mean reading time from 25.8 to 19.3 seconds across sequential batches, suggesting not only immediate efficiency gains but also a progressive learning effect in human--AI collaboration. Similarly, Rajmohamed et al. (2025) observed a reduction in reporting time from 6.1 to 3.43 minutes for complex imaging studies, and Huang et al. (2025) reported a 15.5% decrease in reporting time for plain radiographs, without compromising clinical accuracy. These findings collectively underscore that AI integration can streamline radiology workflows, particularly for routine or high-volume examinations, and may alleviate pressures associated with increasing imaging demands [5-7].

The consistency of efficiency improvements across both standard radiographs and complex imaging modalities indicates that AI can function as an effective adjunct in diverse clinical contexts. Notably, studies reported progressive adaptation, suggesting that radiologists may develop increased trust and familiarity with AI-generated outputs over time, which enhances workflow synergy [12].

Diagnostic accuracy and oversight requirements

While efficiency benefits are clear, the review reveals that the impact of AI on diagnostic accuracy varies depending on case complexity. Hong et al. (2026) found stable agreement and quality scores for normal chest radiographs but significant variability for abnormal cases, highlighting the limits of AI reliability in complex scenarios. Rajmohamed et al. (2025) similarly noted occasional anatomical misinterpretations, which were correctable by radiologists, reinforcing that expert oversight remains essential. Huang et al. (2025) reported no significant difference in peer-reviewed accuracy between AI-assisted and standard reporting, suggesting that AI can maintain clinical quality in controlled settings, but its sensitivity in detecting unexpected urgent findings (e.g., pneumothorax) underscores a complementary, rather than substitutive, role [13,14].

This pattern suggests a central principle: AI serves best as an efficiency and consistency enhancer, particularly for routine or structured tasks, but does not yet obviate the need for human expertise, particularly in complex or abnormal cases [15]. The variability in abnormal case interpretation also highlights the importance of ongoing performance evaluation and model refinement tailored to clinically challenging scenarios.

Report standardization and structured data integration

Several studies emphasized AI's potential to standardize radiology reporting. Sacoransky et al. (2024) highlighted transformer-based language models, such as ChatGPT, in converting unstructured reports into structured formats, extracting key data, and generating impression sections, which improves report uniformity and accessibility. Mehdiratta et al. (2025) further demonstrated technical integration of AI-derived quantitative measurements into radiology reports using Common Data Elements (CDEs), promoting interoperability and consistent communication with referring clinicians [16].

These findings suggest that AI can not only enhance efficiency but also support structured, data-rich reporting that benefits downstream clinical decision-making and research. Standardization may be particularly valuable in large-scale or multicenter settings, where consistency of measurements and terminology is critical.

Patient-centered communication

Park et al. (2024) provide evidence that AI-generated reports can also improve patient comprehension without compromising clinical accuracy. Patient-friendly report formats achieved significantly higher understanding scores compared with conventional reports, highlighting AI's potential to bridge communication gaps between clinicians and patients. However, the occurrence of artificial hallucinations (1.12%) and potentially harmful translations (7.40%) indicates that AI outputs must be validated and carefully reviewed to prevent miscommunication or adverse outcomes [17].

Patterns, inconsistencies, and limitations

A clear pattern emerges wherein AI consistently improves efficiency and standardization, yet its performance is less predictable in complex or abnormal cases. This variability likely reflects limitations in training datasets, model generalizability, and domain-specific knowledge. Transformer-based language models, while effective in structured reporting, raise recurrent concerns regarding reliability, potential medical errors, and data privacy (Sacoransky et al., 2024).

The reviewed studies also differ in methodology and scope. Longitudinal studies (Hong et al., 2026) capture adaptation and learning effects over time, while prospective cohort studies (Huang et al., 2025) assess real-world workflow integration. Comparative studies (Rajmohamed et al., 2025) evaluate AI performance against traditional dictation, providing granular insights into radiologist experience and confidence. Scoping

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

reviews (Feng et al., 2025) and systematic reviews (Sacoransky et al., 2024) synthesize broader trends but may include heterogeneous methodologies, limiting direct comparability. The diversity of imaging modalities, case complexities, and AI architectures adds further variability, which must be considered when generalizing conclusions [18-20].

Implications for clinical practice

Collectively, these findings suggest that AI-generated preliminary reports can enhance radiology efficiency, standardization, and patient-centered communication, while maintaining diagnostic quality under human oversight [21,22]. Adoption of AI should be approached as an augmentative strategy, emphasizing workflow integration, radiologist training, and quality assurance, particularly for abnormal or complex cases. Standardized frameworks for data integration (e.g., CDEs) and patient-friendly reporting formats may further enhance clinical utility [23,24].
However, implementation requires careful attention to potential pitfalls, including data privacy, AI hallucinations, and domain-specific limitations. Regulatory oversight, validation in diverse clinical populations, and continuous monitoring of AI performance will be essential to ensure safe and effective deployment.

Conclusion

The systematic review indicates that AI-assisted radiology reporting represents a valuable adjunct to human expertise, improving efficiency, standardization, and patient comprehension while maintaining diagnostic accuracy. The findings highlight a complementary model of human–AI collaboration, in which AI handles routine, structured, or high-volume tasks and radiologists provide critical oversight, particularly in complex or abnormal cases. Future research should focus on optimizing AI training for complex pathologies, validating patient-centered reporting formats, and developing robust ethical and regulatory frameworks for clinical implementation.

Registration and protocol

This systematic review was not registered in any public registry, and a review protocol was not prepared before commencing the study. The absence of registration and a priori protocol represents a limitation of this review, as it reduces transparency and increases the potential for reporting bias. However, the review adhered strictly to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines throughout the conduct and reporting phases to maintain methodological rigor. No amendments were made to the methodology during the review process, as no protocol existed against which to document changes.

Acknowledgments

The authors would like to thank the Department of Radiology and the Department of Dentistry at PSP Medical College Hospital and Research Institute for their institutional support. The authors also acknowledge the library services staff for their assistance with literature retrieval and the systematic review methodology support provided by the institutional research committee.

List of abbreviations

AI - Artificial Intelligence
AMSTAR-2- A Measurement Tool to Assess Systematic Reviews
CDE - Common Data Element
CT - Computed Tomography
DICOM-SR Digital Imaging and Communications in Medicine - Structured Reporting
GRADE - Grading of Recommendations, Assessment, Development, and Evaluations
MRI - Magnetic Resonance Imaging
PACS - Picture Archiving and Communication System
PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR - Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
QUADAS-2- Quality Assessment of Diagnostic Accuracy Studies

Source of funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was conducted as part of the authors' academic responsibilities at their respective institutions.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

Availability of data

All data analyzed in this systematic review are derived from published articles cited in the reference section. The extracted data, summarized in the results table, are fully presented within this article. Any additional inquiries can be directed to the corresponding author.

Author's contribution

Dr. Priyadarshini Subramani: Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization.
Dr. Karthik Shunmugavelu: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Supervision, Project Administration, Corresponding author.
Both authors read and approved the final manuscript.

Author's biography

Dr. Priyadarshini Subramani, MD, Radiology
Dr. Priyadarshini Subramani is an Assistant Professor in the Department of Radiology at PSP Medical College Hospital and Research Institute, Tamilnadu, India. She completed her MD in Radiology and has clinical and academic expertise in diagnostic imaging, with research interests in artificial intelligence applications in radiology, musculoskeletal imaging, and neuroradiology. She is actively involved in postgraduate teaching and curriculum development in radiology education.
Dr. Karthik Shunmugavelu, BDS, MDS OMFP, MSC LONDON, MFDSRCS ENGLAND, MFDSRCPS GLASGOW, FACULTY AFFILIATE RCS IRELAND, AFFILIATE RCS EDINBURGH, ASSOCIATE FACULTY OF DENTAL TRAINERS EDINBURGH, MCIP, FIBMS USA, MASID AUSTRALIA
Dr. Karthik Shunmugavelu is a Senior Resident in the Department of Dentistry at PSP Medical College Hospital and Research Institute, Tamilnadu, India. He holds multiple international qualifications in oral medicine and facial pathology, including an MSc from the University of London and fellowships from the Royal College of Surgeons of England, Glasgow, and Edinburgh. His research interests include artificial intelligence in healthcare, medical education, and interdisciplinary collaboration between radiology and dentistry. He has authored several publications in peer-reviewed journals and serves as a faculty affiliate for multiple international surgical colleges.

References

1. Gorelik N, Gyftopoulos S. Applications of Artificial Intelligence in Musculoskeletal Imaging: From the Request to the Report. Canadian Association of Radiologists Journal. 2020;72(1):45-59. https://doi.org/10.1177/0846537120947148 PMid:32809857
2. Al-Naser Y, Sharma S, Niure K, Ibach K, Khosa F, Yong-Hing CJ. Geographic prompting and content fidelity in generative Artificial Intelligence: A multi-model study of demographics and imaging equipment in AI-generated videos and images of Canadian medical radiation technologists. Journal of Medical Imaging and Radiation Sciences. 2025 Dec 1;56(6):102122.https://doi.org/10.1016/j.jmir.2025.102122 PMid:41110417
3. Archer H, Reine S, Alshaikhsalama A, Wells J, Kohli A, Vazquez L, et al. Artificial intelligence-generated hip radiological measurements are fast and adequate for reliable assessment of hip dysplasia. Bone Jt Open. 2022 Nov 14;3(11):877-884. https://doi.org/10.1302/2633-1462.311.BJO-2022-0125.R1 PMid:36373773 PMCid:PMC9709495
4. Zhang S, Sun J, Liu C, Fang J, Xie H, Ning B. Clinical application of artificial intelligence-assisted diagnosis using anteroposterior pelvic radiographs in children with developmental dysplasia of the hip. Bone Joint J. 2020 Nov 2;102-B(11):1574-1581. https://doi.org/10.1302/0301-620X.102B11.BJJ-2020-0712.R2 PMid:33135455
5. Hong, E. K., Suh, C. H., Nukala, M., Esfahani, A., Licaros, A., Madan, R., Hunsaker, A., & Hammer, M. (2026). Radiologist Interaction with Artificial Intelligence-Generated Preliminary Reports: A Longitudinal Multireader Study. Journal of the American College of Radiology: JACR, 23(2), 292-298. https://doi.org/10.1016/j.jacr.2025.09.015 PMid:40983179
6. Rajmohamed, R. F., Chapala, S., Shazahan, M. A., Wali, P., & Botchu, R. (2025). Evaluating the Accuracy and Efficiency of AI-Generated Radiology Reports Based on Positive Findings: A Qualitative Assessment of AI in Radiology. Academic radiology, 32(12), 7035-7040.

Student's Journal of Health Research Africa
e-ISSN: 2709-9997, p-ISSN: 3006-1059
Vol.3 No. 12 (2022): December 2022 Issue
https://doi.org/10.51168/sjhrafrica.v3i12.2460
Review Article

https://doi.org/10.1016/j.acra.2025.09.012 PMid:41015710
7. Sacoransky, E., Kwan, B. Y. M., & Soboleski, D. (2024). ChatGPT and assistive AI in structured radiology reporting: A systematic review. Current problems in diagnostic radiology, 53(6), 728-737. https://doi.org/10.1067/j.cpradiol.2024.07.007 PMid:39004580
8. Feng, W., Yazdani, A., Bornet, A., Platon, A., & Teodoro, D. (2025). Synthesising evidence regarding artificial intelligence-generated radiological reports based on medical images: a scoping review protocol. BMJ open, 15(10), e104112. https://doi.org/10.1136/bmjopen-2025-104112 PMid:41043831 PMCid:PMC12496096
9. Huang J, Wittbrodt MT, Teague CN, et al. Efficiency and Quality of Generative AI-Assisted Radiograph Reporting. JAMA Netw Open. 2025;8(6):e2513921. https://doi.org/10.1001/jamanetworkopen.2025.13921 PMid:40471579 PMCid:PMC12142447
10. Mehdiratta, G., Duda, J.T., Elahi, A. et al. Automated Integration of AI Results into Radiology Reports Using Common Data Elements. J Digit Imaging. Inform. med. 38, 2623-2629 (2025). https://doi.org/10.1007/s10278-025-01414-9 PMid:39871037 PMCid:PMC12572421
11. Park, J., Oh, K., Han, K., & Lee, Y. H. (2024). Patient-centered radiology reports with generative artificial intelligence: adding value to radiology reporting. Scientific reports, 14(1), 13218. https://doi.org/10.1038/s41598-024-63824-z PMid:38851825 PMCid:PMC11162416
12. Hartung, M. P., Bickle, I. C., Gaillard, F., & Kanne, J. P. (2020). How to Create a Great Radiology Report. Radiographics: a review publication of the Radiological Society of North America, Inc, 40(6), 1658-1670. https://doi.org/10.1148/rg.2020200020 PMid:33001790
13. Mityul, M. I., Gilcrease-Garcia, B., Mangano, M. D., Demertzis, J. L., & Gunn, A. J. (2018). Radiology Reporting: Current Practices and an Introduction to Patient-Centered Opportunities for Improvement. AJR. American journal of roentgenology, 210(2), 376-385. https://doi.org/10.2214/AJR.17.18721 PMid:29140114
14. Veras Magalhães, G., L. de S. Santos, R., H. S. Vogado, L., Cardoso de Paiva, A., & de Alcântara dos Santos Neto, P. (2024). XRaySwinGen: Automatic medical reporting for X-ray exams with a multimodal model. Heliyon, 10(7), e27516. https://doi.org/10.1016/j.heliyon.2024.e27516 PMid:38560155 PMCid:PMC10979158
15. Hwang, E. J., Goo, J. M., Nam, J. G., Park, C. M., Hong, K. J., & Kim, K. H. (2023). Conventional Versus Artificial Intelligence-Assisted Interpretation of Chest Radiographs in Patients with Acute Respiratory Symptoms in the Emergency Department: A Pragmatic Randomized Clinical Trial. Korean journal of radiology, 24(3), 259-270. https://doi.org/10.3348/kjr.2022.0651 PMid:36788769 PMCid:PMC9971841
16. Haider, S. A., Prabha, S., Gomez-Cabello, C. A., Genovese, A., Collaco, B., Wood, N., Lifson, M. A., Bagaria, S., Tao, C., & Forte, A. J. (2025). Artificial Intelligence Physician Avatars for Patient Education: A Pilot Study. Journal of Clinical Medicine, 14(23), 8595. https://doi.org/10.3390/jcm14238595 PMid:41375898 PMCid:PMC12692777
17. Zalake M. (2023). Doctors' perceptions of using their digital twins in patient care. Scientific reports, 13(1), 21693. https://doi.org/10.1038/s41598-023-48747-5 PMid:38066016 PMCid:PMC10709415
18. Ye, J. Artificial intelligence-generated content (AIGC) in biomedical research, healthcare delivery, and clinical practices: technologies, applications, and regulatory considerations. Artif Intell Rev 59, 86 (2026). https://doi.org/10.1007/s10462-025-11487-1
19. Hong, E. K., Roh, B., Park, B., Jo, J. B., Bae, W., Soung Park, J., & Sung, D. W. (2025). Value of Using a Generative AI Model in Chest Radiography Reporting: A Reader Study. Radiology, 314(3), e241646. https://doi.org/10.1148/radiol.241646 PMid:40067108
20. Lee, J. H., Sun, H. Y., Park, S., Kim, H., Hwang, E. J., Goo, J. M., & Park, C. M. (2020). Performance of a Deep Learning Algorithm Compared with Radiologic Interpretation for Lung Cancer Detection on Chest Radiographs in a Health Screening Population. Radiology, 297(3), 687-696. https://doi.org/10.1148/radiol.2020201240 PMid:32960729

21. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. The Lancet. 2020 May 16;395(10236): 1579-86.doi not available https://doi.org/10.1016/S0140-6736(20)30226-9 PMid:32416782 PMCid:PMC7255280
22. Terranova C, Cestonaro C, Fava L, Cinquetti A. AI and professional liability assessment in healthcare. A revolution in legal medicine. Frontiers in Medicine. 2024 Jan 8; 10:1337335. https://doi.org/10.3389/fmed.2023.1337335 PMid:38259835 PMCid:PMC10800912
23. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. Journal of medical Internet research. 2020 Jun 19;22(6):e15154. https://doi.org/10.2196/15154 PMid:32558657 PMCid:PMC7334754
24. Sandeep Reddy, Sonia Allan, Simon Coghlan, Paul Cooper, A governance model for the application of AI in health care, Journal of the American Medical Informatics Association, Volume 27, Issue 3, March 2020, Pages 491-497, https://doi.org/10.1093/jamia/ocz192 PMid:31682262 PMCid: PMC7647243

PUBLISHER DETAILS

Student's Journal of Health Research (SJHR)
(ISSN 2709-9997) Online
(ISSN 3006-1059) Print
Category: Non-Governmental & Non-profit Organization
Email: studentsjournal2020@gmail.com
WhatsApp: +256 775 434 261
Location: Scholar's Summit Nakigalala, P. O. Box 701432, Entebbe Uganda, East Africa