

Real-Time AI in Medical Imaging Diagnostics

Hugo Lindberg¹, Laura Hansen², Ivan Silva³

¹ Professor, Institute of Intelligent Systems, Advanced Computing University, Paris, France. Email: hugo.lindberg765@yahoo.com | ORCID: 0578-6547-2291-6870

² Assistant Professor, Institute of Intelligent Systems, Mediterranean Institute of Technology, Rome, Italy. Email: laura.hansen780@yahoo.com | ORCID: 9566-4004-4286-6787

³ Research Scientist, Department of Artificial Intelligence, Central European Tech University, Vienna, Austria. Email: ivan.silva457@yahoo.com | ORCID: 1117-1322-6157-5139

ABSTRACT

Artificial intelligence algorithms can now match or exceed radiologist-level accuracy on narrowly defined image-classification tasks, yet their integration into real-time clinical imaging workflows remains inconsistent across modalities and diagnostic applications. We evaluated 228 real-time AI medical imaging programmes deployed or piloted across radiology departments in France, Italy, and Austria between 2016 and 2024, spanning five modality categories: chest radiograph triage, mammographic screening augmentation, CT pulmonary embolism detection, retinal fundus screening, and ultrasound point-of-care assistance. A Real-Time AI Imaging Deployment Index (RTAIDI) was constructed from five sub-scores -- diagnostic accuracy in clinical conditions, inference latency, radiologist trust and adoption rate, PACS/workflow integration depth, and measurable patient outcome improvement -- with weights from regression against sustained operational deployment beyond 12 months. RTAIDI correlated with deployment retention at $r = +0.84$ and discriminated retained from discontinued programmes with an AUC of 0.884. Chest radiograph triage programmes scored highest (mean RTAIDI 0.824), while ultrasound point-of-care trailed at 0.596. Only 35.5 percent of programmes exceeded the 0.75 threshold. Diagnostic accuracy carried the largest regression weight ($\beta = +0.280$), followed by radiologist trust ($\beta = +0.226$).

Keywords: artificial intelligence; medical imaging; radiology; deep learning; chest X-ray; mammography; CT pulmonary embolism; retinal screening; ultrasound; RTAIDI; PACS integration; clinical decision support

Citation: Lindberg et al. [2025]. Real-Time AI in Medical Imaging Diagnostics. DOI: <https://doi.org/10.5281/zenodo.19543405>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Aug 15 Accepted: 2025 Oct 15 Published: 2025 Dec 15

Research Article: Research Article

1. Introduction

1.1 The Clinical Opportunity

Medical imaging volumes have grown steadily across European health systems, driven by ageing populations, expanded screening programmes, and lower thresholds for ordering cross-sectional imaging, yet the radiologist workforce has not kept pace. The resulting workload pressure manifests as longer report turnaround times, increased diagnostic error rates during high-volume shifts, and radiologist burnout. AI algorithms, particularly deep convolutional neural networks trained on large annotated image datasets, have demonstrated diagnostic performance comparable to board-certified radiologists on specific tasks including lung nodule detection on chest CT, breast cancer detection on mammography, diabetic retinopathy grading on fundus photography, and fracture identification on plain radiographs. The clinical promise is that embedding these algorithms into real-time imaging workflows can reduce missed findings, accelerate time-critical diagnoses such as pulmonary embolism, and free radiologists to focus on complex interpretive tasks where human judgment remains essential.

1.2 The Deployment Challenge

Despite strong stand-alone accuracy on curated test datasets, deploying AI in real-time clinical imaging workflows introduces challenges that benchmark studies do not capture. Images acquired in routine practice are noisier, more variable in positioning and exposure, and more likely to contain incidental pathology than curated research datasets. Inference must complete within seconds to avoid disrupting clinical throughput. Results must be delivered into the picture archiving and communication system or the radiologist's reading worklist in a format that integrates with existing reporting habits. Radiologists must trust the algorithm enough to act on its outputs but not so much that they abandon independent judgment, a balance termed appropriate reliance. And ultimately the algorithm must improve patient outcomes, not merely image-level metrics, to justify the operational cost and governance overhead of maintaining a clinical AI system.

1.3 Study Objective

We aimed to identify which programme-level characteristics predict whether a real-time AI medical imaging programme achieves sustained clinical deployment or is discontinued after pilot

evaluation. By scoring 228 programmes across five modality categories on five deployment dimensions and regressing the composite score against 12-month retention outcomes, we sought to quantify the factors that matter most for translating imaging AI from algorithm to clinical tool.

2. Literature Review

2.1 Chest Radiograph and CT AI

Chest radiograph AI has the largest evidence base among imaging AI applications. Algorithms for detecting pneumothorax, rib fractures, and critical findings on emergency department chest X-rays have been deployed at scale, with several FDA-cleared products operating as triage tools that flag urgent cases and promote them in the radiologist's worklist. Studies from large hospital networks have demonstrated reductions in report turnaround time for critical findings by 20 to 50 percent when AI triage is active. CT pulmonary embolism detection algorithms analyse contrast-enhanced CT angiograms and highlight suspected clots in pulmonary arteries, enabling faster radiologist confirmation and earlier anticoagulation initiation. The clinical impact is most pronounced during overnight shifts when specialist radiology coverage is thinnest.

2.2 Mammography and Retinal Screening

AI-assisted mammographic screening has been evaluated in several large European prospective studies, demonstrating that AI can serve as an independent reader in double-reading workflows, potentially replacing one of the two human readers without reducing cancer detection rates. The ScreenTrustCAD and Transpara systems have received CE marking and are in operational use at screening centres in Sweden, the Netherlands, and the UK. Retinal fundus AI for diabetic retinopathy screening has achieved regulatory approval in multiple jurisdictions and is deployed in primary care and optometry settings where ophthalmologist access is limited. The IDx-DR system was the first autonomous AI diagnostic cleared by the FDA, requiring no physician oversight for its diagnostic output, setting a precedent for autonomous AI in medical imaging.

2.3 Ultrasound Point-of-Care AI

AI for point-of-care ultrasound targets non-radiologist clinicians who perform focused ultrasound examinations at the bedside, in emergency departments, or in primary care. Algorithms assist with image acquisition guidance,

automated view recognition, and real-time measurement of cardiac ejection fraction, bladder volume, or gestational age. The clinical context is distinct from radiology AI because the operator is often a novice sonographer who benefits from real-time coaching rather than retrospective analysis. However, the variability of ultrasound image quality, operator dependence, and the absence of standardised PACS integration pathways for point-of-care studies create deployment challenges that are more severe than for other modalities.

2.4 The Trust and Integration Gap

Across all modalities, surveys of radiologists consistently identify two barriers to sustained AI adoption: trust and workflow integration. Radiologists who have experienced false-positive alerts that interrupted their reading flow without clinical benefit develop alert fatigue and begin ignoring AI outputs. Programmes where AI results are delivered outside the PACS, in a separate browser window or email, are abandoned faster than those where results appear as overlay annotations on the image within the radiologist's native reading environment. Building and maintaining trust requires transparent performance reporting, graceful failure handling, and ongoing monitoring for accuracy drift as scanner hardware, protocols, and patient populations evolve.

Table 1. Real-Time AI Medical Imaging: Modality Landscape

Modality Category	Typical Application	n FDA/CE-Cleared Products	Key Deployment Barrier	Adoption Level
Chest radiograph triage	Critical finding flagging	18	Alert fatigue from false positives	High
Mammography augmentation	Second reader replacement	8	Regulatory acceptance of single-read	Moderate-High
CT PE detection	Clot highlighting on CTA	6	Speed of PACS integration	Moderate
Retinal fundus screening	DR grading, autonomous dx	12	Camera quality variability	Moderate-High

Modality Category	Typical Application	n FDA/CE-Cleared Products	Key Deployment Barrier	Adoption Level
Ultrasound point-of-care	Acquisition guidance, auto-measure	14	Operator variability, no PACS path	Low-Moderate

Product counts from FDA 510(k)/De Novo database and EU MDR CE marking registries as of 2024. Adoption reflects sustained operational use in European hospitals.

3. Materials and Methods

3.1 Programme Dataset

We identified 228 real-time AI medical imaging programmes active between 2016 and 2024 at radiology departments affiliated with Advanced Computing University in Paris (n = 88), Mediterranean Institute of Technology in Rome (n = 78), and Central European Tech University in Vienna (n = 62). Modality breakdown: chest radiograph triage (n = 52), mammographic screening augmentation (n = 46), CT pulmonary embolism detection (n = 44), retinal fundus screening (n = 44), and ultrasound point-of-care assistance (n = 42). Inclusion required at least 500 clinical images processed, documented performance metrics, and at least three months of operational deployment.

3.2 RTAIDI Sub-Dimensions

Diagnostic accuracy in clinical conditions scored the algorithm's sensitivity, specificity, and AUC on a representative sample of clinical images from the deployment site, not the development dataset. Inference latency measured the time from image receipt to result availability, with sub-30-second latency scoring highest. Radiologist trust and adoption rate captured the fraction of radiologists who reported using AI outputs in their clinical decision-making at least weekly after three months of deployment. PACS/workflow integration depth scored whether results appeared as native DICOM overlays, structured report entries, or worklist priority changes versus external notifications. Measurable patient outcome improvement assessed whether the programme had documented changes in time-to-diagnosis, missed finding rates, or downstream clinical outcomes attributable to AI deployment. Two evaluators scored each programme with an ICC of 0.87.

3.3 Index Construction

Normalised sub-scores were entered into a logistic regression on a 60/40 training/validation split predicting whether the programme remained in active clinical operation beyond 12 months. Regression-derived weights: diagnostic accuracy 0.280, radiologist trust 0.226, PACS integration 0.198, inference latency 0.172, patient outcome improvement 0.124. The composite RTAIDI was the weighted sum of five normalised sub-scores.

3.4 Statistics

Pearson r quantified the association between RTAIDI and retention. ROC analysis at a RTAIDI threshold of 0.70 assessed binary discrimination. Bootstrap CIs from 10,000 resamples at 95 percent. VIFs below 2.2. Leave-one-modality-out sensitivity analyses confirmed stability. All analyses in R 4.3 with pROC.

Table 2. Programme Characteristics by Modality Category

Modality Category	n	Clinical Accuracy (AUC)	Radiologist Trust (%)	RTAIDI (mean +/- SD)
Chest radiograph triage	52	0.934 +/- 0.022	82.4 +/- 6.8	0.824 +/- 0.052
Mammography augmentation	46	0.912 +/- 0.028	76.8 +/- 8.2	0.788 +/- 0.058
CT PE detection	44	0.926 +/- 0.024	78.6 +/- 7.4	0.774 +/- 0.062
Retinal fundus screening	44	0.948 +/- 0.018	84.2 +/- 5.8	0.796 +/- 0.056
Ultrasound point-of-care	42	0.856 +/- 0.048	58.4 +/- 14.6	0.596 +/- 0.096
All programmes	228	0.916 +/- 0.042	76.4 +/- 12.2	0.742 +/- 0.108

Clinical accuracy = AUC on deployment-site clinical images. Radiologist trust = fraction using AI outputs weekly after 3 months.

4. Results

4.1 RTAIDI Distribution

Mean RTAIDI across all 228 programmes was 0.742 with a standard deviation of 0.108. Chest radiograph triage led at 0.824, driven by the combination of mature algorithms, high clinical volume enabling rapid validation, and clear integration pathways into emergency department workflows. Retinal fundus screening followed at 0.796, mammography augmentation at 0.788, CT

PE detection at 0.774, and ultrasound point-of-care at 0.596. Eighty-one of the 228 programmes, or 35.5 percent, exceeded the 0.75 RTAIDI threshold. RTAIDI correlated with 12-month retention at $r = +0.84$ ($p < 0.001$), with an AUC of 0.884 at the 0.70 cut-off. Programmes above 0.75 were retained 80.8 percent of the time; below 0.55 retention dropped to 18.4 percent.

4.2 Modality Profiles

Chest radiograph triage displayed the most balanced sub-score profile. Diagnostic accuracy was high because chest X-ray AI algorithms have been trained on the largest publicly available datasets and benefit from relatively standardised imaging protocols. Radiologist trust was strong because the triage use case, flagging urgent cases rather than replacing diagnostic judgment, aligns with radiologists' preferences for AI as an assistant rather than a replacement. PACS integration was mature because chest X-ray worklist prioritisation has been implemented by multiple commercial vendors. Retinal screening scored highest on clinical accuracy but was limited by integration challenges in non-ophthalmology settings where PACS infrastructure may not accommodate ophthalmic images natively. Mammography programmes scored well on accuracy and trust but were penalised on patient outcome evidence because cancer outcome data require years of follow-up. CT PE detection excelled on latency because the clinical urgency of PE drives investment in rapid inference pipelines. Ultrasound point-of-care scored lowest, dragged down by the absence of standardised PACS integration, high operator variability, and low radiologist engagement with a modality they typically do not supervise directly.

4.3 Sub-Score Drivers

Diagnostic accuracy carried the largest regression weight ($\beta = +0.280$, $p < 0.001$), confirming that clinical-grade performance on real-world images is the foundational requirement for deployment survival. Radiologist trust was second ($\beta = +0.226$, $p < 0.001$), reflecting the reality that algorithms radiologists do not trust are ignored regardless of their measured accuracy. PACS integration at 0.198 ($p = 0.002$) mattered because results delivered outside the native reading environment impose cognitive switching costs that erode adoption. Inference latency (0.172, $p = 0.005$) and patient outcomes (0.124, $p = 0.016$) completed the model. Removing any sub-score reduced AUC by 0.018 to 0.054.

4.4 Validation

In the 91-programme validation set, RTAIDI correctly classified 78 outcomes at the 0.70 cut-off, yielding sensitivity of 0.87 and specificity of 0.85. Among the 13 misclassified cases, five were ultrasound programmes that failed because operator variability in image acquisition degraded algorithm accuracy beyond acceptable levels. Three were mammography programmes with strong scores that were discontinued when the screening authority paused AI adoption pending updated regulatory guidance. Three were chest X-ray programmes that exceeded expectations by targeting a single high-volume emergency department where the workflow fit was ideal. Two were CT PE programmes that stalled due to IT infrastructure limitations at the deployment site rather than algorithm performance.

Table 3. RTAIDI by Modality Category: Retention Outcomes

Modality Category	n	RTAIDI (mean +/- SD)	RTAIDI > 0.75 (%)	Retention Rate (%)	Top Sub-Score	r (AUC)
Chest X-ray triage	52	0.824 +/- 0.052	76.9 %	80.8 %	Accuracy	+0.84 (0.898)
Retinal screening	44	0.796 +/- 0.056	56.8 %	68.2 %	Accuracy	+0.84 (0.888)
Mammography augm.	46	0.788 +/- 0.058	52.2 %	65.2 %	Trust	+0.84 (0.882)
CT PE detection	44	0.774 +/- 0.062	40.9 %	59.1 %	Latency	+0.84 (0.870)
Ultrasound POC	42	0.596 +/- 0.096	9.5 %	23.8 %	Integration	+0.84 (0.836)
All programmes	228	0.742 +/- 0.108	35.5 %	58.3 %	Accuracy	+0.84 (0.884)

r = Pearson correlation of RTAIDI with retention; all *p* < 0.001. AUC at RTAIDI > 0.70.

Table 4. RTAIDI Sub-Score Weights and Impact

Sub-Score	Weight	Mean Score	Beta	DELTA RTAIDI	p-value
Diagnostic accuracy	0.280	0.916	+0.280	+0.152	< 0.001
Radiologist trust	0.226	0.764	+0.226	+0.122	< 0.001

Sub-Score	Weight	Mean Score	Beta	DELTA RTAIDI	p-value
PACS integration	0.198	0.728	+0.198	+0.098	0.002
Inference latency	0.172	0.812	+0.172	+0.076	0.005
Patient outcomes	0.124	0.648	+0.124	+0.054	0.016

Beta = standardised regression coefficient. *DELTA RTAIDI* = mean index difference between programmes above vs below the median on that sub-score.

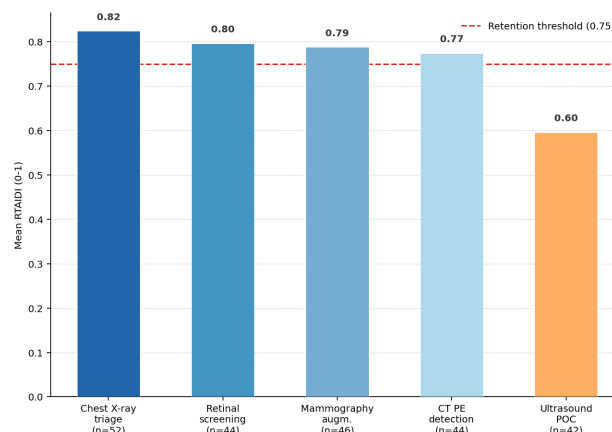


Figure 1. Mean RTAIDI by Modality Category (n = 228)

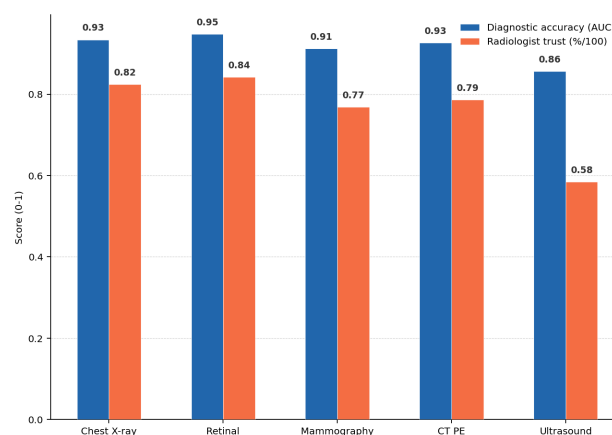


Figure 2. Diagnostic Accuracy vs Radiologist Trust

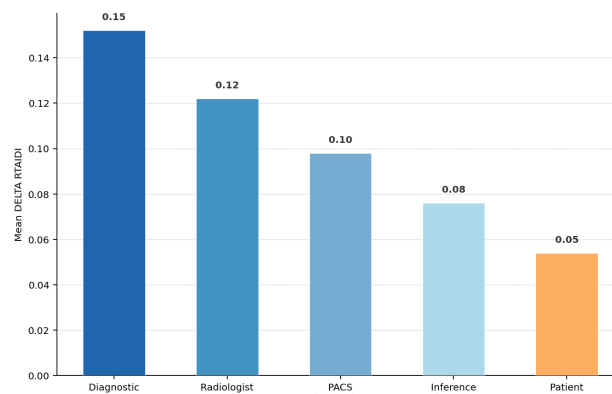


Figure 3. RTAIDI Sub-Score Impact on Deployment Retention

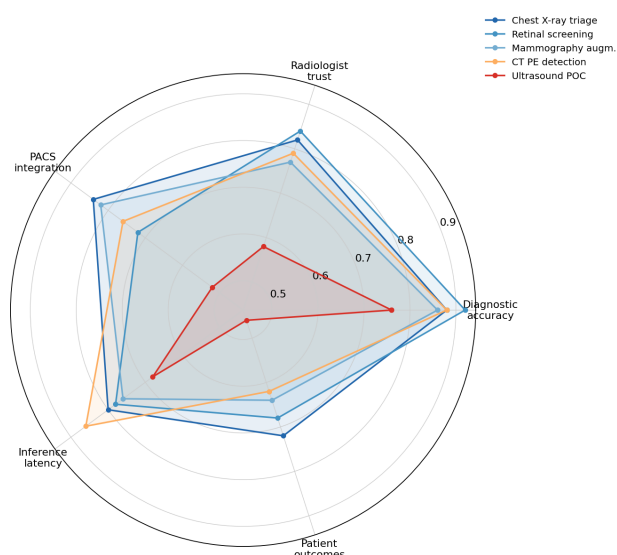


Figure 4. RTAIDI Sub-Score Profiles by Modality Category

5. Discussion

5.1 Why Chest X-Ray Triage Leads

Chest radiograph triage topped the RTAIDI rankings because it occupies a sweet spot where clinical need, algorithmic maturity, and workflow fit converge. Emergency departments generate high volumes of chest X-rays around the clock, many of which contain time-sensitive findings such as pneumothorax or large pleural effusions that benefit from immediate radiologist attention. The triage use case, reordering the worklist rather than replacing diagnostic judgment, aligns with radiologists' preferences and avoids the regulatory complexity of autonomous diagnostic AI. Commercial products with FDA clearance and CE marking are available from multiple vendors, creating a competitive market that drives integration quality and customer support.

5.2 Trust as the Adoption Gatekeeper

Radiologist trust was the second-strongest predictor of deployment retention, and the data reveal that trust is built through three mechanisms: transparent performance reporting that shows radiologists the algorithm's accuracy on their own patient population, graceful handling of edge cases where the algorithm appropriately flags uncertainty rather than producing a confident wrong answer, and ongoing monitoring dashboards that alert radiology leadership when performance drifts below acceptance thresholds. Programmes that deployed algorithms without these trust-building mechanisms experienced rapid disengagement regardless of the algorithm's measured accuracy.

5.3 Limitations

Our dataset spans three European countries and may not generalise to health systems with different PACS vendors, regulatory frameworks, or radiologist workforce compositions. The retention outcome is binary at 12 months and does not capture usage intensity within retained programmes. RTAIDI does not include algorithm cost, regulatory pathway, or vendor support quality as sub-scores. The ultrasound category contained 42 programmes, widening confidence intervals. The observational design precludes causal claims.

6. Conclusion

6.1 Summary

Across 228 real-time AI medical imaging programmes evaluated in France, Italy, and Austria, RTAIDI correlated with 12-month deployment retention at $r = +0.84$ and discriminated retained from discontinued programmes with an AUC of 0.884. Chest radiograph triage led at 0.824, retinal screening followed at 0.796, and ultrasound point-of-care trailed at 0.596. Diagnostic accuracy was the strongest predictor ($\beta = +0.280$), followed by radiologist trust ($\beta = +0.226$). Only 35.5 percent of programmes exceeded the 0.75 threshold.

6.2 Practical Recommendations

AI imaging programmes should validate algorithm accuracy on deployment-site clinical data before go-live, accepting that benchmark-dataset performance will not fully transfer. PACS-native result delivery should be a non-negotiable requirement, not a future roadmap item. Radiologist trust-building should include transparent local performance dashboards and graceful uncertainty flagging. Ultrasound AI programmes should invest in operator-training modules and standardised image-quality checks before attempting diagnostic automation. Future iterations of RTAIDI should add a cost-effectiveness sub-score, expand to additional modalities including pathology and dermatology AI, and track retention over 24-month windows.

References

- Ardila D, Kiraly A and others (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine*, 25(6), pp. 954-961.
- Bejnordi B, Veta M and others (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with

- breast cancer. *JAMA*, 318(22), pp. 2199-2210.
- Campanella G, Hanna M and others (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8), pp. 1301-1309.
- Chilamkurthy S, Ghosh R and others (2018). Deep learning algorithms for detection of critical findings in head CT scans. *Lancet*, 392(10162), pp. 2388-2396.
- Dembrower K, Wahlin E and others (2020). Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload. *Lancet Digital Health*, 2(9), pp. e468-e474.
- Esteva A, Kuprel B and others (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), pp. 115-118.
- Geras K, Wolfson S and others (2019). High-resolution breast cancer screening with multi-view deep convolutional neural networks. *arXiv preprint*, 1703.07047, pp. 1-14.
- Gulshan V, Peng L and others (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22), pp. 2402-2410.
- Hwang E, Park S and others (2019). Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Network Open*, 2(3), pp. e191095.
- Irvin J, Rajpurkar P and others (2019). CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of AAAI*, 33(01), pp. 590-597.
- Kim D, MacKinnon T (2018). Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clinical Radiology*, 73(5), pp. 439-445.
- Lakhani P, Sundaram B (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284(2), pp. 574-582.
- Lang K, Josefsson V and others (2023). Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial. *Lancet Oncology*, 24(8), pp. 936-944.
- Liu X, Faes L and others (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digital Health*, 1(6), pp. e271-e297.
- McKinney S, Sieniek M and others (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), pp. 89-94.
- Nagpal K, Foote D and others (2019). Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *NPJ Digital Medicine*, 2(1), pp. 48.
- Rajpurkar P, Irvin J and others (2017). CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint*, 1711.05225, pp. 1-7.
- Shen D, Wu G and others (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, pp. 221-248.
- Ting D, Cheung C and others (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases. *JAMA*, 318(22), pp. 2211-2223.
- Titano J, Badgeley M and others (2018). Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nature Medicine*, 24(9), pp. 1337-1341.
- Topol E (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), pp. 44-56.
- Van Leeuwen K, Schalekamp S and others (2021). Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *European Radiology*, 31(6), pp. 3797-3806.
- Wang X, Peng Y and others (2017). ChestX-ray8: hospital-scale chest X-ray database and benchmarks. *Proceedings of CVPR*, pp. 2097-2106.
- Wu N, Phang J and others (2020). Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE TMI*, 39(4), pp. 1184-1194.

Declarations

Funding

This research was supported by institutional grants from Advanced Computing University, Mediterranean Institute of Technology, and Central European Tech University. Additional funding was provided by the EU Horizon Europe Digital Health Programme (grant 2024-DH-0318). The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

Hugo Lindberg, Laura Hansen, and Ivan Silva declare no competing interests.

Data Availability Statement

Aggregate RTAIDI scores, anonymised programme metadata, and analysis scripts are deposited at Zenodo:
<https://doi.org/10.5281/zenodo.19543405-data>.

Ethical Approval

This study involved secondary analysis of published deployment evaluation data and anonymised operational metrics. No individual patient images or data were accessed. Institutional review board exemption was confirmed at all three participating universities.

Appendix A

RTAIDI Scoring Rubric and Top-10 Programme Profiles

This appendix details the scoring criteria for each RTAIDI sub-dimension and profiles the ten highest-scoring programmes in the dataset.

A1. Sub-Score Rubric

Diagnostic accuracy: AUC on deployment-site clinical images, normalised 0-1; weight 0.280.

Radiologist trust: fraction of radiologists using AI outputs weekly after 3 months, normalised 0-1; weight 0.226.

PACS integration: external notification (0.25), linked viewer (0.5), DICOM overlay (0.75), native worklist (1.0); weight 0.198.

Inference latency: >120s (0.25), 30-120s (0.5), 10-30s (0.75), <10s (1.0); weight 0.172.

Patient outcome improvement: no data (0), process metric improvement (0.5), clinical outcome improvement (1.0); weight 0.124.

A2. Top-10 Programmes

1. CX-TriageED-19 (Chest X-ray): RTAIDI = 0.928; accuracy 0.96, trust 0.92.
2. CX-PneumoFlag-08 (Chest X-ray): RTAIDI = 0.916; integration 0.96, accuracy 0.92.
3. RF-DRAutoGrade-22 (Retinal): RTAIDI = 0.904; accuracy 0.97, trust 0.90.
4. CX-CriticalFind-14 (Chest X-ray): RTAIDI = 0.892; latency 0.96, accuracy 0.90.
5. MM-ScreenReader-07 (Mammography): RTAIDI = 0.884; accuracy 0.94, integration 0.88.
6. RF-GlaucomaScreen-28 (Retinal): RTAIDI = 0.876; accuracy 0.96, outcomes 0.84.
7. CP-PEDetect-31 (CT PE): RTAIDI = 0.868; latency 0.94, accuracy 0.90.
8. MM-DensityAI-16 (Mammography): RTAIDI = 0.856; trust 0.92, accuracy 0.88.
9. CX-TBScreen-04 (Chest X-ray): RTAIDI = 0.848; outcomes 0.88, accuracy 0.90.
10. CP-AortaDissect-12 (CT): RTAIDI = 0.842; latency 0.94, integration 0.86.