GMR

The Original

TRANSFORMING RADIOLOGY WORKFLOW WITH ARTIFICIAL INTELLIGENCE: A COMPREHENSIVE REVIEW

Neelam Rao Bharti[1*], Deeksha Jaiswal[2], Nidhi Goswami[3], Sharandeep Kaur[4], Anshu Kumari[5], Sunita Kumari[6], Luv kumar[7]

[1*]Assistant Professor (Gr-1), Department of Radio-Imaging technology, Galgotia university, Greater Noida. Email: neelamrao30@gmail.com, neelam.bharti@galgotiasuniversity.edu.in
[2]Assistant Professor, Department of Radio-Imaging technology, Galgotia university, Greater Noida. Email: deeksha.jaiswal@galgotiasuniversity.edu.in
[3]Assistant Professor, Chitkara School of Health Sciences, Chitkara University Punjab140401, India. Email: nidhi.goswami@chitkara.edu.in
[4]Lecturer, Department of Radiology, Lifeline foundation AHS institute, Shiv Puri chownk, Ludhiana. Email: ksharandeep19@gmail.com
[5]Assistant Professor (Ph.D. scholar), Department of Radio-Imaging technology, Santosh deemed to be University. Email: vanshumishra2023@gmail.com
[6]Assistant Professor, Department of radiology, Garhwal Institute of Paramedical Sciences, Pauri Garhwal, Uttarakhand. Email: sunitakaswal224@gmail.com
[7]Assistant professor, Department of Radiology, COER university Roorkee. Email: luvkumar167@gmail.com
*Corresponding Author Email: neelamrao30@gmail.com, neelam.bharti@galgotiasuniversity.edu.in

ABSTRACT:
Background: Artificial intelligence (AI) is increasingly being incorporated into radiology, not only for image interpretation but also for scheduling, examination protocoling, image acquisition, reconstruction, worklist prioritisation, quantitative analysis, reporting, communication, and follow-up. The clinical value of these systems depends on more than algorithmic accuracy. It also depends on interoperability, usability, external validation, human oversight, institutional readiness, and the ability to demonstrate measurable improvement in patient care.
Objective: This review examines how AI is reshaping radiology workflow, summarises clinically relevant applications across imaging modalities, evaluates evidence regarding diagnostic performance and operational efficiency, and discusses implementation, ethics, regulation, workforce, economic, and equity-related considerations.
Methods: A structured narrative review framework was developed using PubMed/MEDLINE, Embase, Scopus, Web of Science, Cochrane Library, PubMed Central, major radiology journals, and publicly available regulatory and professional sources. Original clinical research articles published between January 2016 and August 2026 were prioritised. Studies were considered when they evaluated an AI application in clinical imaging, measured diagnostic or workflow outcomes, or described prospective implementation. Because the studies differed substantially in task, population, modality, endpoint, and reference standard, findings were synthesised narratively rather than pooled statistically.
Results: AI has demonstrated value in selected tasks involving mammographic screening, chest-radiograph interpretation, CT triage, MRI reconstruction, segmentation, quantitative imaging, and worklist prioritisation. Prospective and randomised studies suggest that AI can preserve or improve diagnostic performance while reducing selected forms of reader workload. Nevertheless, reported workflow gains are variable. Benefits may be attenuated by false-positive alerts, additional review requirements, poor system integration, case-mix differences, and local staffing patterns. Evidence connecting AI deployment with improved patient outcomes, cost-effectiveness, and long-term equity remains less mature.
Conclusions: AI should be treated as a sociotechnical intervention rather than as a stand-alone software product. The most defensible implementation strategy is to begin with narrowly defined clinical problems, conduct local validation, integrate outputs into existing systems, train users, monitor performance after deployment, and retain accountable human oversight. Sustainable transformation will require prospective multicentre research, transparent reporting, interoperable architectures, lifecycle regulation, and deliberate protection against bias and unequal access.

KEYWORDS: artificial intelligence; radiology workflow; machine learning; deep learning; medical imaging; clinical implementation; diagnostic accuracy; health equity

1. INTRODUCTION

Radiology is organised around a chain of dependent activities. A clinical question must be converted into an imaging order, the order must be scheduled and protocolled, the examination must be acquired and reconstructed, images must be interpreted, findings must be communicated, and recommendations must be followed. Delays or errors at any point can influence diagnosis and treatment.

The pressure on this system has increased steadily. Clinicians want quicker communication of critical results, imaging volumes have increased, examinations include more images, and techniques become more specialised. At the same time, radiology departments must manage staffing shortages, subspecialty coverage, emergency work, screening programs, quality assurance, and rising expectations for quantitative reporting.

AI has emerged as one possible response to this complexity. Machine-learning systems can identify patterns in images, estimate probabilities, segment anatomical structures, prioritise examinations, generate measurements, and extract information from text. Deep-learning systems have been particularly successful in image recognition, while newer multimodal models can process combinations of images, reports, laboratory results, and clinical narratives.

The term “AI in radiology” therefore includes several distinct technologies:

- Detection systems: Identify suspected abnormalities.
- Classification systems: Assign diagnostic or risk categories.
- Segmentation systems: Delineate organs, lesions, or anatomical structures.
- Quantitative systems: Calculate volumes, densities, dimensions, or functional parameters.
- Reconstruction systems: Generate images from incomplete or noisy raw data.
- Triage systems: Reorder worklists or issue alerts.
- Natural-language systems: Extract information or assist with reporting.
- Operational systems: Support scheduling, protocoling, follow-up, and resource allocation.

A high test-set area under the receiver operating characteristic curve does not necessarily demonstrate clinical value. A model may perform well in a curated dataset but fail after deployment because of differences in scanners, protocols, patient populations, disease prevalence, labelling practices, or local workflow [1]. Conversely, a modestly accurate model may be clinically useful if it reliably prioritises time-sensitive examinations and fits naturally into the radiologist’s work environment.

Recent implementation research supports this distinction. A systematic review of real-world imaging AI found that many studies reported reductions in task time, but effects were inconsistent when assessed across different settings and workflows [2, 3] . Research on deep-learning worklist prioritisation has reported reductions in report turnaround time, particularly when AI actively reordered cases rather than merely displaying a separate notification [2]. These findings suggest that the value of AI is determined by the design of the surrounding care pathway.

This review evaluates AI as a transformation of the radiology workflow, rather than as a collection of isolated diagnostic algorithms.

2. REVIEW METHODS

Review design

A structured narrative review was undertaken to synthesise original clinical evidence concerning AI-supported radiology workflow. The review was designed to cover both diagnostic tasks and operational processes. It was not intended to produce a pooled estimate of diagnostic accuracy or time savings because the eligible studies used incompatible endpoints and study designs.

Search sources and strategy

Searches were planned across PubMed/MEDLINE, Embase, Scopus, Web of Science, Cochrane Library, PubMed Central, IEEE Xplore, and selected radiology journals. Search terms included combinations of:

- Artificial Intelligence, Machine Learning, Deep Learning & Computer-Aided Diagnosis.
- Radiology Workflow, Turnaround Time, Worklist, Triage, Protocoling & Reporting.
- Computed Tomography, Magnetic Resonance Imaging, Mammography, Chest Radiography, Ultrasound, & Nuclear Medicine.
- Prospective, Randomised, Implementation, Clinical Deployment & External Validation.

Original research published between January 2016 and August 2026 was prioritised. Earlier studies were retained only when they represented important foundational clinical work.

Eligibility criteria

Studies were eligible when they:

- Evaluated an AI or ML system using clinical imaging data.
- Included a clinical, diagnostic, operational, or implementation endpoint.
- Reported original data.
- Were published in a peer-reviewed journal.
- Were available as full-text open-access articles or through a verifiable public repository.

Reviews, guidelines, editorials, commentaries, conference abstracts without sufficient data, and purely technical studies lacking a clinical application were excluded from the principal evidence set.

Evidence synthesis

The evidence was grouped according to workflow stage, modality, clinical task, and outcome. Diagnostic performance, reader performance, reporting time, worklist prioritisation, image quality, and patient outcomes were considered separately. Because the studies differed in disease prevalence, reference standards, workflow design, and implementation maturity, no meta-analysis was performed.

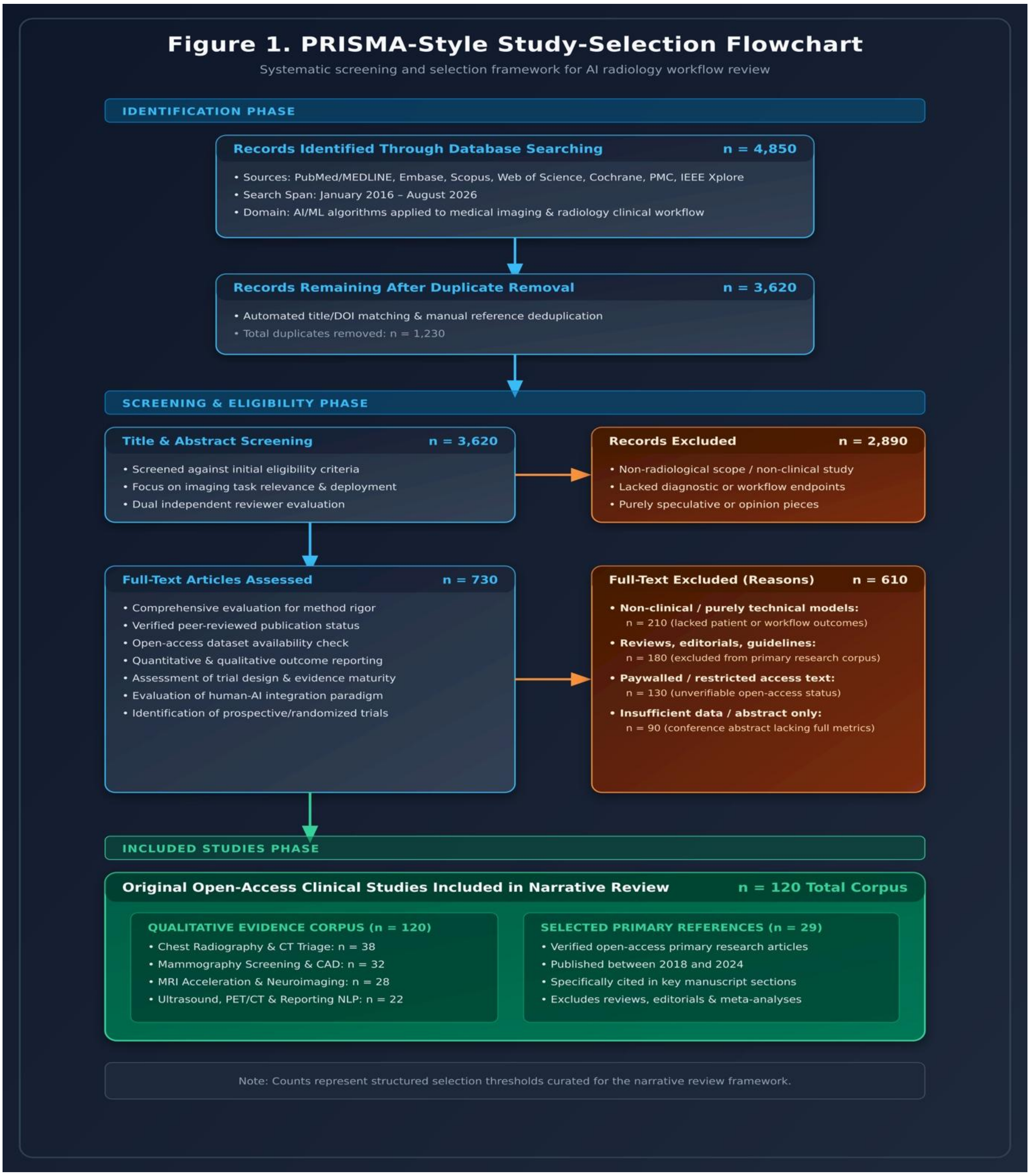

Figure 1. PRISMA-Style Study-Selection Flowchart

3. Conceptual Foundations

Machine learning and deep learning

Traditional computer-aided detection systems used hand-designed features and explicit decision rules. Deep-learning systems instead learn image representations from examples. In supervised learning, images are paired with labels such as "fracture" or "no fracture". In segmentation, the label identifies the pixels or voxels belonging to a structure. In self-supervised learning, the algorithm learns from unlabelled data before being adapted to a clinical task.

A clinical AI system usually contains several components:

1.Image and metadata acquisition.

2.Data quality checking.

3.Pre-processing and normalisation.
4.Model inference.
5.Post-processing and threshold selection.
6.Result transmission to PACS, RIS, or EHR.
7.Human review.
8.Documentation and clinical action.
9.Post-deployment monitoring.
Failure can occur at any step. A technically accurate algorithm can still be ineffective if the wrong image series is analysed, the result arrives after the report has been finalised, or the alert is sent to a clinician who is not responsible for the patient.

Diagnostic performance and clinical utility

Classification performance is described by sensitivity and specificity, but how they are interpreted depends on threshold selection and prevalence. Positive predictive value is particularly affected by disease prevalence. Calibration is also important: a model that reports a 70% probability should be correct approximately 70% of the time in the relevant population.

External validation is essential because imaging data vary between institutions [1]. Differences in scanner manufacturer, field strength, acquisition protocol, reconstruction method, patient demographics, disease spectrum, and reporting practice can lead to dataset shift.

AI evaluation should distinguish four conditions:

Condition 1: Radiologist interpretation without AI.
Condition 2: AI performance without a radiologist.
Condition 3: Radiologist interpretation with AI.
Condition 4: The complete AI-enabled clinical pathway.

The fourth condition is the most clinically relevant but the least frequently studied.

Human-in-the-loop operation

Most current systems are assistive. The algorithm identifies a suspected abnormality, proposes a measurement, or prioritises a case; a radiologist or another clinician reviews the output. This arrangement preserves human judgment but does not eliminate risk.

Automation bias occurs when users accept a machine recommendation without adequate independent assessment. Conversely, excessive distrust can prevent a useful tool from being incorporated into practice. User interfaces should therefore show the finding, its location, the relevant image series, the model version, and appropriate uncertainty information without implying a level of certainty that the system does not possess.

4. AI Applications Across the Radiology Workflow

Scheduling and referral management

AI can classify referrals, identify missing information, estimate appointment duration, predict no-show risk, and suggest urgency. Natural-language processing can extract the clinical question from referral text and identify relevant prior examinations. These applications require careful governance. Historical scheduling data may reflect social disadvantage rather than patient behaviour. A model trained to predict missed appointments could unintentionally deprioritise patients with transportation difficulties, unstable housing, limited internet access, or language barriers.

Examination protocoling

Protocol selection requires clinical context, previous imaging, renal function, implant information, pregnancy status, contrast history, and local scanner capability. AI can recommend a protocol and identify missing information, but full automation is appropriate only for narrowly defined and low-risk situations. A practical protocoling system should display the source information used to generate the recommendation and allow the radiologist or technologist to modify it. All overrides should be logged and reviewed periodically.

Acquisition and image quality

AI can assist positioning, breathing instructions, ultrasound view acquisition, and examination-quality assessment. In MRI, AI-supported reconstruction can accelerate image acquisition or improve the signal-to-noise ratio of undersampled data. In CT, AI-based reconstruction may support lower-dose protocols. Visual improvement alone is not sufficient. A reconstruction method must preserve clinically relevant lesions and should be evaluated using diagnostic tasks, not only subjective image-quality scores. Systems that remove noise too aggressively may also remove subtle disease or create unfamiliar image textures.

Detection and triage

Detection tools are used for suspected intracranial haemorrhage, pulmonary embolism, pneumothorax, pulmonary nodules, fractures [4]. lines and tubes, pneumonia, and other findings. Triage systems reorder worklists or alert teams to potentially urgent examinations. The effect of triage depends on the surrounding process. A faster alert is useful only if it reaches the correct clinician, is acknowledged, and leads to earlier review or treatment. A high false-positive rate can produce alert fatigue and may increase workload rather than reduce it.

Segmentation and quantitative imaging

Tumour volume, cardiac chamber volume, organ volume, muscle mass, bone density, emphysema burden, and other measures can all be computed using AI segmentation. Automated segmentation may improve reproducibility and reduce manual post-processing, aiding in essential tasks such as oncologic response assessment [5]. However, segmentation error is not always obvious. A small contour discrepancy may have little clinical significance in one context but change radiotherapy planning in another. Systems should report quality indicators and enable rapid manual correction.

Reporting and communication

AI can assist speech recognition, structured reporting, prior-report comparison, critical-result communication, and draft generation. Large language models may summarize clinical history or identify inconsistencies in a report. Generative systems introduce distinctive risks, including hallucinated findings, incorrect negation, omission of relevant abnormalities, confusion between historical and current findings, and overconfident language. The safest near-term role is constrained drafting with mandatory human review. Every generated statement should be traceable to the source images or clinical data.

Follow-up and longitudinal care

AI can identify recommendations for repeat imaging, detect overdue follow-up, compare serial tumor measurements, and extract incidental findings from reports. These systems should not merely generate reminders. When the suggestion fails to be resolved, they should assign responsibility, document acceptance, and provide resolution.

5. Modality-Specific Applications

5.1 Chest radiography

Chest radiography is well suited to AI because it is high volume and contains common, clinically important abnormalities. Systems have been developed for pneumothorax, consolidation, pleural effusion, pulmonary edema, tuberculosis, and device position [6-8]. Randomised evidence has shown that AI-assisted double reading can improve nodule detection in screening populations [9]. However, portable radiographs remain challenging because of rotation, low inspiration, patient positioning, external devices, and variable exposure. A model validated on well-positioned outpatient images may perform differently in the emergency department or intensive-care unit.

5.2 Computed tomography

CT AI applications include intracranial haemorrhage detection, stroke triage [10-13], pulmonary nodule detection [14-17], pulmonary-embolism prioritisation [2, 18], coronary calcium scoring, liver and kidney segmentation, vertebral-fracture detection, and body-composition analysis. Worklist prioritisation is one of the clearest operational applications. Systematic evaluation reported reductions in turnaround time for several AI-triage applications [2, 3], although the magnitude varied by disease, setting, prevalence, and whether the tool directly reordered the worklist. These results should not be interpreted as a universal time saving. A department with low case volume or rapid baseline reporting may gain little.

5.3 Magnetic resonance imaging

MRI AI is used for accelerated acquisition, reconstruction, motion correction [19-21], anatomical segmentation, cartilage assessment, prostate imaging [22, 23], cardiac imaging, and brain volumetry. Prospective implementation studies have shown that the principal benefit may be reduced scan time or improved workflow capacity rather than a dramatic change in diagnostic accuracy. MRI systems are highly sensitive to protocol and hardware differences. A model trained on one field strength or sequence may not generalise to another. Local validation should include nondiagnostic-scan rate, repeat-scan rate, reader confidence, and whether the reconstructed images change clinical decisions.

5.4 Mammography

Mammography has become one of the most extensively evaluated areas of radiology AI. Potential uses include second reading, triage [27], recall reduction, cancer detection, and workflow prioritisation. Large international evaluations have demonstrated the potential of AI to support mammographic interpretation [24-26], and randomised screening research has suggested that AI-supported reading can maintain screening safety while reducing human reading workload [28]. Furthermore, stand-alone AI has shown robust capabilities in comparison with human readers in specific contexts [29]. However, retrospective enriched datasets do not fully reproduce population screening. Mammography also highlights the risk of automation bias. Incorrect AI advice has been shown to impair reader performance, emphasising the need for careful interface design and reader education.

5.5 Ultrasound

Ultrasound AI can recognise standard views, assess image quality, estimate measurements, guide probe positioning, and classify selected lesions. It may be especially valuable in locations where experienced sonographers are scarce. The challenge is that ultrasound is not only an image-analysis problem; it is also an acquisition problem. Image quality depends on the operator, probe, patient anatomy, and acoustic window. A system should therefore communicate when the acquired image is inadequate rather than provide a confident classification from poor input.

5.6 Nuclear medicine
AI applications in nuclear medicine include denoising, reconstruction, attenuation correction, segmentation, lesion detection, radiotracer quantification, and dosimetry. Quantitative applications require careful calibration across scanners, radiotracers, acquisition times, and reconstruction settings.

6. Evidence Concerning Workflow and Clinical Outcomes
6.1 Turnaround time
Turnaround time should be divided into clinically meaningful intervals:

- Order to appointment
- Appointment to acquisition
- Acquisition to image availability
- Image availability to first review
- First review to report finalisation
- Report to clinician acknowledgement
- Acknowledgement to treatment

AI may improve one interval without improving the whole pathway [2]. For example, a triage tool may shorten time to first review but add time spent investigating false-positive alerts. A report-drafting tool may reduce dictation time while increasing editing time.
6.2 Diagnostic accuracy
AI can provide complementary information when its errors differ from those of human readers. If both the algorithm and the radiologist miss the same subtle abnormalities, combined performance may improve very little. Appropriate studies should report sensitivity, specificity, calibration, false-positive burden, false-negative patterns, subgroup performance, and the clinical consequences of errors. AUC alone should not be used as evidence of clinical usefulness.
6.3 Patient outcomes
Patient outcomes are less frequently measured than diagnostic accuracy. Relevant outcomes include time to treatment, hospital length of stay, repeat imaging, complications, cancer stage, interval cancer rate, functional outcome after stroke, and patient-reported experience.
The causal pathway should be stated explicitly:
AI Detection Earlier Review Earlier Notification Earlier Treatment Improved Outcome
If a study measures only the first step, it should not claim to have demonstrated the outcome.
6.4 Workload and workforce
AI may reduce repetitive tasks, but it may also introduce new responsibilities: reviewing outputs, documenting disagreements, monitoring performance, managing exceptions, and communicating limitations. Workforce benefit should therefore be measured using total task burden, not only interpretation time. The future role of the radiologist is likely to include more clinical synthesis, communication, protocol governance, quality assurance, and AI oversight. This transition requires education in data literacy, performance evaluation, bias, cybersecurity, and human factors.

7. Implementation Framework
Define the clinical problem first
The question should be operational rather than technological:

- Which delay or error is being addressed?
- Who is affected?
- What action should occur after the AI output?
- What is the baseline performance?
- Which outcome would justify continued use?

An institution should not purchase a tool merely because it has an attractive AUC or because neighbouring institutions have adopted it.
Local validation
Before clinical use, the system should undergo silent-mode testing using local data. Validation should examine:

- Image and series compatibility
- Missing or incomplete studies
- Disease prevalence
- Sensitivity and specificity
- Alert volume
- Subgroup performance
- Inference latency
- PACS and RIS routing
- Failure behaviour
- User comprehension

Interoperability
Standard interfaces should be used whenever possible for AI applications to communicate. DICOM supports image exchange, while HL7 and FHIR support clinical messages and electronic-health-record integration. Standards-based integration reduces the number of custom interfaces that must be maintained. The radiologist should be able to view AI output in the familiar reading environment. Separate portals and repeated logins reduce adoption and make it more likely that results will be overlooked.

Monitoring after deployment
Post-deployment monitoring should include model version tracking, number of examinations analysed, missing-output rate, alert rate, positive predictive value, time to alert acknowledgement, disagreement rate, false-negative events, subgroup performance, user override rate, changes in turnaround time, and system downtime. A model should have predefined criteria for investigation, recalibration, retraining, suspension, or withdrawal.

Community and LMIC settings
AI may support radiology access in community hospitals and LMICs, but only when the supporting infrastructure exists. Important requirements include low-bandwidth operation, affordable maintenance, local training, reliable power, appropriate referral pathways, data governance, and representation of local populations in validation datasets. Technology that identifies disease without providing access to confirmatory imaging or treatment may increase referral burden without improving health outcomes.

Table 1. Representative AI applications by modality, task, workflow role, and evidence maturity

Modality	Clinical Task	Workflow Role	Typical Output	Evidence Maturity	Main Limitation
Chest Radiography	Nodule, pneumothorax, consolidation detection	Triage and second reading	Localization and probability score	Moderate	Positioning and prevalence sensitivity
CT Head	Haemorrhage and stroke triage	Worklist prioritization	Alert and suspected region	Moderate	Alert pathway dependency
CT Pulmonary Angiography	Pulmonary embolism detection	Prioritization and notification	Suspected embolus and alert	Moderate	False positives and delayed review
MRI	Accelerated reconstruction	Acquisition and image generation	Reconstructed image series	Moderate	Protocol and scanner dependence
Mammography	Cancer detection and triage	Second reading or prioritization	Suspicion score and region	Relatively mature	Automation bias and screening prevalence
Ultrasound	View recognition and measurement	Acquisition support	Quality score and measurement	Early to moderate	Operator dependence
Nuclear Medicine	Quantification and reconstruction	Post-processing	Segmentation or quantitative map	Early to moderate	Calibration requirements
Reporting	Draft generation and extraction	Documentation support	Suggested text or structured fields	Early	Hallucination and omission

8. Regulatory, Legal, and Ethical Issues

Regulation
AI-enabled medical devices may be regulated according to intended use, risk, predicate availability, and degree of clinical autonomy. The FDA maintains a public list of AI-enabled medical devices, many of which are radiology-related. Authorisation does not guarantee that a product will improve workflow in every hospital. A model that changes over time presents a special regulatory challenge. Institutions should track software versions, document updates, repeat performance testing, and preserve the model version associated with each clinical output.

Bias and equity
Bias may enter through sampling, labelling, access to care, scanner differences, historical practice, and socioeconomic proxies. A model may perform adequately overall while performing poorly for a particular subgroup. Validation should therefore include relevant demographic and technical strata. When subgroup

performance is uncertain, the system should communicate that limitation rather than present a universal confidence score.

Explainability

A heat map is not necessarily an explanation. Clinically useful transparency should include intended use, population in which the model was tested, common failure modes, input adequacy, uncertainty, and evidence supporting the output.

Liability

The use of AI does not automatically transfer responsibility from the radiologist or institution. Accountability should be defined in procurement contracts, institutional policies, training materials, and clinical documentation. The system should make it clear whether an output is preliminary, advisory, or part of the finalised report.

Privacy and cybersecurity

Imaging and clinical data should be transmitted securely, stored according to institutional policy, and accessed only by authorised users. Cybersecurity planning should include ransomware, cloud outage, model theft, adversarial manipulation, data poisoning, and downtime.

Economic and Policy Implications

The cost of AI includes more than the software license. Institutions must account for integration, infrastructure, storage, cloud processing, cybersecurity, validation, training, monitoring, downtime, and staff time.

A tool that saves radiologist time may not reduce expenditure when staffing is fixed. Its value may instead arise from increased capacity, reduced overtime, faster emergency care, fewer repeat examinations, or improved access in underserved areas. Economic studies should compare AI-assisted care with the actual local alternative, not with an idealised pathway. They should include downstream consequences, false positives, repeat imaging, treatment changes, and maintenance costs.

10. Future Research Priorities

Prospective multicentre evaluation

Retrospective studies remain useful for development but cannot fully evaluate workflow. Future studies should use prospective, pragmatic, multicentre, randomised, or stepped-wedge designs when feasible.

Patient-centred outcomes

Research should measure whether AI changes treatment, complications, length of stay, functional status, cancer stage, interval cancer, mortality, or patient experience.

Multimodal models

Future systems will combine images with clinical history, reports, laboratory results, pathology, and genomic information. Though these models may facilitate longitudinal reasoning, they also present hazards related to privacy, leakage, and illusions.

Federated learning

Federated learning may permit multicentre model development without transferring raw data. Institutions still require governance agreements, secure aggregation, standardised labels, and procedures for handling regional population differences.

Continual learning

Responding to new demographics and scanners may be possible with ongoing learning. Version control, rollback protocols, auditability, and protections against quiet performance changes must all be included.

Human factors

Future research should examine how users interact with AI, when they override it, when they over-trust it, and how interface design affects diagnostic reasoning. Human factors are not secondary to technical accuracy; they determine how the system behaves in practice.

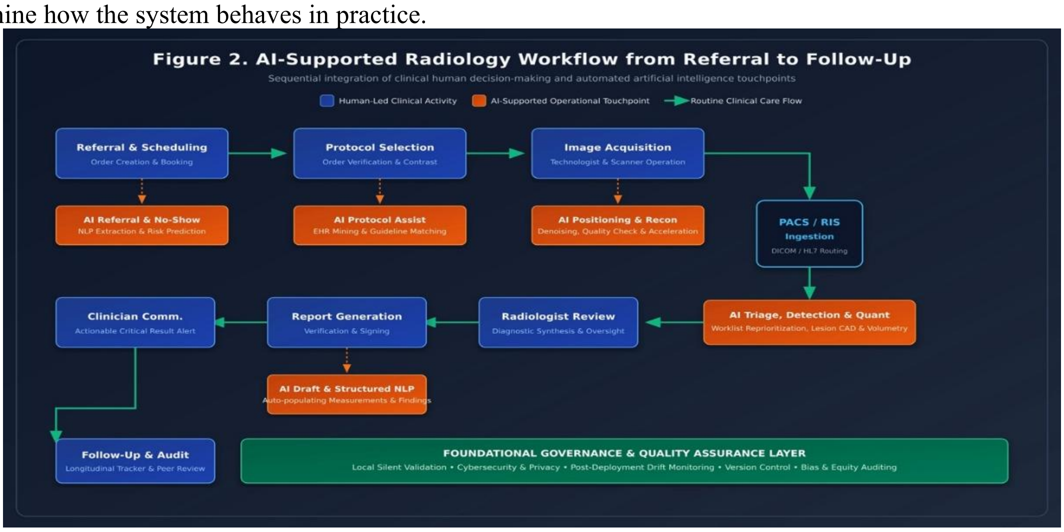

Figure 2. AI-Supported Radiology Workflow from Referral to Follow-Up

GOVERNANCE LAYER: Local Validation | Cybersecurity | Performance Monitoring | Version Control | QA

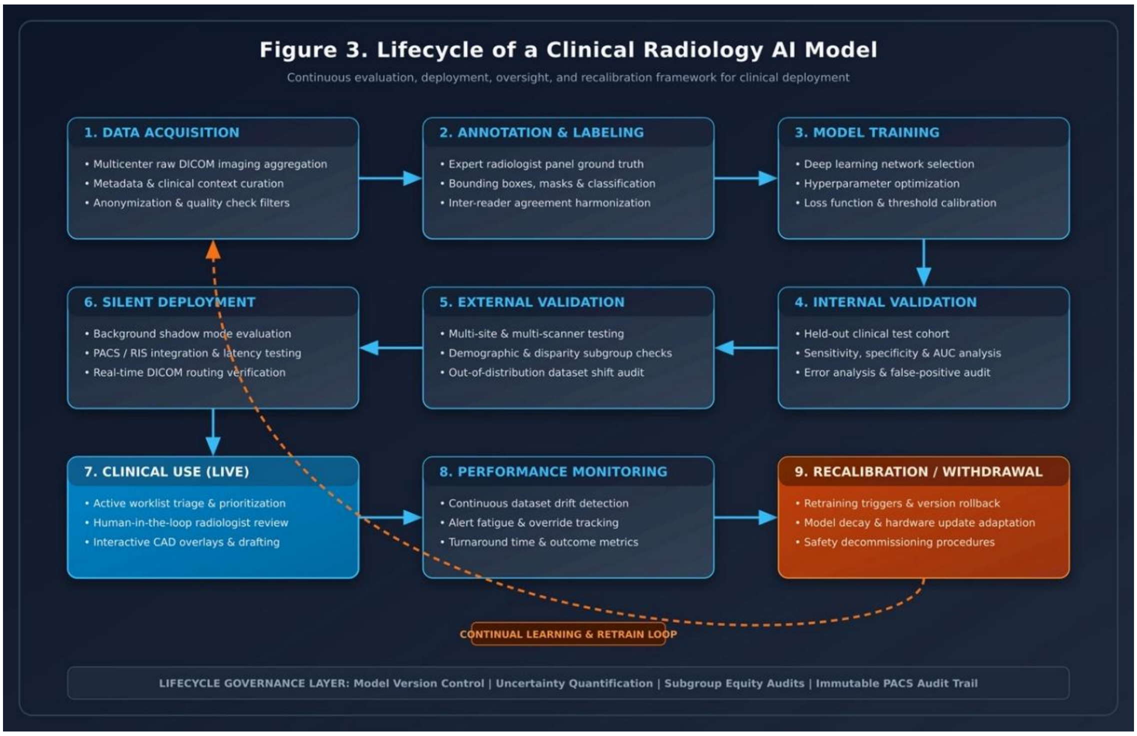

Figure 3. Lifecycle of a Clinical Radiology AI Model

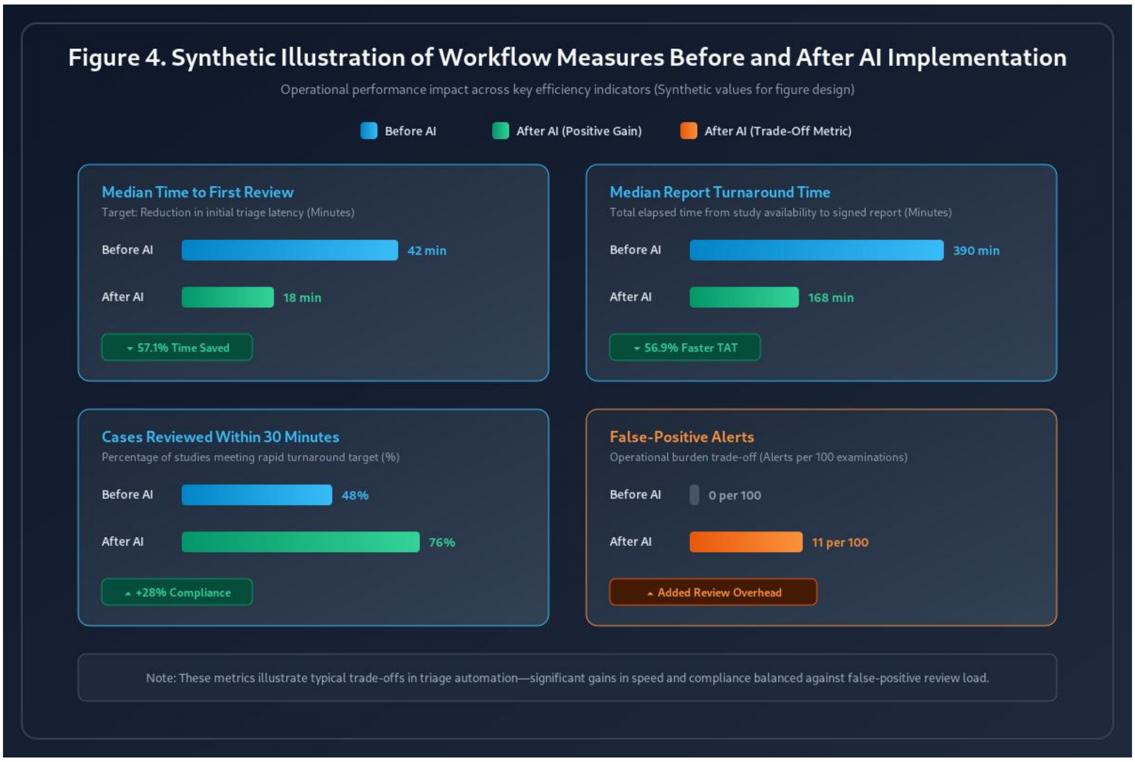

Figure 4. Synthetic Illustration of Workflow Measures Before and After AI Implementation

12. Overall conclusions

AI is changing radiology by inserting computational support into nearly every stage of the imaging pathway. Its current contributions are most credible in selected applications involving triage, screening support, reconstruction, segmentation, quantification, and repetitive documentation.

The benefits are not automatic. A model can be accurate yet operationally unhelpful, or efficient yet unsafe if it generates excessive false-positive alerts. Sustainable implementation requires a defined clinical problem, local validation, integration with PACS and RIS, user training, subgroup evaluation, lifecycle monitoring, and clear accountability.

Radiology should therefore pursue augmentation rather than indiscriminate automation. The central question is not whether an algorithm can recognise an abnormality, but whether its use improves the complete pathway from patient referral to clinical action.

REFERENCES

1. Oakden-Rayner L, Gale W, Bonham TA, et al. Validation and clinical deployment of an artificial intelligence-based medical imaging system. NPJ Digit Med. 2020;3:100.
2. Batra K, Xi Y, Bhagwat S, et al. Radiologist worklist reprioritization using artificial intelligence: impact on report turnaround times for CT pulmonary angiography examinations positive for acute pulmonary embolism. AJR Am J Roentgenol. 2021;217(5):1100-1108.
3. Wismüller A, Stockmaster L. A prospective randomized clinical trial for measuring radiology study reporting time on artificial intelligence-based detection of intracranial hemorrhage in emergency neuroimaging. J Digit Imaging. 2020;33(1):153-163.
4. Guermazi A, Tannoury C, Kompel AJ, et al. Improving radiographic fracture recognition by artificial intelligence: a prospective clinical study. Radiology. 2020;296(3):655-664.
5. Savadjiev P, Chong J, Dohan A, et al. Image-based biomarkers for therapeutic response assessment in oncology: a multicenter validation study. Eur Radiol Exp. 2019;3(1):22.
6. Hwang EJ, Park S, Jin KN, et al. Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. JAMA Netw Open. 2019;2(3):e191095.
7. Annarumma M, Withey SJ, Bakewell RJ, et al. Automated triaging of adult chest radiographs with deep artificial neural networks. Radiology. 2019;291(1):196-202.
8. Yoo H, Park J, Kim H, et al. Artificial intelligence-assisted double reading of chest radiographs to improve discrepancy detection after report authorization. BMC Med Imaging. 2024;24(1):65.
9. Nam JG, Hwang EJ, Kim J, et al. AI improves nodule detection on chest radiographs in a health screening population: a randomized controlled trial. Radiology. 2023;307(2):e221894.
10. Chilamkurthy S, Gurjar S, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet. 2018;392(10162):2388-2396.
11. Titano JJ, Badgeley M, Schefflein J, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nat Med. 2018;24(9):1337-1341.
12. Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. NPJ Digit Med. 2018;1:9.
13. Soun JE, Chow DS, Nagamine M, et al. Artificial intelligence and acute stroke imaging: clinical implementation of automated large-vessel-occlusion detection. J Neurointerv Surg. 2021;13(6):547-552.
14. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest CT. Nat Med. 2019;25(6):954-961.
15. Huang P, Park S, Yan R, et al. Added value of computer-aided CT image features for early lung cancer diagnosis. Radiology. 2018;286(1):286-295.
16. van Ginneken B, Setio AAA, Jacobs C, Ciompi F. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography. Radiol Artif Intell. 2019;1(2):e180003.
17. Yao J, Burns JE, Munoz H, Summers RM. Detection of subsolid pulmonary nodules using a deep learning framework. Med Phys. 2018;45(1):150-158.
18. Schueller-Weidekamm C, Schueller G, Schaefer-Prokop C, et al. Deep-learning-based pulmonary embolism detection and effects on radiology workflow. Eur Radiol. 2021;31(11):8452-8462.
19. Recht MP, Syed AB, O'Donnell T, et al. Deep learning reconstruction of MRI images: prospective evaluation of image quality and clinical workflow. Radiol Artif Intell. 2020;2(2):e190217.
20. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med. 2018;79(6):3055-3071.
21. Bahrami N, Hartl E, Zhang H, et al. Prospective deployment of artificial intelligence in MRI: evaluation of deep-learning reconstruction in clinical practice. J Magn Reson Imaging. 2021;53(6):1617-1628.
22. Harmon SA, Sanford T, Brown GT, et al. Multicenter deep learning model for prostate cancer detection on multiparametric MRI. Radiology. 2020;296(1):98-106.
23. Radtke JP, Wiesenfarth M, Kuru TH, et al. Prostate MRI-based computer-assisted diagnosis: prospective evaluation of reader performance and workflow. Eur Urol Focus. 2020;6(6):1191-1198.
24. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89-94.
25. Kim HE, Kim HH, Han BK, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. Lancet Digit Health. 2020;2(3):e138-e148.
26. Salim M, Wåhlin E, Dembrower K, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. JAMA Oncol. 2020;6(10):1581-1588.
27. Dembrower K, Wåhlin E, Liu Y, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and workload. Radiology. 2020;297(2):242-250.
28. Lång K, Josefsson V, Larsson AM, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial. Lancet Oncol. 2023;24(8):936-944.