**Global Perspective on the Use of Artificial Intelligence in Medical Imaging for Cancer Studies: An Umbrella Review of Systematic Review and Meta-Analyses**

**Author’s: Jiregna Olani Kedida, Tawadod Abdelgader, Corrado Lanera, Dario Gregori, Honoria Ocagli**

**Abstract**

**Introduction:** Artificial intelligence (AI) has emerged as a transformative driving force in healthcare, particularly in medical imaging. With the rapid growth of AI research in cancer imaging and its increasing implementation in clinical environments, it is imperative to consolidate existing knowledge through a systematic and structured synthesis.

**Objectives:** The aim of this umbrella review of systematic and meta-analyses is to systematically identify, categorize, and summarize the current applications of AI in cancer detection.

**Materials and Methods:** Characterization, and treatment response evaluation across various imaging modalities including CT, MRI, X- Ray, PET, CT/ or PET, Ultrasonography, and others. To assess and compare the diagnostic accuracy between AI models and traditional radiological assessments, using performance metrics such as Accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (), Likelihood Ratio, Predictive Value (), Diagnostic Odds Ratio (), heterogeneity (), Dice score (), test (), , and .

**Expected Results:** The methodological quality and reporting transparencies of included studies, examining how AI models were developed, used, validated, implemented will be assessed. This review will inform the development of best practices for integrating AI into radiological workflows. Finally, by combining qualitative and quantitative synthesis with meta-analytic techniques, this study bridges the gap between academic and practical application in studies with medical imaging setting.

**Keywords:** AI, CT, MRI, X-Ray, PET, CT/or PET, Ultrasonography, , , ,

# **Introduction**

## **Background**

Cancer continues to be a global health problem, which ranks among the leading morbidity and mortality worldwide. The evidence from the world Health Organization (WHO), approximately 10 million deaths were attributed to cancer in 2022, with projections indicating greater challenge and much burden due to aging populations, increased environmental exposures, and changing lifestyle factors [1]. Through advancements in oncologic research, therapeutic, treatment approaches, many cancers are still diagnosed at late stages, significantly lower survival rates. Early detection is essential for improving patient outcomes, but current screenings are usually hindered by sensitivity, inconsistent access to imaging resources, and difference in diagnostic accuracy [2].

Interestingly, medical Imaging remains a foundation in the process of cancer detection, staging, therapy planning, and monitoring. A diagnostic tool such as computed tomography (CT), magnetic resonance imaging (MRI), Ultrasonography, X- Ray, positron emission tomography (PET), positron emission tomography and or computed tomography (PET-CT scans) and others are integrated in modern oncology [3]. However, these techniques mainly depend on the knowledge and interpretative skills of radiologists, which introduces challenges such as inter-observer variability, fatigue-related errors, and inconsistencies motivated by the sheer volume of images requiring review [4], [5], [6]. With the global increase in cancer incidence and the increasing complexity of diagnostic protocols, radiology departments face unparalleled workloads and diagnostic bottlenecks [7].

In this context, artificial intelligence (AI) has emerged as a transformative force in healthcare, particularly in medical imaging. The subsets of AI such as; machine learning (ML) deep learning (DL), Computer Vision, Natural Language Processing (NLP) in Radiology Reports, Generative Adversarial Networks (GANs) for Image Enhancement, Reinforcement Learning for Image Interpretation offers the potential to enhance diagnostic precision, better streamline workflows, and minimize human error [8], [9].

We propose an umbrella review of published meta-analyses to better understand the available evidence for the application and use of artificial intelligence in medical imaging. Umbrella reviews of existing meta-analyses, referred to as overviews in the Cochrane Handbook, apply systematic review methodology to existing meta-analyses [10].

In recent years, there has been a substantial and large number of new global research focused on the use of artificial intelligence (AI) in medical imaging, they are also rapidly emerged, originating from all multiple countries. With so much information out there, it is easy unsure about what is actually working and where in which AI is showing the greatest impact in identifying the geographical and clinical settings.

This large spread of information necessitates a systematic approach to merge current knowledge. Therefore, this current study undertakes to do a comprehensive systematic review and meta-analyses and aimed to synthesize high-quality research to provide a clear overview of the global landscape of AI applications in medical imaging. To keep things transparent and thorough, this study was established methodologies, specifically the PRISMA guidelines [11], ensures transparency and methodological aspects, thereby fostering confidence in the presented findings.

# 3. Results

## 3.1 Study data and population description

Current study analyzed identified 4434 articles searched; 9 duplicates removed manually and 1260 duplicates removed by covidence, and 2454 were excluded from title and abstract screening. Of total 532 articles, 451 studies excluded at the stage in the full text screening because of: 411 articles Wrong indication, 15 No full text, 12 Wrong outcomes, 4 Wrong patient population, 3 Wrong study design, 3 unclear inclusion criteria, 2 Wrong setting, 1 Information not full, ex, unavailable DOI (Figure 1).

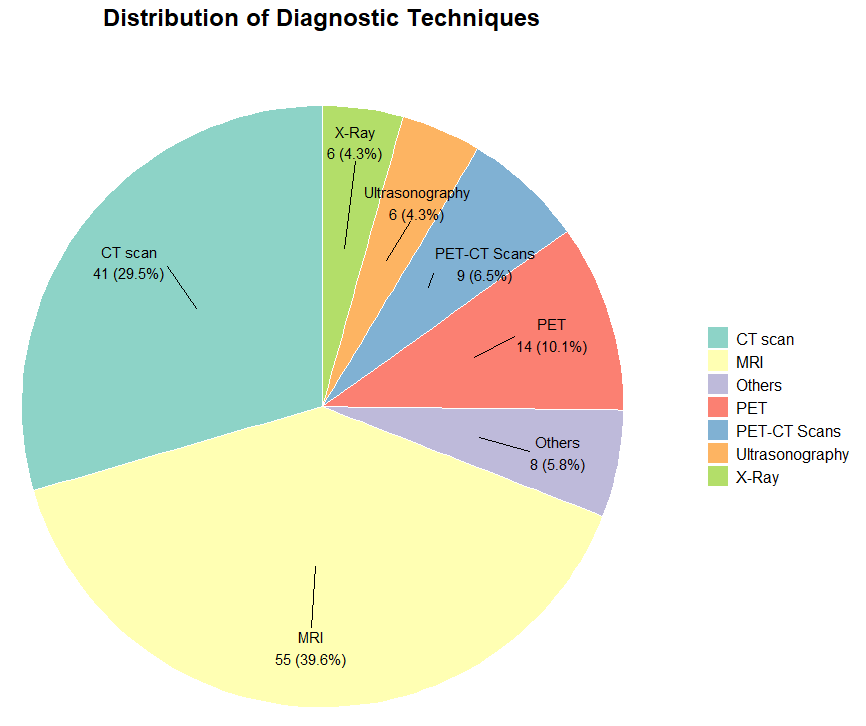
### **Study Characteristics**

The current umbrella review systematically extracted and analyzed published meta data from three major medical journals: Embase, PubMed, and Scopus – as in Table 1. Data is extracted and analyzed based on evidence from the sources on existing meta-analysis, systematic review, and a comprehensive review which focusing on the use and application of artificial intelligence (AI) in medical imaging in cancer study. 80 studies are included, with the vast majority being meta-analyses (91.2%), alongside a smaller proportion of systematic reviews 6(7.5%) and a comprehensive review only 1(1.2%).

|  |  |  |
| --- | --- | --- |
| **Table – 1: Study designs in the umbrella review** | | |
| **Study design** | **Number of articles included (n)** | **Percentage (%)** |
| Meta-analysis | 73 | 91.2 |
| Systematic review | 6 | 7.5 |
| A comprehensive review | 1 | 1.2 |

### **3.1.2 Image modalities**

These articles constituted a wide range of image modalities such as: MRI, CT scan, PET, CT and or PET, ultrasound, mammography, and others, indicating diverse cancer types including wide ranges of cancer types. The included studies encompass multiple countries and healthcare settings, which reflects the nationwide and multidisciplinary engagement in the research are of medical imaging and related.



Moreover, the studies evaluated a range of AI methods, particularly, deep learning (DL) and machine learning (ML) approaches, highlighting both rapid emerging and diversity of AI uses and applications within the field of medical imaging. Consistently, such trends growing with need for systematic evaluation of AI tools in healthcare, as the area shifts from innovation to validation and clinical translation [12], [13].

**Figure – 1**: Numerical distribution of diagnostic techniques (image modalities).

Identification of studies from covidence and data extraction stages

Studies from databases/registers **(n = 4434)**

Embase (n = 2174)

PubMed (n = 1502)

Scopus (n = 758)

References removed **(n = 1269)**

Duplicates identified manually (n = 9)

Duplicates identified by Covidence (n = 1260)

Marked as ineligible by automation tools (n = 0)

Other reasons (n = )

**Identification**

Studies excluded **(n = 2454)**

Studies screened **(n = 3165)**

Studies not retrieved **(n = 0)**

Studies sought for retrieval **(n = 532)**

**Screening**

Studies excluded **(n = 451)**

No full text (n = 15)

Wrong setting (n = 2)

Wrong outcomes (n = 12)

Wrong indication (n = 411)

Wrong study design (n = 3)

Wrong patient population (n = 4)

unclear inclusion criteria (n = 3)

Information not full, ex, unavailable DOI (n = 1)

Studies assessed for eligibility **(n = 532)**

**Included**

Included studies ongoing **(n = 0)**

Studies awaiting classification **(n = 0)**

Studies included in review **(n = 81)**

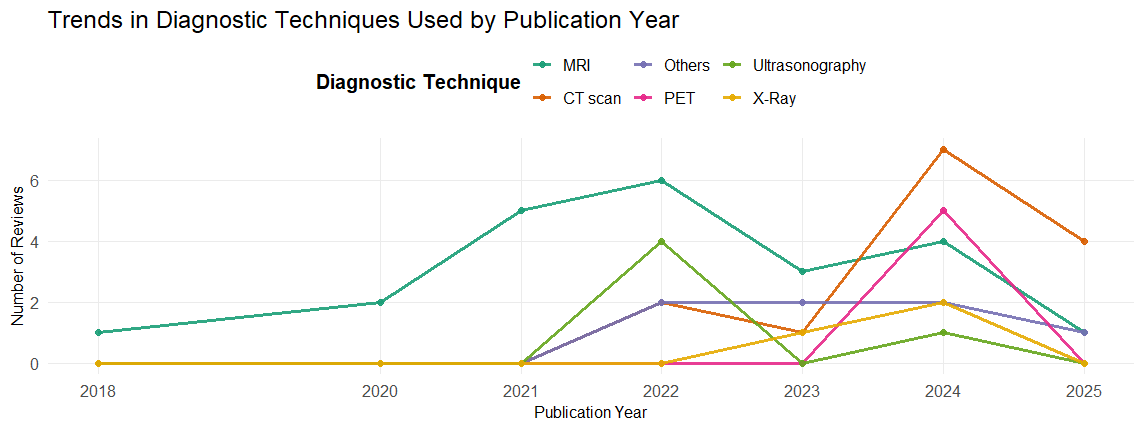
**Figure 1. PRISMA Flow Diagram of the study selection and data extraction process**

Recent advancement in artificial intelligence (AI) is leading to the rapid development and deployment of various AI models in medical imaging, particularly for cancer detection, diagnosis, and management. The figures from this umbrella review show that AI models, especially those utilizing deep learning (DL) models such as convolutional neural networks, achieve high levels of accuracy and sensitivity across multiple imaging modalities, including MRI, CT scan, PET, PET and or CT scan, and mammography so consistently (Figure 1).

### **Trends in Diagnostic techniques used in AI Based Cancer Imaging Studies**

The line graph in Figure – 2 displays the annual number of reviews focusing on various diagnostic imaging modalities in AI-based cancer research from 2018 to 2025. More importantly, MRI has been the most consistently represented techniques, showing a steady rise and peaking in 2022, followed by a moderate decrease. CT scan usage demonstrates a marked increase, with a sharp peak in 2024 (ten reviews), making it the most reviewed image modality for that year. Moreover, PET and ultrasonography also shown increased trends, in the year 2024, when PET reaches seven reviews ultrasonography reach two.

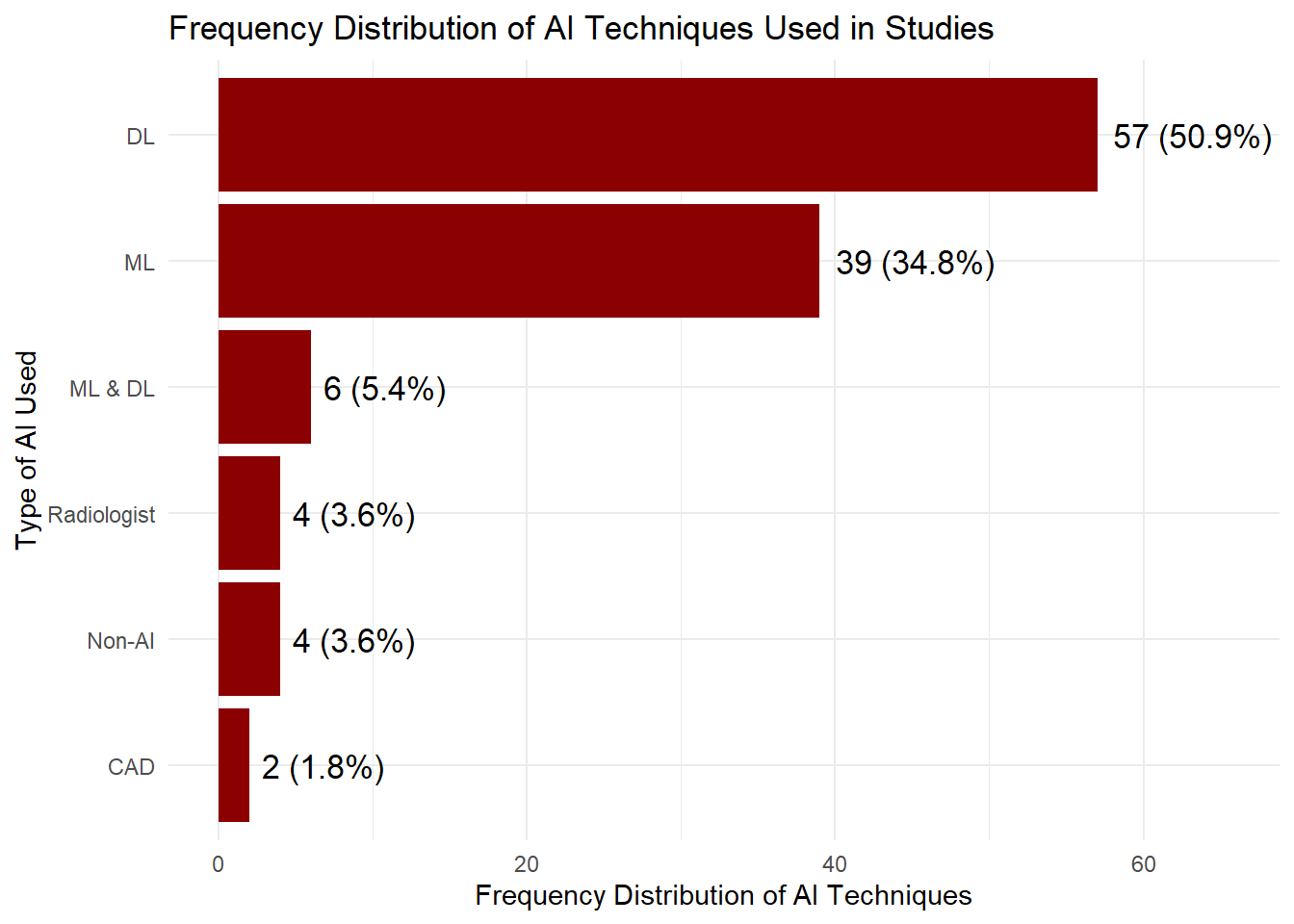
Importantly, the observed decrease in the number of reviews across all modalities in 2025 is not indication of a downward trend but rather reflects the length of data collection period, which occurred only by the second month of 2025. As such, the data for 2025are incomplete and likely not represented the full year’s results. These findings reflected the evolving research focus and technological advancement in AI systems based cancer imaging, with shifting emphasis among modalities over recent years [13], [14], [15].



**Figure – 2**: Trends of Diagnostic modalities type used by publication year

### **AI systems**

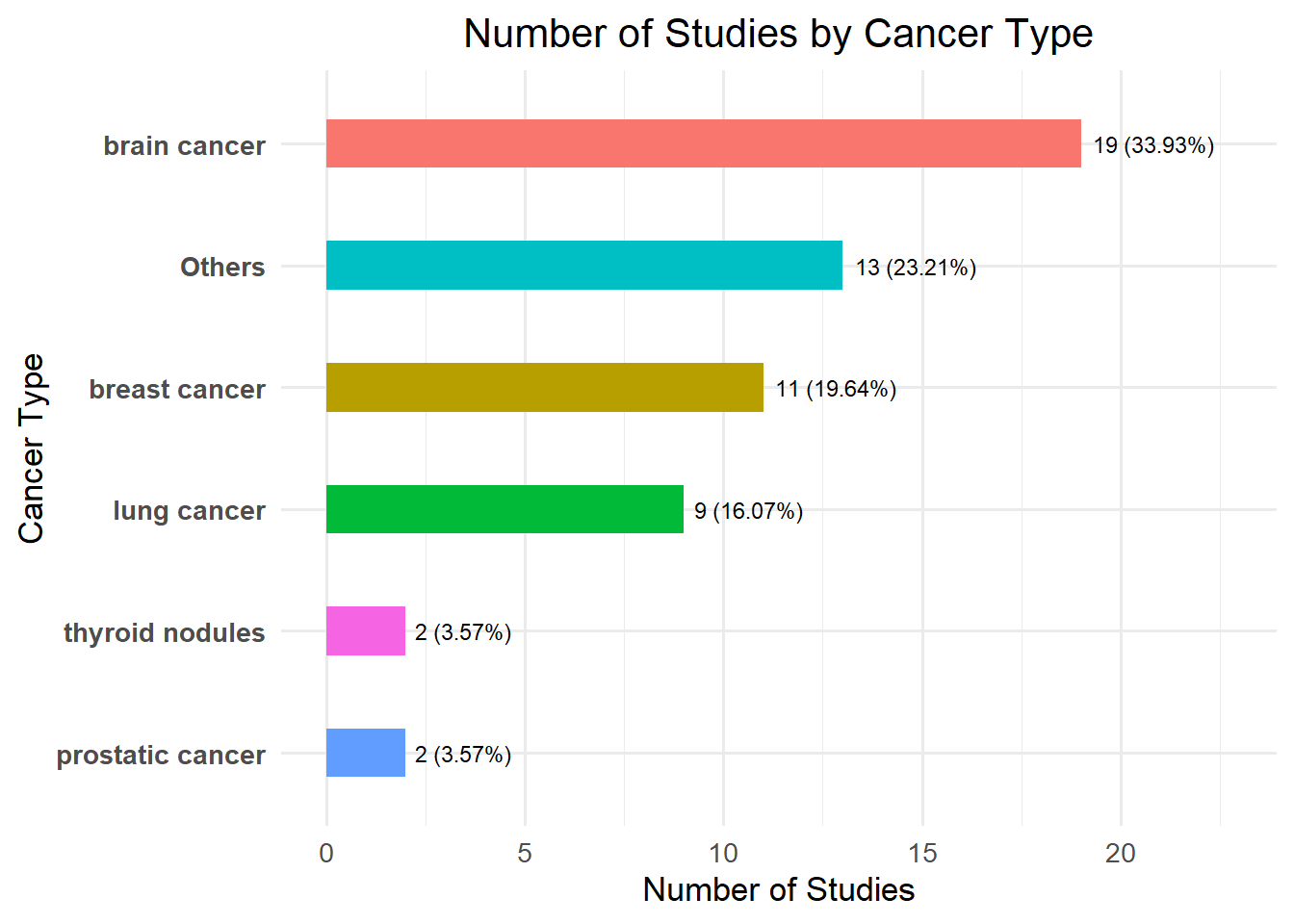
Convolutional neural networks (CNNs) are among the most impactful innovations [12], [13]. Moreover, they are designed well to extract and interpret complex visual features from imaging data. CNN model has been showed a significant high accuracy in detecting, classifying, and characterizing tumors, thereby reducing diagnostic uncertainty and supporting clinical decision making, [16] - [17]. Additionally, advances in explainable AI and hybrid models that combine clinical metadata with imaging inputs are enhancing the interpretability, transparency, trustworthiness of AI systems in clinical settings [[18], [19].



**Figure – 2**: Frequency distribution of type of AI/ML algorithms used.

On the other hand, as many studies report, AI systems are not oly perform comparable to experienced radiologist but also consists the considerable potential to reduce errors in diagnosis and improve early cancer detection rates. AI models have shown promise in automating image segmentation, classification, and prognostic prediction, in that way supporting clinicians in decision-making and improving workflow and procedure efficiency. These findings indicate the transformative AI potential in medical imaging and importance of continued research and validation to ensure safe and effective combination into clinical practice.

### **Number of studies by Cancer Type**

Figure – 3 below displays the distribution of included studies according to cancer type, as identified in this umbrella review. Graphical distribution revealed that brain cancer is the most extensively studied, accounting for 25 studies and comprising 31.25% of the total articles. This implies a strong research needs on the application of artificial intelligence in neuro-oncological imaging, might be reflecting both complexity and clinical significance of brain tumor diagnosis. Breast cancer follows as the second most frequently studied cancer type, with 14 studies (17.5%), indicating the ongoing advancements in breast cancer research possibly due to established role of imaging in breast cancer screening. 

For the category, Others, representing other cancer type of less common cancer, contributes 12 studies (15%), while lung cancer is represented in 10 studies (12.5%), addressing the essential of thoracic imaging and screening in oncological research. Esophageal cancer, thyroid nodules, prostatic cancer, and colorectal cancer each account for 3.75% to 5% of the total, whereas multiple cancers and gastric cancer are the least studied, with only two studies each (2%). Overall, these figures highlighted the greater focuses of AI research medical imaging on a few predominant cancer types, namely brain cancers, and lung cancers while other types of malignancies remain not represented in the review.

## Diagnostic and Prognostic Performance

Essentially, AI’s primary strength in medical imaging lies in its ability to process large amount of data rapidly and to extract complex structures that may unobserved even by the expert in the field or clinicians. In the cancer imaging, the advancement of technology provides and translates improved detection of small lesions, better tumor segmentation, and most risk stratification. Research has demonstrated that use of AI can slower inter-reader variability and improve consistency across institutions [20], [21]. Moreover, AI-based prognostic models that integrate with medical imaging with clinical and genomic data are recently being developed to predict outcomes such as treatment response and survival rates [22], [23].

## AI Workflow Efficiency

AI also provides a consistent and potential to streamline radiology workflows automating image analysis tasks, they also prioritize urgent cases appropriately, and assisting with reporting. Radiology has greater influence AI to become a center of intelligently aggregated, a large amount of quantitative diagnostic information. This might assist to address radiologist shortages and reduce risks, especially in resource-limited setting [24]. In addition, AI-powered tools for automated triage and quality control systems are fundamental to check clinical adoption, further demonstration of AI advantages [25].

## Limitations and Challenges

### Heterogeneity and Bias

Despite the advances in AI, our review identified significant heterogeneity, patient populations, imaging protocols, imaging modalities, with diverse cancer types, and AI implementation methods. This variability influences and complicates direct comparison and the aggregation of results across meta-analyses. As many researches are retrospective, single center, and use non-standardized datasets, raising concerns about selection bias and overfitting [26], [27].

Furthermore, the lack of external validation remains the major problem in clinical studies. Though AI models may perform well on internal or public challenge datasets, their accuracy can decrease significantly when applied to external, real world datasets [28]. This is due to variations in image acquisition, population demographics, and disease occurrence or prevalence.

### Reporting standards and Reproducibility

Inconsistent reporting of model development, training, validation, and performance metrics has large influences on using AI-based models in the are of medical imaging studies. Initiatives such as TRIPOD-AI and CLAIM are good starting point to address this by promoting transparent, often reproducible reporting standards [29], [30]. However, with this large variability of adherences, and majority of publishes studies do not provide a sufficient detail to allow the independent replication of the assessment of clinical applicability.

### Explainability and Clinical implication

Another serious with many AI-based models is the “black box” nature, particularly deep learning models. Clinicians and regulatory bodies often need a better and well trained to having deep insights into the decision-making logic of AI systems before trusting them in clinical settings [31]. Work on interpretable AI, unstructured mapping, and uncertainty quantification is ongoing but not yet routinely incorporated into published studies.

### Regulatory and Ethical Considerations

Another essential part in AI application in clinical setting is contingent approval, ensuring data privacy protection and maintaining confidentiality, and adjusted and improved post-marketing surveillances. Prospective studies, multi-center trials are required maintain real-world efficiency and safety. Ethical consideration regarding data possession, prior consent of the patient, and algorithmic bias must be consistently addressed to ensure balanced benefits from AI advancements [32], [33].

## Discussion

The use of artificial intelligence (AI) in medical imaging is rapidly changing. The adoption of AI in medical imaging for cancer detection, diagnosis, classification, and treatment monitoring rep represents transformative shifts in modern healthcare. Our umbrella review, analyzed 80 systematic reviews and meta-analyses, provides a global perspective on the current state, strengths, challenges, and future directions of AI in medical imaging, specifically oncological research.

AI deployment in clinical practice is also contingent on regulatory approval, ensuring data privacy compliance, and robust post-marketing surveillance. Prospective, multi-center trials are needed to establish real-world effectiveness and safety. Ethical concerns regarding data ownership, patient consent, and algorithmic bias must be proactively addressed to ensure equitable benefits from AI advancements [19,20].

A relevant finding from our review is the overwhelming use of artificial intelligence (AI), particularly use of deep learning (DL) and machine learning (ML) techniques, together accounting for approximately 64% of all reported methods (Figure 2). This distribution shows recent findings, which emphasis the superior potential performances of DL and ML models in cancer detection, segmentation, and classification tasks across various imaging modalities [15], [34], [14]. In particular, DL indicates a remarkable advantage in identification of complex patterns from high-dimensional imaging data, that several studies reporting similar or exceeding diagnostic performance that of expert radiologists, especially in breast cancer and lung cancer screening [14], [15], [35]. Similarly, AI applications in lung cancer CT screening and prostate MRI interpretation have shown promising accuracy and reproducibility [15], [36].

Despite the dominance of DL and ML, the diversity of AI and hybrid techniques remains limited number of uses. As shown in the Figure 2, only a small proportion of studies indicated radiomics, DL and classical machine learning (ML) are mentioned. This indicate s while the field of AI is rapidly growing in technologically, there is still a gap in the evaluation of other potentially important AI techniques, including those that may provide better interpretability and also requires less data [31], [37].

With the growing developments of AI in oncologic imaging various recent cancer studies underscored the potential supports of AI. For instance, in breast cancer, the work by [2] demonstrated that AI system aided in the detection of breast cancer. large-scale randomized trials have shown that AI- assisted screening can increase cancer detection rates when radiologist used AI aid, with improved average area under the receiver operating curve (ROC) from 0.87 to 0.89 and increased sensitivity and specificity. Similarly, in prostate cancer (CsPCa) imaging which used pre-biopsy magnetic resonance (MRI) to detect specious prostate lesions, tools for AI have outperformed traditional radiological evaluations, achieving sensitivity 79% and Lesion Dice 38% compared to 33% by human radiologists [18].

In Pancreatic cancer study, AI has achieved notable progress in the detection and diagnosis of pancreatic cancer. Typically diagnosed at an advanced stage with a poor prognosis, pancreatic cancer has seen significant advancement in high diagnostic accuracy, significantly improving early detection rates and facilitating quicker initiation of treatment. According to [5], AI-based systems have shown remarkable diagnostic performance in identifying melanoma effectively. This breakthrough has greatly increased the precision and speed of skin cancer screening procedures as we as the diagnostic turnaround time.

Also, in the most recent AI role shown liver imaging has been equally promising. A meta-analyses of AI-based models for detecting hepatocellular carcinoma informed pooled estimates of sensitivity and specificity ranging between 84% and 92%, indicating that the potential of AI to improve early detection and reduce diagnostic errors in liver cancer [6]. Different studies of similar topics showed the same successes rate have been observed in diagnosing bone metastases and soft tissue sarcomas, especially in MRI scans where it is difficult for human interpretation [7], [8].

On the other hand, AI is increasingly used in predictive analytics. To estimate tumor grades, forecast treatment responses, and recurrence risks, algorithms have been developed. In rectal cancer study, AI models based on MRI have indicated sensitivity 82% and specificity 84%, respectively, to predict pathological complete response after neoadjuvant chemoradiotherapy [16]. On the other hand, in the study by [17], in non-small cell lung cancer (NSCLC) research, AI algorisms have been successful in predicting lymph node metastases, contributing important insights for pre-surgical planning and personalized therapy.

Moreover, there’s a new model out of Harvard called “Chief” that’s been making waves. Unlike older systems that focused on just one type of cancer, “Chief” can analyze several kinds at once. The results have been pretty impressive so far, with the model correctly diagnosing cases more than 89.5% an overall accuracy of the time. In addition, high prediction performance and generalizability to independent cohorts in tumor origin prediction. This suggests we might finally be moving past narrow, single-task AI tools toward systems that can handle a wider range of jobs in medical diagnostics [38].

Our review also identified important challenges, such as heavily reliance on meta-analyses, although beneficial for summarizing evidence. But, it introduces heterogeneity due to differences in study design, patient populations, imaging protocols, and references standards [12]. Data complexity difficulties encountered during synthesis further highlight the need for standardized datasets, clear and transparent reporting, and external validation, all of which are essential for the deployment of AI in clinical setting [13], [39]. Moreover, the unrepresentative non-AI systems and traditional statistical approaches (Figure 2) may limit opportunities for comparative evaluation and hybrid model deployment.

# Conclusion

To conclude, this umbrella review assessed the use of artificial intelligence in medical imaging based on data extracted from articles from systematic reviews and meta-analyses. Descriptive data analyses showed the large number of data extracted from meta-analyses and small proportion of systematic reviews, and a single comprehensive review.

Our results confirm, DL and ML dominate the landscape of AI-based cancer imaging research, supported by a robust foundation of meta-analytical evidence. While these advances offer significant promises for improved cancer diagnosis, detection, and classifications, with underlined workflow efficiency. Future studies must address current limitations through methodological rigor, standardized reporting, external validation, and a greater focuses on interpretability and ethics [40], [31]. Such efforts will be essential to realizing and implementation in the full clinical potential of AI in medical imaging and ensuring equitable and safe integration into healthcare practice.

# References

[1] “Cancer.” Accessed: May 16, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer

[2] A. Rodríguez-Ruiz *et al.*, “Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System,” *Radiology*, vol. 290, no. 2, pp. 305–314, Feb. 2019, doi: 10.1148/radiol.2018181371.

[3] L. A. R. Reisæter *et al.*, “Assessing Extraprostatic Extension with Multiparametric MRI of the Prostate: Mehralivand Extraprostatic Extension Grade or Extraprostatic Extension Likert Scale?,” *Radiol Imaging Cancer*, vol. 2, no. 1, p. e190071, Jan. 2020, doi: 10.1148/rycan.2019190071.

[4] X. Xu *et al.*, “Systematic review and meta-analysis: diagnostic accuracy of exosomes in pancreatic cancer,” *World Journal of Surgical Oncology*, vol. 23, no. 1, p. 51, Feb. 2025, doi: 10.1186/s12957-025-03666-9.

[5] H. A. Haenssle *et al.*, “Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists,” *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, Aug. 2018, doi: 10.1093/annonc/mdy166.

[6] K. Yasaka, H. Akai, O. Abe, and S. Kiryu, “Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study,” *Radiology*, vol. 286, no. 3, pp. 887–896, Mar. 2018, doi: 10.1148/radiol.2017170706.

[7] S. Han, J. S. Oh, and J. J. Lee, “Diagnostic performance of deep learning models for detecting bone metastasis on whole-body bone scan in prostate cancer,” *Eur J Nucl Med Mol Imaging*, vol. 49, no. 2, pp. 585–595, Jan. 2022, doi: 10.1007/s00259-021-05481-2.

[8] R. Xu *et al.*, “Deep learning-based artificial intelligence for assisting diagnosis, assessment and treatment in soft tissue sarcomas,” *Meta-Radiology*, vol. 2, no. 2, p. 100069, Jun. 2024, doi: 10.1016/j.metrad.2024.100069.

[9] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nat Med*, vol. 25, no. 1, pp. 24–29, Jan. 2019, doi: 10.1038/s41591-018-0316-z.

[10] “Cochrane Handbook for Systematic Reviews of Interventions.” Accessed: May 17, 2025. [Online]. Available: https://training.cochrane.org/handbook

[11] M. J. Page *et al.*, “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *PLoS Med*, vol. 18, no. 3, p. e1003583, Mar. 2021, doi: 10.1371/journal.pmed.1003583.

[12] M. Nagendran *et al.*, “Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies,” *BMJ*, vol. 368, p. m689, Mar. 2020, doi: 10.1136/bmj.m689.

[13] X. Liu *et al.*, “A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis,” *The Lancet Digital Health*, vol. 1, no. 6, pp. e271–e297, Oct. 2019, doi: 10.1016/S2589-7500(19)30123-2.

[14] S. M. McKinney *et al.*, “International evaluation of an AI system for breast cancer screening,” *Nature*, vol. 577, no. 7788, pp. 89–94, Jan. 2020, doi: 10.1038/s41586-019-1799-6.

[15] D. Ardila *et al.*, “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography,” *Nat Med*, vol. 25, no. 6, pp. 954–961, Jun. 2019, doi: 10.1038/s41591-019-0447-x.

[16] Z. Zhang *et al.*, “Improved deep learning for automatic localisation and segmentation of rectal cancer on T2-weighted MRI,” *Journal of Medical Radiation Sciences*, vol. 71, no. 4, pp. 509–518, 2024, doi: 10.1002/jmrs.794.

[17] J. Park *et al.*, “Development of a multi-modal learning-based lymph node metastasis prediction model for lung cancer,” *Clinical Imaging*, vol. 114, p. 110254, Oct. 2024, doi: 10.1016/j.clinimag.2024.110254.

[18] A. Holzinger, C. Biemann, C. S. Pattichis, and D. B. Kell, “What do we need to build explainable AI systems for the medical domain?,” Dec. 28, 2017, *arXiv*: arXiv:1712.09923. doi: 10.48550/arXiv.1712.09923.

[19] S. M. Lundberg *et al.*, “From local explanations to global understanding with explainable AI for trees,” *Nat Mach Intell*, vol. 2, no. 1, pp. 56–67, Jan. 2020, doi: 10.1038/s42256-019-0138-9.

[20] P. Rajpurkar *et al.*, “Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists,” *PLoS Med*, vol. 15, no. 11, p. e1002686, Nov. 2018, doi: 10.1371/journal.pmed.1002686.

[21] P. Tschandl *et al.*, “Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study,” *Lancet Oncol*, vol. 20, no. 7, pp. 938–947, Jul. 2019, doi: 10.1016/S1470-2045(19)30333-X.

[22] P. Lambin *et al.*, “Radiomics: Extracting more information from medical images using advanced feature analysis,” *European Journal of Cancer*, vol. 48, no. 4, pp. 441–446, Mar. 2012, doi: 10.1016/j.ejca.2011.11.036.

[23] P. Kickingereder *et al.*, “Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models,” *Radiology*, Jun. 2016, Accessed: May 23, 2025. [Online]. Available: https://pubs.rsna.org/doi/10.1148/radiol.2016160845

[24] F. Wang, L. P. Casalino, and D. Khullar, “Deep Learning in Medicine—Promise, Progress, and Challenges,” *JAMA Internal Medicine*, vol. 179, no. 3, pp. 293–294, Mar. 2019, doi: 10.1001/jamainternmed.2018.7117.

[25] K. J. Dreyer and J. R. Geis, “When Machines Think: Radiology’s Next Frontier,” *Radiology*, vol. 285, no. 3, pp. 713–718, Dec. 2017, doi: 10.1148/radiol.2017171183.

[26] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, “Key challenges for delivering clinical impact with artificial intelligence,” *BMC Med*, vol. 17, p. 195, Oct. 2019, doi: 10.1186/s12916-019-1426-2.

[27] M. Roberts *et al.*, “Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans,” *Nat Mach Intell*, vol. 3, no. 3, pp. 199–217, Mar. 2021, doi: 10.1038/s42256-021-00307-0.

[28] J. R. Zech, M. A. Badgeley, M. Liu, A. B. Costa, J. J. Titano, and E. K. Oermann, “Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study,” *PLOS Medicine*, vol. 15, no. 11, p. e1002683, Nov. 2018, doi: 10.1371/journal.pmed.1002683.

[29] G. S. Collins and K. G. M. Moons, “Reporting of artificial intelligence prediction models,” *The Lancet*, vol. 393, no. 10181, pp. 1577–1579, Apr. 2019, doi: 10.1016/S0140-6736(19)30037-6.

[30] J. Mongan, L. Moy, and C. E. Kahn, “Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers,” *Radiol Artif Intell*, vol. 2, no. 2, p. e200029, Mar. 2020, doi: 10.1148/ryai.2020200029.

[31] “Explainable AI: Interpreting, Explaining and Visualizing Deep Learning | SpringerLink.” Accessed: May 22, 2025. [Online]. Available: https://link.springer.com/book/10.1007/978-3-030-28954-6

[32] S. Gerke, T. Minssen, and G. Cohen, “Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare,” in *Artificial Intelligence in Healthcare*, A. Bohr and K. Memarzadeh, Eds., Academic Press, 2020, pp. 295–336. doi: 10.1016/B978-0-12-818438-7.00012-5.

[33] D. S. Char, N. H. Shah, and D. Magnus, “Implementing Machine Learning in Health Care — Addressing Ethical Challenges,” *N Engl J Med*, vol. 378, no. 11, pp. 981–983, Mar. 2018, doi: 10.1056/NEJMp1714229.

[34] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nat Med*, vol. 25, no. 1, pp. 24–29, Jan. 2019, doi: 10.1038/s41591-018-0316-z.

[35] H.-E. Kim *et al.*, “Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study,” *The Lancet Digital Health*, vol. 2, no. 3, pp. e138–e148, Mar. 2020, doi: 10.1016/S2589-7500(20)30003-0.

[36] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, “Artificial intelligence in radiology,” *Nat Rev Cancer*, vol. 18, no. 8, pp. 500–510, Aug. 2018, doi: 10.1038/s41568-018-0016-5.

[37] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nat Med*, vol. 25, no. 1, pp. 44–56, Jan. 2019, doi: 10.1038/s41591-018-0300-7.

[38] X. Wang *et al.*, “A pathology foundation model for cancer diagnosis and prognosis prediction,” *Nature*, vol. 634, no. 8035, pp. 970–978, Oct. 2024, doi: 10.1038/s41586-024-07894-z.

[39] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, “Key challenges for delivering clinical impact with artificial intelligence,” *BMC Med*, vol. 17, no. 1, p. 195, Oct. 2019, doi: 10.1186/s12916-019-1426-2.

[40] D. S. Char, N. H. Shah, and D. Magnus, “Implementing Machine Learning in Health Care - Addressing Ethical Challenges,” *N Engl J Med*, vol. 378, no. 11, pp. 981–983, Mar. 2018, doi: 10.1056/NEJMp1714229.