AI for Pulmonary Embolism Detection: A Qualitative Study Exploring Real-World Implementation

Karol Basta (karol.basta1@nhs.net),[1] Anil Mistry,[1] Beatrix Fletcher,[1] Stephan Edey,[2] Carolyn Horst,[2,3] Simon Lambracos,[2] Sze Mun Mak,[2] Charlotte Marriott,[2] Adam Mayers,[2] Tina Mistry,[2] Andrew Refalo,[2] Nishanth Sivarasan,[2] Conn McGrath,[1] Dijana Vilic,[1] Audrey Jacques,[2] Vicky Goh.[2,3]

1. Clinical Scientific Computing, Guy's and St Thomas' NHS Foundation Trust 2. Department of Radiology, Guy's and St Thomas' NHS Foundation Trust 3. Department of Cancer Imaging, School of Biomedical Engineering & Imaging Sciences, King's College London

Background & Aim

Background: Diagnostic AI tools are complex interventions requiring evaluation beyond narrow performance metrics. The WHO,[1] the National Institute for Health and Care Excellence[2] and the Royal College of Radiologists[3] have called for "whole systems" evaluations, including impact on clinical workflows, the human-AI interaction and unintended consequences.

However, most evaluations remain focused on technical performance in controlled settings.[4] Pulmonary embolism (PE) detection AI tools exemplify this gap. Despite growing clinical adoption, no studies have examined implementation or process outcomes, including workflow integration, impact on radiologist practice or organisational adoption.

Aim: This qualitative real-world evaluation explored radiologists' experiences with two commercially developed AI tools for PE detection (Fig.1) using a systems perspective. We aimed to understand how the tools are used in practice, their acceptability, unintended consequences and individual and organisational influences on the human-AI interaction.

Methods

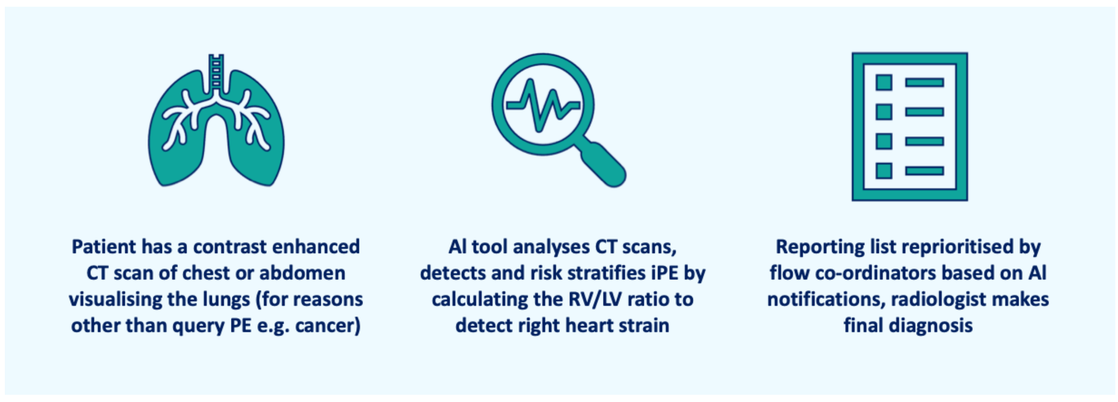
Two AI tools (Fig.1) developed by Aidoc for PE detection were deployed across radiology workstations at a large NHS Trust between March–July 2025.

Purposive sampling recruited radiologists with varied seniority, subspecialty and reporting responsibilities to capture diverse perspectives until thematic saturation was reached.

Semi-structured interviews were conducted with 11 radiologists. Interviews covered understanding and use of the tools, diagnostic and triage value, workflow effects, training implications, trust, unintended consequences and cognitive biases.

Interviews were recorded, transcribed, anonymised and analysed in NVivo 15 using a hybrid thematic approach, combining deductive framework analysis with inductive coding.

1. Incidental Pulmonary Embolism (iPE) Tool



2. Pulmonary Embolism (PE) Tool

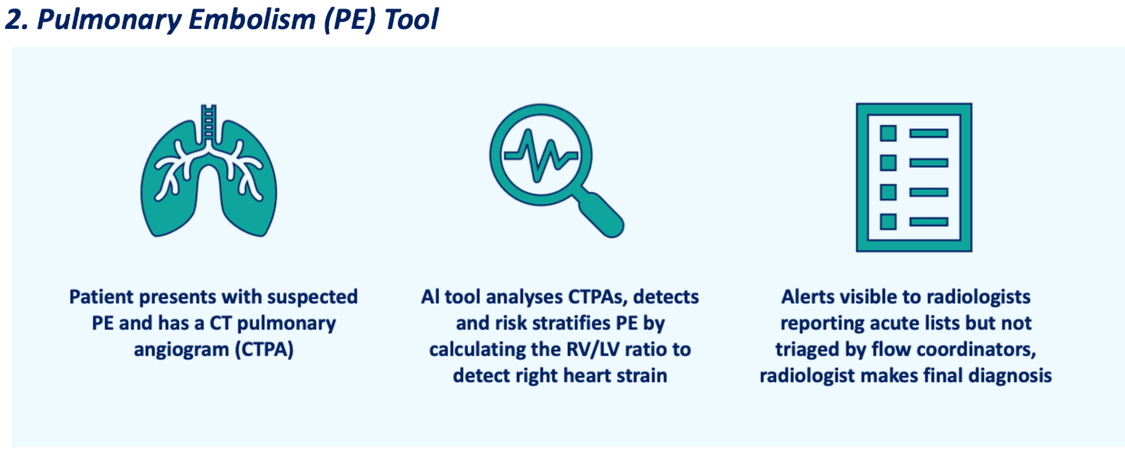


Figure 1: AI tools implemented. RV/LV = right ventricle to left ventricle ratio.

Results & Recommendations

Findings are presented through three questions with themes outlined in Figure 2. Illustrative quotes are shown in Box 1.

Are the AI tools acceptable to radiologists?

Radiologists framed acceptability around clinical utility, opportunity cost, workflow fit and early bottom-up involvement. The iPE triage tool was viewed as useful for backlog management, while the CTPA PE tool offered less value, as acute PEs were perceived to be already well managed. Embedding outputs within existing electronic systems, aligning alerts with escalation pathways and clarifying how single-pathology notifications interact with broader prioritisation processes emerged as key for uptake.

What influences the human–AI interaction?

The order in which prompts appeared, whether scans were reviewed independently and whether AI findings were recorded in reports influenced how critically outputs were interrogated. Confidence in disagreeing with the AI was linked to experience and case volume, raising concern that, as CTPAs are often first reported by juniors, prompts may disproportionately shape early-stage decision-making. Participants described a spectrum of "overcallers" and "undercallers", with radiologists unconsciously aligning with AI outputs that reinforced existing biases. Behaviours were also shaped by a wider culture of low risk tolerance and legal scrutiny, making positive AI findings harder to challenge.

Are unintended consequences emerging?

Radiologists worried AI alerts for a single pathology could divert attention from other unflagged urgent findings. Perceived AI calibration favouring high sensitivity raised fears of overcalling, over-investigation and treatment with anticoagulants. Some felt AI sped up reporting, whilst others thought reviewing AI outputs increased reporting time. Some described cognitive strain, second-guessing and reduced confidence in reporting due to AI. Over time radiologists worried this could alter diagnostic patterns and training. Independent first reads before overlays, retrospective AI use for teaching and structured discussion of biases were suggested safeguards. Unintended consequences were perceived rather than proven and warrant empirical study, alongside testing of proposed mitigations.

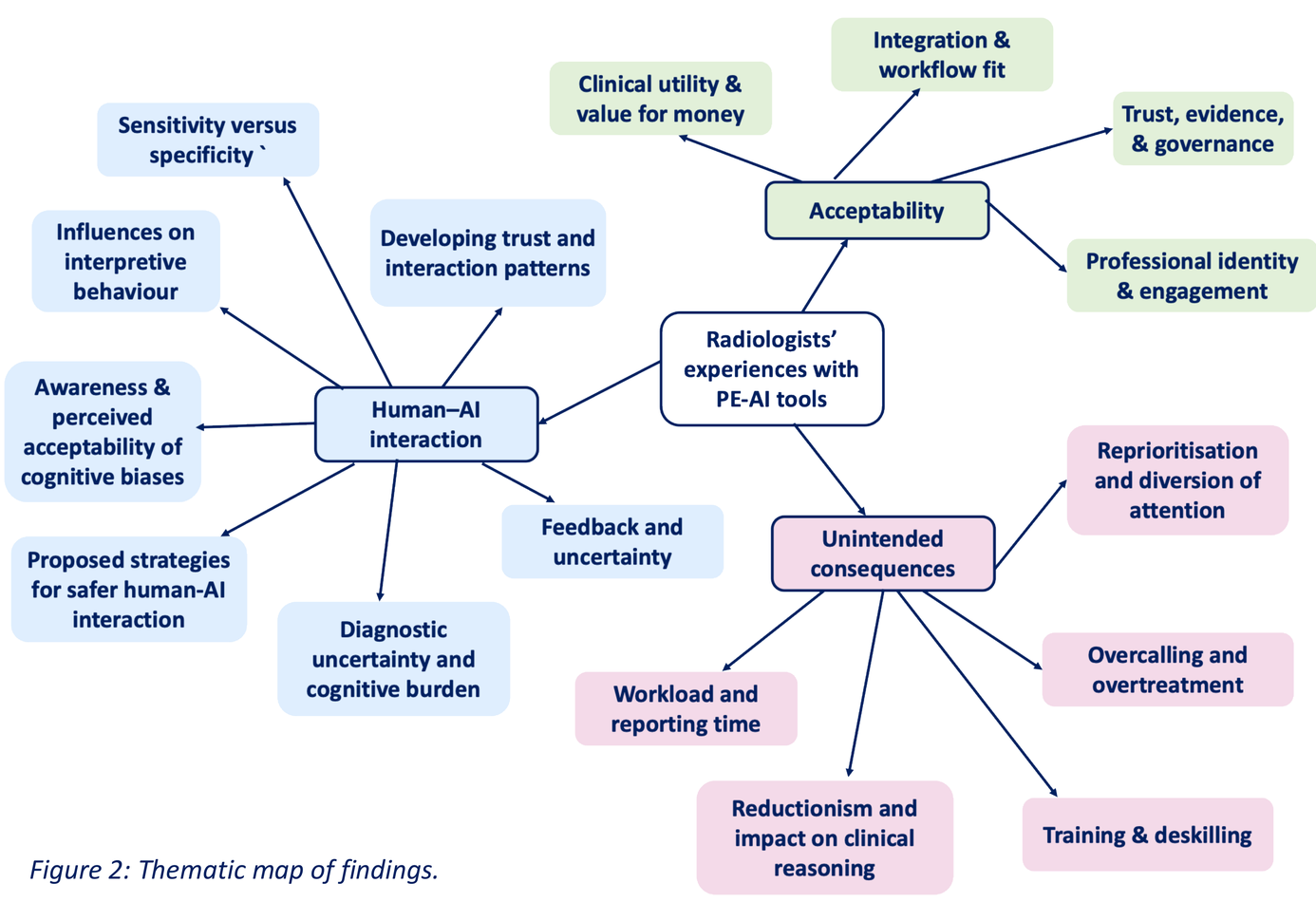


Figure 2: Thematic map of findings.

"Is reducing reporting time from our current baseline actually going to have clinical benefit? Even if it has benefit, is it the best use of money? It has to compete with the cost of having everything reported in faster time e.g. employing more radiologists or outsourcing scan reporting, versus an application that highlights a single risk, which then still has to be properly reviewed and reported anyway." (Consultant 2)

"AI is not purely mathematical. You have human beings afraid of legal consequences, and that fear can make us choose the wrong decision" (Consultant 3).

"I'm worried my reports are going to become inaccurate because I might change what I say because the tool has said it's positive for PE, even though I'd initially thought no PE." (Resident 1)

"If I look at a follow-up CT, it'll take me two seconds, but now there's an additional set of information it might just make it longer." (Consultant 5).

Box 1: Illustrative participant quotes.

Conclusion

Successful AI adoption depends on demonstrable clinical and economic value, seamless integration and transparent governance. Monitoring and potential mitigation of the human-AI interaction and unintended consequences is essential for safe clinical adoption.

References

1. Generating evidence for artificial intelligence-based medical devices: a framework for training, validation and evaluation. World Health Organization. 2021. Licence: CC BY-NC-SA 3.0 IGO.
2. National Institute for Health and Care Excellence (NICE). Evidence standards framework for digital health technologies [ECD7]. 2022.
3. AI deployment fundamentals for medical imaging. The Royal College of Radiologists. 2024.
4. Lawrence R et al. Artificial intelligence for diagnostics in radiology practice: a rapid systematic scoping review. *EClinicalMedicine*. 2025 May 12;83:103228. doi: 10.1016/j.eclinm.2025.103228.

Guy's and St Thomas' NHS Foundation Trust

Fellowship in Clinical AI
NHS Digital Academy Programme