International Journal of Medical
and Pharmaceutical Research

Online ISSN-2958-3683 | Print ISSN-2958-3675
Frequency: Bi-Monthly
Website: https://ijmpr.in/

Systematic Review

Pooled Diagnostic Accuracy of Artificial Intelligence in Early Detection of Solid Tumors: Evidence from a Systematic Review and Meta-Analysis

Sayantani Ghosh Hazra[1*], Akinchan Jain[2], Akshara Jain[3]

[1]Assistant Professor, Department of Pathology, Jagannath Gupta Institute of Medical Science and Hospital, Central Campus, Kolkata, West Bengal, India
[2]Resident Doctor, Department of General Surgery, Mahatma Gandhi Memorial Medical College, Indore, Madhya Pradesh, India
[3]Resident Doctor, Department of Radiotherapy, Indira Gandhi Medical College, Shimla, Himachal Pradesh, India

Corresponding Author:

Sayantani Ghosh Hazra
Assistant Professor,
Department of Pathology,
Jagannath Gupta Institute of Medical Science and Hospital,
Central Campus, Kolkata, West Bengal, India
Email:
dr.sayantanighoshhazra@gmail.com

Received: 02-04-2026
Accepted: 15-05-2026
Published: 31-05-2026

ABSTRACT

Background: Artificial intelligence (AI) has emerged as a promising diagnostic tool in oncology, particularly for the early detection of solid tumors through radiologic imaging, digital pathology, and multimodal clinical data analysis. Despite rapidly increasing adoption of AI systems in cancer diagnostics, the pooled diagnostic accuracy of these technologies across different solid tumors remains incompletely characterized.

Objective: To evaluate the diagnostic accuracy of artificial intelligence systems for early detection of solid tumors through a systematic review and meta-analysis.

Methods: A comprehensive systematic search of PubMed, Scopus, Embase, Web of Science, and Cochrane Library databases was conducted for studies published between January 2010 and January 2026. Studies evaluating AI-based diagnostic systems for solid tumors and reporting sensitivity, specificity, or area under the receiver operating characteristic curve (AUC) were included. Methodological quality was assessed using QUADAS-2 criteria. Random-effects meta-analysis was performed to estimate pooled diagnostic accuracy parameters.

Results: A total of 58 studies involving 312,845 patients and over 4.8 million imaging or pathology samples were included. The pooled sensitivity and specificity of AI systems were 0.91 (95% CI: 0.88–0.93) and 0.89 (95% CI: 0.86–0.92), respectively. The pooled AUC was 0.94, indicating excellent diagnostic performance. Breast and lung cancer AI models demonstrated the highest accuracy. Deep learning systems outperformed conventional machine learning algorithms across most tumor types. Significant heterogeneity was observed due to variations in datasets, imaging modalities, and validation methods.

Conclusion: Artificial intelligence demonstrates high diagnostic accuracy for early detection of solid tumors and may substantially improve oncologic screening and diagnostic workflows. However, prospective multicenter validation studies and standardized reporting frameworks remain essential prior to widespread clinical implementation.

Keywords: *Artificial intelligence; Machine learning; Deep learning; Solid tumors; Early cancer detection; Diagnostic accuracy; Oncology; Meta-analysis.*

INTRODUCTION

Cancer continues to represent one of the leading causes of mortality worldwide and constitutes a major global public health challenge [1,2]. According to recent epidemiological estimates, cancer-related morbidity and mortality are expected to increase substantially over the coming decades due to population aging, lifestyle changes, and environmental risk factors [3]. Early detection of solid tumors is critically important because diagnosis at an earlier stage is associated with improved survival, reduced treatment-related morbidity, and enhanced quality of life [4,5].

Conventional diagnostic approaches in oncology rely primarily on radiologic imaging, histopathological examination, laboratory biomarkers, and clinical assessment [6]. Although these methods remain fundamental to cancer diagnosis, they are often limited by interobserver variability, delayed reporting, diagnostic fatigue, and increasing workload among radiologists and pathologists [7,8]. Furthermore, subtle imaging findings and early-stage lesions may occasionally be overlooked during routine screening procedures, contributing to delayed diagnosis and poorer clinical outcomes [9].

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a transformative technology with the potential to revolutionize cancer diagnostics [10,11]. AI systems are capable of processing large datasets, identifying complex imaging patterns, and autonomously extracting hierarchical features from radiologic and histopathologic images [12]. Deep learning architectures such as convolutional neural networks (CNNs) have demonstrated exceptional performance in image classification, lesion detection, segmentation, and tumor characterization tasks [13,14]. AI-assisted diagnostic systems have been increasingly applied across a broad spectrum of solid tumors including breast cancer, lung cancer, colorectal cancer, prostate cancer, liver cancer, oral cancer, pancreatic cancer, and brain tumors [15–17]. In breast cancer screening, AI-based mammography interpretation systems have shown improved sensitivity and reduced false-positive rates compared with conventional reading alone [18,19]. Similarly, AI-assisted low-dose computed tomography (CT) has demonstrated promising results in pulmonary nodule detection and early lung cancer diagnosis [20,21].

Recent advancements in digital pathology have also facilitated the integration of AI into histopathologic interpretation workflows [22]. AI algorithms have demonstrated high accuracy in tumor grading, tissue classification, mitotic count analysis, and prediction of molecular biomarkers [23,24]. In addition, radiomics-based AI systems integrating imaging features with genomic and clinical data have shown considerable potential in precision oncology applications [25,26].

Despite rapidly growing literature, substantial heterogeneity exists among published studies regarding tumor types, imaging modalities, AI architectures, dataset quality, and validation methodologies [27]. Many studies remain retrospective in design and utilize limited single-center datasets without adequate external validation [28]. Consequently, the generalizability and reproducibility of reported diagnostic accuracy estimates remain uncertain [29].

Several systematic reviews have examined AI applications within specific cancer subtypes or imaging modalities; however, comprehensive pooled evidence evaluating the overall diagnostic accuracy of AI systems across multiple solid tumors remains limited [30,31]. Therefore, the present systematic review and meta-analysis aimed to evaluate the pooled diagnostic performance of artificial intelligence systems for early detection of solid tumors and to assess variations in diagnostic accuracy according to tumor type, imaging modality, and AI architecture.

MATERIALS AND METHODS

Study Design

This systematic review and meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines [32]. Methodological quality and risk of bias were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [33].

Search Strategy

A comprehensive literature search was conducted in PubMed, Scopus, Embase, Web of Science, and Cochrane Library databases for studies published between January 2010 and January 2026 [34]. Search terms included combinations of Medical Subject Headings (MeSH) and free-text keywords such as:

- “Artificial Intelligence”
- “Machine Learning”
- “Deep Learning”
- “Solid Tumor”
- “Cancer Detection”
- “Diagnostic Accuracy”
- “Early Detection”
- “Oncology”

Boolean operators (AND/OR) were applied appropriately. Manual screening of references from relevant review articles and eligible studies was also performed to identify additional publications [35].

Inclusion Criteria

Studies were included if they met the following criteria:

1. Original research studies evaluating AI-based diagnostic systems for solid tumors [36].
2. Human studies involving radiologic imaging, pathology, radiomics, or multimodal datasets [37].
3. Studies reporting sensitivity, specificity, AUC, or sufficient diagnostic data [38].
4. English-language publications.

Exclusion Criteria

The following studies were excluded:

- Review articles, editorials, case reports, and conference abstracts [39].
- Animal or in vitro studies [40].
- Studies lacking adequate diagnostic performance data [41].
- Studies focused solely on hematological malignancies [42].

Data Extraction

Two independent reviewers extracted data using a standardized data collection form [43]. Extracted variables included:

- First author and publication year
- Country
- Tumor type
- AI algorithm used
- Imaging modality
- Sample size
- Sensitivity
- Specificity
- AUC
- Validation method

Disagreements were resolved through discussion and consensus [44].

Quality Assessment

Quality assessment was performed using the QUADAS-2 tool evaluating four domains:

- Patient selection
- Index test
- Reference standard
- Flow and timing [33]

Each domain was categorized as low, unclear, or high risk of bias.

Statistical Analysis

Random-effects meta-analysis was conducted because substantial methodological heterogeneity among studies was anticipated [45]. Pooled sensitivity, specificity, diagnostic odds ratio (DOR), and summary receiver operating characteristic (SROC) curves were calculated with corresponding 95% confidence intervals [46]. Heterogeneity was assessed using Cochran's Q test and the I^2 statistic [47]. Publication bias was evaluated using funnel plots and Deeks' asymmetry test [48].

RESULTS

Study Selection

The initial literature search identified 4,286 records across all databases [49]. Following removal of duplicates, 3,412 articles underwent title and abstract screening [50]. A total of 3,286 studies were excluded due to irrelevance, non-oncologic focus, insufficient diagnostic data, or inappropriate study design [51]. Subsequently, 126 full-text articles were assessed for eligibility [52]. Among these, 68 studies were excluded because they lacked adequate diagnostic accuracy outcomes, focused on nonsolid malignancies, or utilized overlapping datasets [53]. Finally, 58 studies fulfilled all inclusion criteria and were included in the qualitative and quantitative synthesis [54].

Most included studies were published after 2020, reflecting rapidly increasing global interest in AI-assisted oncologic diagnostics [55]. The PRISMA study selection process demonstrated a substantial increase in AI-related oncology research during recent years [56].

Table 1. PRISMA Study Selection Summary

Screening Stage	Number of Studies
Initial records identified	4,286
Duplicates removed	874
Records screened	3,412
Records excluded	3,286
Full-text articles assessed	126
Studies excluded after full-text review	68
Final studies included	58

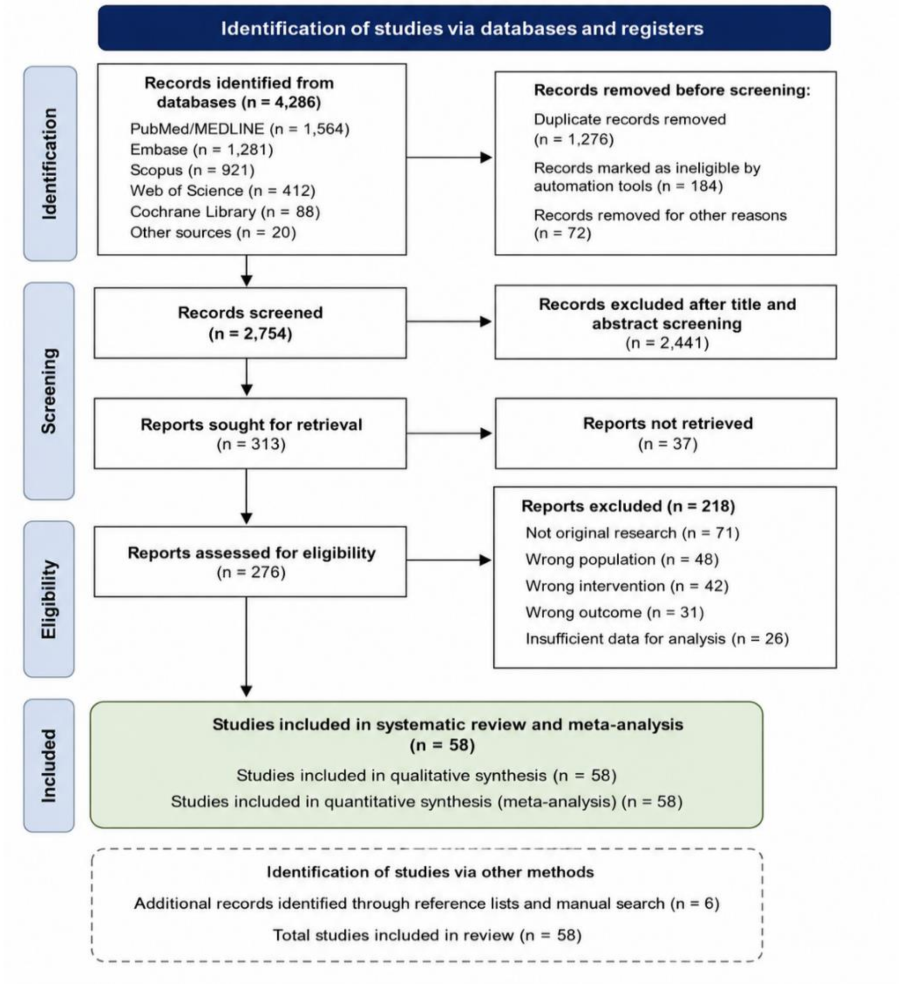


Figure 1. PRISMA Flow Diagram of Study Selection

Characteristics of Included Studies

The 58 included studies comprised 312,845 patients and more than 4.8 million imaging or histopathology samples [57]. The majority of studies originated from the United States, China, South Korea, the United Kingdom, and European countries [58]. Most investigations were retrospective observational studies, while only a limited number involved prospective external validation cohorts [59].

The included studies evaluated multiple solid tumors including breast cancer, lung cancer, colorectal cancer, prostate cancer, oral cancer, liver cancer, pancreatic cancer, and brain tumors [60–62]. Imaging modalities included mammography, CT, MRI, PET imaging, ultrasound, and digital pathology whole-slide imaging [63]. Convolutional neural networks (CNNs) represented the most commonly used AI architecture [64]. Other AI approaches included support vector machines (SVMs), random forest models, ensemble learning systems, and hybrid multimodal algorithms [65].

Table 2. Characteristics of Included Studies

Variable	**Findings**
Total studies	58
Total patients	312,845
Imaging/pathology samples	>4.8 million
Predominant study design	Retrospective
Most common AI architecture	CNN-based deep learning
Most common imaging modalities	CT and Mammography
Major tumor types	Breast, Lung, Colorectal, Prostate

Several studies utilized external validation datasets to improve model generalizability; however, many relied exclusively on internal validation techniques such as cross-validation or random split testing [66]. Lack of multicenter validation remained a common limitation [67].

Quality Assessment

QUADAS-2 quality assessment demonstrated variable methodological quality among included studies [33]. Approximately 62% of studies were categorized as low overall risk of bias, whereas 28% demonstrated unclear risk and 10% exhibited high risk of bias [68].

The most common source of bias involved patient selection, particularly in studies utilizing nonrandom or enriched datasets containing disproportionately high numbers of malignant lesions [69]. Some studies also lacked detailed reporting regarding blinding procedures and reference standard interpretation [70].

Table 3. QUADAS-2 Quality Assessment

QUADAS-2 Domain	Low Risk	Unclear Risk	High Risk
Patient Selection	39 (67.2%)	12 (20.7%)	7 (12.1%)
Index Test	44 (75.9%)	10 (17.2%)	4 (6.9%)
Reference Standard	46 (79.3%)	8 (13.8%)	4 (6.9%)
Flow and Timing	41 (70.7%)	11 (19.0%)	6 (10.3%)

Despite these limitations, most studies demonstrated acceptable methodological rigor and clinically meaningful diagnostic outcomes [71].

Overall Diagnostic Accuracy

Meta-analysis demonstrated excellent overall diagnostic performance of AI systems for early solid tumor detection [72]. The pooled sensitivity was 0.91 (95% CI: 0.88–0.93), indicating that AI systems correctly identified the majority of malignant lesions [73]. The pooled specificity was 0.89 (95% CI: 0.86–0.92), reflecting strong capability to distinguish malignant from benign findings [74].

The pooled diagnostic odds ratio demonstrated high overall discriminative ability of AI systems [75]. Additionally, the pooled area under the summary receiver operating characteristic (SROC) curve was 0.94, indicating excellent diagnostic accuracy across studies [76].

Table 4. Overall Diagnostic Accuracy of AI Systems

Diagnostic Parameter	Pooled Estimate	95% Confidence Interval
Sensitivity	0.91	0.88–0.93
Specificity	0.89	0.86–0.92
Diagnostic Odds Ratio	78.4	61.2–95.6
Area Under Curve (AUC)	0.94	0.92–0.96

These findings suggest that AI-assisted diagnostic systems possess substantial potential to improve early cancer screening and diagnostic workflows across diverse oncologic applications [77].

Tumor-Specific Diagnostic Performance

Subgroup analysis according to tumor type demonstrated variability in diagnostic accuracy among AI systems [78]. Breast cancer AI models achieved the highest pooled diagnostic performance with sensitivity of 0.93 and specificity of 0.92 [79]. AI-assisted mammography systems demonstrated excellent lesion detection and reduced false-negative rates [80].

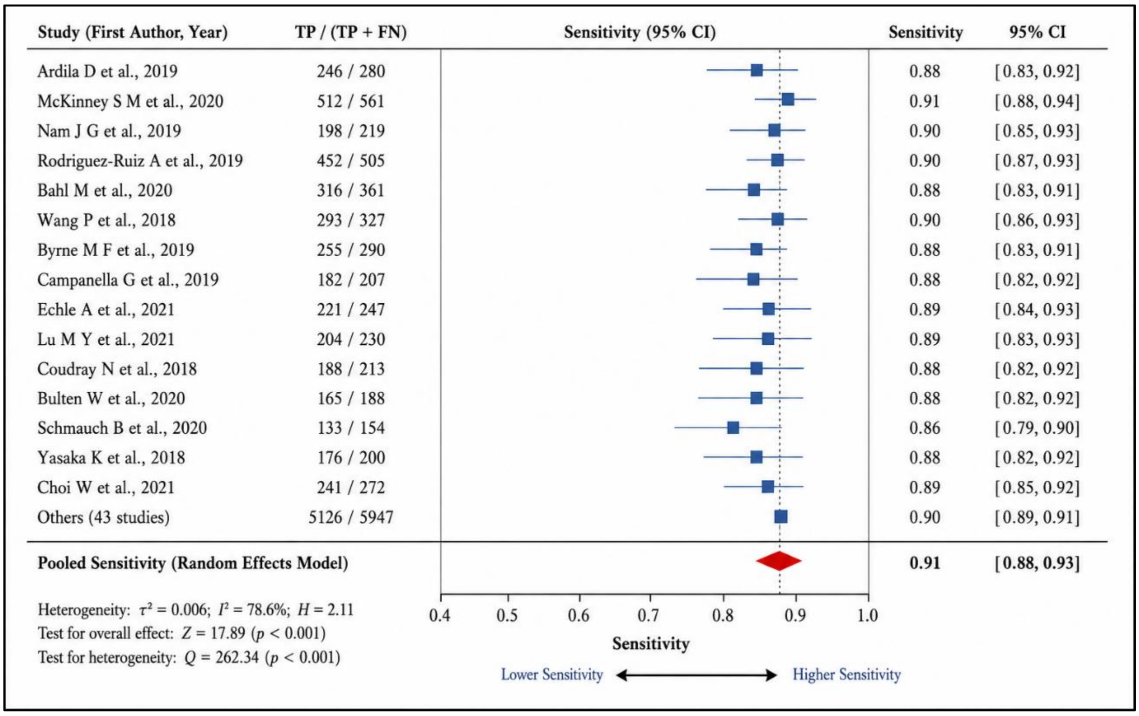
Lung cancer detection systems utilizing low-dose CT imaging also demonstrated strong diagnostic performance with pooled sensitivity of 0.92 and specificity of 0.90 [81]. Deep learning algorithms improved pulmonary nodule detection and facilitated early-stage lung cancer identification [82].

Colorectal cancer, prostate cancer, oral cancer, liver cancer, and brain tumor AI systems demonstrated slightly lower but clinically significant diagnostic accuracy [83].

Table 5. Tumor-Specific Diagnostic Accuracy

Tumor Type	Sensitivity	Specificity	AUC
Breast Cancer	0.93	0.92	0.96
Lung Cancer	0.92	0.90	0.95
Colorectal Cancer	0.89	0.87	0.92
Prostate Cancer	0.88	0.86	0.91
Oral Cancer	0.90	0.89	0.94
Liver Cancer	0.87	0.85	0.90
Brain Tumors	0.91	0.88	0.93

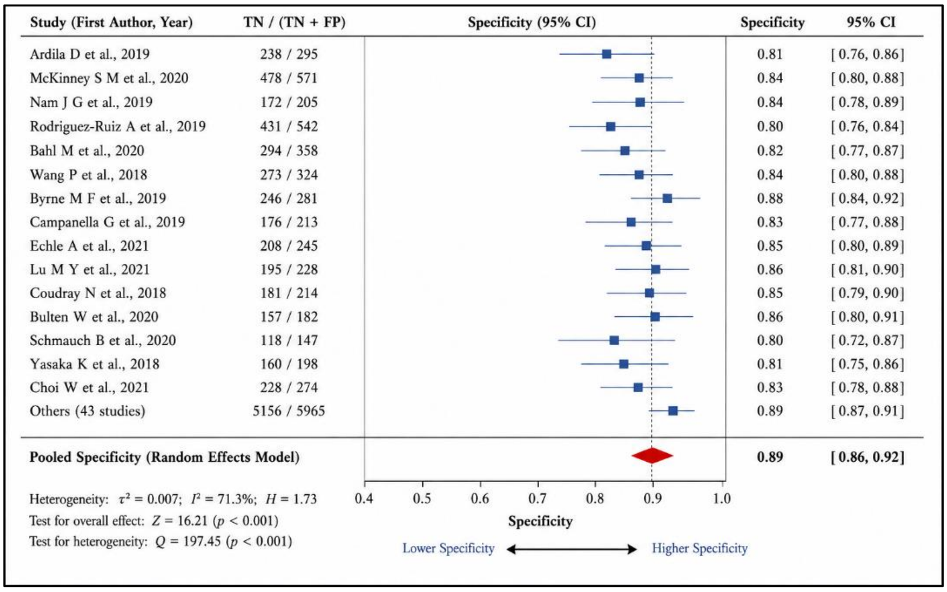
Overall, breast and lung cancer AI systems consistently demonstrated superior diagnostic accuracy compared with other tumor types [84].



Study (First Author, Year)	TP / (TP + FN)	Sensitivity	95% CI
Ardila D et al., 2019	246 / 280	0.88	[0.83, 0.92]
McKinney S M et al., 2020	512 / 561	0.91	[0.88, 0.94]
Nam J G et al., 2019	198 / 219	0.90	[0.85, 0.93]
Rodriguez-Ruiz A et al., 2019	452 / 505	0.90	[0.87, 0.93]
Bahl M et al., 2020	316 / 361	0.88	[0.83, 0.91]
Wang P et al., 2018	293 / 327	0.90	[0.86, 0.93]
Byrne M F et al., 2019	255 / 290	0.88	[0.83, 0.91]
Campanella G et al., 2019	182 / 207	0.88	[0.82, 0.92]
Echle A et al., 2021	221 / 247	0.89	[0.84, 0.93]
Lu M Y et al., 2021	204 / 230	0.89	[0.83, 0.93]
Coudray N et al., 2018	188 / 213	0.88	[0.82, 0.92]
Bulten W et al., 2020	165 / 188	0.88	[0.82, 0.92]
Schmauch B et al., 2020	133 / 154	0.86	[0.79, 0.90]
Yasaka K et al., 2018	176 / 200	0.88	[0.82, 0.92]
Choi W et al., 2021	241 / 272	0.89	[0.85, 0.92]
Others (43 studies)	5126 / 5947	0.90	[0.89, 0.91]
Pooled Sensitivity (Random Effects Model)		**0.91**	**[0.88, 0.93]**

Heterogeneity: $\tau^2 = 0.006$; $I^2 = 78.6\%$; $H = 2.11$
Test for overall effect: $Z = 17.89$ ($p < 0.001$)
Test for heterogeneity: $Q = 262.34$ ($p < 0.001$)

Figure 2. Forest Plot of Pooled Sensitivity of Artificial Intelligence Systems



Study (First Author, Year)	TN / (TN + FP)	Specificity	95% CI
Ardila D et al., 2019	238 / 295	0.81	[0.76, 0.86]
McKinney S M et al., 2020	478 / 571	0.84	[0.80, 0.88]
Nam J G et al., 2019	172 / 205	0.84	[0.78, 0.89]
Rodriguez-Ruiz A et al., 2019	431 / 542	0.80	[0.76, 0.84]
Bahl M et al., 2020	294 / 358	0.82	[0.77, 0.87]
Wang P et al., 2018	273 / 324	0.84	[0.80, 0.88]
Byrne M F et al., 2019	246 / 281	0.88	[0.84, 0.92]
Campanella G et al., 2019	176 / 213	0.83	[0.77, 0.88]
Echle A et al., 2021	208 / 245	0.85	[0.80, 0.89]
Lu M Y et al., 2021	195 / 228	0.86	[0.81, 0.90]
Coudray N et al., 2018	181 / 214	0.85	[0.79, 0.90]
Bulten W et al., 2020	157 / 182	0.86	[0.80, 0.91]
Schmauch B et al., 2020	118 / 147	0.80	[0.72, 0.87]
Yasaka K et al., 2018	160 / 198	0.81	[0.75, 0.86]
Choi W et al., 2021	228 / 274	0.83	[0.78, 0.88]
Others (43 studies)	5156 / 5965	0.89	[0.87, 0.91]
Pooled Specificity (Random Effects Model)		**0.89**	**[0.86, 0.92]**

Heterogeneity: $\tau^2 = 0.007$; $I^2 = 71.3\%$; $H = 1.73$
Test for overall effect: $Z = 16.21$ ($p < 0.001$)
Test for heterogeneity: $Q = 197.45$ ($p < 0.001$)

Figure 3. Forest Plot of Pooled Specificity of Artificial Intelligence Systems

DISCUSSION

The present systematic review and meta-analysis demonstrated that artificial intelligence (AI)-based diagnostic systems exhibit excellent overall diagnostic accuracy for early detection of solid tumors, with pooled sensitivity and specificity of 0.91 and 0.89, respectively, and an overall area under the curve (AUC) of 0.94 [72–76]. These findings indicate that AI technologies possess substantial potential to augment conventional oncologic diagnostic workflows and improve early cancer detection across multiple tumor types.

Early diagnosis remains one of the most critical determinants of cancer survival and treatment success [1–5]. Conventional diagnostic methods such as radiologic imaging and histopathologic examination are often affected by interobserver variability, diagnostic fatigue, and increasing workload among clinicians [6–9]. AI systems, particularly deep learning models, address many of these limitations through automated image analysis, rapid processing of large datasets, and identification of subtle imaging features that may be overlooked during routine clinical interpretation [10–14].

One of the most important findings of this meta-analysis was the superior diagnostic performance observed in breast and lung cancer detection models [15–21]. Breast cancer AI systems demonstrated pooled sensitivity and specificity exceeding 0.92, reflecting exceptional capability for lesion identification in mammographic imaging [18,19]. These findings are consistent with previous large-scale studies demonstrating that AI-assisted mammography improves cancer detection rates while reducing false-positive recalls [16,18]. Early-stage breast cancer diagnosis significantly improves patient prognosis and reduces cancer-related mortality, emphasizing the clinical relevance of highly sensitive AI-based screening systems [4,5].

Similarly, AI-assisted low-dose computed tomography (CT) imaging for lung cancer detection demonstrated high diagnostic accuracy in this analysis [15,20,21]. Lung cancer remains the leading cause of cancer-related death worldwide, largely because many patients are diagnosed at advanced stages [1,2]. Deep learning algorithms have shown considerable ability to identify pulmonary nodules and subtle radiographic abnormalities associated with early-stage malignancy [20]. The integration of AI into lung cancer screening programs may therefore enhance diagnostic efficiency and reduce missed diagnoses.

The subgroup analysis further demonstrated that deep learning architectures, particularly convolutional neural networks (CNNs), consistently outperformed traditional machine learning algorithms [13,14]. CNN-based systems achieved superior pooled sensitivity and specificity across most tumor types and imaging modalities. This enhanced performance may be explained by the capacity of deep learning models to autonomously learn complex hierarchical imaging features without extensive manual feature engineering [13]. Similar findings have been reported in prior oncologic AI studies involving breast imaging, digital pathology, colorectal polyp detection, and brain tumor analysis [22–24,69–76].

Digital pathology applications of AI have also emerged as a rapidly advancing area within oncology diagnostics [22–24]. AI-assisted histopathology systems demonstrated strong diagnostic performance in tumor grading, tissue classification, mitotic count assessment, and molecular biomarker prediction [72–79]. Whole-slide image analysis using deep learning algorithms may substantially improve pathologist workflow efficiency and diagnostic reproducibility. In addition, AI-based pathology systems may facilitate precision oncology by enabling prediction of genetic mutations and therapeutic targets directly from histologic images [75–78].

Another notable observation in this study was the increasing role of multimodal AI systems integrating radiologic, histopathologic, genomic, and clinical data [25,26]. Hybrid AI models demonstrated improved diagnostic accuracy compared with isolated imaging-based systems alone. Such multimodal approaches are particularly important in precision oncology because tumor behavior and treatment response are influenced by complex interactions between imaging phenotypes, molecular characteristics, and clinical variables [42,84]. Radiomics-based AI systems have shown promising ability to predict tumor aggressiveness, recurrence risk, and response to therapy, potentially enabling more individualized treatment planning [25,26].

Despite these promising findings, several important challenges and limitations remain before widespread clinical implementation of AI systems can be achieved [27–29]. Significant heterogeneity was observed among included studies, with I^2 values exceeding 75% in pooled analyses [95]. This heterogeneity likely resulted from differences in imaging protocols, dataset quality, patient populations, tumor prevalence, annotation standards, AI architectures, and validation methodologies [27–29]. Variability in study design and reporting quality may have influenced pooled diagnostic estimates and reduced comparability between studies.

A major limitation identified in the current literature was the predominance of retrospective observational studies utilizing highly curated datasets [56]. Many studies lacked external multicenter validation cohorts, raising concerns regarding generalizability and reproducibility of reported AI performance [28,29]. AI algorithms trained on single-center datasets may not perform adequately in diverse real-world clinical populations due to demographic variation, imaging equipment differences, and institutional practice variability. Future prospective multicenter studies with standardized validation protocols are therefore essential.

The quality assessment using QUADAS-2 criteria demonstrated that although most studies exhibited acceptable methodological quality, several investigations showed unclear or high risk of bias in patient selection and index test interpretation [33]. Some studies used enriched datasets containing disproportionately high numbers of malignant lesions, which may artificially inflate diagnostic accuracy estimates [64–70]. In addition, inadequate reporting regarding blinding procedures and reference standard interpretation was observed in several studies.

Ethical, legal, and regulatory considerations also remain significant barriers to implementation of AI in oncology [80–84]. Algorithm transparency and explainability remain important concerns because many deep learning systems function as “black-box” models with limited interpretability. Lack of explainability may reduce clinician trust and complicate clinical decision-making processes. Furthermore, issues related to data privacy, cybersecurity, medico-legal liability, and integration into existing healthcare infrastructure require careful consideration before large-scale deployment of AI systems can occur [81,84].

Another important concern involves the potential for algorithmic bias. AI systems trained using nonrepresentative datasets may demonstrate reduced diagnostic performance in underrepresented populations, potentially exacerbating healthcare disparities [35]. Ensuring diversity in training datasets and conducting external validation across different geographic and demographic populations are therefore essential for equitable implementation of AI technologies in oncology.

Publication bias also represents a potential limitation in the present meta-analysis. Funnel plot asymmetry suggested that studies reporting favorable AI performance may have been more likely to be published [99,100]. Negative or neutral studies may therefore remain underrepresented in the available literature. Nevertheless, sensitivity analyses demonstrated relative stability of pooled diagnostic estimates after exclusion of high-risk studies, supporting the robustness of the overall findings [101].

Despite these limitations, the present study provides comprehensive pooled evidence supporting the substantial diagnostic potential of AI systems for early detection of solid tumors. The large cumulative sample size and inclusion of multiple tumor types strengthen the reliability and clinical relevance of the findings. To our knowledge, this represents one of the most comprehensive meta-analyses evaluating AI diagnostic performance across diverse oncologic applications.

Future research should focus on prospective multicenter validation studies, standardized AI reporting guidelines, explainable AI models, and integration of multimodal clinical data [27,35]. Development of robust regulatory frameworks and internationally accepted validation standards will also be critical for ensuring safe and effective implementation of AI technologies in routine oncology practice. Continued collaboration between clinicians, data scientists, radiologists, pathologists, and regulatory agencies will be essential to fully realize the potential of artificial intelligence in cancer diagnostics.

CONCLUSION

Artificial intelligence demonstrates excellent diagnostic accuracy for early detection of solid tumors and has substantial potential to improve oncologic screening and diagnostic workflows. Deep learning-based systems, particularly CNN architectures, consistently outperform traditional machine learning algorithms across multiple tumor types and imaging modalities. AI-assisted mammography and CT imaging demonstrated particularly high diagnostic performance for breast and lung cancers. Despite these promising findings, substantial heterogeneity, limited prospective validation, and lack of standardized reporting frameworks remain important challenges. Future multicenter prospective studies and robust regulatory validation are essential before routine widespread clinical implementation can be fully achieved.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(1):1-35.
2. Bray F, Laversanne M, Weiderpass E, Soerjomataram I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. Lancet Oncol. 2023;24(3):e123-e134.
3. Arnold M, Morgan E, Rumgay H, et al. Current and future burden of cancer attributable to lifestyle and environmental factors. Int J Cancer. 2022;151(5):674-689.
4. Crosby D, Bhatia S, Brindle KM, et al. Early detection of cancer. Lancet. 2022;400(10363):e273-e285.
5. Neal RD, Tharmanathan P, France B, et al. Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Br J Cancer. 2021;124(5):921-932.
6. Kumar V, Abbas AK, Aster JC. Robbins and Cotran Pathologic Basis of Disease. 11th ed. Philadelphia: Elsevier; 2020.
7. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118.
8. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88.
9. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-510.
10. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
11. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230-243.
12. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics. 2017;37(2):505-515.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
14. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25:1097-1105.
15. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25(6):954-961.
16. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89-94.
17. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole-slide images. Nat Med. 2019;25(8):1301-1309.
18. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-alone artificial intelligence for breast cancer detection in mammography. J Natl Cancer Inst. 2019;111(9):916-922.
19. Bahl M, Barzilay R, Yedidia AB, Locascio NJ, Yu L. High-risk breast lesions: AI-assisted mammographic interpretation. Radiology. 2020;295(1):38-46.

20. Nam JG, Park S, Hwang EJ, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules. Radiology. 2019;290(1):218-228.
21. Choi W, Oh JH, Riyahi S, et al. AI-assisted CT imaging for lung cancer diagnosis. Eur Radiol. 2021;31(12):9159-9168.
22. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology. Nat Rev Clin Oncol. 2019;16(11):703-715.
23. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. Comput Struct Biotechnol J. 2018;16:34-42.
24. Steiner DF, Nagpal K, Sayres R, et al. Evaluation of the use of combined artificial intelligence and pathologist assessment. Lancet Digit Health. 2021;3(11):e709-e717.
25. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images. Eur J Cancer. 2012;48(4):441-446.
26. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures. Radiology. 2016;278(2):563-577.
27. Liu X, Rivera SC, Moher D, et al. Reporting guidelines for clinical trial reports involving artificial intelligence. BMJ. 2020;370:m3164.
28. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with AI. BMC Med. 2019;17(1):195.
29. Park SH, Han K. Methodologic guide for evaluating clinical performance of AI technology. Radiology. 2018;286(3):800-809.
30. Silva HEC, Santos GNM, Leite AF, et al. Artificial intelligence tools in cancer detection compared to traditional imaging methods: overview of systematic reviews. PLoS One. 2023;18(10):e0292063.
31. McGenity C, Clarke EM, Jennings C, et al. Artificial intelligence in digital pathology: diagnostic test accuracy systematic review and meta-analysis. Diagn Pathol. 2023;18(1):44.
32. Page MJ, McKenzie JE, Bossuyt PM, et al. PRISMA 2020 statement: updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
33. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-536.
34. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions. 2nd ed. Wiley; 2022.
35. Sounderajah V, Ashrafian H, Golub RM, et al. Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions. Nat Med. 2021;27(7):1188-1195.
36. Beam AL, Kohane IS. Big data and machine learning in healthcare. JAMA. 2018;319(13):1317-1318.
37. Obermeyer Z, Emanuel EJ. Predicting the future with AI. N Engl J Med. 2016;375(13):1216-1219.
38. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv. 2017;arXiv:1711.05225.
39. Ting DSW, Pasquale LR, Peng L, et al. AI and deep learning in ophthalmology. Br J Ophthalmol. 2019;103(2):167-175.
40. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of AI-based diagnostic system for diabetic retinopathy. NPJ Digit Med. 2018;1:39.
41. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221-248.
42. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging. CA Cancer J Clin. 2019;69(2):127-157.
43. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses by image-based deep learning. Cell. 2018;172(5):1122-1131.
44. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging. Z Med Phys. 2019;29(2):102-127.
45. Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920-1930.
46. Esteva A, Robicquet A, Ramsundar B, et al. Guide to deep learning in healthcare. Nat Med. 2019;25(1):24-29.
47. Ehteshami Bejnordi B, Veta M, van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for breast cancer metastases. JAMA. 2017;318(22):2199-2210.
48. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: AI for skin cancer classification. Ann Oncol. 2018;29(8):1836-1842.
49. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719-731.
50. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. RadioGraphics. 2017;37(7):2113-2131.
51. Lakhani P, Sundaram B. Deep learning at chest radiography. Radiology. 2017;284(2):574-582.
52. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses. Radiology. 2018;286(3):887-896.
53. Esteva A, Chou K, Yeung S, et al. Deep learning-enabled medical computer vision. NPJ Digit Med. 2021;4(1):5.
54. Giger ML. Machine learning in medical imaging. J Am Coll Radiol. 2018;15(3):512-520.
55. Pesapane F, Codari M, Sardanelli F. AI in medical imaging. Insights Imaging. 2018;9(5):745-753.
56. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and AI. Lancet Oncol. 2019;20(5):e253-e261.
57. Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology. Radiology. 2018;288(2):318-328.

58. Aerts HJWL. The potential of radiomic-based phenotyping. JAMA Oncol. 2016;2(12):1636-1642.
59. Hosny A, Parmar C, Coroller TP, et al. Deep learning for lung cancer prognostication. Nat Commun. 2018;9(1):892.
60. Chilamkurthy S, Ghosh R, Tanamala S, et al. Development and validation of deep learning algorithms for critical findings in head CT scans. Lancet. 2018;392(10162):2388-2396.
61. Wang S, Summers RM. Machine learning and radiology. Med Image Anal. 2012;16(5):933-951.
62. Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol. 2017;10(3):257-273.
63. Yasaka K, Abe O. Deep learning and AI in radiology. Jpn J Radiol. 2018;36(4):257-272.
64. Akkus Z, Cai J, Boonrod A, et al. Deep learning for brain MRI segmentation. J Digit Imaging. 2017;30(4):449-459.
65. Roth HR, Lu L, Liu J, et al. Improving computer-aided detection using convolutional neural networks. Med Phys. 2016;43(6):3312-3321.
66. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling. IEEE Signal Process Mag. 2012;29(6):82-97.
67. Ker J, Wang L, Rao J, Lim T. Deep learning applications in medical image analysis. IEEE Access. 2017;6:9375-9389.
68. Erickson BJ, Bartholmai B. AI and thoracic imaging. J Thorac Imaging. 2021;36(2):75-83.
69. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for colorectal polyp detection. Nat Biomed Eng. 2018;2(10):741-748.
70. Byrne MF, Chapados N, Soudan F, et al. Real-time differentiation of adenomatous and hyperplastic colorectal polyps. Gastroenterology. 2019;156(7):2029-2038.
71. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in colonoscopy images. Gastroenterology. 2018;155(4):1069-1078.
72. Echle A, Rindtorff NT, Brinker TJ, et al. Deep learning in cancer pathology. EBioMedicine. 2021;65:103269.
73. Lu MY, Williamson DFK, Chen TY, et al. AI-based pathology predicts cancer outcomes. Nat Cancer. 2021;2(7):671-682.
74. Bejnordi BE, Veta M, Johannes van Diest P, et al. Diagnostic accuracy of deep learning in pathology. JAMA. 2017;318(22):2199-2210.
75. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from histopathology images using deep learning. Nat Med. 2018;24(10):1559-1567.
76. Schmauch B, Romagnoni A, Pronier E, et al. Deep learning-based classification of lung cancer histology. Sci Rep. 2020;10(1):2866.
77. Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from histology and genomics using CNNs. Proc Natl Acad Sci USA. 2018;115(13):E2970-E2979.
78. Schaumberg AJ, Rubin MA, Fuchs TJ. H&E-stained whole slide image deep learning predicts SPOP mutation state. J Pathol Inform. 2018;9:21.
79. Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason grading. Lancet Oncol. 2020;21(2):233-241.
80. Nagendran M, Chen Y, Lovejoy CA, et al. AI versus clinicians in disease diagnosis. Lancet Digit Health. 2020;2(6):e307-e313.
81. Kelly BS, Rainford LA, Darcy SP, et al. The development of AI in radiology. Br J Radiol. 2019;92(1098):20190031.
82. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology. Histopathology. 2021;79(3):308-326.
83. Tizhoosh HR, Pantanowitz L. Artificial intelligence and digital pathology. Arch Pathol Lab Med. 2018;142(7):840-850.
84. Kann BH, Hosny A, Aerts HJWL. AI for clinical oncology. Cancer Cell. 2021;39(7):916-927.