

We Have Misunderstood What Radiologists Do

Authors:	<p>*Azeem Alam,¹ Eric Y. Auyoung²</p> <p>1. Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, Chelsea and Westminster Hospital, UK</p> <p>2. Princess Alexandra Hospital, Harlow, UK</p> <p>*Correspondence to azeem.alam14@imperial.ac.uk</p>
Conflict of interest:	<p>Auyoung is currently employed as a research scientist at King's College London and as a Clinical Safety Officer at BMS Digital Safety. Alam is a shareholder, Board Member, and Director of BiteWorld Ltd, and has received salary, consulting fees, and support for attending meetings and/or travel; has received consulting fees from Rad AI and Visiba Care, with payments made to BiteWorld Ltd, and consulting fees from Mercor; is an Investment Committee Member at Proximo Ventures and a Venture Scout at Grishin Robotics and EWOR; owns stock or stock options in BiteWorld Ltd; and is an Honorary Clinical Research Fellow at Imperial College London.</p>
Funding statement:	<p>The authors declare they received no funding for this study.</p>
Gen AI use:	<p>During the preparation of this work, the authors used ChatGPT (OpenAI, San Francisco, California, USA) and Claude (Anthropic, San Francisco, California, USA) solely to assist with spelling, grammar, and language editing. The authors reviewed and edited all suggested changes and take full responsibility for the content of the final publication.</p>
Peer review:	<p>This article was accepted following double-blind peer review.</p>
Received:	<p>03.07.26</p>
Accepted:	<p>25.09.26</p>
Keywords:	<p>AI, clinical pathways, clinical decision support, diagnostic imaging radiology, workflow integration.</p>
Citation:	<p>EMJ Radiol. 2026; https://doi.org/10.33590/emjradiol/2XNEWN8I</p>



Abstract

AI is now used in 70% of UK radiology departments, yet only 28% of clinical directors report productivity gains from its implementation. This apparent disconnect may reflect a focus on how algorithms perform in isolation, rather than what happens when they are introduced into clinical practice.

This feature examines evidence from mammography screening, stroke and lung cancer pathways alongside research on implementation, automation bias, deskilling, and accountability. It argues that the clinical value of AI depends on whether it changes decisions, uses clinical time effectively, and improves patient care without creating new risks or unowned work.

As AI becomes embedded in routine practice, developers and healthcare organisations will need to stop treating diagnostic accuracy as a proxy for clinical value and demonstrate how their tools improve care within clinical pathways. Radiologists, meanwhile, must become sceptical adopters who can appraise evidence critically, recognise where local conditions may limit performance and help govern how machine outputs influence patient care.

Key Points

1. AI is used in 70% of UK radiology departments, yet only 28% report productivity gains.
2. This article gives an evidence-informed perspective examining why radiology AI does not always translate into clinical benefit.
3. AI must be evaluated by its impact on clinical pathways and patient outcomes, not just diagnostic accuracy, and radiologists must become sceptical adopters who can evaluate and govern its use.

INTRODUCTION

An imaging finding does not end when the radiology report is verified; it has just started its clinical journey. The report itself informs subsequent patient care. It compares the present with the past, translates uncertainty into usable language, and helps another clinician decide what to do next. Much of the debate about AI in radiology has missed this point. It has treated the image, or the algorithm analysing it, as the natural unit of value. In practice, the relevant unit is the pathway in which a report becomes a decision.

That mistake is not merely academic. Radiology is already under severe strain. The most recent Royal College of Radiologists (RCR) census reported a 32% shortfall in consultant radiologists and more than one million diagnostic imaging studies not reported within the 28-day target.¹ This is despite AI tools now being used in 70% of radiology departments,² yet widespread adoption has seemingly not translated into widespread productivity gains. Only 28% of radiology clinical directors reported that AI implementation had increased productivity, while 61% reported no impact; by contrast, 57% reported productivity gains from process improvements.¹ The lesson is uncomfortable but necessary: a technically capable tool, inserted at the wrong point in a pathway, can redistribute work rather than reduce it.

This is why the familiar replacement narrative, crystallised by Geoffrey Hinton's, Professor Emeritus, University of Toronto, Canada, prediction that AI would outperform radiologists within 5 years, was always

too narrow. The question is not whether an algorithm can detect a haemorrhage, identify a nodule, or assign a malignancy score, though some can in defined settings. The more important question is whether that output reaches the right person, at the right time, in the right form, and changes the right decision without creating new risks. The usefulness of AI in radiology will not be determined at the level of the image; it will be at the level of the clinical pathway.

THE RADIOLOGY REPORT IS WHERE DETECTION BECOMES CARE

In reporting practice, identifying an abnormality is rarely the sole rate-limiting step. The clinical work is everything around that moment: understanding the question being asked, reviewing prior imaging, distinguishing new from established change, deciding how much uncertainty to express, and communicating findings in a way that alters management rather than merely recording an observation.

The same finding can mean different things in different pathways. A small pulmonary nodule, an equivocal subsegmental pulmonary embolus, or a subtle intracranial haemorrhage is not merely present or absent. Its significance depends on the patient, the indication, prior imaging, treatment options, and the clinician who must act on the report. Noticing the finding may be straightforward, but deciding what it means, and how to express it, is often the real work.

This is where accuracy-centred AI discourse becomes clinically thin. Sensitivity, specificity, positive predictive value, and area under the receiver operating characteristic curve are necessary, but insufficient. They do not tell us whether a tool reduces uncertainty, changes treatment, improves outcomes, is cost-effective, or works in a department with its own scanners, population, staffing, IT systems, and reporting pressures.

This concept should not be new to radiology. In 1991, Fryback and Thornbury³ published a now widely used framework that placed technical performance and diagnostic accuracy below diagnostic thinking, therapeutic impact, patient outcomes, and societal value. Recent radiology AI frameworks, however, have had to rediscover the same hierarchy. A tool may detect abnormalities reliably yet still fail where it matters locally, whether it changes decisions, outcomes, cost-effectiveness, or the practical functioning of a particular department.

THE BEST AI DOES NOT JUST SEE, IT MOVES THE PATIENT FORWARD

There are now credible examples of radiology AI producing clinical benefit. In Sweden, the MASAI randomised trial showed that AI-assisted mammography triage improved cancer detection while reducing reader workload within that protocol.⁴ In England, a study associated AI with faster treatment times and higher rates of mechanical thrombectomy.⁵ In Scotland, the NHS Grampian's Radiology Assisted Chest X-ray Evaluation project demonstrated a 9-day reduction in time from imaging to lung cancer treatment and increased detection of treatable cancers.⁶

These examples are not important because the algorithms produced impressive outputs in isolation. They are important because, in each case, the algorithm was not the destination: its output entered a pathway capable of turning information into action. They were able to reallocate scarce reader attention in screening; support acute stroke logistics so the right scan reaches the right

specialist sooner; and reprioritise the chest radiograph worklist so suspicious cases could enter a lung cancer pathway more rapidly. The tools that matter clinically are not merely better classifiers, they are pathway interventions.

THE ADOPTION GAP IS NOT A MYSTERY

Drawing on their experience as clinicians, procurement and clinical safety advisers, investors, and venture scouts, the authors have observed a gap between the metrics that win attention and those that determine successful deployment. Early financing, marketing, and purchasing discussions tend to focus on area under the receiver operating characteristic curve, benchmark sensitivity, and favourable comparisons with human readers under controlled conditions. These are useful proxies for algorithmic capability. However, they do not capture whether a tool integrates into picture archiving and communication systems, generates alert fatigue, performs on local hardware, requires extra clicks, creates unowned work, or leaves clinicians carrying accountability for systems they did not select or configure. These issues are not apparent until it's too late: after procurement.

The evidence base shows the same imbalance. A 2025 review of 173 Conformité Européenne (CE)-certified radiological AI products found that products with peer-reviewed evidence increased from 36% in 2020 to 66% in 2023.⁷ Yet the evidence base remained skewed towards diagnostic accuracy studies, with relatively little growth in higher-level, vendor-independent, multinational, or prospective evaluations.⁷ The field has become better at producing evidence, but much of that evidence still sits too close to the algorithm and too far from the pathway.

The Non-adoption, Abandonment, Scale-up, Spread, Sustainability (NASSS) framework by Greenhalgh et al.⁸ suggests that technologies fail not only because they are inaccurate, but because the clinical problem, users, organisations, regulations, and wider systems into which they are introduced are

often more complex than deployment plans assume. If an AI tool interrupts reporting rhythm, requires duplicate data entry, or generates alerts that nobody owns, busy departments will route around it. Such workarounds emerge naturally when clinicians seek to maintain patient care despite solutions that are misaligned with the realities of clinical practice.

BAD PATHWAY DESIGN TURNS ASSISTANCE INTO HARM

High-sensitivity systems may shorten time to treatment for true positives, but if thresholds generate large numbers of false positives, the pathway pays the price through alert fatigue, cognitive burden, and loss of trust. This is an operational problem. Every false positive creates additional work: another case for a radiologist to open, another notification for a clinician to process, and another decision about whether action is required.

Automation bias creates the opposite hazard. A prospective mammography experiment showed that incorrect suggestions significantly affected inexperienced, moderately experienced, and very experienced radiologists. This should end any assumption that AI assistance automatically improves judgement. Human performance can become tethered to machine output. If the model is wrong, out of distribution, poorly calibrated, or presented too authoritatively, the pathway may amplify error rather than catch it.

The design questions are therefore clinical, not cosmetic. Should alerts interrupt, queue, or annotate? Should uncertainty be visible? Should thresholds vary by pathway? How are disagreements audited? These are patient safety decisions.

WHAT DOES A FUTURE RADIOLOGIST NEED TO LEARN?

If AI changes the pathway, radiologists must help govern the pathway. The future radiologist should not be imagined as a faster reporter who also supervises

machines. The more important role is to decide when algorithmic output should change care: when to escalate, recommend further imaging, trigger a multidisciplinary team discussion, call a clinician directly, or reject an apparently confident machine output because it does not fit the patient.

This governance role carries more risks, and thus requires training. The 2025 EU-funded ACCEPT trial, a prospective study of colonoscopy performance before, during, and after a period of AI assistance, found that adenoma detection rates fell from 28.4% to 22.4% when AI support was withdrawn after 3 months of routine use.⁹ Colonoscopy is not radiology, but the warning is relevant: clinicians expected to govern sophisticated systems must remain capable of independent judgement when those systems fail, are unavailable, or are wrong.

Therefore, what the specialty needs is the sceptical adopter: a radiologist with deep clinical expertise and enough technical literacy to read validation studies critically, distinguish accuracy from utility, recognise when a model performs differently in local populations or clinical settings (dataset shift), challenge vendors, interpret local efficacy, and participate in governance. The specialty should ensure that the doctors accountable for imaging decisions can evaluate the systems increasingly shaping those decisions. If AI becomes part of routine reporting, AI literacy cannot remain extracurricular.

ACCOUNTABILITY FOLLOWS DESIGN

As radiologists take on broader roles in AI-enabled pathways, accountability frameworks will need to be developed. The reporting radiologist will usually remain professionally accountable for the final report, while vendor accountability is often indirect and distant from day-to-day clinical decision-making. That tension sharpens when radiologists are asked to use tools they did not select, validate, configure, or monitor.

A 2026 study found that when a radiologist reviewed AI output before overriding it and

patient harm followed, 74.7% of mock jurors sided with the plaintiff, compared with only 52.9% in a double-read design where the radiologist formed an independent view first.¹⁰ Although the legal context differs from European practice, the study highlights an important principle: the sequence in which AI enters clinical judgement may affect how responsibility is perceived.

Thus, how we design the technical architecture is vital to successful AI implementation. Whether alerts are interruptive or passive, whether outputs are auditable, and whether there is named clinical ownership for monitoring performance are all pathway design choices. They shape both clinical safety and perceived responsibility.

CONCLUSION

The last decade did not simply show that radiologists are harder to replace than predicted. It showed that the debate had chosen the wrong unit of analysis. An algorithm may identify an abnormality, but clinical value only emerges if that information changes what happens next.

A technically impressive tool can still fail if it creates false-positive work, interrupts reporting, lacks local validation, shifts accountability without authority, or produces outputs that do not change management. Conversely, a narrower tool can be valuable if it accelerates the right step in a stroke, cancer, or screening pathway.

The future radiologist therefore needs to be more than a user of AI. The specialty needs sceptical adopters who can evaluate evidence beyond accuracy, recognise when local conditions invalidate general claims, challenge vendors, monitor drift, and govern the interface between machine output and clinical consequence.

There are reasons for optimism about what AI can offer radiology, but less optimism about the pace at which radiologists and health systems are being prepared to govern these pathways. The question for radiology is no longer whether AI can read an image; it is whether radiologists and health systems can govern the pathway in which that reading becomes care.

References

1. The Royal College of Radiologists. Clinical radiology workforce census 2025. 2025. Available at: <https://www.rcr.ac.uk/news-policy/workforce-censuses/>. Last accessed: 2 July 2026.
2. The Royal College of Radiologists. The future of AI in healthcare: public perceptions of AI in radiology. 2025. Available at: <https://www.rcr.ac.uk/media/poelyzlr/rcr-reports-the-future-of-ai-in-healthcare-public-perceptions-of-ai-in-radiology.pdf>. Last accessed: 3 July 2026.
3. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making*. 1991;11(2):88-94.
4. Lång K 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. *Lancet Oncol*. 2023;24(8):936-44.
5. Nagaratnam K et al. Artificial intelligence imaging decision support for acute stroke treatment in England: a prospective observational study. *Lancet Digit Health*. 2025;7(12):100927.
6. Scottish Health Technologies Group. Artificial intelligence supported clinician review of chest X-rays from patients with suspected lung cancer. 2025. Available at: <https://shtg.scot/our-advice/artificial-intelligence-supported-clinician-review-of-chest-x-rays-from-patients-with-suspected-lung-cancer/>. Last accessed: 2 July 2026.
7. Antonissen N et al. Artificial intelligence in radiology: 173 commercially available products and their scientific evidence. *Eur Radiol*. 2026;36(1):526-36.
8. Greenhalgh T et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res*. 2017;19(11):e367.
9. Budzyń K et al. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *Lancet Gastroenterol Hepatol*. 2025;10(10):896-903.
10. Bernstein MH et al. The radiologist-AI workflow and the risk of medical malpractice claims. *Nat Health*. 2026;1(4):386-9.