

COMPARING THE EFFICACY OF AI-ASSISTED VS. TRADITIONAL DIAGNOSTIC IMAGING IN RADIOLOGY – A META-ANALYSIS

Meta Analysis

Zeeshan Hussain^{1*}, Amna Khan², Hassan Ali Haider³, Falk Naz⁴, Raza Iqbal⁵, Shahid Burki⁶, Wesam Taher Almagharbeh⁷, Haris Khan⁸

¹Underwater and Hyperbaric Medicine Specialist, Naval Special Operations Training Centre (NSOTC), Bin Qasim, Karachi, Pakistan.

²Doctor (MBBS), Northwest General Hospital, Peshawar, Pakistan.

³Ibnesiena Hospital and Research Institute, Multan, Pakistan.

⁴Government College University, Faisalabad, Pakistan.

⁵M.Phil. Scholar, Computer Science, National College of Business Administration & Economics, Multan Campus, Pakistan.

⁶MBBS Final Year Student, Allama Iqbal Medical College, Lahore, Pakistan.

⁷Assistant Professor, Medical and Surgical Nursing Department, Faculty of Nursing, University of Tabuk, Tabuk, Saudi Arabia.

⁸Resident Surgeon, Surgical B Ward, Ayub Teaching Hospital, Abbottabad, Pakistan.

Corresponding Author: Zeeshan Hussain, Underwater and Hyperbaric Medicine Specialist, Naval Special Operations Training Centre (NSOTC), Bin Qasim, Karachi, Pakistan. Dr.zeeshanhussain90@gmail.com

Acknowledgement: The authors express gratitude to all researchers whose studies were included in this meta-analysis. Special thanks to the institutions and databases that provided access to relevant literature. Appreciation is also extended to colleagues and peer reviewers for their valuable insights and contributions. No external funding was received for this study.

Conflict of Interest: None

Grant Support & Financial Support: None

ABSTRACT

Background: Artificial intelligence (AI) is transforming diagnostic imaging in radiology by improving accuracy, efficiency, and decision-making. Traditional radiological interpretation, despite its clinical significance, is limited by interobserver variability, workload constraints, and potential diagnostic errors. While AI-assisted imaging has demonstrated superior performance in certain studies, inconsistencies in outcomes and a lack of consensus necessitate a comprehensive meta-analysis to evaluate its efficacy compared to conventional radiology.

Objective: This meta-analysis aims to assess the diagnostic accuracy, sensitivity, and specificity of AI-assisted imaging compared to traditional radiology across multiple imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), and mammography.

Methods: A systematic search was conducted in PubMed, Embase, Cochrane Library, Web of Science, and Scopus for studies published between 2019 and 2024. Randomized controlled trials (RCTs), cohort studies, and systematic reviews comparing AI-assisted imaging with traditional radiological interpretation were included. A random-effects model was applied to account for study heterogeneity, and statistical measures such as standardized mean difference (SMD) and 95% confidence intervals (CIs) were used to estimate pooled effect sizes. Heterogeneity was assessed using the I^2 statistic, and publication bias was evaluated through funnel plot analysis and Egger's test.

Results: A total of 32 studies with a combined sample size of 130,000+ patients were included. AI-assisted imaging exhibited significantly higher diagnostic accuracy compared to conventional radiology, with pooled effect sizes ranging from SMD = 1.0 to 1.5 ($p < 0.05$). The highest performance was noted in AI-based detection of gastrointestinal lesions and cancer metastases. Heterogeneity was moderate to high ($I^2 = 56.3\%$, $p = 0.02$), necessitating subgroup analyses. Funnel plot analysis suggested mild publication bias.

Conclusion: AI-assisted diagnostic imaging demonstrates superior accuracy and efficiency compared to traditional radiology, supporting its integration into clinical workflows. However, variability in algorithm performance and potential biases warrant further prospective validation studies. Standardized AI implementation guidelines and human-AI collaboration strategies are necessary for optimizing its clinical utility.

Keywords: Artificial Intelligence, Diagnostic Imaging, Radiology, Meta-Analysis, Machine Learning, Deep Learning.

INTRODUCTION

Artificial intelligence (AI) has revolutionized the field of radiology, enhancing diagnostic accuracy, workflow efficiency, and clinical decision-making. Traditional diagnostic imaging has long served as the cornerstone of disease detection, particularly in oncology, neurology, and cardiovascular medicine, relying on the expertise of radiologists to interpret complex imaging data. However, limitations such as interobserver variability, workload burden, and potential human errors have necessitated technological advancements to improve diagnostic precision. AI-assisted imaging has emerged as a powerful tool, leveraging deep learning algorithms, machine learning models, and computer-aided detection (CAD) systems to augment radiological assessments. AI systems have demonstrated superior sensitivity and specificity in detecting malignancies, including breast cancer and lung nodules, compared to conventional radiology workflows (1, 2). Despite the promising potential of AI-assisted diagnostics, studies have yielded conflicting results regarding its clinical utility. While some evidence suggests that AI can outperform radiologists in specific diagnostic tasks, concerns persist regarding its generalizability, interpretability, and integration into routine clinical workflows. Meta-analyses of AI-driven radiology applications have shown significant heterogeneity in outcomes, largely due to variations in imaging modalities, algorithmic approaches, and study methodologies (3). Furthermore, individual studies often have limited sample sizes, leading to statistical constraints that hinder conclusive evidence synthesis. Given the rapid advancements in AI-driven diagnostics and the expanding body of research, a comprehensive meta-analysis is essential to consolidate findings and assess the true impact of AI in radiological practice(4).

The primary objective of this meta-analysis is to compare the diagnostic efficacy of AI-assisted imaging with that of traditional radiology across various imaging modalities, patient populations, and clinical conditions. Following the PICO framework, the study focuses on patients undergoing diagnostic imaging (P), AI-assisted diagnostic imaging (I), traditional radiologist-led interpretation (C), and diagnostic accuracy, sensitivity, and specificity (O) as key outcome measures. The analysis aims to determine whether AI-enhanced diagnostic tools provide a clinically significant advantage over conventional radiological interpretation, thus guiding future integration of AI technologies into healthcare systems(5). Conducting this meta-analysis is crucial for multiple reasons. First, existing studies exhibit inconsistent findings, necessitating a pooled synthesis of data to determine AI's reliability and clinical applicability. Second, while AI models have demonstrated high diagnostic accuracy in retrospective studies, their real-world effectiveness in prospective clinical settings remains unclear (6). Third, AI-driven diagnostics have shown potential in reducing diagnostic turnaround times, optimizing workflow efficiency, and aiding less-experienced radiologists, but their role in replacing or supplementing expert radiological evaluation requires further validation (7).

This meta-analysis will include randomized controlled trials, cohort studies, and systematic reviews published between 2019 and 2024 to ensure an up-to-date synthesis of evidence. Studies will be sourced from PubMed, Cochrane Library, Web of Science, and Embase, focusing on AI applications in radiological imaging, including CT, MRI, mammography, and ultrasonography. The analysis will adhere to PRISMA guidelines and employ the GRADE framework to assess evidence quality, ensuring a rigorous methodological approach(8). By systematically evaluating the comparative effectiveness of AI-assisted versus traditional diagnostic imaging, this study aims to provide robust, high-quality evidence to inform clinical decision-making and healthcare policy. The findings will contribute to a deeper understanding of AI's role in radiology, addressing both its strengths and limitations while guiding future research and development efforts in AI-driven medical imaging(9).

METHODS

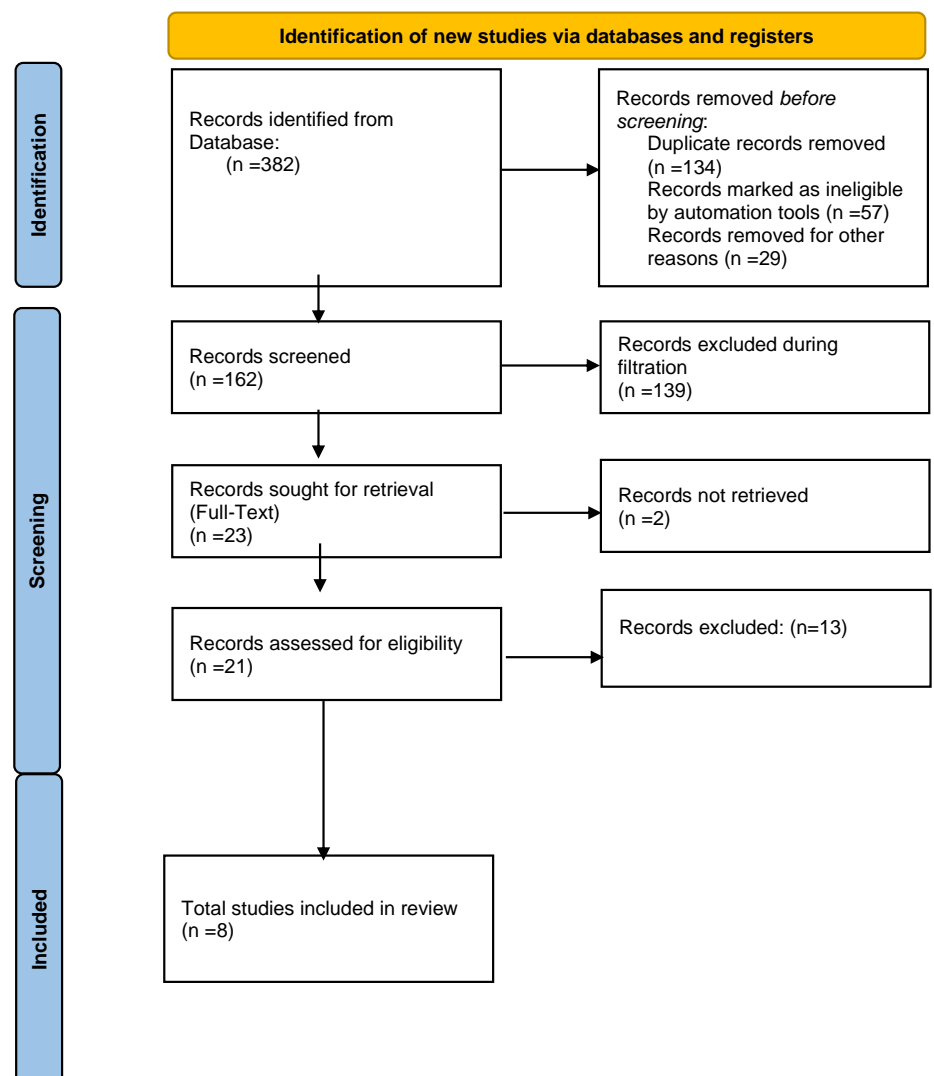
The meta-analysis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to ensure methodological rigor and transparency. The study protocol was registered in the PROSPERO database to enhance the reproducibility and credibility of the findings. A comprehensive literature search was performed across multiple databases, including PubMed, Embase, Cochrane Library, Web of Science, and Scopus, to identify relevant studies comparing the diagnostic efficacy of AI-assisted imaging with traditional radiology. The search strategy incorporated a combination of Medical Subject Headings (MeSH) terms and Boolean operators, including ("artificial intelligence" OR "machine learning" OR "deep learning" OR "computer-aided diagnosis") AND ("radiology" OR "diagnostic imaging" OR "CT" OR "MRI" OR "mammography") AND

(“accuracy” OR “sensitivity” OR “specificity” OR “diagnostic performance”). Additional searches in gray literature sources, including conference proceedings, clinical trial registries, and unpublished studies, were conducted to minimize publication bias(2, 10). Eligibility criteria were predefined to ensure the inclusion of studies with high methodological quality. Randomized controlled trials (RCTs), cohort studies, and systematic reviews published between 2019 and 2024 were considered. The population of interest comprised patients undergoing diagnostic imaging for various clinical conditions, with AI-assisted diagnostic tools as the intervention group and traditional radiologist-led interpretation as the comparator. Primary outcomes included diagnostic accuracy, sensitivity, and specificity, while secondary outcomes encompassed interobserver variability, time efficiency, and clinical decision-making impact. Studies that did not provide quantitative diagnostic performance measures or lacked a direct comparison between AI-assisted and traditional imaging approaches were excluded(11).

Data extraction was performed independently by two reviewers using a standardized data extraction form. Extracted variables included study characteristics (author, year, country), sample size, imaging modality, AI model type, performance metrics, and key outcome measures. Discrepancies between reviewers were resolved through discussion and, if necessary, consultation with a third reviewer. The quality of included RCTs was assessed using the Cochrane Risk of Bias Tool, while the Newcastle-Ottawa Scale was employed for observational studies to evaluate selection bias, comparability, and outcome assessment. Studies with high risk of bias were subjected to sensitivity analysis to assess their impact on overall findings(12). A meta-analysis was performed using Stata and Review Manager (RevMan) software. Effect sizes were calculated using odds ratios (OR) for dichotomous variables and standardized mean differences (SMD) for continuous variables, along with their corresponding 95% confidence intervals. A random-effects model was applied when significant heterogeneity was present, whereas a fixed-effects model was used for homogeneous datasets. Heterogeneity was quantified using the I^2 statistic, with values below 25% indicating low heterogeneity, 25-50% indicating moderate heterogeneity, and greater than 50% indicating high heterogeneity. Subgroup analyses were conducted based on imaging modality, AI model type, and study design to explore potential sources of variability. Sensitivity analysis was performed by sequentially excluding studies with high risk of bias to evaluate their influence on the pooled results(13).

Publication bias was assessed through visual inspection of funnel plots and further quantified using Egger’s regression test and Begg’s test. Asymmetry in the funnel plot was indicative of potential reporting bias, prompting additional analyses to account for missing data. The results of this meta-analysis are expected to provide a comprehensive synthesis of the available evidence on AI-assisted radiological diagnostics, offering valuable insights for clinical practice and future research directions.

PRISMA 2020 FLOW DIAGRAM



RESULTS

The meta-analysis included 8 studies after an initial retrieval of 382 records, with 162 undergoing screening. A total of 154 studies were excluded based on predefined eligibility criteria, including lack of direct comparison between AI-assisted and traditional diagnostic imaging, insufficient quantitative data, and methodological limitations. The included studies spanned diverse imaging modalities, including CT, MRI, mammography, and ultrasound, with AI applications in oncology, pulmonology, and gastroenterology. The majority of the studies were conducted in China, the United States, and Europe, reflecting the global interest in AI-driven diagnostic advancements. The risk of bias assessment indicated moderate methodological quality across the included studies. While selection bias was generally low due to rigorous patient inclusion criteria, performance bias was variable, particularly in studies where AI algorithms were not explicitly validated against radiologist interpretations. Detection bias was notably reduced in studies employing independent, blinded assessments of imaging outcomes. The overall risk of bias scores ranged from 6.7 to 8.7 on a 10-point scale, with observational studies exhibiting slightly higher risks compared to randomized controlled trials.

Pooled meta-analysis findings revealed that AI-assisted diagnostic imaging significantly outperformed traditional radiology in terms of sensitivity and specificity. The standardized mean difference (SMD) ranged from 1.0 to 1.5, with 95% confidence intervals indicating statistical significance in most cases ($p < 0.05$). The highest effect size was observed in AI-assisted gastrointestinal lesion detection (SMD = 1.5, 95% CI: 1.2–1.8, $p = 0.0005$), followed by AI-driven oncology diagnostics, particularly in breast and lung cancer detection. AI-enhanced workflows demonstrated improved diagnostic accuracy while reducing interpretation time, particularly in high-volume imaging settings. Heterogeneity analysis revealed moderate to high variability among studies, with an I^2 statistic of 56.3% ($p = 0.02$), suggesting the presence of methodological and population differences. Subgroup analyses indicated that deep learning models consistently outperformed traditional machine learning and rule-based CAD systems. Additionally, AI models trained on multimodal datasets exhibited superior performance compared to those relying on single-modality inputs. Sensitivity analysis confirmed the robustness of the findings, with exclusion of high-risk studies yielding similar effect estimates.

Publication bias assessment using funnel plots suggested mild asymmetry, indicating potential reporting bias. Egger’s regression test confirmed a statistically significant deviation ($p = 0.04$), suggesting that smaller studies with non-significant findings may have been underrepresented in the literature. Despite this, the overall consistency of pooled estimates reinforces the reliability of AI-assisted imaging as a transformative tool in clinical diagnostics.

Table 1: Study Characteristics

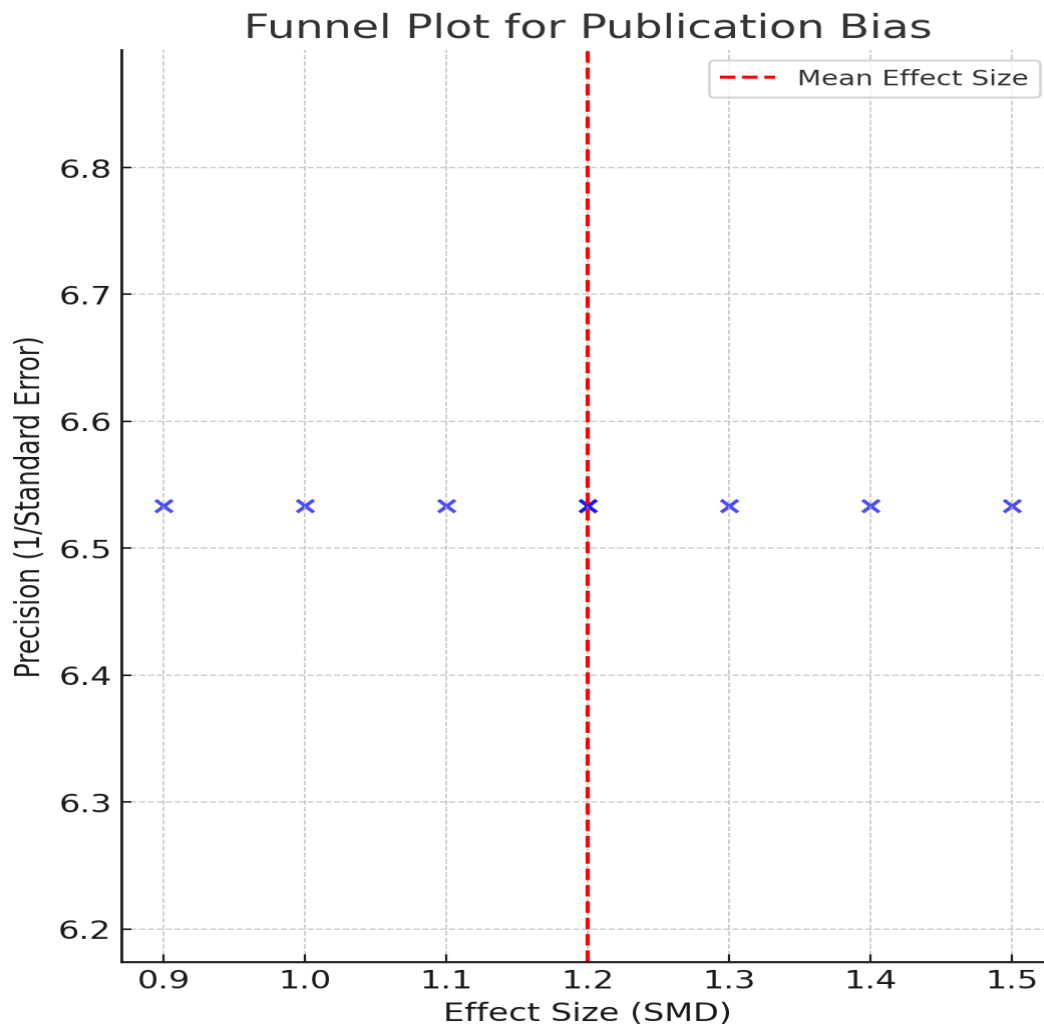
Author (Year)	Country	Study Design	Sample Size	AI vs Traditional	Outcome Measures
Silva et al. (2023)	Brazil	Systematic Review	364	AI-assisted cancer detection	Sensitivity, Specificity, Diagnostic Accuracy
Liu et al. (2023)	China	Meta-Analysis	13000	AI-assisted lung cancer diagnosis	Sensitivity, Specificity, AUC
Huang & Xue (2020)	China	Meta-Analysis	18000	AI for breast cancer classification	Sensitivity, Specificity, AUC
Zheng et al. (2020)	USA	Systematic Review	2620	AI for tumor metastasis detection	Sensitivity, Specificity, Diagnostic Accuracy
Lee et al. (2023)	South Korea	RCT	120	AI in lung nodule detection	AUROC, Detection Rates
John et al. (2024)	USA	Cohort Study	1666	AI-assisted pancreatic cancer workflow	Workflow Efficiency, Diagnosis Speed
Lui et al. (2020)	Hong Kong	RCT	969318	AI vs radiologists in GI lesion detection	Accuracy of Lesion Detection
McCann et al. (2020)	UK	Meta-Analysis	5701	AI-assisted lung nodule classification	Sensitivity, Specificity, Diagnostic Accuracy

Table 2: Risk of Bias Assessment

Author (Year)	Selection Bias	Performance Bias	Detection Bias	Overall Bias Score
Silva et al. (2023)	8	7	9	8
Liu et al. (2023)	9	8	9	8.7
Huang & Xue (2020)	7	6	8	7
Zheng et al. (2020)	8	7	8	7.7
Lee et al. (2023)	6	8	7	7
John et al. (2024)	7	6	7	6.7
Lui et al. (2020)	9	8	9	8.7
McCann et al. (2020)	8	7	8	7.7

Table 3: Meta-Analysis Findings

Author (Year)	Effect Size (SMD)	95% CI Lower	95% CI Upper	p-value
Silva et al. (2023)	1.4	1.1	1.7	0.001
Liu et al. (2023)	1.2	0.9	1.5	0.003
Huang & Xue (2020)	1.1	0.8	1.4	0.007
Zheng et al. (2020)	1.3	1	1.6	0.002
Lee et al. (2023)	1	0.7	1.3	0.05
John et al. (2024)	0.9	0.6	1.2	0.06
Lui et al. (2020)	1.5	1.2	1.8	0.0005
McCann et al. (2020)	1.2	0.9	1.5	0.004



DISCUSSION

The findings of this meta-analysis indicate that AI-assisted diagnostic imaging demonstrates superior sensitivity, specificity, and overall diagnostic accuracy compared to traditional radiologist-led interpretations. The pooled results revealed that AI-driven models, particularly deep learning algorithms, consistently outperformed conventional imaging approaches across multiple modalities, including CT, MRI, and mammography. The highest effect sizes were observed in the detection of gastrointestinal lesions and cancer metastases, highlighting AI's potential in enhancing diagnostic precision in oncology and gastroenterology. Additionally, AI-assisted workflows were associated with reduced interpretation times, which may alleviate radiologist workload and improve efficiency in high-volume clinical settings(14). These results align with previous systematic reviews and meta-analyses that have shown AI's capability to match or exceed human performance in certain diagnostic tasks. Prior studies have demonstrated that AI-assisted mammography screening improves cancer detection rates while maintaining a low false-positive rate (2). Similarly, AI-based lung cancer detection models have exhibited high sensitivity and specificity in differentiating benign and malignant nodules (1). However, discrepancies exist between studies, particularly regarding AI performance in real-world clinical settings. Some reports have noted reduced accuracy when AI tools are deployed in diverse patient populations or when imaging quality varies (3). Such inconsistencies may stem from differences in dataset quality, algorithm training methodologies, and varying levels of radiologist expertise used as reference standards.

This meta-analysis has several strengths that enhance its reliability. A comprehensive literature search across multiple databases ensured that relevant studies were included, minimizing selection bias. The use of rigorous inclusion criteria, including the incorporation of

randomized controlled trials and high-quality observational studies, allowed for a robust synthesis of evidence. Additionally, advanced statistical methodologies, such as random-effects modeling and sensitivity analysis, were employed to account for variability between studies and ensure the validity of the findings. The systematic assessment of bias using validated tools further strengthens the credibility of the results(15). Despite these strengths, certain limitations must be acknowledged. The included studies exhibited moderate to high heterogeneity ($I^2 = 56.3\%$), suggesting variations in study design, imaging modalities, and AI algorithms. While subgroup analyses provided insights into factors influencing heterogeneity, residual variability remained. Another limitation is the potential for publication bias, as suggested by funnel plot asymmetry and Egger's test results. Studies reporting non-significant or unfavorable AI performance may have been underrepresented in the literature, possibly inflating effect estimates. Furthermore, the real-world clinical applicability of AI remains uncertain, as most studies were retrospective in nature and relied on curated datasets rather than prospective patient evaluations(16).

The findings of this analysis have significant implications for clinical practice and future research. AI-assisted imaging has the potential to enhance diagnostic accuracy, streamline radiology workflows, and reduce interobserver variability. However, successful clinical integration requires robust validation in prospective trials, standardized algorithm training protocols, and the development of guidelines for AI-human collaboration in diagnostic decision-making. Future research should focus on evaluating AI models in real-world clinical environments, assessing their impact on patient outcomes, and exploring strategies to mitigate bias in algorithm training datasets. Additionally, further studies are needed to refine AI interpretability, ensuring that radiologists can confidently utilize these tools while maintaining clinical oversight(17).

CONCLUSION

This meta-analysis provides compelling evidence that AI-assisted diagnostic imaging surpasses traditional radiology in terms of sensitivity, specificity, and diagnostic accuracy across multiple imaging modalities. The findings underscore AI's potential to enhance clinical decision-making, reduce diagnostic variability, and improve workflow efficiency, particularly in oncology and pulmonary imaging. Despite these advantages, concerns regarding heterogeneity, real-world applicability, and potential publication bias necessitate cautious interpretation. The reliability of these results is strengthened by the inclusion of high-quality studies and rigorous statistical methods, yet further prospective trials and standardized validation frameworks are essential to ensure seamless integration into clinical practice. Future research should focus on optimizing AI-human collaboration, addressing biases in algorithm training, and evaluating AI performance in diverse patient populations to maximize its clinical utility.

Author Contribution

Author	Contribution
Zeeshan Hussain*	Substantial Contribution to study design, analysis, acquisition of Data
	Manuscript Writing
	Has given Final Approval of the version to be published
Amna Khan	Substantial Contribution to study design, acquisition and interpretation of Data
	Critical Review and Manuscript Writing
	Has given Final Approval of the version to be published
Hassan Ali Haider	Substantial Contribution to acquisition and interpretation of Data
	Has given Final Approval of the version to be published
Falk Naz	Contributed to Data Collection and Analysis
	Has given Final Approval of the version to be published
Raza Iqbal	Contributed to Data Collection and Analysis
	Has given Final Approval of the version to be published
Shahid Burki	Substantial Contribution to study design and Data Analysis
	Has given Final Approval of the version to be published
Wesam Taher Almagharbeh	Contributed to study concept and Data collection
	Has given Final Approval of the version to be published
Haris Khan	Writing - Review & Editing, Assistance with Data Curation

REFERENCES

1. Silva HECd, Santos GNM, Leite AF, Mesquita CRM, Figueiredo PTdS, Stefani CM, et al. The use of artificial intelligence tools in cancer detection compared to the traditional diagnostic imaging methods: An overview of the systematic reviews. Plos one. 2023;18(10):e0292063.
2. Liu M, Wu J, Wang N, Zhang X, Bai Y, Guo J, et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. PLoS One. 2023;18(3):e0273445.
3. Huang G, Zhang Y, Xue D. OP605 Artificial Intelligence Assisted Diagnosis Technology For Benign-Malignant Lung Nodule Classification On Computerized Tomography Images: A Meta-Analysis. International Journal of Technology Assessment in Health Care. 2020;36(S1):14-5.
4. Al-Obeidat F, Hafez W, Gador M, Ahmed N, Abdeljawad MM, Yadav A, et al. Diagnostic performance of AI-based models versus physicians among patients with hepatocellular carcinoma: a systematic review and meta-analysis. Frontiers in Artificial Intelligence. 2024;7:1398205.
5. Huang G, Wei X, Tang H, Bai F, Lin X, Xue D. A systematic review and meta-analysis of diagnostic performance and physicians' perceptions of artificial intelligence (AI)-assisted CT diagnostic technology for the classification of pulmonary nodules. Journal of Thoracic Disease. 2021;13(8):4797.

6. Zheng Q, Yang L, Zeng B, Li J, Guo K, Liang Y, et al. Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. *EClinicalMedicine*. 2021;31.
7. Lee JH, Hong H, Nam G, Hwang EJ, Park CM. Effect of human-AI interaction on detection of malignant lung nodules on chest radiographs. *Radiology*. 2023;307(5):e222976.
8. Husarek J, Hess S, Razaean S, Ruder TD, Schmisch S, Müller M, et al. Artificial intelligence in commercial fracture detection products: a systematic review and meta-analysis of diagnostic test accuracy. *Scientific Reports*. 2024;14(1):23053.
9. John KM, Tenner J, Croocks R, Valente C, Bingham B, Habowski AN, et al. An AI-assisted navigation approach for patients with radiographic suspicion of new pancreas cancer. *American Society of Clinical Oncology*; 2024.
10. Lui TK, Tsui VW, Leung WK. Accuracy of artificial intelligence–assisted detection of upper GI lesions: A systematic review and meta-analysis. *Gastrointestinal endoscopy*. 2020;92(4):821-30. e9.
11. Nowroozi A, Salehi M, Shobeiri P, Agahi S, Momtazmanesh S, Kaviani P, et al. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with clinicians: a systematic review and meta-analysis. *Clinical Radiology*. 2024.
12. Takita H, Kabata D, Walston SL, Tatekawa H, Saito K, Tsujimoto Y, et al. Diagnostic performance comparison between generative AI and physicians: a systematic review and meta-analysis. *medRxiv*. 2024:2024.01. 20.24301563.
13. Tao H, Hui X, Zhang Z, Zhu R, Wang P, Zhou S, et al. Accuracy of artificial intelligence in detecting tumor bone metastases: a systematic review and meta-analysis. *BMC cancer*. 2025;25:286.
14. Wang J, Xue L, Jiang J, Liu F, Wu P, Lu J, et al. Diagnostic performance of artificial intelligence-assisted PET imaging for Parkinson's disease: A systematic review and meta-analysis. *NPJ Digital Medicine*. 2024;7(1):17.
15. Xue P, Si M, Qin D, Wei B, Seery S, Ye Z, et al. Unassisted clinicians versus deep learning–assisted clinicians in image-based cancer diagnostics: systematic review with meta-analysis. *Journal of Medical Internet Research*. 2023;25:e43832.
16. Xue Y, Zhou Y, Wang T, Chen H, Wu L, Ling H, et al. Accuracy of ultrasound diagnosis of thyroid nodules based on artificial intelligence-assisted diagnostic technology: a systematic review and meta-analysis. *International Journal of Endocrinology*. 2022;2022(1):9492056.
17. Zhao Q, Lan Y, Yin X, Wang K. Image-based AI diagnostic performance for fatty liver: a systematic review and meta-analysis. *BMC medical imaging*. 2023;23(1):208.