



Artificial intelligence in radiology: Enhancing diagnostic accuracy and early disease detection

Ananya Mehra

Department of Radiodiagnosis and Imaging, Institute of Medical Sciences and Research, Jaipur, Rajasthan, India

Abstract

Radiology is central to modern diagnosis, yet growing scan volumes and uneven access to specialists strain its capacity, particularly in India. This paper examines how artificial intelligence (AI), mainly deep learning, is being used to improve diagnostic accuracy and to detect disease earlier. Drawing on randomised, prospective and systematic evidence published largely between 2023 and 2025, it considers three areas: breast cancer screening, chest imaging for lung cancer, and tuberculosis screening. Large studies in Sweden and Germany indicate that AI-supported mammography reading can raise cancer detection while reducing radiologist workload and without inflating recall. Reviews of lung nodule detection report high sensitivity, although much of the evidence is retrospective. Indian implementation work suggests AI can add measurable tuberculosis yield and appears cost-effective, but deployment is hampered by infrastructure and workflow problems. Key barriers include automation bias, dataset shift, limited external validation and unclear accountability. The paper concludes that AI performs best as a supervised second reader within monitored workflows, and that prospective, locally validated trials should be the priority for responsible adoption.

Keywords: Artificial intelligence, radiology, deep learning, diagnostic accuracy, early detection, mammography screening, tuberculosis India

Introduction

Imaging now underpins decisions in oncology, infectious disease, cardiology and emergency care. As the number of examinations grows, radiologists face a widening gap between the images that must be read and the time available to read them. Prolonged, repetitive interpretation can erode accuracy, and subtle findings such as small lung nodules or early breast cancers are easily missed. In India the pressure is sharper because radiology services are concentrated in cities, while screening needs for conditions like tuberculosis and cancer are spread across the whole population.

Artificial intelligence, and in particular deep learning, has been proposed as a way to relieve this pressure. Algorithms trained on large image sets can flag suspicious regions, prioritise urgent studies and provide a consistent second opinion. Earlier research was dominated by retrospective tests on curated data, which said little about real clinical benefit. Since 2023, however, randomised and large prospective screening studies have begun to appear, allowing a more grounded assessment. This paper asks a focused question: to what extent does AI improve diagnostic accuracy and early disease detection in radiology, and what stands in the way of safe implementation, especially in the Indian context?

Review Approach

This is a narrative review rather than a formal systematic review. Literature was identified through targeted searches of publicly indexed sources using combinations of the terms artificial intelligence, deep learning, radiology, mammography, lung nodule, tuberculosis, automation bias and India. Priority was given to publications from 2023 to 2025, and to randomised trials, large prospective implementation studies and systematic reviews, since these carry more weight than small retrospective model comparisons. Studies were retained if they reported measurable outcomes such as cancer detection rate, recall,

sensitivity, workload or cost-effectiveness. Every finding below is paraphrased from the cited publications. Because the search was selective and not exhaustive, the conclusions should be read as an informed synthesis of key evidence and not as a complete census of the field.

Major Applications and Practices

Most clinical AI in radiology relies on convolutional neural networks and, increasingly, transformer-based architectures that learn image patterns directly from labelled examinations. Their outputs generally take one of three forms: detection of a lesion, classification of its likely nature, or a case-level risk score used to prioritise reading. Three application areas have produced the most useful recent evidence.

1. AI-supported breast cancer screening

Mammography screening is the clearest test case because it involves very large volumes, a low cancer prevalence and a well-defined outcome. The Mammography Screening with Artificial Intelligence (MASAI) trial in Sweden randomised 105,934 women to AI-supported screening or standard double reading. In this design the algorithm estimated the risk of each examination, so that low-risk cases received a single reading and high-risk cases received a double reading with the AI output displayed. Interim results published in 2023 suggested roughly a fifth more cancers were detected with AI support (Lång et al., 2023). The fuller 2025 report found that AI-supported screening detected 338 cancers versus 262 in the control arm, about 29% more, including a higher number of early-stage invasive cancers, while cutting the screen-reading workload of breast radiologists by around 44% (Hernström et al., 2025).

A complementary picture comes from Germany. The PRAIM study followed 463,094 women across 12 screening sites, read by 119 radiologists, and compared AI-supported double reading with standard double reading in routine

practice. The detection rate was 6.7 per 1,000 with AI against 5.7 per 1,000 without, a relative increase of 17.6%, while the recall rate was not worse (37.4 versus 38.3 per 1,000). The positive predictive value of recall and of biopsy was also higher in the AI group (Eisemann et al., 2025).

2. Lung nodule detection and chest CT

Chest CT is the main tool for finding lung nodules, but reading hundreds of thin slices is slow and tiring. A 2025 systematic review following PRISMA 2020 examined 1,448 articles published between 2015 and 2024 and reported that deep learning systems commonly achieved nodule detection rates above 95% with at most four false positives per scan, with very high classification accuracy in the best-performing models (Abdullahi et al., 2025). The same authors stressed that variability in data, limited generalisability and the difficulty of fitting tools into clinical routines remain obstacles. Such figures describe performance on research datasets and should not be assumed to carry over unchanged to a hospital with different scanners and patients.

3. Tuberculosis screening and the Indian experience

India carries a very large share of the global tuberculosis burden, and chest radiography is a practical triage tool, but trained readers are scarce in many districts. In a programme in Nagpur run by PATH, an AI system (qXR) was installed in eight private radiography laboratories, and 10,481 people with presumptive TB were screened. Of these, 2,303 were flagged by a radiologist or by the software, and about 15.8% of the overall yield was attributable to AI alone, meaning cases the radiologists had not considered presumptive (Vijayan et al., 2023). A rapid health technology assessment commissioned in India later evaluated two AI tools, qXR and Genki, against conventional interpretation using 2023 screening data. It reported good sensitivity and specificity, and incremental cost-effectiveness ratios for both tools that fell below India's per capita GDP for 2022 (Raval et al., 2025).

Evidence on Effectiveness

Taken together, the strongest evidence concerns screening mammography, where randomised and population-scale data converge. Both MASAI and PRAIM point towards higher detection without a penalty in recall, and MASAI adds a substantial reduction in reading workload, which matters in services that already lack radiologists. The consistency of direction between a randomised trial and a real-world implementation study increases confidence, although the two used different AI products and different workflows, so their numerical results are not interchangeable.

For lung imaging, the evidence is encouraging on technical accuracy but thinner on patient outcomes. Review-level findings show that algorithms can find and characterise nodules reliably under study conditions (Abdullahi et al., 2025), but they do not by themselves show that earlier detection translates into lower mortality. In tuberculosis, the Indian work is especially relevant because it was conducted in real service settings and linked AI use to additional detected cases and acceptable cost (Raval et al., 2025; Vijayan et al., 2023). A common thread is that benefits are largest when AI is used to support screening and triage, where volume is high and the target condition is well

defined, and weaker where cases are complex and require integrated clinical judgement.

Barriers and Challenges to Implementation

Human factors are an early concern. In an experiment with 27 radiologists reading 50 mammograms, wrong AI suggestions pulled readers towards incorrect assessments at every experience level. Accuracy was around 80% when the AI was right but fell to roughly 20% to 45% when it was wrong, and the least experienced readers were the most easily misled (Dratsch et al., 2023). This automation bias suggests that how AI output is displayed, and how readers are trained to challenge it, can matter as much as the algorithm.

Technical and data issues form a second group. Models tuned on one population or scanner can perform worse elsewhere, and the reviews of lung nodule detection explicitly point to data variability and limited generalisability (Abdullahi et al., 2025). The PRAIM investigators also noted that radiologists chose for themselves whether to use the AI, which makes it harder to rule out selection effects (Eisemann et al., 2025). Practical deployment adds further friction: the Nagpur programme encountered problems with software installation, workflow integration, network connectivity and limited staff available to act on the AI findings (Vijayan et al., 2023). In many Indian settings these infrastructure gaps are as limiting as algorithm quality.

Finally, governance remains unsettled. Questions about who is accountable when an AI-assisted reading is wrong, how patient data are protected, how tools are approved and monitored after release, and how training data represent the population being served are still being worked out. These issues are not unique to India, but regulatory and data-protection frameworks there are still developing, which raises the stakes for careful local validation.

Discussion and Implications

The reviewed evidence supports a cautious but optimistic position. AI in radiology is most convincing as an assistant that widens the net and saves reading time, not as a substitute for the radiologist. The MASAI and PRAIM findings suggest that well-integrated AI can raise detection while protecting specificity, and the Indian tuberculosis studies show that the same logic applies where specialist access is limited. At the same time, the automation-bias findings are a reminder that human oversight has to be designed deliberately and not assumed.

Several implications follow. First, hospitals and programmes should validate any tool on their own population before routine use and continue to monitor its performance after deployment, including recall, false positives and subgroup differences. Second, training for radiologists and technicians should address how to interpret AI prompts critically, particularly for early-career readers who appear most susceptible to over-reliance. Third, investment in connectivity, interoperable systems and workflow design is a precondition for benefit in resource-limited settings. Fourth, future research should move beyond detection rates to outcomes that matter to patients, such as stage at diagnosis, treatment delay and mortality, and should include Indian populations and public-sector facilities more fully. Policymakers can support this by funding prospective evaluations and by clarifying accountability rules.

Conclusion

Recent randomised and real-world studies show that artificial intelligence can improve the detection of breast cancer, assist in finding lung nodules and strengthen tuberculosis screening, often while easing the workload of radiologists. These gains are most reliable when AI acts as a supervised second reader within a monitored workflow. Automation bias, poor generalisability, infrastructure limits and unresolved governance questions continue to restrain safe adoption, particularly in India. The most useful next steps are locally validated, prospective studies that track patient outcomes, together with training and system design that keep the radiologist firmly in control of the final decision.

References

1. Abdullahi K, Ramakrishnan K, Ali A. Deep learning techniques for lung cancer diagnosis with computed tomography imaging: A systematic review for detection, segmentation, and classification. *Information*,2025;16(6):451.
<https://doi.org/10.3390/info16060451>
2. Dratsch T, Chen X, Rezazade Mehrizi M, Kloeckner R, Mähringer-Kunz A, Püsken M, et al. Automation bias in mammography: The impact of artificial intelligence BI-RADS suggestions on reader performance. *Radiology*,2023;307(4):e222176.
<https://doi.org/10.1148/radiol.222176>
3. Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nature Medicine*,2025;31:917–924.
<https://doi.org/10.1038/s41591-024-03408-6>
4. Hernström V, Josefsson V, Sartor H, et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI), a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. *The Lancet Digital Health*,2025;7(3):e175–e183.
5. Lång K, Josefsson V, Larsson A-M, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI), a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *The Lancet Oncology*,2023;24(8):936–944.
6. Raval, Parmar, Saha, Sarkar, Wadhwa, Pandya, et al. Cost-effectiveness analysis of AI-assisted chest X-ray interpretation tools for TB screening: A rapid HTA in India. *Frontiers in Digital Health*,2025.
<https://doi.org/10.3389/fdgth.2025.1629127>